Numerical Methods for Deterministic and Stochastic Phase Field Models of Phase Transition and Related Geometric Flows

A Dissertation Presented for the

Doctor of Philosophy

Degree

The University of Tennessee, Knoxville

Yukun Li

August 2015

© by Yukun Li, 2015
All Rights Reserved.

To my beloved parents Yangyu Li and Dinglan Hu, for their continuous understanding

Acknowledgements

I would like to express my deepest appreciation to my advisor Professor Xiaobing Feng for his careful guidance through my whole graduate studies. I could not go this far without his guidance and encouragement. I am very grateful to Professors Ohannes Karakashian, Vasilios Alexiades, Steven Wise, Yulong Xing and Vasileios Maroulas for their teachings and for being always kind and patient to answer my questions. I would also like to thank Professers Ohannes Karakashian, Vasileios Maroulas and Xiaopeng Zhao for serving on my doctoral committee. Moreover, I would like to thank my collaborators, Professors Andreas Prohl, Yulong Xing, Zhihao Ge, and Dr. Yi Zhang, for their teachings and stimulating discussions. Furthermore, I would like to thank Professor Balram Rajput for his valuable suggestions on my teaching. My gratitude extends to my fellow graduate students and friends Zhen Guan, Yi Zhang, Liguo Wang and Wenqiang Feng for their friendship, help and for sharing knowledge with me. Finally, I would like to acknowledge the support of the NSF grants DMS-1016173 and DMS-1318486 for my research from August of 2013 to May of 2015.

"Learning without thought is labor lost; thought without learning is perilous."— Confucius

Abstract

This dissertation consists of three integral parts with each part focusing on numerical approximations of several partial differential equations (PDEs). The goals of each part are to design, to analyze and to implement continuous or discontinuous Galerkin finite element methods for the underlying PDE problem.

Part One studies discontinuous Galerkin (DG) approximations of two phase field models, namely, the Allen-Cahn and Cahn-Hilliard equations, and their related curvature-driven geometric problems, namely, the mean curvature flow and the Hele-Shaw flow. We derive two discrete spectrum estimates, which play an important role in proving the sharper error estimates which only depend on a negative power of the singular perturbation parameter ϵ [epsilon] instead of an exponential power. It is also proved that the zero level sets of the numerical solutions of the Allen-Cahn equation and the Cahn-Hilliard equation approximate the mean curvature flow and the Hele-Shaw flow respectively. Numerical experiments are carried out to verify the theoretical results and to compare the zero level sets of the numerical solutions and the geometric motions.

Part Two focuses on finite element approximations of stochastic geometric PDEs including the phase field formulation of a stochastic mean curvature flow and the level set formulation of the stochastic mean curvature flow. Both formulations give PDEs with gradient-type multiplicative noises. We establish PDE energy laws and the Hölder [Holder] continuity in time for the exact solutions. Moreover, optimal error

estimates are derived, and various numerical experiments are carried out to study the interplay of the geometric evolution and gradient-type noises.

Part Three studies finite element methods for a quasi-static model of poroelasticity, which is a fluid-solid interaction multiphysics system at pore scale. We reformulate the original multiphysics system into a new system which explicitly reveals the diffusion process and has a built-in mechanism to overcome the "locking phenomenon". Fully discrete finite element methods are proposed for approximating the new system. We derive a discrete energy law and optimal error estimates for our finite element methods. Numerical experiments are also provided to verify the theoretical results and to confirm that the "locking phenomenon" has indeed been overcome.

Table of Contents

1 Introduction **1**

1.1 Background 1

1.1.1 The moving interface problem 1

1.1.2 The level set method and the phase field method 2

1.1.3 The mean curvature flow and the Hele-Shaw flow 5

1.1.4 The Allen-Cahn equation and the Cahn-Hilliard equation . . . 7

1.1.5 Stochastic mean curvature flow and stochastic Hele-Shaw flow 10

1.2 Scope of the dissertation research 11

1.3 Summary of main contributions 13

2 Discontinuous Galerkin Methods for the Allen-Cahn Equation **16**

2.1 Introduction 16

2.2 Preliminaries 18

2.3 Fully discrete interior penalty discontinuous Galerkin approximations 28

2.3.1 Formulations 28

2.3.2 Discrete energy laws and well-posedness 30

2.3.3 Discrete discontinuous Galerkin spectrum estimate 35

2.3.4 Polynomial order in ϵ^{-1} error estimates 40

2.4 Convergence of the numerical interface to the mean curvature flow . . 49

2.5 Numerical experiments 53

3 Discontinuous Galerkin Methods for the Cahn-Hilliard Equation 61

3.1 Introduction . . . 61

3.2 Preliminaries . . . 63

3.3 Fully discrete mixed interior penalty discontinuous Galerkin approximations . . . 71

3.3.1 Formulations of the mixed interior penalty discontinuous Galerkin method . . . 72

3.3.2 Discrete energy law and well-posedness . . . 75

3.3.3 Discrete spectrum estimate on the discontinuous Galerkin space 78

3.3.4 Error analysis . . . 85

3.4 Convergence of numerical interfaces . . . 99

3.5 Numerical experiments . . . 103

4 Finite Element Methods for the Stochastic Mean Curvature Flow of Planer Curves of Graphs 109

4.1 Introduction . . . 109

4.2 Preliminaries and error estimates for a partial differential equation regularization . . . 112

4.3 Finite element methods . . . 116

4.3.1 Semi-discretization in space . . . 116

4.3.2 Full discretization in space and in time . . . 124

4.4 Numerical experiments . . . 130

4.4.1 Verifying the rate of convergence of time discretization . . . 131

4.4.2 Dynamics of the stochastic mean curvature flow . . . 131

4.4.3 Verifying energy dissipation . . . 134

4.4.4 Thresholding for colored noise . . . 135

4.4.5 Thresholding for white noise . . . 138

5 Finite Element Methods for the Stochastic Allen-Cahn Equation with Gradient Multiplicative Noises 142

5.1 Introduction . . . 142
5.2 Preliminary results . . . 147
5.3 Finite element methods . . . 158
5.4 Numerical results . . . 169

6 Multiphysics Finite Element Methods for a Quasi-static Poroelasticity Model 181

6.1 Introduction . . . 181
6.2 Partial differential equation model and its analysis . . . 183
6.2.1 Preliminaries . . . 183
6.2.2 Partial differential equation model and its multiphysics reformulation . . . 185
6.2.3 Analysis of the partial differential equation model . . . 188
6.3 Fully discrete finite element methods . . . 199
6.3.1 Basic time-stepping algorithm . . . 199
6.3.2 Fully discrete finite element methods . . . 200
6.3.3 Stability analysis of fully discrete finite element methods . . . 203
6.3.4 Convergence analysis . . . 207
6.4 Numerical experiments . . . 214

7 Future Works 223

Bibliography 225

Vita 236

List of Tables

2.1 Spatial errors and convergence rates of Test 1. 54

2.2 Spatial errors and convergence rates of Test 2. 57

2.3 Spatial errors and convergence rates of Test 3. 60

3.1 Spatial errors and convergence rates of Test 1 with $\epsilon = 0.1$. 104

3.2 Spatial errors and convergence rates of Test 2 with $\epsilon = 0.1$. 105

3.3 Spatial errors and convergence rates of Test 3 with $\epsilon = 0.1$. 108

4.1 Computed time discretization errors and convergence rates. 131

5.1 Computed time discretization errors and convergence rates. 171

6.1 Spatial errors and convergence rates of Test 1. 216

List of Figures

2.1 Decay of the numerical energy $J_\epsilon^h(u_h^\ell)$ of Test 1. 54

2.2 Test 1: Snapshots of the zero-level set of $u^{\epsilon,h,k}$ at time $t = 0, 2 \times 10^{-2}, 3.2 \times 10^{-2}, 4 \times 10^{-2}$ and $\epsilon = 0.125, 0.025, 0.005, 0.001$. 55

2.3 Decay of the numerical energy $J_\epsilon^h(u_h^\ell)$ of Test 2. 57

2.4 Test 2: Snapshots of the zero-level set of $u^{\epsilon,h,k}$ at time $t = 0, 0.06, 0.09, 0.2$ and $\epsilon = 0.125, 0.025, 0.005, 0.001$. 58

2.5 Test 3: Snapshots of the zero-level set of $u^{\epsilon,h,k}$ at time $t = 0, 6 \times 10^{-3}, 1.2 \times 10^{-2}, 2 \times 10^{-2}$ and $\epsilon = 0.125, 0.025, 0.005, 0.001$. 59

2.6 Decay of the numerical energy $J_\epsilon^h(u_h^\ell)$ of Test 3. 60

3.1 Decay of the numerical energy $E_h(U^\ell)$ of Test 1. 104

3.2 Test 1: Snapshots of the zero-level set of $u^{\epsilon,h,k}$ at time $t = 0, 0.005, 0.015, 0.03$ and $\epsilon = 0.125, 0.025, 0.005, 0.001$. 105

3.3 Decay of the numerical energy $E_h(U^\ell)$ of Test 2. 106

3.4 Test 2: Snapshots of the zero-level set of $u^{\epsilon,h,k}$ at time $t = 0, 0.001, 0.04, 0.09$ and $\epsilon = 0.125, 0.025, 0.005, 0.001$. 107

3.5 Decay of the numerical energy $E_h(U^\ell)$ of Test 3. 108

3.6 Test 3: Snapshots of the zero-level set of $u^{\epsilon,h,k}$ at time $t = 0, 0.006, 0.012, 0.02$ and $\epsilon = 0.125, 0.025, 0.005, 0.001$. 108

4.1 Plot of the errors in Table 4.1. 132

4.2 Comparison of the computed solution with (blue line) and without (red line) the regularization term (Color figure online). 133
4.3 Surface plots of computed solution at a fixed stochastic sample on the space time domains $(0,1)\times(0,0.1)$ (left) and $(0,1)\times(0,2^8\times10^{-5})$ (right). $u_0(x)=\sin(\pi x)$ and $\epsilon=0.1$. 133
4.4 Surface plots of computed solution at a fixed stochastic sample on the space time domains $(0,1)\times(0,0.1)$ (left) and $(0,1)\times(0,2^8\times10^{-5})$ (right). $u_0(x)=\sin(\pi x)$ and $\epsilon=1$. 134
4.5 Surface plots of computed solution at a fixed stochastic sample on the space time domains $(0,1)\times(0,0.1)$ (left) and $(0,1)\times(0,2^8\times10^{-5})$ (right). $u_0(x)=\sin(\pi x)$ and $\epsilon=\sqrt{2}$. 134
4.6 Surface plots of computed solution at a fixed stochastic sample on the space time domains $(0,1)\times(0,0.1)$ (left) and $(0,1)\times(0,2^8\times10^{-5})$ (right). $u_0(x)=\sin(\pi x)$ and $\epsilon=5$. 135
4.7 Surface plots of computed solution at a fixed stochastic sample on the space time domains $(0,1)\times(0,0.1)$ (left) and $(0,1)\times(0,2^8\times10^{-5})$ (right). u_0 is given in (4.43) and $\epsilon=0.1$. 135
4.8 Surface plots of computed solution at a fixed stochastic sample on the space time domains $(0,1)\times(0,0.1)$ (left) and $(0,1)\times(0,2^8\times10^{-5})$ (right). u_0 is given in (4.43) and $\epsilon=1$. 136
4.9 Surface plots of computed solution at a fixed stochastic sample on the space time domains $(0,1)\times(0,0.1)$ (left) and $(0,1)\times(0,2^8\times10^{-5})$ (right). u_0 is given in (4.43) and $\epsilon=\sqrt{2}$. 136
4.10 Surface plots of computed solution at a fixed stochastic sample on the space time domains $(0,1)\times(0,0.1)$ (left) and $(0,1)\times(0,2^8\times10^{-5})$ (right). u_0 is given in (4.43) and $\epsilon=5$. 137
4.11 Decay of the energy $J(t)$ on the interval $(0,0.1)$. 137
4.12 Thresholding for colored noise: Trajectories for $\varepsilon=0.1$ (top left), $\varepsilon=0.5$ (top right), $\varepsilon=\sqrt{2}$ (bottom). 138

4.13 Geometric evolution vs colored noise evolution ($q_j^{\frac{1}{2}} = j^{-0.6}$, $J = 20$): 1st row: single trajectory for $n \mapsto \|\partial_x u_h^{\delta,n}(\omega)\|_{L^2}^2$ and $\varepsilon = 0.1$ (left), $\varepsilon = 0.5$ (right); 2nd row: $n \mapsto \mathbb{E}\big[\|\partial_x u_h^{\delta,n}\|_{L^2}^2\big]$ for $\varepsilon = 0.1$ (left), $\varepsilon = 0.5$ (right). 139

4.14 Geometric evolution vs colored noise evolution ($q_j^{\frac{1}{2}} = j^{-1}$, $J = 50$): 1st row: single trajectory for $n \mapsto \|\partial_x u_h^{\delta,n}(\omega)\|_{L^2}^2$ and $\varepsilon = 0.1$ (left), $\varepsilon = 0.5$ (right); 2nd row: $n \mapsto \mathbb{E}\big[\|\partial_x u_h^{\delta,n}\|_{L^2}^2\big]$ for $\varepsilon = 0.1$ (left), $\varepsilon = 0.5$ (right). 140

4.15 Thresholding and white noise: $\varepsilon = 0.1$ (top left), $\varepsilon = 0.5$ (top right) $\varepsilon = \sqrt{2}$ (bottom). 141

5.1 Snapshots of the zero-level set of $\bar{u}^{\delta,\epsilon,h,\tau}$ at time $t = 0, 0.020, 0.040, 0.043$ with $\delta = 0.1$ and $\epsilon = 0.01$. 172

5.2 Snapshots of the zero-level set of $\bar{u}^{\delta,\epsilon,h,\tau}$ at time $t = 0.005, 0.020, 0.040, 0.043$ with $\delta = 1$ and $\epsilon = 0.01$. 173

5.3 Snapshots of the zero-level set of $\bar{u}^{\delta,\epsilon,h,\tau}$ at time $t = 0.0025, 0.0050, 0.0100,$ $0.0200, 0.0250, 0.0280$ with $\delta = 10$ and $\epsilon = 0.01$. 174

5.4 Snapshots of the zero-level set of $\bar{u}^{\delta,\epsilon,h,\tau}$ at time $t = 0.0010, 0.0050, 0.0125,$ 0.0175 with $\delta = 0.1$ and $\epsilon = 0.01, 0.011, 0.02$. 175

5.5 Snapshots of the zero-level set of $\bar{u}^{\delta,\epsilon,h,\tau}$ at time $t = 0, 0.040, 0.200, 0.456$ with $\delta = 0.1$ and $\epsilon = 0.01$. 177

5.6 Snapshots of the zero-level set of $\bar{u}^{\delta,\epsilon,h,\tau}$ at time $t = 0.004, 0.040, 0.200, 0.456$ with $\delta = 1$ and $\epsilon = 0.01$. 178

5.7 Snapshots of the zero-level set of $\bar{u}^{\delta,\epsilon,h,\tau}$ at time $t = 0.004, 0.040, 0.080, 0.140,$ $0.180, 0.216$ with $\delta = 10$ and $\epsilon = 0.01$. 179

5.8 Snapshots of the zero-level set of $\bar{u}^{\delta,\epsilon,h,\tau}$ at time $t = 0.008, 0.040, 0.080,$ 0.120. 180

6.1 Test 1: Surface plot of the Computed pressure p at the terminal time T. 217

6.2 Test 1: Computed pressure p (color plot) and displacement $\mathbf{u}$ (arrow plot) at T. 218

6.3 Test 2: boundary conditions. 219
6.4 Test 2: Surface plot of the computed pressure p at the terminal time T. 219
6.5 Test 2: Computed pressure p (color plot) and displacement (arrow plot) at T. 220
6.6 Test 3: boundary conditions. 221
6.7 Test 3: Computed pressure p: surface plot (left) and color plot (right). 221
6.8 Test 3: Arrow plot of the computed displacement (left) and deformation of $\mathcal{D}$ (right). 222

Chapter 1

Introduction

1.1 Background

This section introduces some background materials related to mathematical models (i.e. partial differential equations) to be studied in this dissertation.

1.1.1 The moving interface problem

An interface in this dissertation refers to a hypersurface in $\mathbb{R}^d$. A moving interface means that the location and/or shape of the interface vary in time. How an interface moves is often described by its (pointwise) velocity V or the normal velocity $V_n := V \cdot n$. Such a formula for V or V_n is called a geometric law. It may depend on intrinsic features (such as curvatures) of the interface and on external factors (such as flow velocity) of the environment where the interface exists. The moving interface problem arises in many scientific and engineering fields such as fluid mechanics, materials science and biology. The pioneer work on the moving interface was done by Jozef Stefan around 1890 when he studied the problem about melting of the polar ice cap. The simple version of the Stefan problem is the melting model of a sheet of ice in the water at an initial temperature 0 °C. The interface between the water and the ice is raised at a temperature above zero through the whole process, and then the

interface moves toward the ice sheet. Other applications of the moving interface problem include two-phase flow problems in fluid mechanics, the shock waves in gas dynamics, and the phase transition problems in materials science. There are some direct methods for the moving interface problem, i.e., the parametrization method [95], the front tracking method [91] and the immersed interface method [64], and some indirectly methods, i.e., the level set method [72] and the phase field method [81, 92]. The direct methods are visual, qualitative, and the computational cost is smaller than using the indirect methods. However, direct methods have difficulties in handling with the topological changes, such as pinches, splits and merging. On the other hand, the topological changes can be easily handled by indirect methods, although the computational cost is higher. In this dissertation, we only consider numerical methods based on the indirect approach for the moving interface problem.

1.1.2 The level set method and the phase field method

The level set method was introduced by Stanley Osher and James A. Sethian [72] to compute and analyze the moving interface problem. For example, consider a closed hypersurface Γ_t in $\mathbb{R}^d$. Let Ω^+ denote the outside of the hypersurface Γ_t and Ω^- denote the inside of the hypersurface Γ_t. The idea of the level set method is to implicitly represent Γ_t as the zero level set of a function $u(\cdot, t)$ in $\mathbb{R}^d$, that is,

$$\Gamma_t := \big\{x(t) \in \mathbb{R}^d : u(x(t), t) = 0\big\}. \tag{1.1}$$

Taking the time derivative on both sides of the equation $u(x(t), t) = 0$, we get

$$u_t + \nabla u \cdot \frac{dx}{dt} = 0 \qquad \text{on} \quad \Gamma_t. \tag{1.2}$$

Since $V := \frac{dx}{dt}$ is the velocity of the surface, then

$$u_t + \nabla u \cdot V = 0 \qquad \text{on} \quad \Gamma_t. \tag{1.3}$$

Equation (1.3) is often called the level set equation, and it is determined by the velocity field V and the initial condition u_0 such that $\Gamma_0 = \{x \in \mathbb{R}^d; u_0(x) = 0\}$. To illustrate, we consider the following mean curvature flow as an example:

$$V_n(t,\cdot) = -H(t,\cdot), \tag{1.4}$$

here H denotes the mean curvature of Γ_t. By the facts from differential geometry

$$n = \frac{\nabla u}{|\nabla u|} \qquad \text{and} \quad H = \operatorname{div}(n), \tag{1.5}$$

the level set equation (1.3) becomes

$$0 = u_t + \nabla u \cdot V = u_t + |\nabla u| V_n = u_t - |\nabla u| H = u_t - |\nabla u| \operatorname{div}\left(\frac{\nabla u}{|\nabla u|}\right), \tag{1.6}$$

or

$$u_t - |\nabla u| \operatorname{div}\left(\frac{\nabla u}{|\nabla u|}\right) = 0. \tag{1.7}$$

Equation (1.7) is the famous level set formulation of the mean curvature flow [34, 52, 71].

The phase field method is another important method for the moving interface problem proposed by Lord Rayleigh [81] and Van der Waals [92]. This method was originally developed as a model of solidification, and it is also useful in many other applications, such as crack propagation, electromigration, crystal and tumor growth. The main idea of phase field method is to seek a phase field function u^ϵ such that the interface lies in the narrow region (called the diffuse interface)

$$\Gamma_t \subset Q_t^\epsilon := \left\{x(t) \in \mathbb{R}^d : |u^\epsilon(x(t),t)| \le 1 - \mathcal{O}(\epsilon)\right\}. \tag{1.8}$$

Here ϵ is a small positive constant, which controls the width of Q_t^ϵ. The phase field function takes two distinct value $+1$ and -1, which represent two distinct phases, with a smooth change between -1 and $+1$ in Q_t^ϵ. The zero level set $\Gamma_t^\epsilon := \{x(t) \in \mathbb{R}^d; u^\epsilon(x(t), t) = 0\}$ of u^ϵ, which is contained in the diffuse interface Q_t^ϵ, is often chosen to represent Γ_t approximately. The diffuse interface approach provides a convenient mathematical formalism for numerically approximating the moving interface problems because explicitly tracking the interface is not needed in the diffuse interface formulation. The main advantage of the diffuse interface method is its ability to handle with ease singularities of the interfaces. Like many singular perturbation problems, the main computational issue is to resolve the (small) scale introduced by the parameter ϵ in the equation. Computationally, the problem could become intractable, especially in three-dimensional cases if uniform meshes are used. This difficulty is often overcome by exploiting the predictable (at least for small ϵ) PDE solution profile and by using adaptive mesh techniques (cf. [49] and the references therein), so fine meshes are only used in the diffuse interface region to reduce the computational cost.

There is no general phase field equation for all moving interface problem and the formulation is problem-dependent. Notice that the difficulty is due to the fact that the interface lies inside Q_t^ϵ, but the specified location of the interface is unknown, so the curvature at the interface can not be calculated exactly as in the level set method. Again, we consider the phase field formulation of the mean curvature flow (1.4) as an example. To the end, let $d(x)$ denote the signed distance function between point x and the interface Γ_t, and consider the fact that the solution approximates the $\tanh(\cdot)$ function, we heuristically assume

$$u^\epsilon(x, t) := \tanh\left(\frac{d(x)}{\sqrt{2}\epsilon}\right). \tag{1.9}$$

Then we have

$$\tanh'(s) = 1 - \tanh^2(s), \tag{1.10}$$

$$\tanh''(s) = -2\tanh(s)(1 - \tanh^2(s)), \tag{1.11}$$

and

$$\nabla u^\epsilon(x) = \frac{\tanh\left(\frac{d(x)}{\sqrt{2}\epsilon}\right)}{\sqrt{2}\epsilon}\nabla d(x), \tag{1.12}$$

$$\Delta d(x) = \frac{\sqrt{2}\epsilon}{1-(u^\epsilon(x))^2}\left(\Delta u^\epsilon(x) + \frac{2u^\epsilon(x)}{1-(u^\epsilon(x))^2}\nabla u^\epsilon(x)\otimes\nabla u^\epsilon(x)\right). \tag{1.13}$$

Notice

$$|\nabla d(x)| = 1, \tag{1.14}$$

$$|\nabla u^\epsilon(x)|^2 = \frac{1}{2\epsilon^2}\left(1-(u^\epsilon(x))^2)^2\right), \tag{1.15}$$

then

$$H = \mathrm{tr}(\Delta d(x)) = \frac{\sqrt{2}\epsilon}{1-(u^\epsilon(x))^2}\left(\Delta u^\epsilon(x) + \frac{1}{\epsilon^2}(u^\epsilon(x) - (u^\epsilon(x))^3)\right). \tag{1.16}$$

Hence, by (1.6) and (1.16), we obtain the phase field equation

$$u_t^\epsilon - \Delta u^\epsilon + \frac{1}{\epsilon^2}\left((u^\epsilon)^3 - u^\epsilon\right) = 0. \tag{1.17}$$

It was proved in [33] that Γ_t^ϵ converges to Γ_t defined in (1.1) as $\epsilon \to 0$.

1.1.3 The mean curvature flow and the Hele-Shaw flow

The mean curvature flow (MCF) refers to a one-parameter family of hypersurfaces $\{\Gamma_t\}_{t\geq 0} \subset \mathbf{R}^d$ which starts from a given initial surface Γ_0 and evolves according to the geometric law in (1.4). The MCF is the best known curvature-driven geometric

flow which finds many applications in differential geometry, geometric measure theory, image processing and materials science and has been extensively studied both analytically and numerically (cf. [27, 52, 71, 87, 96] and the references therein).

As a geometric problem, the MCF can be described using different formulations. Among them, we mention the classical parametric formulation [56], Brakke's varifold formulation [11], De Giorgi's barrier function formulation [53, 8, 9], the variational formulation [5], the level set formulation [72, 34, 21], and the phase field formulation [33, 57]. We remark that different formulations often lead to different solution concepts and also lead to developing different analytical (and numerical) concepts and techniques to analyze and approximate the MCF. However, all these formulations of the MCF give rise to difficult but interesting nonlinear geometric partial differential equations (PDEs), and the resolution of the MCF then depends on the solutions of these nonlinear geometric PDEs. One interesting feature of the MCF is the development of singularities, in particular singularities which may occur in finite time, even when the initial hypersurface is smooth. The singularities may appear in different forms such as self-intersection, pinch-off, merging, and fattening. To understand and characterize these singularities have been the focus of the analytical and numerical research on the MCF (cf. [21, 27, 34, 45, 71, 87, 96], and the references therein).

The Hele-Shaw flow was originally defined as the Stokes flow between two parallel flat plates separated by an infinitesimally small gap. It describes the intricate patterns that appear in the gas-liquid interface problems when the upper plate is lifted slowly. It has significant applications since many problems in fluid mechanics can be considered as the approximations of the Hele-Shaw flow, and it has connections with the asymptotic behavior of phase field models. The mathematical model of the

Hele-Shaw problem is given as follows

$$\Delta w = 0 \qquad \text{in } \mathcal{D} \setminus \Gamma_t,\ t \in [0,T], \tag{1.18}$$

$$\frac{\partial w}{\partial n} = 0 \qquad \text{on } \partial\mathcal{D},\ t \in [0,T], \tag{1.19}$$

$$w = \sigma H \qquad \text{on } \Gamma_t,\ t \in [0,T], \tag{1.20}$$

$$V_n = \frac{1}{2}\left[\frac{\partial w}{\partial n}\right]_{\Gamma_t} \qquad \text{on } \Gamma_t,\ t \in [0,T], \tag{1.21}$$

$$\Gamma_0 = \Gamma_{00}, \qquad \text{when } t = 0. \tag{1.22}$$

Here $\left[\frac{\partial w}{\partial n}\right]_{\Gamma_t}$ represents the jump of the outward normal derivatives across the interface Γ_t, and

$$\sigma = \int_{-1}^{1} \sqrt{\frac{F(s)}{2}}\, ds. \tag{1.23}$$

It was proven by Xinfu Chen [18] that the solution of the Hele-Shaw problem exists locally if the initial interface Γ_0 is smooth, and the solution exists globally if Γ_0 is close to a circle. If (1.18) is replaced by a heat equation $w_t - \Delta w = 0$, the Hele-Shaw problem is called the Stefan problem with Gibbs-Thomson relation for the equilibrium of the solid-liquid interface. It was proven by Fred Almgren and Lihe Wang [6] and Stephan Luckhaus [65] that the global weak solution exists, and by E. Radkevitch [80] that the local classical solution exists.

1.1.4 The Allen-Cahn equation and the Cahn-Hilliard equation

The first PDE to be considered in this dissertation is the following singularly perturbed heat equation

$$u_t - \Delta u + \frac{1}{\epsilon^2} f(u) = 0 \qquad \text{in } \mathcal{D}_T := \mathcal{D} \times (0,T), \tag{1.24}$$

where $\mathcal{D} \subseteq \mathbf{R}^d\,(d = 2, 3)$ is a bounded domain, $f = F'$ for some double well potential density function F and ϵ, which is called the interaction length, is a small positive number. Here we focus on the following widely used quartic density function:

$$F(u) = \frac{1}{4}(u^2 - 1)^2. \tag{1.25}$$

Equation (1.24), which is known as the Allen-Cahn equation in the literature, was originally introduced by Samuel M. Allen and John W. Cahn in [4] as a model to describe the phase separation process of a binary alloy at a fixed temperature. In the equation u denotes the concentration of one of the two species of the alloy. We remark that equation (1.24) differs from the original Allen-Cahn equation in the scaling of the time, t here represents $\frac{t}{\epsilon^2}$ in the original formulation, hence, it is a fast time. The Allen-Cahn equation is not mass-conserved because $\int_{\mathcal{D}} u\,dx$ is not a constant in t. To completely describe the physical (and mathematical) problem, equation (1.24) must be complemented with appropriate initial and boundary conditions. The following boundary and initial conditions will be considered in this dissertation:

$$\frac{\partial u}{\partial n} = 0 \qquad \text{in } \partial\mathcal{D}_T, \tag{1.26}$$

$$u = u_0 \qquad \text{in } \mathcal{D} \times \{t = 0\}. \tag{1.27}$$

The Allen-Cahn equation is a phase field formulation of the mean curvature flow [33, 57], and it is well known [33, 57] that the Allen-Cahn equation (1.24) can be interpreted as the L^2-gradient flow for the following Cahn-Hilliard energy functional

$$J_\epsilon(v) := \int_{\mathcal{D}} \left(\frac{1}{2}|\nabla v|^2 + \frac{1}{\epsilon^2} F(v) \right) dx. \tag{1.28}$$

The second PDE to be considered in this dissertation is the following singularly perturbed fourth order PDE

$$u_t + \Delta\big(\epsilon\Delta u - \frac{1}{\epsilon}f(u)\big) = 0 \quad \text{in } \mathcal{D}_T. \tag{1.29}$$

Equation (1.29) is known as the Cahn-Hilliard equation. It was originally introduced by John W. Cahn and John E. Hilliard in [16] to describe the process of phase separation, by which the two components of a binary fluid spontaneously separate and form domains. Here u and $1-u$ denote respectively the concentrations of the two fluids, with $\{u = \pm 1\}$ indicating domains of the two components. It is often written as the following system of two second order PDEs along with two boundary conditions and one initial condition [42]

$$u_t - \Delta w = 0 \quad \text{in } \mathcal{D}_T, \tag{1.30}$$

$$w + \epsilon\Delta u - \frac{1}{\epsilon}f(u) = 0 \quad \text{in } \mathcal{D}_T, \tag{1.31}$$

$$\frac{\partial u}{\partial n} = \frac{\partial w}{\partial n} = 0 \quad \text{on } \partial\mathcal{D}_T, \tag{1.32}$$

$$u = u_0 \quad \text{in } \mathcal{D}\times\{t=0\}. \tag{1.33}$$

Here w is called the chemical potential in the literature. Notice that equation (1.29) differs from the original Cahn-Hilliard equation in the scaling of the time, and t here corresponds to $\frac{t}{\epsilon}$ in the original formulation.

The Cahn-Hilliard equation is the phase field formulation of the Hele-Shaw flow, and it is well known [3, 74, 20] that the Cahn-Hilliard equation (1.29) can be interpreted as the H^{-1}-gradient flow for the Cahn-Hilliard energy functional

$$J_\epsilon(v) := \int_{\mathcal{D}} \Big(\frac{\epsilon}{2}|\nabla v|^2 + \frac{1}{\epsilon}F(v)\Big)\, dx. \tag{1.34}$$

In addition to their important roles in materials phase transition, the Allen-Cahn equation [40] and the Cahn-Hilliard equation [42] have emerged as fundamental

equations as well as a building block in the phase field methodology (or the diffuse interface methodology) for moving boundary and free boundary problems arising from various applications such as fluid dynamics, materials science, image processing and biology (cf. [36, 67] and the references therein).

1.1.5 Stochastic mean curvature flow and stochastic Hele-Shaw flow

For application problems of the mean curvature flow and the Hele-Shaw flow, there may exist uncertainty which arises from various sources such as thermal fluctuation, impurities of the materials and the intrinsic instabilities of the deterministic evolutions. Therefore, it is interesting and necessary to consider the stochastic effects, and to study the impact of the noises on regularities of the solutions and their long-time behaviors. This motivates us to consider the following stochastically perturbed mean curvature flow [41]:

$$V_n(t,\cdot) = H(t,\cdot) + \epsilon \overset{\circ}{W}_t, \tag{1.35}$$

where $\overset{\circ}{W}$ denotes a R-valued white in time noise, $\epsilon > 0$ is a constant, and '$\circ$' refers to the Stratonovich derivative of the Wiener process W_t.

We also consider a general space dependent noise [59, 83, 43]:

$$V_n(t,\cdot) = H(t,\cdot) + \delta X(x) \cdot n \overset{\circ}{W}_t, \tag{1.36}$$

where δ is a (small) positive constant, n is the outward normal vector to the interface Γ_t and $X(x)$ is a R^d-valued smooth function with compact support.

Similarly, to obtain our stochastic Hele-Shaw flow, instead of modifying the PDE, a noise is added to the velocity of the underlying moving interface [44], that is

$$V_n = \frac{1}{2}\left[\frac{\partial w}{\partial n}\right]_{\Gamma_t} + \delta X(x)\cdot n \overset{\circ}{W}_t \,. \tag{1.37}$$

The reason and motivation for introducing the above form of noises were briefly explained in [83].

1.2 Scope of the dissertation research

The focuses of this dissertation are design, analysis, and implementation of efficient continuous and discontinuous Galerkin finite element methods for solving deterministic and stochastic nonlinear PDEs which arise from materials science, fluid and solid mechanics, and differential geometry. The reason why we favor discontinuous Galerkin methods is due to their advantages compared to the classical finite element method in regard to designing adaptive mesh methods and algorithms, which is an indispensable strategy with the diffuse interface methodology. Below we shall outline the main specific issues to be addressed in this dissertation.

Let $\Gamma_t^\epsilon := \{x \in \mathcal{D}; u(x,t) = 0.\}$ be the zero level set of the Allen-Cahn problem (1.24)-(1.27). It was proved in [33] that Γ_t^ϵ converges to Γ_t (the solution of the MCF) as $\epsilon \to 0$ and $u \to \pm 1$ uniformly in $\mathcal{D} \setminus \Gamma_t$. The connection between the Allen-Cahn equation and the MCF opens a door for approximating and computing the latter via the former. Indeed, such a connection is the basis for the phase field methodology for approximating curvature-driven moving interfaces. It has been widely used in many practical simulations. Let $u_{h,\tau}^\epsilon$ denote the (post-processed if needed) numerical solution and $\Gamma_{t,h,\tau}^\epsilon := \{x \in \mathcal{D};\, u_{h,\tau}^\epsilon(x,t) = 0\}$ be the numerical interface. The first rigorous proof of the convergence $\Gamma_{t,h,\tau}^\epsilon$ to Γ_t as $h, \tau, \epsilon \to 0$ ($u_{h,\tau}^\epsilon$ happens to be a continuous finite element solution) was given by Feng and Prohl in [45]. A natural question is that whether the result of [45] still holds for nonconforming and

discontinuous Galerkin method (DG). This question has been open in the past ten years even some unsuccessful attempts were made. To settle down this open question is one of the main goals of this dissertation.

Analogously, it was proved in [3] that the zero level set Γ_t^ϵ of u, which is the solution of the Cahn-Hilliard problem (1.30)-(1.33), converges to the sharp interface Γ_t (the solution of the Hele-Shaw flow) and $\mu^\epsilon := -\epsilon\Delta u + \frac{1}{\epsilon}f(u)$ converges to w as $\epsilon \to 0$, provided that w and Γ_t are smooth. Such a convergence result serves the theoretical basis for using the former to approximate the latter. This approach has been used in many practical applications. Let $u_{h,\tau}^\epsilon$ denote the (post-processed if needed) numerical solution and $\Gamma_{t,h,\tau}^\epsilon := \{x \in \mathcal{D};\, u_{h,\tau}^\epsilon(x,t) = 0\}$ be the numerical interface. The first rigorous proof of the convergence $\Gamma_{t,h,\tau}^\epsilon$ to Γ_t as $h, \tau, \epsilon \to 0$ ($u_{h,\tau}^\epsilon$ happens to be a mixed finite element solution) was given by Feng and Prohl in [47]. Again, a natural question is that whether the result of [47] can be extended to nonconforming and discontinuous Galerkin methods. This question has been open in the past ten years. To resolve this open question is another main goal of this dissertation.

As is mentioned early, the physical environment is seldom a deterministic system, and the noises could come from the thermal fluctuation, impurities of the materials and the intrinsic instabilities of the deterministic evolutions, so there is a need to consider the stochastic effects. The white-in-time noise is the most basic and fundamental noise, which can be the start point of the research. Moreover, in the dissertation, the noise is added to the velocity of the interface, instead of being added to the equation directly. This leads to considering the stochastic Allen-Cahn and Cahn-Hilliard equations with gradient-type multiplicative noises. There is no numerical analysis result in the literature for such nonlinear stochastic PDEs partly because most numerical analysis techniques for deterministic PDEs do not work for stochastic PDEs. In addition to the goal of developing efficient numerical methods for these stochastic PDEs, another overreaching goal of this dissertation is to use these nonlinear stochastic PDEs as a testbed for developing new numerical analysis techniques which hopefully are applicable to many other stochastic PDEs.

A poroelastic material is a fluid-solid system, and the displacement-pressure formulation for linear poroelasticity can be found in [75]. Its dynamic can be described by a multiphysics fluid-solid interaction process at pore scale. Unlike standard (macroscopic) fluid-solid interaction systems, some physical phenomena of the multiphysics process of a poroelastic material may not be explicitly revealed in its mathematical model. Instead, they are hidden in the model. Numerical methods based on the displacement-pressure formulation have been proposed earlier [75]. However, those methods require to solve large linear systems which have no "good" structure, especially, they are not symmetric positive definite (SPD). Moreover, those numerical methods often suffer a "locking phenomenon" so that the computed pressure exhibits some oscillations, especially, for small time t. To develop multiphysics finite element methods which can avoid these limitations and drawbacks and better capture the multiphysics (deformation and diffusion) of the poroelastic system is the final goal of this dissertation.

1.3 Summary of main contributions

This dissertation is comprised of several research projects in the area of numerical PDEs. Based on the research topics, it can be divided into three parts.

Part one (Chapter 2 and Chapter 3) of the dissertation studies interior penalty discontinuous Galerkin (IPDG) approximation of the Allen-Cahn equation [40] (reps. mixed interior penalty discontinuous Galerkin (MIPDG) approximation of the Cahn-Hilliard equation [42]), and its related curvature-driven geometric problem the mean curvature flow (resp. the Hele-Shaw flow). Two fully discrete (interior penalty) discontinuous Galerkin (DG) methods for the Allen-Cahn problem (1.24)-(1.27) based on two different time-stepping schemes are proposed. One is fully implicit with $f^{m+1} := (u^{m+1})^3 - u^{m+1}$ and the other is convex-splitting with $f^{m+1} := (u^{m+1})^3 - u^m$. It is shown that the fully implicit method is conditionally stable but the convex-splitting method is unconditionally stable. The discrete spectrum estimates for these

two equations are proven separately based on a perturbation argument. With the help of the discrete spectrum estimates, the sharper error which only depend on the negative power of ϵ instead of the exponential power, are derived. Furthermore, it is proven that the zero-level set of the IPDG solution for the Allen-Cahn equation (resp. the Cahn-Hilliard equation) converges to the mean curvature flow (resp. the Hele-Shaw flow). It should be pointed out that the analysis for the Cahn-Hilliard equation is much more involved than that for the Allen-Cahn equation.

Part two (Chapter 4 and Chapter 5) of the dissertation concerns with numerical methods for nonlinear stochastic partial differential equations (SPDEs). The focuses of the study are on the stochastic Allen-Cahn equation [43], the stochastic Cahn-Hilliard equation [44], and the level set equation of the stochastic mean curvature flow [41]. These SPDEs all contain gradient type multiplicative (white-in-time) noises, which belong to the strongest forms of noises for second order quasilinear PDEs. In this dissertation, several fully discrete finite element methods for approximating these three nonlinear stochastic PDEs are proposed and analyzed. In each case, a discrete energy law which mimics the corresponding PDE energy law is derived, and strong convergence is proven by establishing optimal order error estimates for the proposed finite element methods. Those are first strong convergence results for SPDEs with gradient type multiplicative (white-in-time) noises, the analysis techniques developed in the dissertation are quite involved and new (in particular, in comparison with deterministic techniques), they will certainly be useful and adaptable for studying other SPDEs.

Part three (Chapter 6) of the dissertation studies numerical methods for a quite different PDE problem from poroelasticity [37]. It is well known that deformation and diffusion are two common physical processes involved for an isothermal system. The subtlety and difficulty for numerical approximations of poroelasticity are caused by the fact that the diffusion process is not explicitly described in popular poroelasticity PDE models, instead, it is hidden in the models. Moreover, the peculiar structure of

poroelasticity PDE models often leads to so-called "locking phenomenon" for all direct numerical approximations of these PDE models. In the dissertation, a prototypical quasi-static poroelasticity model is considered and a novel reformulation of the model is introduced. The reformulated model consists of a generalized Stokes problem for the displacement vector and a "pseudo elastic pressure" coupled with a diffusion problem for another "pseudo elastic pressure", hence, the diffusion process is explicitly revealed in the reformulated model. Based on this reformulation, some fully discrete multiphysics finite element methods are constructed and the convergence with optimal rates in the energy norm is shown. Moreover, the troublesome "locking phenomenon" is completed avoided because the reformulated model has a built-in mechanism to automatically enforce nearly divergence-free constraint for the displacement vector near $t = 0$, which is verified by the extensive numerical tests.

The dissertation is completed with a list of future research projects presented in Chapter 7.

Chapter 2

Discontinuous Galerkin Methods for the Allen-Cahn Equation

2.1 Introduction

The Allen-Cahn equation (1.24) not only plays an important role in materials phase transition, it has also been well-known and intensively studied in the past thirty years due to its connection to the celebrated curvature driven geometric flow known as *the mean curvature flow* or *the motion by mean curvature* (cf. [33, 57] and the references therein). It was proved that [33] the zero-level set $\Gamma_t^\epsilon := \{x \in \mathcal{D}; u(x,t) = 0\}$ of the solution u to the problem (1.24)–(1.27) converges to *the mean curvature flow* which refers to the evolution of a curve/surface governed by the geometric law $V_n = -H$, where V_n and H respectively stand for the outward normal velocity and the mean curvature of the curve/surface.

Numerical approximations of the Allen-Cahn equation have been extensively investigated in the past thirty years (cf. [7, 30, 45] and the references therein). However, most of these works were carried out for a fixed parameter ϵ. The error estimates, which are obtained using the standard Gronwall inequality technique, show an exponential dependence on $\frac{1}{\epsilon}$. Such an estimate is clearly not useful for small ϵ,

in particular, in addressing the issue whether the flow of the computed numerical interfaces converge to the original sharp interface model: the mean curvature flow. Better error estimates should only depend on $\frac{1}{\epsilon}$ in some (low) polynomial orders because they can be used to provide an answer to the above convergence issue. In fact, such an estimate is the best result (in terms of ϵ) one can expect. The first such polynomial order in $\frac{1}{\epsilon}$ a priori estimate was obtained by Feng and Prohl in [45] for standard finite element approximations of the Allen-Cahn problem (1.24)–(1.27). Extensions of the results of [45], in particular, the sensitivity of the eigenvalue to the topology was later considered, and some numerical tests were also given by Bartels et al. in [7]. In addition, polynomial order in $\frac{1}{\epsilon}$ a posteriori error estimates were obtained in [7, 48, 60]. One of the key ideas employed in all these works is to use a nonstandard error estimate technique which is based on establishing a discrete spectrum estimate (using its continuous counterpart) for the linearized Allen-Cahn operator. An immediate application of the polynomial order in $\frac{1}{\epsilon}$ a priori and a posteriori error estimates is to prove the convergence of the numerical interfaces of the underlying finite element approximations to the mean curvature flow as ϵ and mesh sizes h and τ all tend to zero, and to establish rates of convergence (in powers of ϵ) for the numerical interfaces before the onset of singularities of the mean curvature flow.

The primary objectives of this chapter are twofold: First, we want to develop some interior penalty discontinuous Galerkin (IP-DG) methods and to establish polynomial order in $\frac{1}{\epsilon}$ a priori error estimates as well as to prove convergence and rates of convergence for the IP-DG numerical interfaces. This goal is motivated by the advantages of DG methods in regard to designing adaptive mesh methods and algorithms, which is an indispensable strategy with the diffuse interface methodology. Second, we use the Allen-Cahn equation as a prototype to develop new analysis techniques for analyzing convergence of numerical interfaces to the sharp interface for DG (and nonconforming finite element) discretizations of phase field models. To the best of our knowledge, no such convergence result and analysis technique is available

in the literature. The main obstacle for adapting the techniques of [45] is that the DG (and nonconforming finite element) spaces are not subspaces of $H^1(\mathcal{D})$. As a result, whether the desired discrete spectrum estimate holds becomes a key question to answer.

The remainder of this chapter is organized as follows. In section 2.2 we first recall some facts about the Allen-Cahn equation. In particular, we cite the spectrum estimate for the linearized Allen-Cahn operator from [19] and a nonlinear discrete Gronwall inequality from [73]. In section 2.3 we present two fully nonlinear IP-DG methods for problem (1.24)–(1.27) with the implicit Euler time stepping for the linear terms. The two methods differ in how the nonlinear term is discretized in time. The first is fully implicit and the second uses a well-known energy splitting idea due to Eyre [35]. The rest of section 2.3 devotes to the convergence analysis of the proposed IP-DG methods. The highlights of analysis include establishing a discrete spectrum estimate for the linearized Allen-Cahn operator in DG spaces and deriving optimal order (in h and τ) and polynomial order in $\frac{1}{\epsilon}$ a priori error estimates for the proposed IP-DG methods. In section 2.4, using the error estimates of section 2.3 we prove the convergence and rates of convergence for the numerical interfaces of the IP-DG solutions to the sharp interface of the mean curvature flow. Finally, we present some numerical experiment results in section 2.5 to gauge the performance of the proposed fully discrete IP-DG methods.

2.2 Preliminaries

In this section, we first recall a few facts about the solution of the problem (1.24)–(1.27) which can be found in [19, 45]. These facts will be used in the analysis of section 2.3 and 2.4. We then cite a lemma which provides an upper bound for discrete sequences that satisfy a Bernoulli-type inequality, and this lemma is crucially used in our error analysis in section 2.3. Standard function and space notations are adopted in this chapter. $(\cdot,\cdot)_{\mathcal{D}}$ denotes the standard inner product on $L^2(\mathcal{D})$, C and c denote

generic positive constants which is independent of ϵ, space and time step sizes h and τ.

In order to derive a priori solution estimates, as in [45] we assume the initial condition $u_0 \in H^1(\mathcal{D}) \cap H^2(\mathcal{D})$ with $\|u_0\| \leq 1$ satisfies the following assumptions:
General Assumption (GA)

(1) There exists a nonnegative constant σ_1 such that

$$J_\epsilon(u_0) \leq C\epsilon^{-2\sigma_1}. \tag{2.1}$$

(2) There exists a nonnegative constant σ_2 such that

$$\|\Delta u_0 - \epsilon^{-2} f(u_0)\|_{L^2(\mathcal{D})} \leq C\epsilon^{-\sigma_2}. \tag{2.2}$$

(3) There exists nonnegative constant σ_3 such that

$$\lim_{s\to 0^+} \|\nabla u_t(s)\|_{L^2(\mathcal{D})} \leq C\epsilon^{-\sigma_3}. \tag{2.3}$$

The following solution estimates and their proofs can be found in [45].

Proposition 2.2.1. *Suppose that (2.1) and (2.2) hold. Then the solution u of problem (1.24)–(1.27) satisfies the following estimates:*

(i) $\operatorname{ess\,sup}_{t\in[0,\infty)} \|u(t)\|_{L^\infty(\mathcal{D})} \leq 1,$

(ii) $\operatorname{ess\,sup}_{t\in[0,\infty)} J_\epsilon(u) + \int_0^\infty \|u_t(s)\|^2_{L^2(\mathcal{D})}\, ds \leq C\epsilon^{-2\sigma_1},$

(iii) $\int_0^T \|\Delta u(s)\|^2\, ds \leq C\epsilon^{-2(\sigma_1+1)},$

(iv) $\operatorname{ess\,sup}_{t\in[0,\infty)} \left(\|u_t\|^2_{L^2(\mathcal{D})} + \|u\|^2_{H^2(\mathcal{D})}\right) + \int_0^\infty \|\nabla u_t(s)\|^2_{L^2(\mathcal{D})}\, ds \leq C\epsilon^{-2\max\{\sigma_1+1,\sigma_2\}},$

(v) $\int_0^\infty \left(\|u_{tt}(s)\|^2_{H^{-1}(\mathcal{D})} + \|\Delta u_t(s)\|^2_{H^{-1}(\mathcal{D})}\right) ds \leq C\epsilon^{-2\max\{\sigma_1+1,\sigma_2\}}.$

In addition to (2.1) *and* (2.2), *suppose that* (2.3) *holds, then* u *also satisfies*

$$\text{(vi)} \quad \operatorname*{ess\,sup}_{t\in[0,\infty)} \|\nabla u_t\|^2_{L^2(\mathcal{D})} + \int_0^\infty \|u_{tt}(s)\|^2_{L^2}\, ds \le C\epsilon^{-2\max\{\sigma_1+2,\sigma_3\}},$$

$$\text{(vii)} \quad \int_0^\infty \|\Delta u_t(s)\|^2_{L^2(\mathcal{D})}\, ds \le C\epsilon^{-2\max\{\sigma_1+2,\sigma_3\}}.$$

Proof. (i). Define a function v as

$$v := (u-1)^+ := \begin{cases} u-1, & u-1 \ge 0, \\ 0, & u-1 < 0. \end{cases} \tag{2.4}$$

Then we have

$$\nabla v = \begin{cases} \nabla u, & u-1 \ge 0, \\ 0, & u-1 < 0, \end{cases} \tag{2.5}$$

and

$$u_t(u-1)^+ = \frac{1}{2}\frac{d}{dt}|(u-1)^+|^2. \tag{2.6}$$

Then

$$\nabla u \cdot \nabla v = (\nabla(u-1)^+)^2, \tag{2.7}$$

and

$$f(u)v = u(u+1)|(u-1)^+|^2. \tag{2.8}$$

Testing (1.24) with v, we get

$$\big(u_t, v\big)_{\mathcal{D}} + \big(\nabla u, \nabla v\big)_{\mathcal{D}} + \big(\frac{1}{\epsilon^2} f(u), v\big)_{\mathcal{D}} = 0 \qquad \forall v \in H^1(\mathcal{D}). \tag{2.9}$$

Then we have

$$\frac{d}{dt}\int_{\mathcal{D}} |(u-1)^+|^2 dx \le 0. \tag{2.10}$$

Based on the assumption $|u_0| \leq 1$ and (2.10), we can easily prove $u \leq 1$.
Similarly, when the function v is defined by

$$v = (u+1)^- = \begin{cases} 0, & u+1 \geq 0, \\ -u-1, & u+1 < 0, \end{cases} \tag{2.11}$$

we can prove $u \geq -1$. (i) is proved.
(ii). Taking the derivative in t on both sides of (1.28), we get

$$\begin{aligned} \frac{d\, J_\epsilon(u)}{dt} &= \int_{\mathcal{D}} \left(\nabla u \cdot \nabla u_t + \frac{1}{\epsilon^2} f(u) u_t \right) dx, \\ &= \int_{\mathcal{D}} \left(-\Delta u u_t + \frac{1}{\epsilon^2} f(u) u_t \right) dx, \\ &= -\, ||u_t||^2_{L^2(\mathcal{D})}. \end{aligned} \tag{2.12}$$

Suppose the maximum of $J_\epsilon(u(t))$ is obtained at $t = t_1$, then integrating (2.12) over $[0, t_1]$ and $[0, \infty]$ respectively, we get

$$\operatorname*{ess\,sup}_{t \in [0,\infty)} \; J_\epsilon(u) \leq C\epsilon^{-2\sigma_1}, \tag{2.13}$$

and

$$\int_0^\infty ||u_t(s)||^2_{L^2} ds \leq C\epsilon^{-2\sigma_1}. \tag{2.14}$$

Plus (2.13) and (2.14), we get (ii).
(iii). Testing (1.24) with $-\Delta u$, we get

$$\|\Delta u\|^2_{L^2(\mathcal{D})} = \big(u_t, \Delta u\big)_{\mathcal{D}} - \frac{1}{\epsilon^2}\big(f'(u), |\nabla u|^2\big)_{\mathcal{D}}. \tag{2.15}$$

Using Schwarz inequality to the first term on the right of the above formula, we get

$$||\Delta u||^2_{L^2(\mathcal{D})} \leq ||u_t||^2_{L^2(\mathcal{D})} - \frac{2}{\epsilon^2}(f'(u), |\nabla u|^2)_{\mathcal{D}}. \tag{2.16}$$

Integrating over $[0, T]$ on both sides of the above equation, we have

$$\int_0^T ||\Delta u(s)||^2_{L^2(\mathcal{D})} ds \leq \int_0^T ||u_t(s)||^2_{L^2(\mathcal{D})} ds - \int_0^T \frac{2}{\epsilon^2}(f'(u), |\nabla u|^2)_{\mathcal{D}} ds. \tag{2.17}$$

(iii) is immediately obtained by (2.17), (i) and (ii).

(iv). Taking the derivative in t on both sides of (1.24), we get

$$u_{tt} - \Delta u_t + \frac{1}{\epsilon^2} f'(u) u_t = 0. \tag{2.18}$$

Testing (2.18) with u_t, then

$$\frac{1}{2}\frac{d}{dt}||u_t||^2_{L^2(\mathcal{D})} + ||\nabla u_t||^2_{L^2(\mathcal{D})} \leq \frac{C}{\epsilon^2}||u_t||^2_{L^2(\mathcal{D})}. \tag{2.19}$$

Suppose the maximum of $||u_t||^2_{L^2}$ is obtained when $t = t_1$, then integrating above equality over $[0, t_1]$ and $[0, \infty]$ respectively, we get

$$\begin{aligned} \frac{1}{2}||u_t(t_1)||^2_{L^2(\mathcal{D})} + \int_0^{t_1} ||\nabla u_t(s)||^2_{L^2(\mathcal{D})} ds & \\ \leq \frac{1}{2}||u_t(0)||^2_{L^2(\mathcal{D})} + \frac{C}{\epsilon^2}\int_0^{t_1} ||u_t(s)||^2_{L^2(\mathcal{D})} ds & \end{aligned} \tag{2.20}$$

and

$$\int_0^{\infty} ||\nabla u_t(s)||^2_{L^2(\mathcal{D})} ds \leq \frac{C}{\epsilon^2}\int_0^{\infty} ||u_t(s)||^2_{L^2(\mathcal{D})} ds. \tag{2.21}$$

By (2.2) and (ii), we have

$$\frac{1}{2}||u_t(t_1)||^2_{L^2(\mathcal{D})} + \int_0^{t_1} ||\nabla u_t(s)||^2_{L^2(\mathcal{D})} ds \leq C(\epsilon^{-2\sigma_2} + \epsilon^{-2\sigma_1-2}) \leq C\epsilon^{2\min\{-\sigma_1-1,\sigma_2\}} \tag{2.22}$$

and

$$\int_0^{\infty} ||\nabla u_t(s)||^2_{L^2(\mathcal{D})} ds \leq C\epsilon^{-2\sigma_1-2}. \tag{2.23}$$

By (2.22) and (ii), we have

$$\begin{aligned} ||\Delta u||^2_{L^2(\mathcal{D})} &\leq C(\epsilon^{2\min\{-\sigma_1-1,\sigma_2\}} + \epsilon^{-2\sigma_1-2}) \\ &\leq C\epsilon^{2\min\{-\sigma_1-1,\sigma_2\}}. \end{aligned} \tag{2.24}$$

Then (iv) can proved by (2.24), (i) and (ii).

(v). By (2.18), we have

$$\begin{aligned} ||u_{tt}||_{H^{-1}(\mathcal{D})} \leq & ||\nabla u_t||_{L^2(\mathcal{D})} + \frac{1}{\epsilon^2} \sup_{\phi \in H^1(\mathcal{D})} \frac{(f'(u)u_t, \phi)_{\mathcal{D}}}{||\phi||_{H^1(\mathcal{D})}} \\ \leq & ||\nabla u_t||_{L^2(\mathcal{D})} + \frac{C}{\epsilon^2} ||f'(u)||_{L^\infty(\mathcal{D})} ||u_t||_{L^2(\mathcal{D})} \\ \leq & C\left\{ ||\nabla u_t||_{L^2(\mathcal{D})} + \frac{1}{\epsilon^2} ||u_t||_{L^2(\mathcal{D})} \right\}. \end{aligned} \tag{2.25}$$

By (ii) and (iv), we get

$$\begin{aligned} \int_0^{\infty} ||u_{tt}(s)||^2_{H^{-1}(\mathcal{D})} ds \leq & 2C\left\{ \int_0^{\infty} ||u_t(s)||^2_{L^2(\mathcal{D})} ds + \frac{1}{\epsilon^4} \int_0^{\infty} ||\nabla u_t(s)||^2_{L^2(\mathcal{D})} ds \right\} \\ \leq & C\epsilon^{2\min\{-\sigma_1-1,\sigma_2\}} + C\frac{1}{\epsilon^4}\epsilon^{-2\sigma_1} \\ \leq & C\epsilon^{2\min\{-\sigma_1-2,\sigma_2\}}. \end{aligned} \tag{2.26}$$

Using (iv), we have

$$\int_0^\infty ||\Delta u_t(s)||^2_{H^{-1}(\mathcal{D})} ds \leq \int_0^\infty ||\nabla u_t||^2_{L^2(\mathcal{D})} ds \leq C\epsilon^{2\min\{-\sigma_1-1,\sigma_2\}}. \tag{2.27}$$

Then (v) is proved.

(vi). Testing (2.18) with u_{tt}, we get

$$||u_{tt}||^2_{L^2(\mathcal{D})} + \frac{1}{2}\frac{d}{dt}||\nabla u_t||^2_{L^2(\mathcal{D})} + \frac{1}{\epsilon^2}(f'(u)u_t, u_{tt})_{\mathcal{D}} = 0. \tag{2.28}$$

Then

$$\begin{aligned} ||u_{tt}||^2_{L^2(\mathcal{D})} + \frac{1}{2}\frac{d}{dt}||\nabla u_t||^2_{L^2(\mathcal{D})} =& -\frac{1}{\epsilon^2}(f'(u)u_t, u_{tt})_{\mathcal{D}} \\ \leq& \frac{1}{2}||u_{tt}||^2_{L^2(\mathcal{D})} + \frac{1}{2\epsilon^4}||f'(u)||^2_{L^\infty(\mathcal{D})}||u_t||^2_{L^2(\mathcal{D})}. \end{aligned} \tag{2.29}$$

That is,

$$\frac{1}{2}||u_{tt}||^2_{L^2(\mathcal{D})} + \frac{1}{2}\frac{d}{dt}||\nabla u_t||^2_{L^2(\mathcal{D})} \leq \frac{1}{2\epsilon^4}||f'(u)||^2_{L^\infty(\mathcal{D})}||u_t||^2_{L^2(\mathcal{D})}. \tag{2.30}$$

Integrating both sides of (2.30) over $[0,\infty]$ and using (ii), we have

$$\int_0^\infty ||u_{tt}(s)||^2_{L^2(\mathcal{D})} ds \leq C\epsilon^{2\min\{-\sigma_1-2,-\sigma_3\}}. \tag{2.31}$$

Suppose the maximum of $||\nabla u_t||^2_{L^2(\mathcal{D})}$ is obtained at $t = t_1$, then integrating (2.30) over $[0, t_1]$, and using (2.3) and (ii), we get

$$\operatorname*{ess\,sup}_{t\in[0,\infty)} ||\nabla u_t||^2_{L^2(\mathcal{D})} \leq C\epsilon^{2\min\{-\sigma_1-2,-\sigma_3\}}. \tag{2.32}$$

Adding (2.31) and (2.32), (vi) is derived.

(vii). Testing (2.18) with Δu_t, and integrating over $[0, \infty]$, we get

$$\int_0^\infty ||\Delta u_t||^2_{L^2(\mathcal{D})} dt = \int_0^\infty (u_{tt}, \Delta u_t)_{\mathcal{D}} ds + \int_0^\infty (\frac{1}{\epsilon^2} f'(u) u_t, \Delta u_t)_{\mathcal{D}} ds. \tag{2.33}$$

Using Young's inequality, (2.31) and (ii), we have

$$\int_0^\infty ||\Delta u_t(s)||^2_{L^2(\mathcal{D})} ds \leq C\epsilon^{2\min\{-\sigma_1-2,-\sigma_3\}}. \tag{2.34}$$

Then (vii) is proved.

□

Next, we quote a lower bound estimate for the principal eigenvalue of the following linearized Allen-Cahn operator:

$$\mathcal{L}_{\text{AC}} := -\Delta + f'(u)I, \tag{2.35}$$

where I stands for the identity operator.

Proposition 2.2.2. *Suppose that (2.1) and (2.2) hold. Given a smooth initial curve/surface Γ_0, let u_0 be a smooth function satisfying $\Gamma_0 = \{x \in \mathcal{D}; u_0(x) = 0\}$ and some profile as described in [19]. Let u denote the solution of problem (1.24)–(1.27). Then there exists a positive ϵ-independent constant C_0 such that the principle eigenvalue of the linearized Allen-Cahn operator $\mathcal{L}_{AC}$ satisfies for $0 < \epsilon << 1$*

$$\lambda_{AC} \equiv \inf_{\substack{\psi \in H^1(\mathcal{D}) \\ \psi \neq 0}} \frac{\|\nabla\psi\|^2_{L^2(\mathcal{D})} + \epsilon^{-2}\left(f'(u)\psi, \psi\right)}{\|\psi\|^2_{L^2(\mathcal{D})}} \geq -C_0. \tag{2.36}$$

Remark 2.2.3. *(a) A proof of Proposition 2.2.2 can be found in [19]. A discrete generalization of (2.36) on C^0 finite element spaces was proved in [45]. It plays a pivotal role in the nonstandard convergence analysis of [45]. In the next section, we shall prove another discrete generalization of (2.36) on DG finite element spaces.*

(b) The restriction on the initial function u_0 is needed to guarantee that the solution $u(t)$ satisfies certain profile at later time $t > 0$ which is required in the proof of [19]. One example of admissible initial functions is $u_0 = tanh(\frac{d_0(x)}{\epsilon})$, where $d_0(x)$ stands for the signed distance function to the initial interface Γ_0. Such a u_0 is smooth when Γ_0 is smooth.

The classical Gronwall lemma derives an estimate for any function which satisfies a first order linear differential inequality. It is a main technique for deriving error estimates for continuous-in-time semi-discrete discretizations of many initial-boundary value PDE problems. Similarly, the discrete counterpart of Gronwall lemma is a main technical tool for deriving error estimates for fully discrete schemes. However, for many nonlinear PDE problems, the classical Gronwall lemma does not apply because of nonlinearity, instead, some nonlinear generalization must be used. In case of the power (or Bernoulli-type) nonlinearity, a generalized Gronwall lemma was proved in [48]. In the following we state a discrete counterpart of the lemma in [48], and the proof of a similar lemma can be found in [73]. This lemma will be utilized crucially in the next section and in Chapter 3.

Lemma 2.2.4. *Let $\{S_\ell\}_{\ell\geq 1}$ be a positive nondecreasing sequence and $\{b_\ell\}_{\ell\geq 1}$ and $\{k_\ell\}_{\ell\geq 1}$ be nonnegative sequences, and $p > 1$ be a constant. If*

$$S_{\ell+1} - S_\ell \leq b_\ell S_\ell + k_\ell S_\ell^p \qquad \text{for } \ell \geq 1, \tag{2.37}$$

$$S_1^{1-p} + (1-p)\sum_{s=1}^{\ell-1} k_s a_{s+1}^{1-p} > 0 \qquad \text{for } \ell \geq 2, \tag{2.38}$$

then

$$S_\ell \leq \frac{1}{a_\ell}\left\{ S_1^{1-p} + (1-p)\sum_{s=1}^{\ell-1} k_s a_{s+1}^{1-p} \right\}^{\frac{1}{1-p}} \qquad \text{for } \ell \geq 2, \tag{2.39}$$

where

$$a_\ell := \prod_{s=1}^{\ell-1} \frac{1}{1+b_s} \qquad \text{for } \ell \geq 2. \tag{2.40}$$

Proof.

$$\begin{aligned} a_{\ell+1} - a_\ell &= \prod_{s=1}^{\ell}[1+b_s]^{-1} - \prod_{s=1}^{\ell-1}[1+b_s]^{-1} \\ &= \prod_{s=1}^{\ell}[1+b_s]^{-1}(1-1-b_\ell) \\ &= -\, a_{\ell+1}b_\ell. \end{aligned} \tag{2.41}$$

Define

$$a_1 = 1. \tag{2.42}$$

Multiplying (2.37) by $a_{\ell+1}$ and using (2.41), we have

$$a_{\ell+1}S_{\ell+1} - [1+b_\ell]a_{\ell+1}S_\ell \le k_\ell a_{\ell+1}S_\ell^p. \tag{2.43}$$

That is

$$a_{\ell+1}S_{\ell+1} - a_\ell S_\ell \le k_\ell a_{\ell+1}S_\ell^p. \tag{2.44}$$

By Mean Value Theorem,

$$(a_{\ell+1}S_{\ell+1})^{1-p} - (a_\ell S_\ell)^{1-p} = (1-p)x^{-p}(a_{\ell+1}S_{\ell+1} - a_\ell S_\ell), \tag{2.45}$$

where x lies between $a_{\ell+1}S_{\ell+1}$ and $a_\ell S_\ell$. It is easy to see that a_ℓ is nonincreasing and S_ℓ is nondecreasing, so we have

$$x > a_\ell S_\ell \ge a_{\ell+1}S_\ell \qquad \text{if} \quad a_\ell S_\ell < a_{\ell+1}S_{\ell+1}, \tag{2.46}$$

$$x > a_{\ell+1}S_{\ell+1} \ge a_{\ell+1}S_\ell \qquad \text{if} \quad a_\ell S_\ell > a_{\ell+1}S_{\ell+1}. \tag{2.47}$$

Using (2.44) and (2.45), we get in both cases,

$$(a_{\ell+1}S_{\ell+1})^{1-p} - (a_\ell S_\ell)^{1-p} \geq (1-p)k_\ell a_{\ell+1}^{1-p}. \tag{2.48}$$

Now setting $\ell = s$ in (2.48) and taking the sum over s from 1 to $\ell - 1$, we obtain

$$(a_\ell S_\ell)^{1-p} \geq (a_1 S_1)^{1-p} + (1-p)\sum_{s=1}^{\ell-1} k_s a_{s+1}^{1-p}. \tag{2.49}$$

Taking q^{th} roots, we have

$$a_\ell S_\ell \leq \left\{ (S_1)^{1-p} + (1-p)\sum_{s=1}^{\ell-1} k_s a_{s+1}^{1-p} \right\}^{1/(1-p)}. \tag{2.50}$$

Moving a_ℓ into right-hand side, we get our conclusion. □

2.3 Fully discrete interior penalty discontinuous Galerkin approximations

2.3.1 Formulations

Let $\mathcal{T}_h$ be a quasi-uniform "triangulation" of $\mathcal{D}$ such that $\overline{\mathcal{D}} = \bigcup_{K\in\mathcal{T}_h} \overline{K}$. Let h_K denote the diameter of $K \in \mathcal{T}_h$ and $h := \max\{h_K; K \in \mathcal{T}_h\}$. We recall that the standard broken Sobolev space $H^s(\mathcal{T}_h)$ and DG finite element space V_h are defined as

$$H^s(\mathcal{T}_h) := \prod_{K\in\mathcal{T}_h} H^s(K), \qquad V_h := \prod_{K\in\mathcal{T}_h} P_r(K),$$

where $P_r(K)$ denotes the set of all polynomials whose degrees do not exceed a given positive integer r. Let $\mathcal{E}_h^I$ denote the set of all interior faces/edges of $\mathcal{T}_h$, $\mathcal{E}_h^B$ denote the set of all boundary faces/edges of $\mathcal{T}_h$, and $\mathcal{E}_h := \mathcal{E}_h^I \cup \mathcal{E}_h^B$. The L^2-inner product

for piecewise functions over the mesh $\mathcal{T}_h$ is naturally defined by

$$(v, w)_{\mathcal{T}_h} := \sum_{K \in \mathcal{T}_h} \int_K vw \, dx,$$

and for any set $\mathcal{S}_h \subset \mathcal{E}_h$, the L^2-inner product over $\mathcal{S}_h$ is defined by

$$\langle v, w \rangle_{\mathcal{S}_h} := \sum_{e \in \mathcal{S}_h} \int_e vw \, ds.$$

Let $K, K' \in \mathcal{T}_h$ and $e = \partial K \cap \partial K'$ and assume global labeling number of K is smaller than that of K'. We choose $n_e := n_K|_e = -n_{K'}|_e$ as the unit normal on e and define the following standard jump and average notations across the face/edge e:

$$\begin{aligned} [v] &:= v|_K - v|_{K'} \quad \text{on } e \in \mathcal{E}_h^I, \qquad [v] := v \quad \text{on } e \in \mathcal{E}_h^B, \\ \{v\} &:= \frac{1}{2}\big(v|_K + v|_{K'}\big) \quad \text{on } e \in \mathcal{E}_h^I, \qquad \{v\} := v \quad \text{on } e \in \mathcal{E}_h^B \end{aligned}$$

for $v \in V_h$.

Let M be a (large) positive integer. Define $\tau := T/M$ and $t_m := m\tau$ for $m = 0, 1, 2, \cdots, M$ be a uniform partition of $[0, T]$. For a sequence of functions $\{v^m\}_{m=0}^M$, we define the (backward) difference operator

$$d_t v^m := \frac{v^m - v^{m-1}}{k}, \qquad m = 1, 2, \cdots, M.$$

We are now ready to introduce our fully discrete DG finite element methods for problem (1.24)–(1.27). They are defined by seeking $u_h^m \in V_h$ for $m = 0, 1, 2, \cdots, M$ such that

$$\big(d_t u_h^{m+1}, v_h\big)_{\mathcal{T}_h} + a_h(u_h^{m+1}, v_h) + \frac{1}{\epsilon^2}\big(f^{m+1}, v_h\big)_{\mathcal{T}_h} = 0 \quad \forall v_h \in V_h, \tag{2.51}$$

where

$$a_h(w_h, v_h) := \big(\nabla w_h, \nabla v_h\big)_{\mathcal{T}_h} - \big\langle \{\partial_n w_h\}, [v_h]\big\rangle_{\mathcal{E}_h^I} + \lambda\big\langle [w_h], \{\partial_n v_h\}\big\rangle_{\mathcal{E}_h^I} + j_h(w_h, v_h), \tag{2.52}$$

$$j_h(w_h, v_h) := \sum_{e\in\mathcal{E}_h^I} \frac{\sigma_e}{h_e}\big\langle [w_h], [v_h]\big\rangle_e, \tag{2.53}$$

$$f^{m+1} := (u_h^{m+1})^3 - u_h^m \quad \text{or} \quad f^{m+1} := (u_h^{m+1})^3 - u_h^{m+1}, \tag{2.54}$$

where $\lambda = 0, \pm 1$ and σ_e is a positive piecewise constant function on $\mathcal{E}_h^I$, which will be chosen later (see Lemma 2.3.4). In addition, we need to supply u_h^0 to start the time-stepping, whose choice will be clear (and will be specified) later when we derive the error estimates in section 2.3.4.

We conclude this subsection with a few remarks to explain the above IP-DG methods.

Remark 2.3.1. *(a) The mesh-dependent bilinear form $a_h(\cdot,\cdot)$ is a well-known IP-DG discretization of the negative Laplace operator $-\Delta$, see [82].*

(b) Different choices of λ give different schemes. In this chapter we only focus on the symmetric case with $\lambda = -1$. Also, σ_e is called the penalty constant.

(c) The time discretization is the simple backward Euler method for the linear terms. However, we shall prove in section 2.3.2 that the treatment of the nonlinear term results in two implicit schemes which have different stability properties with respect to ϵ. We also note that only fully implicit scheme (i.e., $f^{m+1} = (u_h^{m+1})^3 - u_h^{m+1}$) was considered in [45], and the resulted finite element method was proved only conditionally stable there.

2.3.2 Discrete energy laws and well-posedness

As a gradient flow, problem (1.24)–(1.27) enjoys an energy law which leads to the estimate (ii) and then the subsequent estimates given in Proposition 2.2.1. One

simple criterion for building a numerical method for problem (1.24)–(1.27) is whether the method satisfies a discrete energy law which mimics the continuous energy law [39, 45]. The goal of this subsection is to show that the IP-DG methods proposed in the previous subsection are either unconditionally energy stable when $f^{m+1} = (u_h^{m+1})^3 - u_h^m$ or conditionally energy stable when $f^{m+1} = (u_h^{m+1})^3 - u_h^{m+1}$.

First, we introduce three mesh-dependent energy functionals which can be regarded as DG counterparts of the continuous Cahn-Hilliard energy J_ϵ defined in (1.28).

$$\Phi^h(v) := \frac{1}{2}\|\nabla v\|^2_{L^2(\mathcal{T}_h)} - \big\langle \{\partial_n v\}, [v]\big\rangle_{\mathcal{E}_h^I} + \frac{1}{2} j_h(v,v) \qquad \forall v \in H^2(\mathcal{T}_h), \tag{2.55}$$

$$J_\epsilon^h(v) := \Phi^h(v) + \frac{1}{\epsilon^2}\big(F(v), 1\big)_{\mathcal{T}_h} \qquad \forall v \in H^2(\mathcal{T}_h), \tag{2.56}$$

$$I_\epsilon^h(v) := \Phi^h(v) + \frac{1}{\epsilon^2}\big(F_c^+(v), 1\big)_{\mathcal{T}_h} \qquad \forall v \in H^2(\mathcal{T}_h), \tag{2.57}$$

where $F(v) = \frac{1}{4}(v^2-1)^2$ and $F_c^+(v) := \frac{1}{4}(v^4+1)$.

If we define $F_c^-(v) := \frac{1}{2}v^2$, then there holds the convex decomposition $F(v) = F_c^+(v) - F_c^-(v)$. It is easy to check that Φ^h and I_ϵ^h are convex functionals but J_ϵ^h is not because F is not convex. Moreover, we have

Lemma 2.3.2. *Let $\lambda = -1$ in (2.52), then there holds for all $v_h, w_h \in V_h$*

$$\Big(\frac{\delta \Phi^h(v_h)}{\delta v_h}, w_h\Big)_{\mathcal{T}_h} := \lim_{s\to 0} \frac{\Phi^h(v_h + s w_h) - \Phi^h(v_h)}{s} = a_h(v_h, w_h), \tag{2.58}$$

$$\begin{aligned}\Big(\frac{\delta J_\epsilon^h(v_h)}{\delta v_h}, w_h\Big)_{\mathcal{T}_h} &:= \lim_{s\to 0} \frac{J_\epsilon^h(v_h + s w_h) - J_\epsilon^h(v_h)}{s} \\ &= a_h(v_h, w_h) + \frac{1}{\epsilon^2}\big(F'(v_h), w_h\big)_{\mathcal{T}_h},\end{aligned} \tag{2.59}$$

$$\begin{aligned}\Big(\frac{\delta I_\epsilon^h(v_h)}{\delta v_h}, w_h\Big)_{\mathcal{T}_h} &:= \lim_{s\to 0} \frac{I_\epsilon^h(v_h + s w_h) - I_\epsilon^h(v_h)}{s} \\ &= a_h(v_h, w_h) + \frac{1}{\epsilon^2}\big((F_c^+)'(v_h), w_h\big)_{\mathcal{T}_h}.\end{aligned} \tag{2.60}$$

The proofs are straightforward. They can be proved by the definition of $\Phi^h(v)$, $J_\epsilon^h(v)$ and $I_\epsilon^h(v)$.

Remark 2.3.3. *We remark that (2.58)–(2.60) provide respectively the representations of the Fréchet derivatives of the energy functionals Φ^h, J_ϵ^h and I_ϵ^h in V^h. This simple observation is very helpful, it allows us to recast our DG formulations in (2.51)–(2.54) as a minimization/variation problem at each time step. It is also a deeper reason why the proposed DG methods satisfy some discrete energy laws to be proved below.*

Lemma 2.3.4. *There exist constants $\sigma_0, \alpha > 0$ such that for $\sigma_e > \sigma_0$ for all $e \in \mathcal{E}_h$ there holds*

$$\Phi^h(v_h) \geq \alpha \|v_h\|_{1,DG}^2 \qquad \forall v_h \in V_h, \tag{2.61}$$

where

$$\|v_h\|_{1,DG}^2 := \|\nabla v_h\|_{L^2(\mathcal{T}_h)}^2 + j_h(v_h, v_h). \tag{2.62}$$

Proof. Inequality (2.61) follows immediately from the following observation

$$2\Phi^h(v_h) = a_h(v_h, v_h) \qquad \forall v_h \in V_h, \tag{2.63}$$

and the well-known coercivity property of the DG bilinear form $a_h(\cdot,\cdot)$ (cf. [82]). □

We now are ready to state our discrete energy/stability estimates.

Theorem 2.3.5. *Let $\{u_h^m\}$ be a solution of scheme (2.51)–(2.54). Then there exists $\sigma_0' > 0$ such that for $\sigma_e > \sigma_0', \forall e \in \mathcal{E}_h$*

$$J_\epsilon^h(u_h^\ell) + k\sum_{m=0}^{\ell} R_{\epsilon,h}^m \leq J_\epsilon^h(u_h^0) \qquad \text{for } 0 \leq \ell \leq M, \tag{2.64}$$

where

$$\begin{aligned} R_{\epsilon,h}^m := &\Big(1 \pm \frac{k}{2\epsilon^2}\Big)\|d_t u_h^{m+1}\|_{L^2(\mathcal{T}_h)}^2 + \frac{k}{4}\|\nabla d_t u_h^{m+1}\|_{L^2(\mathcal{T}_h)}^2 \\ &+ \frac{k}{4} j_h\big(d_t u_h^{m+1}, d_t u_h^{m+1}\big) + \frac{k}{4\epsilon^2}\|d_t(|u_h^{m+1}|^2 - 1)\|_{L^2(\mathcal{T}_h)}^2, \end{aligned} \tag{2.65}$$

and the "+" sign in the first term is taken when $f^{m+1} = (u_h^{m+1})^3 - u_h^m$ and "-" sign is taken when $f^{m+1} = (u_h^{m+1})^3 - u_h^{m+1}$.

Proof. Setting $v = d_t u_h^{m+1}$ in (2.51) we get

$$\|d_t u_h^{m+1}\|^2_{L^2(\mathcal{T}_h)} + a_h(u_h^{m+1}, d_t u_h^{m+1}) + \frac{1}{\epsilon^2}\big(f^{m+1}, d_t u_h^{m+1}\big)_{\mathcal{T}_h} = 0. \tag{2.66}$$

By the algebraic identity $a(a-b) = \frac{1}{2}(a^2-b^2) + \frac{1}{2}(a-b)^2$ we have

$$\begin{aligned} a_h(u_h^{m+1}, \nabla d_t u_h^{m+1}) = \frac{1}{2} d_t a_h(u_h^{m+1}, u_h^{m+1}) + \frac{k}{2}\Big(&\|\nabla d_t u_h^{m+1}\|^2_{L^2(\mathcal{T}_h)} \\ &+ 2\big\langle\{d_t\partial_n u_h^{m+1}\}, [d_t u_h^{m+1}]\big\rangle_{\mathcal{E}_h^I} + j_h\big(d_t u_h^{m+1}, d_t u_h^{m+1}\big)\Big). \end{aligned} \tag{2.67}$$

It follows from the trace and Schwarz inequalities that

$$\begin{aligned} 2\big\langle\{d_t\partial_n u_h^{m+1}\}, [d_t u_h^{m+1}]\big\rangle_{\mathcal{E}_h^I} &\geq -2\|\{d_t\partial_n u_h^{m+1}\}\|_{L^2(\mathcal{E}_h^I)}\|[d_t u_h^{m+1}]\|_{L^2(\mathcal{E}_h^I)} \\ &\geq -Ch^{-\frac{1}{2}}\|d_t\nabla u_h^{m+1}\|_{L^2(\mathcal{T}_h)}\|[d_t u_h^{m+1}]\|_{L^2(\mathcal{E}_h^I)} \\ &\geq -\frac{1}{2}\|d_t\nabla u_h^{m+1}\|^2_{L^2(\mathcal{T}_h)} - Ch^{-1}\|[d_t u_h^{m+1}]\|^2_{L^2(\mathcal{E}_h^I)}. \end{aligned} \tag{2.68}$$

Then there exists $\sigma_1 > 0$ such that for $\sigma_e > \sigma_1$

$$\begin{aligned} a_h(u_h^{m+1}, \nabla d_t u_h^{m+1}) \geq \frac{1}{2} d_t a_h(u_h^{m+1}, u_h^{m+1}) \\ + \frac{k}{4}\Big(\|\nabla d_t u_h^{m+1}\|^2_{L^2(\mathcal{T}_h)} + j_h\big(d_t u_h^{m+1}, d_t u_h^{m+1}\big)\Big). \end{aligned} \tag{2.69}$$

We now bound the third term on the left-hand side of (2.66) from below. We first consider the case $f^{m+1} = (u_h^{m+1})^3 - u_h^m$. To the end, we write

$$\begin{aligned} f^{m+1} &= u_h^{m+1}\big(|u_h^{m+1}|^2 - 1\big) + k d_t u_h^{m+1} \\ &= \frac{1}{2}\big((u_h^{m+1} + u_h^m) + k d_t u_h^{m+1}\big)\big(|u_h^{m+1}|^2 - 1\big) + k d_t u_h^{m+1}. \end{aligned}$$

A direct calculation then yields

$$\frac{1}{\epsilon^2}\big(f^{m+1}, d_t u_h^{m+1}\big)_{\mathcal{T}_h} \geq \frac{1}{4\epsilon^2} d_t \||u_h^{m+1}|^2 - 1\|^2_{L^2(\mathcal{T}_h)} + \frac{k}{4\epsilon^2}\|d_t(|u_h^{m+1}|^2 - 1)\|^2_{L^2(\mathcal{T}_h)} + \frac{k}{2\epsilon^2}\|d_t u_h^{m+1}\|^2_{L^2(\mathcal{T}_h)}. \tag{2.70}$$

On the other hand, when $f^{m+1} = f(u_h^{m+1}) = (u_h^{m+1})^3 - u_h^{m+1}$, we have (cf. [45])

$$\frac{1}{\epsilon^2}\big(f^{m+1}, d_t u_h^{m+1}\big)_{\mathcal{T}_h} \geq \frac{1}{4\epsilon^2} d_t \||u_h^{m+1}|^2 - 1\|^2_{L^2(\mathcal{T}_h)} + \frac{k}{4\epsilon^2}\|d_t(|u_h^{m+1}|^2 - 1)\|^2_{L^2(\mathcal{T}_h)} - \frac{k}{2\epsilon^2}\|d_t u_h^{m+1}\|^2_{L^2(\mathcal{T}_h)}. \tag{2.71}$$

It follows from (2.66), (2.69), (2.63) and (2.70) (resp. (2.71)) that

$$\Big(1 \pm \frac{k}{2\epsilon^2}\Big)\|d_t u_h^{m+1}\|^2_{L^2(\mathcal{T}_h)} + d_t\Big(\Phi^h(u_h^{m+1}) + \frac{1}{4\epsilon^2}\||u_h^{m+1}|^2 - 1\|^2_{L^2(\mathcal{T}_h)}\Big) + \frac{k}{4}\Big(\|\nabla d_t u_h^{m+1}\|^2_{L^2(\mathcal{T}_h)} + j_h\big(d_t u_h^{m+1}, d_t u_h^{m+1}\big) + \frac{1}{\epsilon^2}\|d_t(|u_h^{m+1}|^2 - 1)\|^2_{L^2(\mathcal{T}_h)}\Big) \leq 0.$$

Finally, applying the summation operator $k\sum_{m=0}^{M-1}$ and using the definition of J_ϵ^h we obtain the desired estimate (2.64). The proof is complete. □

The above theorem immediately infers the following corollary.

Corollary 2.3.6. *The scheme (2.51)–(2.54) is stable for all $h, k > 0$ when $f^{m+1} = (u_h^{m+1})^3 - u_h^m$ and is stable for $h > 0, 2\epsilon^2 > k > 0$ when $f^{m+1} = (u_h^{m+1})^3 - u_h^{m+1}$, provided that $\sigma_e > \max\{\sigma_0, \sigma_0'\}$ for every $e \in \mathcal{E}_h$.*

Theorem 2.3.7. *Suppose that $\sigma_e > \max\{\sigma_0, \sigma_1\}$ for every $e \in \mathcal{E}_h$. Then there exists a unique solution u_h^{m+1} to the scheme (2.51)–(2.54) at every time step t_{m+1} for $h, k > 0$ in the case $f^{m+1} = (u_h^{m+1})^3 - u_h^m$. The conclusion still holds provided that $h > 0, 2\epsilon^2 > k > 0$ in the case $f^{m+1} = (u_h^{m+1})^3 - u_h^{m+1}$.*

Proof. Define the following functionals

$$G(v) := k\Phi^h(v) + \frac{k}{\epsilon^2}\big(F(v),1\big)_{\mathcal{T}_h} + \frac{1}{2}\|v\|^2_{L^2(\mathcal{T}_h)} - \big(u_h^m, v\big)_{\mathcal{T}_h},$$
$$H(v) := k\Phi^h(v) + \frac{k}{\epsilon^2}\big(F_c^+(v),1\big)_{\mathcal{T}_h} + \frac{1}{2}\|v\|^2_{L^2(\mathcal{T}_h)} - \Big(\frac{k}{\epsilon^2}+1\Big)\big(u_h^m, v\big)_{\mathcal{T}_h}.$$

Clearly, H is strictly convex for all $h, k > 0$. G is not always convex, however, it becomes strictly convex when $k < 2\epsilon^2$. To see this, we write $F(v) = F_c^+(v) - F_c^-(v)$ in the definition of $G(v)$ and notice that

$$-\frac{k}{\epsilon^2}\big(F_c^-(v),1\big)_{\mathcal{T}_h} + \frac{1}{2}\|v\|^2_{L^2(\mathcal{T}_h)} = \frac{1}{2}\Big(1-\frac{k}{\epsilon^2}\Big)\|v\|^2_{L^2(\mathcal{T}_h)},$$

which is strictly convex when $k < 2\epsilon^2$.

Using (2.58)–(2.60), it is easy to check that problem (2.51)–(2.54) is equivalent to the following minimization/variation problems:

$$u_h^{m+1} = \operatorname*{argmin}_{v_h\in V_h} G(v_h), \qquad \text{when } f^{m+1} = (u_h^{m+1})^3 - u_h^{m+1},$$
$$u_h^{m+1} = \operatorname*{argmin}_{v_h\in V_h} H(v_h), \qquad \text{when } f^{m+1} = (u_h^{m+1})^3 - u_h^{m}.$$

Thus, the conclusions of the theorem follow from the standard theory of finite-dimensional convex minimization problems. The proof is complete. □

2.3.3 Discrete discontinuous Galerkin spectrum estimate

In this subsection, we shall establish a discrete counterpart of the spectrum estimate (2.36) for the DG approximation. Such an estimate will play a vital role in our error analysis to be given in the next subsection. We recall that the desired spectrum estimate was obtained in [45] for the standard finite element approximation and it plays a vital role in the error analysis of [45]. Compared with the standard finite element approximation, the main additional difficulty for the DG approximation

is caused by the nonconformity of the DG finite element space V_h and its mesh-dependent bilinear form $a_h(\cdot,\cdot)$.

First, we introduce the DG elliptic projection operator $P_r^h : H^s(\mathcal{T}_h) \to V_h$ by

$$a_h(v - P_r^h v, w_h) + \left(v - P_r^h v, w_h\right)_{\mathcal{T}_h} = 0 \quad \forall w_h \in V_h \tag{2.72}$$

for any $v \in H^s(\mathcal{T}_h)$.

Next, we quote the following well known error estimate results from [22, 82].

Lemma 2.3.8. *Let $v \in W^{s,\infty}(\mathcal{T}_h)$, then there hold*

$$\|v - P_r^h v\|_{L^2(\mathcal{T}_h)} + h\|\nabla(v - P_r^h v)\|_{L^2(\mathcal{T}_h)} \le Ch^{\min\{r+1,s\}}\|u\|_{H^s(\mathcal{T}_h)}, \tag{2.73}$$

$$\frac{1}{|\ln h|^{\overline{r}}}\|v - P_r^h v\|_{L^\infty(\mathcal{T}_h)} + h\|\nabla(u - P_r^h u)\|_{L^\infty(\mathcal{T}_h)} \le Ch^{\min\{r+1,s\}}\|u\|_{W^{s,\infty}(\mathcal{T}_h)}. \tag{2.74}$$

where $\overline{r} := \min\{1, r\} - \min\{1, r-1\}$.

Let

$$C_1 = \max_{|\xi|\le 2}|f''(\xi)|. \tag{2.75}$$

and $\widehat{P}_r^h$, corresponding to P_r^h, denote the elliptic projection operator on the finite element space $S_h := V_h \cap C^0(\overline{\mathcal{D}})$, there holds the following estimate [45]:

$$\|u - \widehat{P}_r^h u\|_{L^\infty} \le Ch^{2-\frac{d}{2}}||u||_{H^2}. \tag{2.76}$$

We now state our discrete spectrum estimate for the DG approximation.

Proposition 2.3.9. *Suppose there exists a positive number $\gamma > 0$ such that the solution u of problem (1.24)–(1.27) satisfies*

$$\operatorname*{ess\,sup}_{t\in[0,T]} \|u(t)\|_{W^{r+1,\infty}(\mathcal{D})} \le C\epsilon^{-\gamma}. \tag{2.77}$$

Then there exists an ϵ-independent and h-independent constant $c_0 > 0$ such that for $\epsilon \in (0,1)$ and a.e. $t \in [0,T]$

$$\lambda_h^{DG}(t) := \inf_{\substack{\psi_h \in V_h \\ \psi_h \not\equiv 0}} \frac{a_h(\psi_h,\psi_h) + \frac{1}{\epsilon^2}\Big(f'\big(P_r^h u(t)\big)\psi_h,\psi_h\Big)_{\mathcal{T}_h}}{\|\psi_h\|^2_{L^2(\mathcal{T}_h)}} \geq -c_0, \tag{2.78}$$

provided that h satisfies the constraint

$$h^{2-\frac{d}{2}} \leq C_0(C_1C_2)^{-1}\epsilon^{\max\{\sigma_1+3,\sigma_2+2\}}, \tag{2.79}$$

$$h^{\min\{r+1,s\}}|\ln h|^{\overline{r}} \leq C_0(C_1C_2)^{-1}\epsilon^{\gamma+2}, \tag{2.80}$$

where C_2 arises from the following inequality:

$$\|u - P_r^h u\|_{L^\infty((0,T);L^\infty(\mathcal{D})} \leq C_2 h^{\min\{r+1,s\}}|\ln h|^{\overline{r}}\epsilon^{-\gamma}, \tag{2.81}$$

$$\|u - \widehat{P}_r^h u\|_{L^\infty((0,T);L^\infty(\mathcal{D})} \leq C_2 h^{2-\frac{d}{2}}\epsilon^{-\max\{\sigma_1+1,\sigma_2\}}. \tag{2.82}$$

Proof. Let $S_h := V_h \cap C^0(\overline{\mathcal{D}})$. For any $\psi_h \in V_h$, we define its finite element (elliptic) projection $\psi_h^{\mathrm{FE}} \in S_h$ by

$$\widetilde{a}_h(\psi_h^{\mathrm{FE}},\varphi_h) = \widetilde{a}_h(\psi_h,\varphi_h) \qquad \forall \varphi_h \in S_h, \tag{2.83}$$

where

$$\widetilde{a}_h(\psi,\varphi) = a_h(\psi,\varphi) + \beta(\psi,\varphi)_{\mathcal{T}_h} \qquad \forall \psi,\varphi \in H^2(\mathcal{T}_h),$$

and β is a positive constant to be specified later.

By Proposition 8 of [45] we have under the mesh constraint (2.79) that

$$\|f'(\widehat{P}_r^h u) - f'(u)\|_{L^\infty((0,T);L^\infty(\mathcal{D}))} \leq C_0\epsilon^2. \tag{2.84}$$

Similarly, under the mesh constraint (2.80) we can show that

$$\|f'(P_r^h u) - f'(u)\|_{L^\infty((0,t);L^\infty(\mathcal{D}))} \leq C_0\epsilon^2. \tag{2.85}$$

Then

$$\|f'(P_r^h u) - f'(\widehat{P}_r^h u)\|_{L^\infty((0,T);L^\infty(\mathcal{D}))} \leq 2C_0\epsilon^2. \tag{2.86}$$

Therefore,

$$f'(P_r^h u) \geq f'(\widehat{P}_r^h u) - 2C_0\epsilon^2. \tag{2.87}$$

By the definition of ψ_h^{FE} we have

$$a_h(\psi_h,\psi_h) = a_h(\psi_h^{\mathrm{FE}},\psi_h^{\mathrm{FE}}) + a_h(\psi_h-\psi_h^{\mathrm{FE}},\psi_h-\psi_h^{\mathrm{FE}}) - 2\beta(\psi_h-\psi_h^{\mathrm{FE}},\psi_h^{\mathrm{FE}})_{\mathcal{T}_h}.$$

Using the above inequality and equality we get

$$\begin{aligned} &a_h(\psi_h,\psi_h) + \frac{1}{\epsilon^2}\Big(f'\big(P_r^h u(t)\big)\psi_h,\psi_h\Big)_{\mathcal{T}_h} \\ &\quad\geq a_h(\psi_h^{\mathrm{FE}},\psi_h^{\mathrm{FE}}) + \frac{1}{\epsilon^2}\Big(f'\big(\widehat{P}_r^h u(t)\big),(\psi_h^{\mathrm{FE}})^2\Big)_{\mathcal{T}_h} \\ &\qquad + a_h(\psi_h-\psi_h^{\mathrm{FE}},\psi_h-\psi_h^{\mathrm{FE}}) - 2\beta\big(\psi_h-\psi_h^{\mathrm{FE}},\psi_h^{\mathrm{FE}}\big)_{\mathcal{T}_h} \\ &\qquad + \frac{1}{\epsilon^2}\Big(f'\big(\widehat{P}_r^h u(t)\big),(\psi_h)^2-(\psi_h^{\mathrm{FE}})^2\Big)_{\mathcal{T}_h} - 2C_0\|\psi_h\|^2_{L^2(\mathcal{T}_h)}. \end{aligned} \tag{2.88}$$

We now bound the fourth and fifth terms on the right-hand side of (2.88) from below. For the fourth term we have

$$\begin{aligned} -2\beta\big(\psi_h-\psi_h^{\mathrm{FE}},\psi_h^{\mathrm{FE}}\big)_{\mathcal{T}_h} &\geq 2\beta\|\psi_h^{\mathrm{FE}}\|^2_{L^2(\mathcal{T}_h)} - 2\beta\|\psi_h^{\mathrm{FE}}\|_{L^2(\mathcal{T}_h)}\|\psi_h\|_{L^2(\mathcal{T}_h)} \\ &\geq \beta\|\psi_h^{\mathrm{FE}}\|^2_{L^2(\mathcal{T}_h)} - \beta\|\psi_h\|^2_{L^2(\mathcal{T}_h)}. \end{aligned} \tag{2.89}$$

To bound the fifth term, by (2.54) and the L^∞-norm estimate for $u(t) - \widehat{P}_r^h u(t)$ we have that under the mesh constraint (2.79)

$$\begin{aligned}\|f'(\widehat{P}_r^h u(t))\|_{L^\infty(\mathcal{D})} &\le \|f'(u(t))\|_{L^\infty(\mathcal{D})} + \|f'(u(t)) - f'(\widehat{P}_r^h u(t))\|_{L^\infty(\mathcal{D})} \\ &\le \|f'(u(t))\|_{L^\infty(\mathcal{D})} + C\|u(t) - \widehat{P}_r^h u(t)\|_{L^\infty(\mathcal{D})} \le C.\end{aligned}$$

Thus, by the algebraic formula $|a^2 - b^2| \le |a-b|^2 + 2|ab|$, we get for some $C > 0$

$$\begin{aligned}\frac{1}{\epsilon^2}\Big(f'\big(\widehat{P}_r^h u(t)\big), (\psi_h)^2 - (\psi_h^{\mathrm{FE}})^2\Big)_{\mathcal{T}_h} &\ge -\frac{C}{\epsilon^2}\|(\psi_h)^2 - (\psi_h^{\mathrm{FE}})^2\|_{L^1(\mathcal{T}_h)} \\ &\ge -\frac{C}{\epsilon^2}\Big(\|\psi_h - \psi_h^{\mathrm{FE}}\|_{L^2(\mathcal{T}_h)}^2 + 2\|\psi_h - \psi_h^{\mathrm{FE}}\|_{L^2(\mathcal{T}_h)}\|\psi_h^{\mathrm{FE}}\|_{L^2(\mathcal{T}_h)}\Big) \\ &\ge -\frac{C}{\epsilon^2}\Big(\big(1+\epsilon^{-2}\big)\|\psi_h - \psi_h^{\mathrm{FE}}\|_{L^2(\mathcal{T}_h)}^2 + \epsilon^2\|\psi_h^{\mathrm{FE}}\|_{L^2(\mathcal{T}_h)}^2\Big).\end{aligned} \tag{2.90}$$

Now it comes to a key idea in bounding $\|\psi_h - \psi_h^{\mathrm{FE}}\|_{L^2(\mathcal{T}_h)}$, which is to use the duality argument to bound it from above by the energy norm $a_h(\psi_h - \psi_h^{\mathrm{FE}}, \psi_h - \psi_h^{\mathrm{FE}})^{\frac{1}{2}}$ To the end, we consider the following auxiliary problem: find $\phi \in H^1(\mathcal{D}) \cap H^2_{\mathrm{loc}}(\mathcal{D})$ such that

$$\widetilde{a}_h(\phi, \chi) = \big(\psi_h - \psi_h^{\mathrm{FE}}, \chi\big)_{\mathcal{T}_h} \qquad \forall \chi \in H^1(\mathcal{D}).$$

We assume the above variational problem is $H^{1+\theta}$-regular for some $\theta \in (0, 1]$, that is, there exists a unique $\phi \in H^{1+\theta}(\mathcal{D})$ such that

$$\|\phi\|_{H^{1+\theta}(\mathcal{D})} \le C\|\psi_h - \psi_h^{\mathrm{FE}}\|_{L^2(\mathcal{D})}.$$

It should be noted that $C(> 0)$ can be made independent of β.

By the definition of ψ_h^{FE} in (2.83), we immediately get the Galerkin orthogonality

$$\widetilde{a}_h\big(\psi_h - \psi_h^{\mathrm{FE}}, \chi_h\big) = 0 \qquad \forall \chi_h \in S_h.$$

The above orthogonality allows us easily to obtain by the duality argument (cf. [82] for a general duality argument for DG methods)

$$\|\psi_h - \psi_h^{\mathrm{FE}}\|^2_{L^2(\mathcal{T}_h)} \le C h^{2\theta}\, a_h(\psi_h - \psi_h^{\mathrm{FE}}, \psi_h - \psi_h^{\mathrm{FE}}) \tag{2.91}$$

Again, the constant C can be made independent of β.

By Proposition 8 of [45] we also have the following spectrum estimate in the finite element space S_h:

$$a_h(\psi_h^{\mathrm{FE}}, \psi_h^{\mathrm{FE}}) + \frac{1}{\epsilon^2}\Big(f'\big(\widehat{P}_r^h u(t)\big), (\psi_h^{\mathrm{FE}})^2\Big)_{\mathcal{T}_h} \ge -2C_0 \|\psi_h^{\mathrm{FE}}\|^2_{L^2(\mathcal{T}_h)}. \tag{2.92}$$

Finally, combining (2.88)–(2.92) we get

$$\begin{aligned} a_h(\psi_h, \psi_h) + \frac{1}{\epsilon^2}\Big(f'\big(P_r^h u(t)\big)\psi_h, \psi_h\Big)_{\mathcal{T}_h} & \\ \ge \big(1 - Ch^{2\theta}\epsilon^{-4}\big) a_h(\psi_h - \psi_h^{\mathrm{FE}}, \psi_h - \psi_h^{\mathrm{FE}}) & \\ + \big(\beta - C - 2C_0\big)\|\psi_h^{\mathrm{FE}}\|^2_{L^2(\mathcal{T}_h)} - \big(\beta + 2C_0\big)\|\psi_h\|^2_{L^2(\mathcal{T}_h)} & \\ \ge -\big(\beta + 2C_0\big)\|\psi_h\|^2_{L^2(\mathcal{T}_h)} \qquad \forall \psi_h \in V_h, & \end{aligned} \tag{2.93}$$

provided that β is chosen large enough such that $\beta - C - 2C_0 > 0$ and $1 - Ch^{2\theta}\epsilon^{-4} > 0$, under the mesh constraint (2.80). The proof is complete after setting $c_0 = \beta + 2C_0$. □

Remark 2.3.10. *The proof actually is constructive in finding the ϵ- and h-independent constant c_0. As expected, $c_0 > 2C_0$. We also note that inequality (2.93) is a Gärding-type inequality for the non-coercive elliptic operator $\mathcal{L}_{AC}$.*

2.3.4 Polynomial order in ϵ^{-1} error estimates

The goal of this subsection is to derive optimal order error estimates for the global error $u(t_m) - u_h^m$ of the fully discrete scheme (2.51)–(2.54) under some reasonable mesh constraints on h, k and regularity assumptions on u_0. This will be achieved

by adapting the nonstandard error estimate technique with a help of the generalized Gronwall lemma (Lemma 2.2.4) and the discrete spectrum estimate (2.78).

The main result of this subsection is the following error estimate theorem.

Theorem 2.3.11. *suppose* $\sigma_e > \max\{\sigma_0, \sigma_0'\}$. *Let* u *and* $\{u_h^m\}_{m=1}^M$ *denote respectively the solutions of problems* (1.24)–(1.27) *and* (2.51)–(2.54). *Assume* $u \in H^2((0,T); L^2(\mathcal{D})) \cap L^2((0,T); W^{s,\infty}(\mathcal{D}))$ *and suppose (GA) and* (2.77) *hold. Then, under the following mesh and initial value constraints:*

$$
\begin{aligned}
h^{2-\frac{d}{2}} &\le C_0(C_1C_2)^{-1}\epsilon^{\max\{\sigma_1+3,\sigma_2+2\}},\\
h^{\min\{r+1,s\}}|\ln h|^{\overline{r}} &\le C_0(C_1C_2)^{-1}\epsilon^{\gamma+2},\\
k^2 + h^{2\min\{r+1,s\}} &\le \epsilon^{4+d+2(\sigma_1+2)},\\
k &\le C\epsilon^{\frac{8+2d+4\sigma_1}{4-d}},\\
u_h^0 \in S_h \text{ such that} \quad \|u_0 - u_h^0\|_{L^2(\mathcal{T}_h)} &\le Ch^{\min\{r+1,s\}},
\end{aligned}
$$

there hold

$$
\begin{aligned}
\max_{0\le m\le M}\|u(t_m)-u_h^m\|_{L^2(\mathcal{T}_h)} + \Big(k^2\sum_{m=1}^M \|d_t(u(t_m)-u_h^m)\|_{L^2(\mathcal{T}_h)}^2\Big)^{\frac12} &\\
\le C(k+h^{\min\{r+1,s\}})\epsilon^{-(\sigma_1+2)}, &\qquad (2.94)\\
\Big(k\sum_{m=1}^M \|u(t_m)-u_h^m\|_{H^1(\mathcal{T}_h)}^2\Big)^{\frac12} \le C(k+h^{\min\{r+1,s\}-1})\epsilon^{-(\sigma_1+3)}, &\qquad (2.95)\\
\max_{0\le m\le M}\|u(t_m)-u_h^m\|_{L^\infty(\mathcal{T}_h)} \le Ch^{\min\{r+1,s\}}|\ln h|^{\overline{r}}\epsilon^{-\gamma} &\qquad (2.96)\\
+ Ch^{-\frac{d}{2}}(k+h^{\min\{r+1,s\}})\epsilon^{-(\sigma_1+2)}. &
\end{aligned}
$$

Proof. We only give a proof for the case $f^{m+1} = (u_h^{m+1})^3 - u_h^m$ because its proof is slightly more difficult than that for the case $f^{m+1} = (u_h^{m+1})^3 - u_h^{m+1}$. Since the proof is long, we divide it into four steps.

Step 1: We begin with introducing the following error decompositions:

$$u(t_m) - u_h^m = \eta^m + \xi^m, \quad \eta^m := u(t_m) - P_r^h u(t_m), \quad \xi^m := P_r^h u(t_m) - u_h^m.$$

It is easy to check that the exact solution u satisfies

$$\big(d_t u(t_{m+1}), v_h\big)_{\mathcal{T}_h} + a_h(u(t_{m+1}), v_h) + \frac{1}{\epsilon^2}\big(f(u(t_{m+1})), v_h\big)_{\mathcal{T}_h} = \big(R_{m+1}, v_h\big)_{\mathcal{T}_h} \tag{2.97}$$

for all $v_h \in V_h$, where

$$R_{m+1} := -\frac{1}{k}\int_{t_m}^{t_{m+1}} (t - t_m) u_{tt}(t)\, dt.$$

Hence

$$\begin{aligned} k\sum_{m=0}^{\ell} \|R_{m+1}\|_{L^2(\mathcal{D})}^2 &\le \frac{1}{k}\sum_{m=0}^{\ell}\Big(\int_{t_m}^{t_{m+1}} (s - t_m)^2 ds\Big)\Big(\int_{t_m}^{t_{m+1}} \|u_{tt}\|_{L^2(\mathcal{D})}^2\, ds\Big) \\ &\le C k^2 \epsilon^{-2\max\{\sigma_1+2, \sigma_3\}}. \end{aligned} \tag{2.98}$$

Subtracting (2.51) from (2.97) and using the definitions of η^m and ξ^m we get the following error equation:

$$\begin{aligned} &\big(d_t \xi^{m+1}, v_h\big)_{\mathcal{T}_h} + a_h(\xi^{m+1}, v_h) + \frac{1}{\epsilon^2}\big(f(u(t_{m+1})) - f^{m+1}, v_h\big)_{\mathcal{T}_h} \\ &\quad = \big(R_{m+1}, v_h\big)_{\mathcal{T}_h} - \big(d_t \eta^{m+1}, v_h\big)_{\mathcal{T}_h} - a_h(\eta^{m+1}, v_h) \\ &\quad = \big(R_{m+1}, v_h\big)_{\mathcal{T}_h} - \big(d_t \eta^{m+1}, v_h\big)_{\mathcal{T}_h} + (\eta^{m+1}, v_h)_{\mathcal{T}_h}. \end{aligned} \tag{2.99}$$

Setting $v_h = \xi^{m+1}$ and using Schwarz inequality yield

$$\begin{aligned} &\frac{1}{2}\Big(d_t \|\xi^{m+1}\|_{L^2(\mathcal{T}_h)}^2 + k\|d_t \xi^{m+1}\|_{L^2(\mathcal{T}_h)}^2\Big) + a_h(\xi^{m+1}, \xi^{m+1}) \\ &\qquad + \frac{1}{\epsilon^2}\big(f(u(t_{m+1})) - f^{m+1}, \xi^{m+1}\big)_{\mathcal{T}_h} \\ &\le \Big(\|R_{m+1}\|_{L^2(\mathcal{T}_h)} + \|d_t \eta^{m+1}\|_{L^2(\mathcal{T}_h)} + \|\eta^{m+1}\|_{L^2(\mathcal{T}_h)}\Big)\|\xi^{m+1}\|_{L^2(\mathcal{T}_h)}. \end{aligned}$$

Summing in m (after having lowered the index by 1) from 1 to $\ell\,(\leq M)$ and using (2.73) and (2.98) we get

$$\begin{aligned}
&\|\xi^{\ell}\|^2_{L^2(\mathcal{T}_h)} + 2k\sum_{m=1}^{\ell} k\|d_t\xi^m\|^2_{L^2(\mathcal{T}_h)} + 2k\sum_{m=1}^{\ell} a_h(\xi^m,\xi^m) \\
&\qquad + 2k\sum_{m=1}^{\ell}\frac{1}{\epsilon^2}\big(f(u(t_m)) - f^m, \xi^m\big)_{\mathcal{T}_h} \\
&\leq k\sum_{m=1}^{\ell}\|\xi^m\|^2_{L^2(\mathcal{T}_h)} + 2\|\xi^0\|^2_{L^2(\mathcal{T}_h)} + C\Big(k^2\epsilon^{-2\max\{\sigma_1+2,\sigma_3\}} \\
&\qquad + h^{2\min\{r+1,s\}}\,\|u\|^2_{H^1((0,T);H^s(\mathcal{D}))}\Big).
\end{aligned} \tag{2.100}$$

Step 2: We now bound the fourth term on the left-hand side of (2.100). By the definition of f^m we have

$$\begin{aligned}
f(u(t_m)) - f^m &= f(u(t_m)) - f\big(P_r^h u(t_m)\big) + f\big(P_r^h u(t_m)\big) - f^m \\
&= -\big[f(u(t_m)) - f\big(P_r^h u(t_m)\big)\big] + \big(P_r^h u(t_m)\big)^3 - P_r^h u(t_m) - (u_h^m)^3 + u_h^{m-1} \\
&= -\big[f(u(t_m)) - f\big(P_r^h u(t_m)\big)\big] + \Big(\big(P_r^h u(t_m)\big)^2 + P_r^h u(t_m)\,u_h^m + (u_h^m)^2\Big)\xi^m \\
&\quad - \xi^m - k d_t u_h^m \\
&= -\big[f(u(t_m)) - f\big(P_r^h u(t_m)\big)\big] + \Big(3\big(P_r^h u(t_m)\big)^2 - 1\Big)\xi^m - 3P_r^h u(t_m)\,(\xi^m)^2 \\
&\quad + (\xi^m)^3 - k d_t u_h^m \\
&= -\big[f(u(t_m)) - f\big(P_r^h u(t_m)\big)\big] + f'\big(P_r^h u(t_m)\big)\,\xi^m - 3P_r^h u(t_m)\,(\xi^m)^2 \\
&\quad + (\xi^m)^3 - k d_t u_h^m.
\end{aligned}$$

Hence, on noting that $-\big[f(u(t_m)) - f\big(P_r^h u(t_m)\big)\big] \geq -C|\eta^m|$, we have

$$
\begin{aligned}
&2k\sum_{m=1}^{\ell}\frac{1}{\epsilon^2}\big(f(u(t_m)) - f^m, \xi^m\big)_{\mathcal{T}_h} \\
&\geq -\frac{Ck}{\epsilon^2}\sum_{m=1}^{\ell}\||\eta^m\|_{L^2(\mathcal{T}_h)}\|\xi^m\|_{L^2(\mathcal{T}_h)} + 2k\sum_{m=1}^{\ell}\frac{1}{\epsilon^2}\Big(f'\big(P_r^h u(t_m)\big), (\xi^m)^2\Big)_{\mathcal{T}_h} \\
&\quad - \frac{Ck}{\epsilon^2}\sum_{m=1}^{\ell}\|\xi^m\|_{L^3(\mathcal{T}_h)}^3 + \frac{2k}{\epsilon^2}\sum_{m=1}^{\ell}\|\xi^m\|_{L^4(\mathcal{T}_h)}^4 - \frac{2k}{\epsilon^2}\sum_{m=1}^{\ell}k\|d_t u_h^m\|_{L^2(\mathcal{T}_h)}\,\|\xi^m\|_{L^2(\mathcal{T}_h)} \\
&\geq 2k\sum_{m=1}^{\ell}\frac{1}{\epsilon^2}\Big(f'\big(P_r^h u(t_m)\big), (\xi^m)^2\Big)_{\mathcal{T}_h} + \frac{2k}{\epsilon^2}\sum_{m=1}^{\ell}\|\xi^m\|_{L^4(\mathcal{T}_h)}^4 - \frac{Ck}{\epsilon^2}\sum_{m=1}^{\ell}\|\xi^m\|_{L^3(\mathcal{T}_h)}^3 \\
&\quad - k\sum_{m=1}^{\ell}\|\xi^m\|_{L^2(\mathcal{T}_h)}^2 - C\Big(h^{2\min\{r+1,s\}}\epsilon^{-4}\|u\|_{L^2((0,T);H^s(\mathcal{D})}^2 + k^2\epsilon^{-4}J_\epsilon^h(u_h^0)\Big).
\end{aligned}
$$

Here we have used the fact that $|P_r^h u(t_m)| \leq C$ and (2.64).

Substituting the above estimate into (2.100) yields

$$
\begin{aligned}
&\|\xi^\ell\|_{L^2(\mathcal{T}_h)}^2 + 2k\sum_{m=1}^{\ell}k\|d_t\xi^m\|_{L^2(\mathcal{T}_h)}^2 + \frac{2}{\epsilon^2}k\sum_{m=1}^{\ell}\|\xi^m\|_{L^4(\mathcal{T}_h)}^4 \\
&\qquad + 2k\sum_{m=1}^{\ell}\left(a_h(\xi^m,\xi^m) + \frac{1}{\epsilon^2}\Big(f'\big(P_r^h u(t_m)\big), (\xi^m)^2\Big)_{\mathcal{T}_h}\right) \\
&\leq 2k\sum_{m=1}^{\ell}\|\xi^m\|_{L^2(\mathcal{T}_h)}^2 + \frac{Ck}{\epsilon^2}\sum_{m=1}^{\ell}\|\xi^m\|_{L^3(\mathcal{T}_h)}^3 \\
&\qquad + 2\|\xi^0\|_{L^2(\mathcal{T}_h)}^2 + Ck^2\Big(\epsilon^{-2\max\{\sigma_1+2,\sigma_3\}} + \epsilon^{-4}J_\epsilon^h(u_h^0)\Big) \\
&\qquad + Ch^{2\min\{r+1,s\}}\Big(\|u\|_{H^1((0,T);H^s(\mathcal{D}))}^2 + \epsilon^{-4}\|u\|_{L^2((0,T);H^s(\mathcal{D})}^2\Big).
\end{aligned}
\tag{2.101}
$$

Step 3: To control the second term on the right-hand side of (2.101) we use the following Gagliardo-Nirenberg inequality [2]:

$$
\|v\|_{L^3(K)}^3 \leq C\|\nabla v\|_{L^2(K)}^{\frac{d}{2}}\big\|v\big\|_{L^2(K)}^{\frac{6-d}{2}} \qquad \forall K \in \mathcal{T}_h
$$

to get

$$\frac{Ck}{\epsilon^2}\sum_{m=1}^{\ell}\|\xi^m\|_{L^3(\mathcal{T}_h)}^3 \le \epsilon^2\alpha k\sum_{m=1}^{\ell}\|\nabla\xi^m\|_{L^2(\mathcal{T}_h)}^2 + C\epsilon^{-\frac{2(4+d)}{4-d}}k\sum_{m=1}^{\ell}\sum_{K\in\mathcal{T}_h}\left\|\xi^m\right\|_{L^2(K)}^{\frac{2(6-d)}{4-d}} \tag{2.102}$$

$$\le \epsilon^2\alpha k\sum_{m=1}^{\ell}\|\nabla\xi^m\|_{L^2(\mathcal{T}_h)}^2 + C\epsilon^{-\frac{2(4+d)}{4-d}}k\sum_{m=1}^{\ell}\left\|\xi^m\right\|_{L^2(\mathcal{T}_h)}^{\frac{2(6-d)}{4-d}}. \tag{2.103}$$

Finally, for the fourth term on the left-hand side of (2.101) we utilize the discrete spectrum estimate (2.78) to bound it from below as follows:

$$\begin{aligned}
&2k\sum_{m=1}^{\ell}\left(a_h(\xi^m,\xi^m)+\frac{1}{\epsilon^2}\Big(f'\big(P_r^h u(t_m)\big),(\xi^m)^2\Big)_{\mathcal{T}_h}\right) \qquad (2.104)\\
&= 2(1-\epsilon^2)k\sum_{m=1}^{\ell}\left(a_h(\xi^m,\xi^m)+\frac{1}{\epsilon^2}\Big(f'\big(P_r^h u(t_m)\big),(\xi^m)^2\Big)_{\mathcal{T}_h}\right)\\
&\quad + 2\epsilon^2 k\sum_{m=1}^{\ell}\left(a_h(\xi^m,\xi^m)+\frac{1}{\epsilon^2}\Big(f'\big(P_r^h u(t_m)\big),(\xi^m)^2\Big)_{\mathcal{T}_h}\right)\\
&\ge -2(1-\epsilon^2)c_0 k\sum_{m=1}^{\ell}\|\xi^m\|_{L^2(\mathcal{T}_h)}^2 + 4\epsilon^2\alpha k\sum_{m=1}^{\ell}\|\xi^m\|_{1,\mathrm{DG}}^2 - Ck\sum_{m=1}^{\ell}\|\xi^m\|_{L^2(\mathcal{T}_h)}^2,
\end{aligned}$$

where we have used (2.63) and (2.61) to get the second term on the right-hand side.

Step 4: Substituting (2.102) and (2.104) into (2.101) we get

$$\begin{aligned}
&\|\xi^\ell\|^2_{L^2(\mathcal{T}_h)} + k\sum_{m=1}^{\ell}\Big(2k\|d_t\xi^m\|^2_{L^2(\mathcal{T}_h)} + 3\epsilon^2\alpha\|\xi^m\|^2_{1,\mathrm{DG}}\Big) \qquad (2.105)\\
&\quad\le C(1+c_0)k\sum_{m=1}^{\ell}\|\xi^m\|^2_{L^2(\mathcal{T}_h)} + C\epsilon^{-\frac{2(4+d)}{4-d}}k\sum_{m=1}^{\ell}\big\|\xi^m\big\|^{\frac{2(6-d)}{4-d}}_{L^2(\mathcal{T}_h)}\\
&\qquad + 2\|\xi^0\|^2_{L^2(\mathcal{T}_h)} + Ck^2\Big(\epsilon^{-2\max\{\sigma_1+2,\sigma_3\}} + \epsilon^{-4}J^h_\epsilon(u^0_h)\Big)\\
&\qquad + Ch^{2\min\{r+1,s\}}\Big(\|u\|^2_{H^1((0,T);H^s(\mathcal{D}))} + \epsilon^{-4}\|u\|^2_{L^2((0,T);H^s(\mathcal{D}))}\Big).
\end{aligned}$$

At this point, notice that there are two terms on the right-hand side of (2.105) that involve the approximated initial datum u^0_h. On one hand, we need to choose u^0_h such that $\|\xi^0\|_{L^2(\mathcal{T}_h)} = O(h^{\min\{r+1,s\}})$ to maintain the optimal rate of convergence in h. Clearly, both the L^2 and the elliptic projection of u_0 will work. In fact, in the latter case, $\xi^0 = 0$. On the other hand, we want $J^h_\epsilon(u^0_h)$ to be uniformly bounded in h. but the jump term in $J^h_\epsilon(u^0_h)$ always depend on h unless it vanishes. To satisfy this requirement, we ask $u^0_h \in S_h$. Therefore, we are led to choose u^0_h to be the L^2 or the elliptic projection of u_0 into the finite element space S_h.

It then follows from (2.105) and (ii), (iv), (vii) in Proposition 2.2.1 that

$$\begin{aligned}
&\|\xi^\ell\|^2_{L^2(\mathcal{T}_h)} + k\sum_{m=1}^{\ell}\Big(2k\|d_t\xi^m\|^2_{L^2(\mathcal{T}_h)} + 3\epsilon^2\alpha\|\xi^m\|^2_{1,\mathrm{DG}}\Big) \qquad (2.106)\\
&\quad\le C(1+c_0)k\sum_{m=1}^{\ell}\|\xi^m\|^2_{L^2(\mathcal{T}_h)} + C\epsilon^{-\frac{2(4+d)}{4-d}}k\sum_{m=1}^{\ell}\big\|\xi^m\big\|^{\frac{2(6-d)}{4-d}}_{L^2(\mathcal{T}_h)}\\
&\qquad + Ck^2\epsilon^{-2(\sigma_1+2)} + Ch^{2\min\{r+1,s\}}\epsilon^{-2(\sigma_1+2)}.
\end{aligned}$$

On noting that u^ℓ_h can be written as

$$u^\ell_h = k\sum_{m=1}^{\ell} d_t u^m_h + u^0_h, \qquad (2.107)$$

then by (2.1) and (2.64), we get

$$\|u_h^\ell\|_{L^2(\mathcal{T}_h)} \le k\sum_{m=1}^{\ell}\|d_t u_h^m\|_{L^2(\mathcal{T}_h)} + \|u_h^0\|_{L^2(\mathcal{T}_h)} \le C\epsilon^{-2\sigma_1}. \tag{2.108}$$

By the boundedness of the projection, we have

$$\|\xi^\ell\|^2_{L^2(\mathcal{T}_h)} \le C\epsilon^{-2\sigma_1}. \tag{2.109}$$

Then (2.106) can be reduced to

$$\|\xi^\ell\|^2_{L^2(\mathcal{T}_h)} + k\sum_{m=1}^{\ell}\Big(2k\|d_t\xi^m\|^2_{L^2(\mathcal{T}_h)} + 3\epsilon^2\alpha\|\xi^m\|^2_{1,\mathrm{DG}}\Big) \le M_1 + M_2, \tag{2.110}$$

where

$$\begin{aligned} M_1 :&= C(1+c_0)k\sum_{m=1}^{\ell-1}\|\xi^m\|^2_{L^2(\mathcal{T}_h)} + C\epsilon^{-\frac{2(4+d)}{4-d}}k\sum_{m=1}^{\ell-1}\big\|\xi^m\big\|^{\frac{2(6-d)}{4-d}}_{L^2(\mathcal{T}_h)} \\ &\quad + Ck^2\epsilon^{-2(\sigma_1+2)} + Ch^{2\min\{r+1,s\}}\epsilon^{-2(\sigma_1+2)}, \end{aligned} \tag{2.111}$$

$$M_2 := C(1+c_0)k\|\xi^\ell\|^2_{L^2(\mathcal{T}_h)} + C\epsilon^{-\frac{2(4+d)}{4-d}}k\big\|\xi^\ell\big\|^{\frac{2(6-d)}{4-d}}_{L^2(\mathcal{T}_h)}. \tag{2.112}$$

It is easy to check that

$$M_2 < \frac{1}{2}\|\xi^\ell\|^2_{L^2(\mathcal{T}_h)} \qquad \text{provided that} \quad k < C\epsilon^{\frac{8+2d+4\sigma_1}{4-d}}. \tag{2.113}$$

By (2.110) we have

$$\begin{aligned}
&\|\xi^\ell\|^2_{L^2(\mathcal{T}_h)} + k\sum_{m=1}^{\ell}\Big(2k\|d_t\xi^m\|^2_{L^2(\mathcal{T}_h)} + 3\epsilon^2\alpha\|\xi^m\|^2_{1,\mathrm{DG}}\Big) \le 2M_1 \qquad (2.114)\\
&\quad = 2C(1+c_0)k\sum_{m=1}^{\ell-1}\|\xi^m\|^2_{L^2(\mathcal{T}_h)} + 2C\epsilon^{-\frac{2(4+d)}{4-d}}k\sum_{m=1}^{\ell-1}\big\|\xi^m\big\|^{\frac{2(6-d)}{4-d}}_{L^2(\mathcal{T}_h)}\\
&\qquad + 2Ck^2\epsilon^{-2(\sigma_1+2)} + Ch^{2\min\{r+1,s\}}\epsilon^{-2(\sigma_1+2)}\\
&\quad \le C(1+c_0)k\sum_{m=1}^{\ell-1}\|\xi^m\|^2_{L^2(\mathcal{T}_h)} + C\epsilon^{-\frac{2(4+d)}{4-d}}k\sum_{m=1}^{\ell-1}\big\|\xi^m\big\|^{\frac{2(6-d)}{4-d}}_{L^2(\mathcal{T}_h)}\\
&\qquad + Ck^2\epsilon^{-2(\sigma_1+2)} + Ch^{2\min\{r+1,s\}}\epsilon^{-2(\sigma_1+2)}.
\end{aligned}$$

Let $d_\ell \ge 0$ be the slack variable such that

$$\begin{aligned}
&\|\xi^\ell\|^2_{L^2(\mathcal{T}_h)} + k\sum_{m=1}^{\ell}\Big(2k\|d_t\xi^m\|^2_{L^2(\mathcal{T}_h)} + 3\epsilon^2\alpha\|\xi^m\|^2_{1,\mathrm{DG}}\Big) + d_\ell \qquad (2.115)\\
&\quad = C(1+c_0)k\sum_{m=1}^{\ell-1}\|\xi^m\|^2_{L^2(\mathcal{T}_h)} + C\epsilon^{-\frac{2(4+d)}{4-d}}k\sum_{m=1}^{\ell-1}\big\|\xi^m\big\|^{\frac{2(6-d)}{4-d}}_{L^2(\mathcal{T}_h)}\\
&\qquad + Ck^2\epsilon^{-2(\sigma_1+2)} + Ch^{2\min\{r+1,s\}}\epsilon^{-2(\sigma_1+2)},
\end{aligned}$$

and define for $\ell \ge 1$

$$S_{\ell+1} := \|\xi^\ell\|^2_{L^2(\mathcal{T}_h)} + k\sum_{m=1}^{\ell}\Big(2k\|d_t\xi^m\|^2_{L^2(\mathcal{T}_h)} + 3\epsilon^2\alpha\|\xi^m\|^2_{1,\mathrm{DG}}\Big) + d_\ell, \qquad (2.116)$$

$$S_1 := Ck^2\epsilon^{-2(\sigma_1+2)} + Ch^{2\min\{r+1,s\}}\epsilon^{-2(\sigma_1+2)}, \qquad (2.117)$$

then we have

$$S_{\ell+1} - S_\ell \le C(1+c_0)kS_\ell + C\epsilon^{-\frac{2(4+d)}{4-d}}kS_\ell^{\frac{6-d}{4-d}} \qquad \text{for } \ell \ge 1. \qquad (2.118)$$

Applying Lemma 2.2.4 to $\{S_\ell\}_{\ell\geq 1}$ defined above, we obtain for $\ell \geq 1$

$$S_\ell \leq a_\ell^{-1}\left\{ S_1^{-\frac{2}{4-d}} - \frac{2Ck}{4-d}\sum_{s=1}^{\ell-1} \epsilon^{-\frac{2(4+d)}{4-d}} a_{s+1}^{-\frac{2}{4-d}} \right\}^{-\frac{4-d}{2}} \tag{2.119}$$

provided that

$$\frac{1}{2} S_1^{-\frac{2}{4-d}} - \frac{2Ck}{4-d}\sum_{s=1}^{\ell-1} \epsilon^{-\frac{2(4+d)}{4-d}} a_{s+1}^{-\frac{2}{4-d}} > 0. \tag{2.120}$$

We note that $a_s\,(1 \leq s \leq \ell)$ are all bounded as $k \to 0$, therefore, (2.120) holds under the mesh constraint stated in the theorem. It follows from (2.119) and (2.120) that

$$S_\ell \leq 2a_\ell^{-1} S_1 \leq Ck^2\epsilon^{-2(\sigma_1+2)} + Ch^{2\min\{r+1,s\}}\epsilon^{-2(\sigma_1+2)}. \tag{2.121}$$

Finally, using the above estimate and the properties of the operator P_r^h we obtain (2.94) and (2.95). The estimate (2.96) follows from (2.95) and the inverse inequality bounding the L^∞-norm by the L^2-norm and (2.81). The proof is complete. □

2.4 Convergence of the numerical interface to the mean curvature flow

In this section, we establish the convergence and rate of convergence of the numerical interface $\Gamma_t^{\epsilon,h,k}$, which is defined as the zero-level set of the numerical solution $\{u_h^n\}$ (see the precise definition below), to the sharp interface limit (the mean curvature flow) of the Allen-Cahn equation. The key ingredient of the proof is the $L^\infty(J;L^\infty)$ error estimate obtained in the previous section, which depends on ϵ^{-1} in a low polynomial order. It is proved that the numerical interface converges with the rate $O(\epsilon^2|\ln\epsilon|^2)$ before the singularities appear. We note that the proof to be given below essentially follows the same lines as in the proof of [45]. For the reader's convenience, we provide here a self-contained proof. Throughout this section, u^ϵ denotes the solution of the Allen-Cahn problem (1.24)–(1.27).

We notice that, unlike in the finite element case, the DG solution u_h^n is discontinuous in space (and in time). As a result, the zero-level set of u_h^n may not be well defined. To circumvent this technicality, we introduce the finite element approximation $\widehat{u}_h^m$ of u_h^m which is defined using the averaged degrees of freedom of u_h^n as the degrees of freedom for determining $\widehat{u}_h^m$ (cf. [58]). The following approximation result was proved in Theorem 2.1 in [58].

Theorem 2.4.1. *Let $\mathcal{T}_h$ be a conforming mesh consisting of triangles when $d = 2$, and tetrahedra when $d = 3$. For $v_h \in V_h$, let $\widehat{v}_h$ be the finite element approximation of v_h as defined above. Then for any $v_h \in V_h$ and $i = 0, 1$ there holds*

$$\sum_{K\in\mathcal{T}_h} \|v_h - \widehat{v}_h\|^2_{H^i(K)} \le C \sum_{e\in\mathcal{E}_h^I} h_e^{1-2i} \|[v_h]\|^2_{L^2(e)}, \tag{2.122}$$

where $C > 0$ is a constant independent of h and v_h but may depend on r and the minimal angle θ_0 of the triangles in $\mathcal{T}_h$.

Using the above approximation result we can show that the error estimates of Theorem 2.3.11 also hold for $\widehat{u}_h^n$.

Theorem 2.4.2. *Let u_h^m denote the solution of the DG scheme (2.51)–(2.54) and $\widehat{u}_h^m$ denote its finite element approximation as defined above. Then under the assumptions of Theorem 2.3.11 the error estimates for u_h^m given in Theorem 2.3.11 are still valid for $\widehat{u}_h^m$, in particular, there holds*

$$\begin{aligned}\max_{0\le m\le M} \|u(t_m) - \widehat{u}_h^m\|_{L^\infty(\mathcal{T}_h)} \le{}& Ch^{\min\{r+1,s\}} |\ln h|^{\overline{r}} \epsilon^{-\gamma} \\ &+ Ch^{-\frac{d}{2}} (k + h^{\min\{r+1,s\}}) \epsilon^{-(\sigma_1+2)}.\end{aligned} \tag{2.123}$$

Proof. We only give a proof for (2.123) because other estimates can be proved likewise. By the triangle inequality we have

$$\|u(t_m) - \widehat{u}_h^m\|_{L^\infty(\mathcal{T}_h)} \le \|u(t_m) - u_h^m\|_{L^\infty(\mathcal{T}_h)} + \|u_h^m - \widehat{u}_h^m\|_{L^\infty(\mathcal{T}_h)}. \tag{2.124}$$

Hence, it suffices to show that the second term on the right-hand side is an equal or higher order term compared to the first one.

Let $u^I(t)$ denote the finite element interpolation of $u(t)$ into S_h. It follows from (2.122) and the trace inequality that

$$\begin{aligned}\|u_h^m - \widehat{u}_h^m\|_{L^2(\mathcal{T}_h)}^2 &\le C\sum_{e\in\mathcal{E}_h^I} h_e\|[u_h^m]\|_{L^2(e)}^2 \\ &= C\sum_{e\in\mathcal{E}_h^I} h_e\|[u_h^m - u^I(t_m)]\|_{L^2(e)}^2 \\ &\le C\sum_{K\in\mathcal{T}_h} h_e h_K^{-1}\|u_h^m - u^I(t_m)\|_{L^2(K)}^2 \\ &\le C\big(\|u_h^m - u(t_m)\|_{L^2(\mathcal{T}_h)}^2 + \|u(t_m) - u^I(t_m)\|_{L^2(\mathcal{T}_h)}^2\big).\end{aligned} \tag{2.125}$$

Substituting (2.125) into (2.124) after using the inverse inequality yields

$$\begin{aligned}\|u(t_m) - \widehat{u}_h^m\|_{L^\infty(\mathcal{T}_h)} &\le \|u(t_m) - u_h^m\|_{L^\infty(\mathcal{T}_h)} + Ch^{-\frac{d}{2}}\|u_h^m - \widehat{u}_h^m\|_{L^2(\mathcal{T}_h)} \\ &\le \|u(t_m) - u_h^m\|_{L^\infty(\mathcal{T}_h)} \\ &\quad + Ch^{-\frac{d}{2}}\big(\|u_h^m - u(t_m)\|_{L^2(\mathcal{T}_h)} + \|u(t_m) - u^I(t_m)\|_{L^2(\mathcal{T}_h)}\big),\end{aligned}$$

which together with (2.94) implies the desired estimate (2.123). The proof is complete. □

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

Theorem 2.4.3. *Let $\{\Gamma_t\}$ denote the (generalized) mean curvature flow defined in [33], that is, Γ_t is the zero-level set of the solution w of the following initial value problem:*

$$w_t = \Delta w - \frac{D^2 w Dw\cdot Dw}{|Dw|^2} \qquad \text{in } \mathbf{R}^d\times(0,\infty), \tag{2.126}$$

$$w(\cdot,0) = w_0(\cdot) \qquad \text{in } \mathbf{R}^d. \tag{2.127}$$

Let $u^{\epsilon,h,k}$ denote the piecewise linear interpolation in time of the numerical solution $\{\widehat{u}_h^m\}$ defined by

$$u^{\epsilon,h,k}(x,t) := \frac{t-t_m}{k}\widehat{u}_h^{m+1}(x) + \frac{t_{m+1}-t}{k}\widehat{u}_h^m(x), \quad t_m \le t \le t_{m+1} \tag{2.128}$$

for $0 \le m \le M-1$. Let $\{\Gamma_t^{\epsilon,h,k}\}$ denote the zero-level set of $u^{\epsilon,h,k}$, namely,

$$\Gamma_t^{\epsilon,h,k} = \{x \in \mathcal{D};\, u^{\epsilon,h,k}(x,t) = 0\}. \tag{2.129}$$

Suppose $\Gamma_0 = \{x \in \overline{\mathcal{D}}; u_0(x) = 0\}$ is a smooth hypersurface compactly contained in $\mathcal{D}$, and $k = O(h^2)$. Let t_* be the first time at which the mean curvature flow develops a singularity, then there exists a constant $\epsilon_1 > 0$ such that for all $\epsilon \in (0, \epsilon_1)$ and $0 < t < t_*$ there holds

$$\sup_{x \in \Gamma_t^{\epsilon,h,k}} \{\mathrm{dist}(x, \Gamma_t)\} \le C\epsilon^2 |\ln \epsilon|^2.$$

Proof. We note that since $u^{\epsilon,h,k}(x,t)$ is continuous in both t and x, then $\Gamma_t^{\epsilon,h,k}$ is well defined. Let I_t and O_t denote the inside and the outside of Γ_t defined by

$$I_t := \{x \in \mathbf{R}^d;\, w(x,t) > 0\}, \qquad O_t := \{x \in \mathbf{R}^d;\, w(x,t) < 0\}. \tag{2.130}$$

Let $d(x,t)$ denote the signed distance function to Γ_t which is positive in I_t and negative in O_t. By Theorem 6.1 of [9], there exist $\widehat{\epsilon}_1 > 0$ and $\widehat{C}_1 > 0$ such that for all $t \ge 0$ and $\epsilon \in (0, \widehat{\epsilon}_1)$ there hold

$$u_\epsilon(x,t) \ge 1 - \epsilon \qquad \forall x \in \{x \in \overline{\mathcal{D}};\, d(x,t) \ge \widehat{C}_1\epsilon^2 |\ln \epsilon|^2\}, \tag{2.131}$$

$$u_\epsilon(x,t) \le -1 + \epsilon \qquad \forall x \in \{x \in \overline{\mathcal{D}};\, d(x,t) \le -\widehat{C}_1\epsilon^2 |\ln \epsilon|^2\}. \tag{2.132}$$

Since for any fixed $x \in \Gamma_t^{\epsilon,h,k}$, $u^{\epsilon,h,k}(x,t) = 0$, by (2.123) with $k = O(h^2)$, we have

$$\begin{aligned} |u^\epsilon(x,t)| &= |u^\epsilon(x,t) - u^{\epsilon,h,k}(x,t)| \\ &\leq \tilde{C}\Big(h^{\min\{r+1,s\}}|\ln h|^{\overline{r}}\epsilon^{-\gamma} + h^{-\frac{d}{2}}(k + h^{\min\{r+1,s\}})\epsilon^{-(\sigma_1+2)}\Big). \end{aligned}$$

Then there exists $\widetilde{\epsilon}_1 > 0$ such that for $\epsilon \in (0, \widetilde{\epsilon}_1)$ there holds

$$|u^\epsilon(x,t)| < 1 - \epsilon. \tag{2.133}$$

Therefore, the assertion follows from setting $\epsilon_1 = \min\{\widehat{\epsilon}_1, \widetilde{\epsilon}_1\}$. The proof is complete. □

2.5 Numerical experiments

In this section, we present three two-dimensional numerical tests to gauge the performance of the proposed fully discrete IP-DG method with $r = 1$. All tests are done on the square domain $\mathcal{D} = [-1,1]^2$ and $u_0(x) = \tanh\Big(\frac{d_0(x)}{\sqrt{2}\epsilon}\Big)$, where $d_0(x)$ stands for the signed distance from x to the initial curve Γ_0.

The first test uses a smooth initial curve Γ_0, hence the requirements for u_0 are satisfied. Consequently, the results established in this chapter apply to this test example. In the test we first verify the spatial rate of convergence given in (2.94) and (2.95), and the decay of the energy $J_\epsilon^h(u_h^\ell)$ defined in (2.64) using $\epsilon = 0.1$. As expected, the energy decreases monotonically during the whole evolution. We then compute the evolution of the zero-level set of the solution of the Allen-Cahn problem with $\epsilon = 0.125, 0.025, 0.005, 0.001$ and at various time instances.

On the other hand, the second and third tests use non-smooth initial curve Γ_0, so u_0 defined above is not smooth anymore, hence the theoretical results of this chapter may not apply to these two cases. Nevertheless, we still use our DG method to compute the solutions, the energy decay as well as the evolution of the zero-level sets

of the solutions of these two test problems. The numerical results suggest that the proposed DG method still works well in these two cases where a convergence theory is missing.

Test 1. Consider the Allen-Cahn problem with the following initial condition:

$$u_0(x) = \begin{cases} \tanh(\frac{d(x)}{\sqrt{2}\epsilon}), & \text{if } \frac{x_1^2}{0.36} + \frac{x_2^2}{0.04} \geq 1, \\ \tanh(\frac{-d(x)}{\sqrt{2}\epsilon}), & \text{if } \frac{x_1^2}{0.36} + \frac{x_2^2}{0.04} < 1. \end{cases}$$

Here $d(x)$ stands for the distance function to the ellipse $\frac{x_1^2}{(0.6)^2} + \frac{x_2^2}{(0.2)^2} = 1$.

Table 2.1: Spatial errors and convergence rates of Test 1.

	$L^\infty(L^2)$ error	$L^\infty(L^2)$ order	$L^2(H^1)$ error	$L^2(H^1)$ order
$h = 0.4\sqrt{2}$	0.28704		1.22726	
$h = 0.2\sqrt{2}$	0.10847	1.4040	0.69576	0.8188
$h = 0.1\sqrt{2}$	0.02146	2.3376	0.32855	1.0825
$h = 0.05\sqrt{2}$	0.00511	2.0703	0.16448	0.9982
$h = 0.025\sqrt{2}$	0.00129	1.9860	0.08230	0.9989

Table 2.1 shows the spatial L^2 and H^1-norm errors and convergence rates, which are consistent with what are proved for the linear element in the convergence theorem. Figure 2.1 plots the change of the discrete energy $J_\epsilon^h(u_h^\ell)$ in time. This graph clearly confirms the energy decay property.

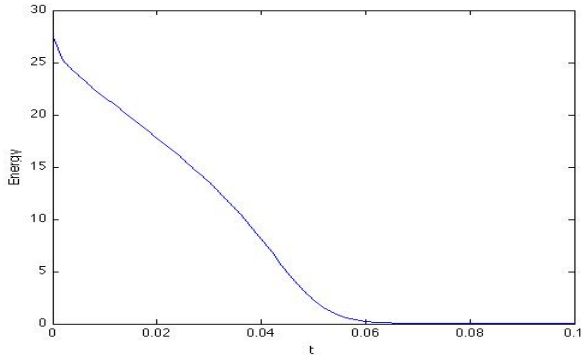

Figure 2.1: Decay of the numerical energy $J_\epsilon^h(u_h^\ell)$ of Test 1.

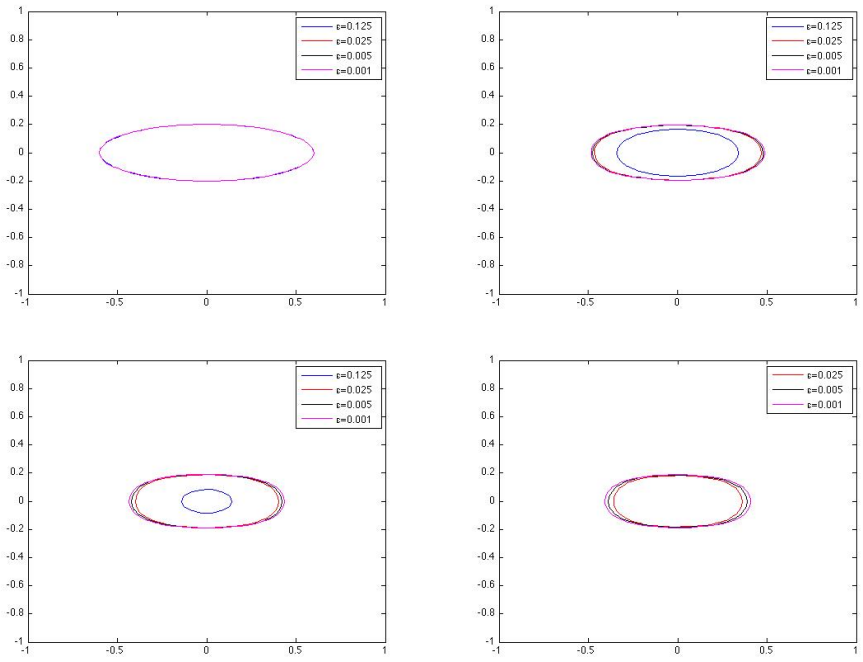

Figure 2.2: Test 1: Snapshots of the zero-level set of $u^{\epsilon,h,k}$ at time $t = 0,2 \times 10^{-2}, 3.2 \times 10^{-2}, 4 \times 10^{-2}$ and $\epsilon = 0.125, 0.025, 0.005, 0.001$.

Figure 2.2 displays four snapshots at four fixed time points of the zero-level set of the numerical solution $u^{\epsilon,h,k}$ with four different ϵ. Once again, we observe that at each time point the zero-level set converges to the mean curvature flow Γ_t as ϵ tends to zero, and the zero-level set evolves faster in time for larger ϵ.

Test 2. This test considers a case with nonsmooth initial curve Γ_0 which encloses a dumbbell-shaped domain. To explicitly define the desired initial function, we introduce the following functions:

$$\tanh(x) := \frac{e^x - e^{-x}}{e^x + e^{-x}}, \qquad \psi_1(y) := \frac{-1 + \sqrt{0.8y + 0.04}}{2},$$
$$\psi_2(y) := \frac{1 - \sqrt{1.92y + 0.2304}}{2}, \qquad \psi_3(y) := \frac{-1 + \sqrt{-0.8y + 0.04}}{2},$$
$$\psi_4(y) := \frac{1 - \sqrt{-1.92y + 0.2304}}{2}, \qquad \psi_5(y) := -\sqrt{\frac{1 - 0.2451y^2}{0.0049}}.$$

We then consider the Allen-Cahn problem (1.24)–(1.27) with the following initial condition:

$$u_0(x,y) = \begin{cases} \tanh(\frac{1}{\sqrt{2}\epsilon}(-\sqrt{(x-0.14)^2+(y-0.15)^2})), & \text{if } x > 0.14, 0 \le y < -\frac{5}{12}(x-0.5), \\ \tanh(\frac{1}{\sqrt{2}\epsilon}(-\sqrt{(x-0.14)^2+(y+0.15)^2})), & \text{if } x > 0.14, \frac{5}{12}(x-0.5) < y < 0, \\ \tanh(\frac{1}{\sqrt{2}\epsilon}(-\sqrt{(x+0.3)^2+(y-0.15)^2})), & \text{if } x < -0.3, 0 \le y < \frac{3}{4}(x+0.5), \\ \tanh(\frac{1}{\sqrt{2}\epsilon}(-\sqrt{(x+0.3)^2+(y+0.15)^2})), & \text{if } x < -0.3, -\frac{3}{4}(x+0.5) < y < 0, \\ \tanh(\frac{1}{\sqrt{2}\epsilon}(\sqrt{(x-0.5)^2+y^2}-0.39)), & \text{if } x > 0.14, y \ge -\frac{5}{12}(x-0.5) \\ & \text{or } y \le \frac{5}{12}(x-0.5), \\ \tanh(\frac{1}{\sqrt{2}\epsilon}(\sqrt{(x+0.5)^2+y^2}-0.25)), & \text{if } x < -0.3, y \ge -\frac{3}{4}(x+0.5) \\ & \text{or } y \le -\frac{3}{4}(x+0.5), \\ \tanh(\frac{1}{\sqrt{2}\epsilon}(|y|-0.15)), & \text{if } -0.3 \le x \le 0.14, \\ & \psi_1(y) \le x \le \psi_2(y) \\ & \text{and } \psi_3(y) \le x \le \psi_4(y), \\ \tanh(\frac{1}{\sqrt{2}\epsilon}(\sqrt{(x-0.5)^2+y^2}-0.39)), & \text{if } -0.3 \le x \le 0.14,\ x \ge \psi_2(y) \\ & \text{and } x \ge \psi_5(y), \\ \tanh(\frac{1}{\sqrt{2}\epsilon}(\sqrt{(x-0.5)^2+y^2}-0.39)), & \text{if } -0.3 \le x \le 0.14,\ x \ge \psi_4(y) \\ & \text{and } x \ge \psi_5(y), \\ \tanh(\frac{1}{\sqrt{2}\epsilon}(\sqrt{(x+0.5)^2+y^2}-0.25)), & \text{if } -0.3 \le x \le 0.14,\ x \le \psi_1(y) \\ & \text{and } x \le \psi_5(y), \\ \tanh(\frac{1}{\sqrt{2}\epsilon}(\sqrt{(x+0.5)^2+y^2}-0.25)), & \text{if } -0.3 \le x \le 0.14,\ x \le \psi_3(y) \\ & \text{and } x \le \psi_5(y). \end{cases}$$

We note that u_0 can be rewritten as

$$u_0 = \tanh\Big(\frac{d_0(x)}{\sqrt{2}\epsilon}\Big).$$

Since Γ_0 contains conner points, it is only Lipschitz. Then u_0 is not smooth, hence, it does not satisfy the assumptions of Proposition 2.2.2. As a result, the convergence theorem of this chapter may not apply to this case. Nevertheless, the numerical results given in Table 2.2 show that the spatial L^2 and H^1-norm errors and convergence rates are still consistent with what are proved for the linear element in the convergence theorem.

Table 2.2: Spatial errors and convergence rates of Test 2.

	$L^\infty(L^2)$ error	$L^\infty(L^2)$ order	$L^2(H^1)$ error	$L^2(H^1)$ order
$h = 0.4\sqrt{2}$	0.20604		0.95123	
$h = 0.2\sqrt{2}$	0.04598	2.1638	0.49633	0.9385
$h = 0.1\sqrt{2}$	0.01330	1.7896	0.25199	0.9779
$h = 0.05\sqrt{2}$	0.00372	1.8381	0.12686	0.9901
$h = 0.025\sqrt{2}$	0.00098	1.9244	0.06350	0.9984

Figure 2.3 plots the change of the discrete energy $J_\epsilon^h(u_h^\ell)$ in time, which should decrease according to (2.64). This graph clearly confirms this decay property.

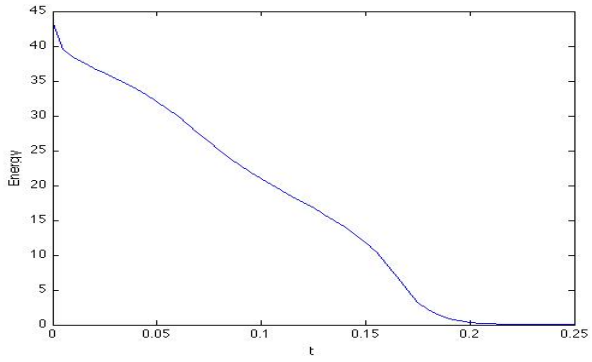

Figure 2.3: Decay of the numerical energy $J_\epsilon^h(u_h^\ell)$ of Test 2.

Figure 2.4 displays four snapshots at four fixed time points of the zero-level set of the numerical solution $u^{\epsilon,h,k}$ with four different ϵ. They clearly indicate that at each time point the zero-level set converges to the mean curvature flow Γ_t as ϵ tends to zero. It also shows that the zero-level set evolves faster in time for larger ϵ.

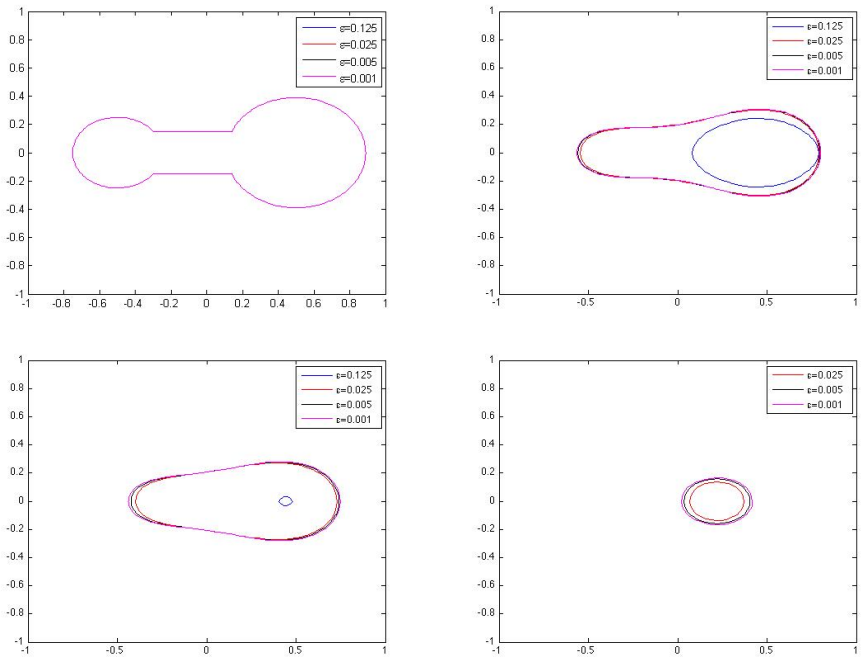

Figure 2.4: Test 2: Snapshots of the zero-level set of $u^{\epsilon,h,k}$ at time $t = 0, 0.06, 0.09, 0.2$ and $\epsilon = 0.125, 0.025, 0.005, 0.001$.

Test 3. Consider the Allen-Cahn problem (1.24)–(1.27) with the following initial condition:

$$u_0(x) = \begin{cases} \tanh(\frac{1}{\sqrt{2}\epsilon}(\min\{d_1(x), d_2(x)\})), & \text{if } \frac{x_1^2}{0.04} + \frac{x_2^2}{0.36} \geq 1, \frac{x_1^2}{0.36} + \frac{x_2^2}{0.04} \geq 1, \\ & \text{or } \frac{x_1^2}{0.04} + \frac{x_2^2}{0.36} \leq 1, \frac{x_1^2}{0.36} + \frac{x_2^2}{0.04} \leq 1, \\ \tanh(\frac{1}{\sqrt{2}\epsilon}(-\min\{d_1(x), d_2(x)\})), & \text{if } \frac{x_1^2}{0.04} + \frac{x_2^2}{0.36} < 1, \frac{x_1^2}{0.36} + \frac{x_2^2}{0.04} > 1, \\ & \text{or } \frac{x_1^2}{0.04} + \frac{x_2^2}{0.36} > 1, \frac{x_1^2}{0.36} + \frac{x_2^2}{0.04} < 1. \end{cases}$$

Here $d_1(x)$ and $d_2(x)$ stand for, respectively, the distance functions to the two ellipses. Obviously, the above Γ_0 is not smooth, moreover, it contains four self-intersection points. A topological change (i.e., a singularity) is expected to occur instantaneously in such a case. Figure 2.5 displays four snapshots at four fixed time points of the zero-level set of the numerical solution $u^{\epsilon,h,k}$ with four different ϵ. It clearly shows how the pinch-off occurs for this self-intersected curve under the mean curvature flow.

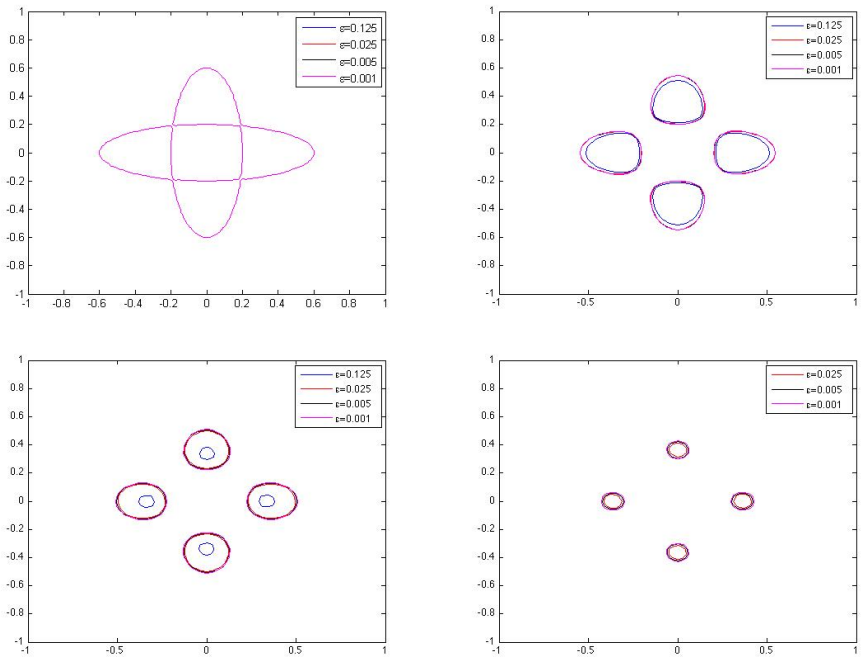

Figure 2.5: Test 3: Snapshots of the zero-level set of $u^{\epsilon,h,k}$ at time $t = 0,6 \times 10^{-3}, 1.2 \times 10^{-2}, 2 \times 10^{-2}$ and $\epsilon = 0.125, 0.025, 0.005, 0.001$.

We also compute the spatial L^2 and H^1-norm errors and convergence rates in Table 2.3, they are consistent with what are proved for the linear element in the convergence theorem although the theorem does not cover this case. Figure 2.6 plots the change of the discrete energy $J_\epsilon^h(u_h^\ell)$ in time. The graph not only confirms the energy decay property but also reveals the rapid decay of the energy at the beginning of the evolution, which is caused by the singularity.

Table 2.3: Spatial errors and convergence rates of Test 3.

	$L^\infty(L^2)$ error	$L^\infty(L^2)$ order	$L^2(H^1)$ error	$L^2(H^1)$ order
$h = 0.4\sqrt{2}$	0.09186		0.29686	
$h = 0.2\sqrt{2}$	0.03670	1.3237	0.16331	0.8622
$h = 0.1\sqrt{2}$	0.00911	2.0103	0.07603	1.1030
$h = 0.05\sqrt{2}$	0.00276	1.7228	0.03740	1.0235
$h = 0.025\sqrt{2}$	0.00071	1.9588	0.01846	1.0186

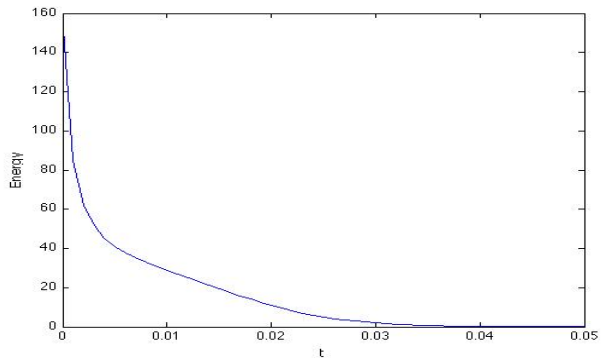

Figure 2.6: Decay of the numerical energy $J_\epsilon^h(u_h^\ell)$ of Test 3.

Chapter 3

Discontinuous Galerkin Methods for the Cahn-Hilliard Equation

3.1 Introduction

The Cahn-Hilliard equation plays an important role in materials phase transition, and it also has been extensively studied due to its close relation with the Hele-Shaw problem. It was first formally proved by Pego [74] that the chemical potential $w := -\epsilon\Delta u + \frac{1}{\epsilon}f(u)$ tends to a limit which satisfies the following free boundary problem known as the Hele-Shaw problem (1.18)-(1.21). A rigorous justification that $u \to \pm 1$ in the interior or exterior of Γ_t for all $t \in [0, T]$ as $\epsilon \searrow 0$ was given by Stoth [88] for the radially symmetric case, and by Alikakos, Bates and Chen [3] for the general case. In addition, Chen [20] established the convergence of the weak solution of the Cahn-Hilliard problem to a weak (or varifold) solution of the Hele-Shaw problem.

Numerical approximations of the Cahn-Hilliard equation have been extensively carried out in the past thirty years (cf. [28, 31, 46] and the references therein). On the other hand, the majority of these works were done for a fixed parameter ϵ. The error bounds, which are obtained using the standard Gronwall inequality technique, show an exponential dependence on $1/\epsilon$. Such an estimate is clearly not useful

for small ϵ, in particular, in addressing the issue whether the computed numerical interfaces converge to the original sharp interface of the Hele-Shaw problem. Better and practical error bounds should only depend on $1/\epsilon$ in some (low) polynomial orders because they can be used to provide an answer to the above convergence question, which in fact is the best result (in terms of ϵ) one can expect. The first such polynomial order in $1/\epsilon$ a priori estimate was obtained in [47] for mixed finite element approximations of the Cahn-Hilliard problem (1.30)–(1.33). In addition, polynomial order in $1/\epsilon$ a posteriori error estimates were obtained in [49] for the same mixed finite element methods. One of the key ideas employed in all these works is to use a nonstandard error estimate technique which is based on establishing a discrete spectrum estimate (using its continuous counterpart) for the linearized Cahn-Hilliard operator. An immediate corollary of the polynomial order in $1/\epsilon$ a priori and a posteriori error estimates is the convergence of the numerical interfaces of the underlying mixed finite element approximations to the Hele-Shaw flow before the onset of singularities of the Hele-Shaw flow as ϵ and mesh sizes h and k all tend to zero.

The objectives of this chapter are twofold: Firstly, we develop some MIP-DG methods and establish polynomial order in $1/\epsilon$ a priori error bounds, as well as prove convergence of numerical interfaces for the MIP-DG methods. This goal is motivated by the advantages of DG methods in regard to designing adaptive mesh methods and algorithms, which is an indispensable strategy with the diffuse interface methodology. Secondly, we use the Cahn-Hilliard equation as another prototypical model problem [40] to develop new analysis techniques for analyzing convergence of numerical interfaces to the underlying sharp interface for DG (and nonconforming finite element) discretizations of phase field models. To the best of our knowledge, no such convergence result and analysis technique is available in the literature for fourth order PDEs. The main obstacle for improving the finite element techniques of [47] is that the DG (and nonconforming finite element) spaces are not subspaces of

$H^1(\mathcal{D})$. As a result, whether the needed discrete spectrum estimate holds becomes a key question to answer.

This chapter consists of four additional sections. In section 3.2 we first collect some a priori error estimates for problem (1.30)-(1.33), which show the explicit dependence on the parameter ϵ. We then cite an important technical lemma to be used in the later sections. This states the spectral estimate for the linearized Cahn-Hilliard operator. In section 3.3, we propose two fully discrete MIP-DG schemes for problem (1.30)–(1.33), they differ only in their treatment of the nonlinear term. The first main result of this section is to establish a discrete spectrum estimate in the DG space, which mimics the spectral estimates for the differential operator and its finite element counterpart. The second main result of this section is to derive optimal error bounds which depends on $1/\epsilon$ only in low polynomial orders for both fully discrete MIP-DG methods. In section 3.4, using the refined error estimates of section 3.3, we prove the convergence of the numerical interfaces of the fully discrete MIP-DG methods to the interface of the Hele-Shaw flow before the onset of the singularities as ϵ, h and k all tend to zero. Finally, in section 3.5 we provide some numerical experiments to gauge the performance of the proposed fully discrete MIP-DG methods.

3.2 Preliminaries

In this section, we shall collect some known results about problem (1.30)–(1.33) from [19, 46, 47], which will be used in sections 3.3 and 3.4. Some general assumptions on the initial condition, as well as some energy estimates based on these assumptions, will be cited. Standard function and space notations are adopted in this chapter [2, 13]. We use $(\cdot, \cdot)$ and $\|\cdot\|_{L^2}$ to denote the standard inner product and norm on $L^2(\mathcal{D})$. Throughout this chapter, C denotes a generic positive constant independent of ϵ, space and time step sizes h and k, which may have different values at different occasions.

The following assumptions on the initial datum u_0 were made in [46, 42], they were used to derive a priori estimates for the solution of problem (1.30)–(1.33).

General Assumption (GA)

(1) Assume that $m_0 \in (-1, 1)$ where

$$m_0 := \frac{1}{|\mathcal{D}|} \int_{\mathcal{D}} u_0(x) dx. \tag{3.1}$$

(2) There exists a nonnegative constant σ_1 such that

$$J_\epsilon(u_0) \leq C\epsilon^{-2\sigma_1}. \tag{3.2}$$

(3) There exists nonnegative constants σ_2, σ_3 and σ_4 such that

$$\left\| -\epsilon\Delta u_0 + \epsilon^{-1} f(u_0) \right\|_{H^\ell(\mathcal{D})} \leq C\epsilon^{-\sigma_{2+\ell}}, \quad \ell = 0, 1, 2. \tag{3.3}$$

Under the above assumptions, we formally prove the following a priori solution estimates, and the brief proofs could also be found in [46, 47].

Proposition 3.2.1. *The solution u of problem (1.30)–(1.33) satisfies the following energy estimates:*

(i) $$\operatorname*{ess\,sup}_{t\in[0,\infty)} \left(\frac{\epsilon}{2} \|\nabla u\|_{L^2}^2 + \frac{1}{\epsilon} \|F(u)\|_{L^1} \right) + \begin{cases} \int_0^\infty \|u_t(s)\|_{H^{-1}}^2 \, ds \\ \int_0^\infty \|\nabla w(s)\|_{L^2}^2 \, ds \end{cases} \leq J_\epsilon(u_0),$$

(ii) $$\operatorname*{ess\,sup}_{t\in[0,\infty)} \|u\|_{L^4}^4 \leq C(1 + J_\epsilon(u_0)),$$

(iii) $$\operatorname*{ess\,sup}_{t\in[0,\infty)} \|u^2 - 1\|_{L^2}^2 \leq C\epsilon J_\epsilon(u_0).$$

Moreover, suppose that (3.1)–(3.3) hold, $u_0 \in H^4(\mathcal{D})$ and $\partial\mathcal{D} \in C^{2,1}$, then u satisfies the additional estimates:

$$\text{(iv)} \quad \frac{1}{|\mathcal{D}|}\int_{\mathcal{D}} u(x,t)\,dx = m_0 \quad \forall t \geq 0,$$

$$\text{(v)} \quad \int_0^T \|\Delta u\|_{L^2}^2 ds \leq C\epsilon^{-(2\sigma_1+3)},$$

$$\text{(vi)} \quad \int_0^T \|\nabla\Delta u\|_{L^2}^2 ds \leq C\epsilon^{-(2\sigma_1+5)},$$

$$\text{(vii)} \quad \begin{cases} \operatorname*{ess\,sup}_{t\in[0,\infty)} \|u_t\|_{H^{-1}}^2 \\ \operatorname*{ess\,sup}_{t\in[0,\infty)} \|\nabla w\|_{L^2}^2 \end{cases} + \epsilon\int_0^\infty \|\nabla u_t\|_{L^2}^2 ds \leq C\epsilon^{-\max\{2\sigma_1+3,2\sigma_3\}},$$

$$\text{(viii)} \quad \operatorname*{ess\,sup}_{t\in[0,\infty)} \|\Delta u\|_{L^2} \leq C\epsilon^{-\max\{\sigma_1+\frac{5}{2},\sigma_3+1\}},$$

$$\text{(ix)} \quad \operatorname*{ess\,sup}_{t\in[0,\infty)} \|\nabla\Delta u\|_{L^2} \leq C\epsilon^{-\max\{\sigma_1+\frac{5}{2},\sigma_3+1\}},$$

$$\text{(x)} \quad \begin{cases} \int_0^T \|u_t\|_{L^2}^2 ds \\ \int_0^T \|\Delta w\|_{L^2}^2 ds \end{cases} + \operatorname*{ess\,sup}_{t\in[0,T]} \epsilon\|\Delta u\|_{L^2}^2 \leq C\epsilon^{-\max\{2\sigma_1+\frac{7}{2},2\sigma_3+\frac{1}{2},2\sigma_2+1\}},$$

$$\text{(xi)} \quad \operatorname*{ess\,sup}_{t\in[0,T]} \|u_t\|_{L^2}^2 + \epsilon\int_0^T \|\Delta u_t\|_{L^2}^2 ds \leq C\epsilon^{-\max\{2\sigma_1+\frac{13}{2},2\sigma_3+\frac{7}{2},2\sigma_2+4,2\sigma_4\}}.$$

Furthermore, if there exists $\sigma_5 > 0$ such that

$$\lim_{s\to 0^+} \|\nabla u_t(s)\|_{L^2} \leq C\epsilon^{-\sigma_5}, \tag{3.4}$$

then there hold for $d = 2, 3$,

$$\text{(xii)} \quad \operatorname*{ess\,sup}_{t\in[0,T]} \|\nabla u_t\|_{L^2}^2 + \epsilon\int_0^T \|\nabla\Delta u_t\|_{L^2}^2 ds \leq C\rho_0(\epsilon,d),$$

$$\text{(xiii)} \quad \int_0^T \|u_{tt}\|_{H^{-1}}^2 ds \leq C\rho_1(\epsilon,d),$$

$$\text{(xiv)} \quad \operatorname*{ess\,sup}_{t\in[0,T]} \|\Delta^2 u\|_{L^2} \leq C\rho_2(\epsilon),$$

where

$$\begin{aligned}\rho_0(\epsilon,d) &:= \epsilon^{-\frac{2}{6-d}\max\{2\sigma_1+5,2\sigma_3+2\}-\max\{2\sigma_1+\frac{13}{2},2\sigma_3+\frac{7}{2},2\sigma_2+4\}} + \epsilon^{-2\sigma_5} \\ &\quad + \epsilon^{-\max\{2\sigma_1+7,2\sigma_3+4\}}, \\ \rho_1(\epsilon,d) &:= \epsilon\rho_0(\epsilon,d), \\ \rho_2(\epsilon) &:= \epsilon^{-\max\{\sigma_1+5,\sigma_3+\frac{7}{2},\sigma_2+\frac{5}{2},\sigma_4+1\}}.\end{aligned}$$

Proof. (i). Taking the time derivative of the Cahn-Hilliard energy functional, we get

$$\frac{d}{dt}J_\epsilon(u(t)) = \begin{cases} -\|u_t(t)\|^2_{H^{-1}}, \\ -\|\nabla w(t)\|^2_{L^2}. \end{cases} \tag{3.5}$$

Assume the maximum of $J_\epsilon(u(t))$ is obtained at $t = t_0$, then integrating both sides of (3.5) over $(0, t_0)$ and $(0, \infty)$ respectively, (i) is obtained.

(ii). Using Young's inequality to $2u^2$, we get

$$\int_{\mathcal{D}} F(u) = \frac{1}{4}\int_{\mathcal{D}}(u^4 - 2u^2 + 1)dx, \tag{3.6}$$

$$\geq \frac{1}{4}\int_{\mathcal{D}}(\frac{1}{2}u^4 - 1)dx. \tag{3.7}$$

By (i), we get (ii) immediately.

(iii). It is an immediate corollary of (i).

(iv). Integrating both sides of (1.30) over $(0, t)$, and using (3.1), (iv) is obtained.

(v). Testing (1.30) with u, we get

$$\big(u_t, u\big) + \big(\epsilon\Delta u - \frac{1}{\epsilon}f(u), \Delta u\big) = 0. \tag{3.8}$$

Integrating both sides over $(0, T)$, we get

$$\epsilon\int_0^T \|\Delta u\|^2_{L^2}dt + \frac{1}{\epsilon}\int_0^T \big(f'(u)\nabla u, \nabla u\big)dt \leq \frac{1}{2}\|u(0)\|^2_{L^2}. \tag{3.9}$$

By Lemma 2.2 in [47], (i), and the following fact:

$$-\big(f'(u)v, v\big) \leq \|v\|_{L^2}^2, \tag{3.10}$$

we can get (v).

(vi). Rewrite (1.31) as

$$\Delta u = \frac{1}{\epsilon^2} f(u) - \frac{1}{\epsilon} w. \tag{3.11}$$

Then applying the operator ∇ to (3.11), taking the L^2 norm and integrating both sides of (3.11) over $(0, T)$, we get

$$\int_0^T \|\nabla \Delta u\|_{L^2}^2 ds \leq 2 \int_0^T \|\frac{1}{\epsilon^2} f'(u) \nabla u\|_{L^2}^2 ds + 2 \int_0^T \|\frac{1}{\epsilon} \nabla w\|_{L^2}^2 ds. \tag{3.12}$$

By Lemma 2.2 in [47] and (i), we can get (vi).

(vii). Taking the derivative with respect to t on both sides of (1.30), we get

$$u_{tt} + \epsilon \Delta^2 u_t - \frac{1}{\epsilon} \Delta\big(f'(u) u_t\big) = 0. \tag{3.13}$$

Testing (3.13) with $-\Delta^{-1} u_t$, and integrating over $(0, m)$, we get

$$\int_0^m \frac{d}{dt} \|\nabla w\|_{L^2}^2 ds + \epsilon \int_0^m \|\nabla u_t\|_{L^2}^2 ds + \frac{1}{\epsilon} \int_0^m (f'(u) u_t, u_t) ds = 0. \tag{3.14}$$

Interpolating $\|u_t\|_{L^2}$ between $\|u_t\|_{H^1}$ and $\|u_t\|_{H^{-1}}$, and using Young's inequality, we get

$$\|\nabla w(m)\|_{L^2}^2 + \frac{\epsilon}{2} \int_0^m \|\nabla u_t\|_{L^2}^2 ds \leq \|\nabla w(0)\|_{L^2}^2 + \frac{C}{\epsilon^3} \int_0^m \|u_t\|_{H^{-1}}^2 ds. \tag{3.15}$$

By (GA) and (i), together with the fact $\|\nabla w\|_{L^2} = \|u_t\|_{H^{-1}}$, we obtain (vii).

(viii). By (3.11),

$$\operatorname*{ess\,sup}_{t\in[0,\infty)} \|\Delta u\|_{L^2} \le \frac{1}{\epsilon^2} \operatorname*{ess\,sup}_{t\in[0,\infty)} \|f(u)\|_{L^2} + \frac{1}{\epsilon} \operatorname*{ess\,sup}_{t\in[0,\infty)} \|w\|_{L^2}. \tag{3.16}$$

By Lemma 2.2, Lemma 2.3 in [47], we get

$$\operatorname*{ess\,sup}_{t\in[0,\infty)} \|\Delta u\|_{L^2} \le \frac{C}{\epsilon^2} + \frac{C}{\epsilon} \operatorname*{ess\,sup}_{t\in[0,\infty)} J_\epsilon(u(t)) + \operatorname*{ess\,sup}_{t\in[0,\infty)} \|\nabla w\|_{L^2}. \tag{3.17}$$

Using (i) and (vii), (viii) is proved.

(ix). Apply the operator ∇ to both sides of (3.11), we get

$$\operatorname*{ess\,sup}_{t\in[0,\infty)} \|\nabla\Delta u\|_{L^2} \le \frac{1}{\epsilon^2} \operatorname*{ess\,sup}_{t\in[0,\infty)} \|f'(u)\nabla u\|_{L^2} + \frac{1}{\epsilon} \operatorname*{ess\,sup}_{t\in[0,\infty)} \|\nabla w\|_{L^2}. \tag{3.18}$$

By Lemma 2.2 in [47], (i) and (vii),

$$\operatorname*{ess\,sup}_{t\in[0,\infty)} \|\nabla\Delta u\|_{L^2} \le \epsilon^{-\sigma_1-\frac{5}{2}} + \frac{1}{\epsilon}\epsilon^{-\max\{\sigma_1+\frac{3}{2},\sigma_3\}}. \tag{3.19}$$

Then (ix) is proved.

(x). Testing (1.30) with u_t, we get

$$\|u_t\|_{L^2}^2 + \frac{1}{2}\frac{d}{dt}\epsilon\|\Delta u\|_{L^2}^2 + \frac{1}{\epsilon}\big(f'(u)\nabla u, \nabla u_t\big) = 0. \tag{3.20}$$

Integrating (3.20) over $(0,T)$, and using Schwarz's inequality, we get

$$\begin{aligned} &2\int_0^T \|u_t\|_{L^2}^2 ds + \operatorname*{ess\,sup}_{t\in[0,T]} \epsilon\|\Delta u(t)\|_{L^2}^2 \\ &\quad\le \operatorname*{ess\,sup}_{t\in[0,T]} \epsilon\|\Delta u(0)\|_{L^2}^2 + \epsilon^{-\frac{5}{2}}\int_0^T \|f'(u)\nabla u\|_{L^2}^2 ds + \epsilon^{\frac{1}{2}}\int_0^T \|\nabla u_t\|_{L^2}^2 ds. \end{aligned} \tag{3.21}$$

Using Lemma 2.2 in [47], (i) and (vii), we can get (x).

(xi). Testing (3.13) with u_t, we get

$$\frac{d}{dt}\|u_t\|_{L^2}^2 + \epsilon\|\Delta u_t\|_{L^2}^2 - \frac{1}{\epsilon}\big(\Delta(f'(u)u_t), u_t\big) = 0. \tag{3.22}$$

Integrating both sides of (3.22), and using integration by part, we get

$$\begin{aligned} &\operatorname*{ess\,sup}_{t\in[0,T]} \|u_t(t)\|_{L^2}^2 + \epsilon\int_0^T \|\Delta u_t\|_{L^2}^2 ds \\ &\quad\le \frac{1}{\epsilon}\int_0^T \big(f'(u)u_t, \Delta u_t\big) + \operatorname*{ess\,sup}_{t\in[0,T]} \|u_t(0)\|_{L^2}^2 \\ &\quad\le \frac{1}{2\epsilon^3}\int_0^T \|f'(u)u_t\|_{L^2}^2 ds + \frac{\epsilon}{2}\int_0^T \|\Delta u_t\|_{L^2}^2 ds + \operatorname*{ess\,sup}_{t\in[0,T]} \|u_t(0)\|_{L^2}^2. \end{aligned} \tag{3.23}$$

By Lemma 2.2 in [47] and (x), we get (xi).

(xii). Multiplying (3.13) by $-\Delta u_t$, we get

$$\begin{aligned} \frac{1}{2}\frac{d}{dt}\|\nabla u_t\|_{L^2}^2 + \epsilon\|\nabla\Delta u_t\|_{L^2}^2 &= \frac{1}{\epsilon}\big(\nabla(f'(u)u_t), \nabla\Delta u_t\big) \\ &= \frac{1}{\epsilon}\big(f''(u)\nabla u u_t + f'(u)\nabla u_t, \nabla\Delta u_t\big) \\ &\le \frac{\epsilon}{2}\|\nabla\Delta u_t\|_{L^2}^2 + \frac{C}{\epsilon^3}\big(\|\nabla u\|_{L^\infty}^2\|u_t\|_{L^2}^2 + \|\nabla u_t\|_{L^2}^2\big). \end{aligned} \tag{3.24}$$

Using the following Gagliardo-Nirenberg inequality [2],

$$\|\nabla u\|_{L^\infty} \le C\left(\|\nabla\Delta u\|_{L^2}^{\frac{2}{6-d}}\|u\|_{L^\infty}^{\frac{4-d}{6-d}} + \|u\|_{L^\infty}\right), \tag{3.25}$$

we have

$$\begin{aligned} &\frac{1}{2}\frac{d}{dt}\|\nabla u_t\|_{L^2}^2 + \frac{\epsilon}{2}\|\nabla\Delta u_t\|_{L^2}^2 \\ &\quad\le \frac{C}{\epsilon^3}\left[\|\nabla u_t\|_{L^2}^2 + \left(\|\nabla\Delta u\|_{L^2}^{\frac{4}{6-d}} + 1\right)\|u_t\|_{L^2}^2\right]. \end{aligned} \tag{3.26}$$

Using (3.26), (vii), (ix) and (x), we get (xii).

(xiii). From (3.13), we have

$$\begin{aligned}\|u_{tt}\|_{H^{-1}} &= \sup_{0\not\equiv\phi\in H^1} \frac{< u_{tt}, \phi >}{\|\phi\|_{H^1}} \\ &\leq \epsilon\|\nabla\Delta u_t\|_{L^2} + \frac{1}{\epsilon}\|\nabla(f'(u)u_t)\|_{L^2}.\end{aligned} \tag{3.27}$$

Then

$$\|u_{tt}\|^2_{H^{-1}} \leq 2\epsilon^2\|\nabla\Delta u_t\|^2_{L^2} + \frac{2}{\epsilon^2}\bigg(\|\nabla u\|^2_{L^\infty}\|u_t\|^2_{L^2} + \|\nabla u_t\|^2_{L^2}\bigg). \tag{3.28}$$

Integrating (3.28) over $(0, T)$, and using (3.25) and (xii), we get (xiii).

(xiv). Rewrite (1.30) as

$$\begin{aligned}\Delta^2 u &= \frac{1}{\epsilon^2}\Delta f(u) - \frac{1}{\epsilon}u_t \\ &= \frac{1}{\epsilon^2}\bigg[f''(u)\Delta u + f'(u)|\nabla u|^2\bigg] - \frac{1}{\epsilon}u_t.\end{aligned} \tag{3.29}$$

By triangle inequality, Lemma 2.2 in [47], (viii) and (xi), we can obtain (xiv). □

The next lemma concerns with a lower bound estimate for the principal eigenvalue of the linearized Cahn-Hilliard operator, a proof of this lemma can be found in [19].

Lemma 3.2.2. *Suppose that (3.1)–(3.3) hold. Given a smooth initial curve/surface Γ_0, let u_0 be a smooth function satisfying $\Gamma_0 = \{x \in \mathcal{D}; u_0(x) = 0\}$ and some profile described in [19]. Let u be the solution to problem (1.30)–(1.33). Define $\mathcal{L}_{CH}$ as*

$$\mathcal{L}_{CH} := \Delta\left(\epsilon\Delta - \frac{1}{\epsilon}f'(u)I\right). \tag{3.30}$$

Then there exists $0 < \epsilon_0 << 1$ *and a positive constant* C_0 *such that the principle eigenvalue of the linearized Cahn-Hilliard operator* $\mathcal{L}_{CH}$ *satisfies*

$$\lambda_{CH} := \inf_{\substack{0\neq\psi\in H^1(\mathcal{D}) \\ \Delta w=\psi}} \frac{\epsilon\|\nabla\psi\|_{L^2}^2 + \frac{1}{\epsilon}(f'(u)\psi,\psi)}{\|\nabla w\|_{L^2}^2} \geq -C_0 \tag{3.31}$$

for $t \in [0,T]$ *and* $\epsilon \in (0,\epsilon_0)$.

Remark 3.2.3. *(a) A discrete generalization of* (3.31) *on* C^0 *finite element spaces was proved in [46, 47]. It plays a pivotal role in the nonstandard convergence analysis of [46, 47]. In the next section, we shall prove another discrete generalization of* (3.31) *on the DG finite element space.*

(b) The restriction on the initial function u_0 *is needed to guarantee that the solution* $u(t)$ *satisfies certain profile at later time* $t > 0$ *which is required in the proof of [19]. One example of admissible initial functions is* $u_0 = \tanh(\frac{d_0(x)}{\epsilon})$, *where* $d_0(x)$ *stands for the signed distance function to the initial interface* Γ_0. *Such a* u_0 *is smooth when* Γ_0 *is smooth.*

3.3 Fully discrete mixed interior penalty discontinuous Galerkin approximations

In this section we present and analyze two fully discrete MIP-DG methods for the Cahn-Hilliard problem (1.30)–(1.33). The primary goal of this section is to derive error estimates for the DG solutions that depend on ϵ^{-1} only in low polynomial orders, instead of exponential orders. As in the finite element case (cf. [47]), the crux is to establish a discrete spectrum estimate for the linearized Cahn-Hilliard operator on the DG space.

3.3.1 Formulations of the mixed interior penalty discontinuous Galerkin method

Let $\mathcal{T}_h = \{K\}_{K\in\mathcal{D}}$ be a quasi-uniform triangulation of $\mathcal{D}$ parameterized by $h > 0$. For any triangle/tetrahedron $K \in \mathcal{T}_h$, we define h_K to be the diameter of K, and $h := \max_{K\in\mathcal{T}_h} h_K$. The standard broken Sobolev space is defined as

$$H^s(\mathcal{T}_h) := \big\{v \in L^2(\mathcal{D});\, \forall K \in \mathcal{T}_h,\, v|_K \in H^s(K)\big\}. \tag{3.32}$$

For any $K \in \mathcal{T}_h$, $P_r(K)$ denotes the set of all polynomials of degree at most $r(\geq 1)$ on the element K, and the DG finite element space V_h is defined as

$$V_h := \big\{v \in L^2(\mathcal{D});\, \forall K \in \mathcal{T}_h,\, v|_K \in P_r(K)\big\}. \tag{3.33}$$

Let L_0^2 denote the set of functions in $L^2(\mathcal{D})$ with zero mean, and let $\mathring{V}_h := V_h \cap L_0^2$. We also define $\mathcal{E}_h^I$ to be the set of all interior edges/faces of $\mathcal{T}_h$, $\mathcal{E}_h^B$ to be the set of all boundary edges/faces of $\mathcal{T}_h$ on $\Gamma = \partial\mathcal{D}$, and $\mathcal{E}_h := \mathcal{E}_h^I \cup \mathcal{E}_h^B$. Let e be an interior edge shared by two elements K_1 and K_2. For a scalar function v, define

$$\{v\} = \frac{1}{2}(v|_K + v|_{K'}), \quad [v] = v|_K - v|_{K'}, \quad \text{on } e \in \mathcal{E}_h^I,$$

where K is K_1 or K_2, whichever has the bigger global labeling and K' is the other. The L^2-inner product for piecewise functions over the mesh $\mathcal{T}_h$ is naturally defined by

$$(v, w)_{\mathcal{T}_h} := \sum_{K\in\mathcal{T}_h} \int_K vw dx.$$

Let $0 \leq t_0 < t_1 < \cdots < t_M = T$ be a partition of the interval $[0, T]$ with time step $k = t_{n+1} - t_n$. Our fully discrete MIP-DG methods are defined as follows: for

any $1 \le m \le M$, $(U^m, W^m) \in V_h \times V_h$ are given by

$$(d_t U^m, \eta) + a_h(W^m, \eta) = 0 \qquad \forall\, \eta \in V_h, \tag{3.34}$$

$$\epsilon a_h(U^m, v) + \frac{1}{\epsilon}(f^m, v) - (W^m, v) = 0 \qquad \forall\, v \in V_h, \tag{3.35}$$

where

$$\begin{aligned} a_h(u, v) = &\sum_{K \in \mathcal{T}_h} \int_K \nabla u \cdot \nabla v \, dx - \sum_{e \in \mathcal{E}_h^I} \int_e \{\nabla u \cdot \mathbf{n}_e\}[v] \, ds \\ &- \sum_{e \in \mathcal{E}_h^I} \int_e \{\nabla v \cdot \mathbf{n}_e\}[u] \, ds + \sum_{e \in \mathcal{E}_h^I} \int_e \frac{\sigma_e^0}{h_e}[u][v] \, ds, \end{aligned} \tag{3.36}$$

and $\sigma_e^0 > 0$ is the penalty parameter. There are two choices of f^m considered in this chapter, namely

$$f^m = (U^m)^3 - U^{m-1} \qquad \text{and} \qquad f^m = (U^m)^3 - U^m,$$

which lead to the energy-splitting scheme and fully implicit scheme respectively. d_t is the (backward) difference operator defined by $d_t U^m := (U^m - U^{m-1})/k$ and $U^0 := \widehat{P}_h u_0$ (or $\widehat{Q}_h u_0$) is the starting value, with the continuous finite element projection $\widehat{P}_h$ (or $\widehat{Q}_h$) to be defined below. We refer to [40] for the details why the continuous projection is needed for the initial condition. We remark that only the fully implicit case was considered in [46, 47] for the mixed finite element method.

In order to analyze the stability of (3.34)–(3.35), we need some preparations. First, we introduce three projection operators that will be needed to derive the error estimates in section 3.4. $P_h : H^s(\mathcal{T}_h) \to V_h$ denotes the elliptic projection operator defined by

$$a_h(u - P_h u, v_h) + (u - P_h u, v_h) = 0 \qquad \forall\, v_h \in V_h, \tag{3.37}$$

which has the following approximation properties (see [22]):

$$\|v - P_h v\|_{L^2(\mathcal{T}_h)} + h\|\nabla(v - P_h v)\|_{L^2(\mathcal{T}_h)} \leq Ch^{\min\{r+1,s\}}\|u\|_{H^s(\mathcal{T}_h)}, \tag{3.38}$$

$$\frac{1}{|\ln h|^{\overline{r}}}\|v - P_h v\|_{L^\infty(\mathcal{T}_h)} + h\|\nabla(u - P_h u)\|_{L^\infty(\mathcal{T}_h)} \leq Ch^{\min\{r+1,s\}}\|u\|_{W^{s,\infty}(\mathcal{T}_h)}. \tag{3.39}$$

Here $\overline{r} := \min\{1, r\} - \min\{1, r-1\}$.

Let $\widehat{P}_h : H^s(\mathcal{T}_h) \to S_h := V_h \cap C^0(\overline{\mathcal{D}})$ denote the standard continuous finite element elliptic projection, which is the counterpart of projection P_h. It has the following well-known property [46, 47]:

$$\|u - \widehat{P}_h u\|_{L^\infty} \leq Ch^{2-\frac{d}{2}}\|u\|_{H^2}. \tag{3.40}$$

Next, for any DG function $\Psi_h \in V_h$, we define its continuous finite element projection $\Psi_h^{FE} \in S^h$ by

$$\tilde{a}_h(\Psi_h^{FE}, v_h) = \tilde{a}_h(\Psi_h, v_h) \qquad \forall\, v_h \in S_h, \tag{3.41}$$

where

$$\tilde{a}_h(u, v) = a_h(u, v) + \alpha(u, v),$$

and α is a parameter that will be specified later in section 3.3.

A mesh-dependent H^{-1} norm will also be needed. To the end, we introduce the inverse discrete Laplace operator $\Delta_h^{-1} : V_h \to \mathring{V}_h$ as follows: given $\zeta \in V_h$, let $\Delta_h^{-1}\zeta \in \mathring{V}_h$ such that

$$a_h(-\Delta_h^{-1}\zeta, w_h) = (\zeta, w_h) \qquad \forall\, w_h \in \mathring{V}_h. \tag{3.42}$$

We note that Δ_h^{-1} is well defined provided that $\sigma_e^0 > \sigma_*^0$ for some positive number σ_*^0 and for all $e \in \mathcal{E}_h$ because this condition ensures the coercivity of the DG bilinear form $a_h(\cdot, \cdot)$.

We then define "-1" inner product by

$$(\zeta,\xi)_{-1,h} := a_h(-\Delta_h^{-1}\zeta,-\Delta_h^{-1}\xi) = (\zeta,-\Delta_h^{-1}\xi) = (-\Delta_h^{-1}\zeta,\xi), \tag{3.43}$$

and the induced mesh-dependent H^{-1} norm is given by

$$\|\zeta\|_{-1,h} := \sqrt{(\zeta,\zeta)_{-1,h}} = \sup_{0\neq\xi\in\mathring{V}_h} \frac{(\zeta,\xi)}{|||\xi|||_a}, \tag{3.44}$$

where $|||\xi|||_a := \sqrt{a_h(\xi,\xi)}$. The following properties can be easily verified (cf. [1]):

$$|(\zeta,\xi)| \le \|\zeta\|_{-1,h}|||\xi|||_a \qquad \forall\,\xi\in V_h,\ \zeta\in\mathring{V}_h, \tag{3.45}$$

$$\|\zeta\|_{-1,h} \le C\|\zeta\|_{L^2} \qquad \forall\,\zeta\in\mathring{V}_h, \tag{3.46}$$

and, if $\mathcal{T}_h$ is quasi-uniform, then

$$\|\zeta\|_{L^2} \le Ch^{-1}\|\zeta\|_{-1,h} \qquad \forall\,\zeta\in\mathring{V}_h. \tag{3.47}$$

3.3.2 Discrete energy law and well-posedness

In this subsection we first establish a discrete energy law, which mimics the differential energy law, for both fully discrete MIP-DG methods defined in (3.34)–(3.35). Based on this discrete energy law, we prove the existence and uniqueness of solutions to the MIP-DG methods by recasting the schemes as convex minimization problems at each time step. It turns out that the energy-splitting scheme is unconditionally stable but the fully implicit scheme is only conditionally stable.

Theorem 3.3.1. *Let $(U^m, W^m) \in V_h \times V_h$ be a solution to scheme (3.34)–(3.35). The following energy law holds for any $h, k > 0$:*

$$E_h(U^\ell) + k\sum_{m=1}^{\ell} \|d_t U^m\|_{-1,h}^2 + k^2 \sum_{m=1}^{\ell} \left\{ \frac{\epsilon}{2} |||d_t U^m|||_a^2 + \frac{1}{4\epsilon}\|d_t (U^m)^2\|_{L^2}^2 \right. \tag{3.48}$$
$$\left. + \frac{1}{2\epsilon}\|U^m d_t U^m\|_{L^2}^2 \pm \frac{1}{2\epsilon}\|d_t U^m\|_{L^2}^2 \right\} = E_h(U^0)$$

for all $1 \le \ell \le M$, where

$$E_h(U) := \frac{1}{4\epsilon}\|U^2 - 1\|_{L^2}^2 + \frac{\epsilon}{2}|||U|||_a^2. \tag{3.49}$$

The sign "$\pm$" in (3.48) takes "$+$" when $f^m = (U^m)^3 - U^{m-1}$ and "$-$" when $f^m = (U^m)^3 - U^m$.

Proof. Taking $\eta = -\Delta_h^{-1} d_t U^m$ in (3.34) and $v = d_t U^m$ in (3.35) to get

$$\|d_t U^m\|_{-1,h}^2 + (W^m, d_t U^m) = 0, \tag{3.50}$$
$$\epsilon a_h(U^m, d_t U^m) + \frac{1}{\epsilon}(f^m, d_t U^m) - (W^m, d_t U^m) = 0. \tag{3.51}$$

Besides, noticing the following inequalities:

$$a_h(U^m, d_t U^m) = \frac{1}{2}\left[d_t |||U^m|||_a^2 + k|||d_t U^m|||_a^2 \right], \tag{3.52}$$
$$((U^m)^3 - U^{m-1}, d_t U^m) = \frac{1}{4} d_t \|(U^m)^2 - 1\|_{L^2}^2 + \frac{k}{4}\|d_t (U^m)^2\|_{L^2}^2 \tag{3.53}$$
$$+ \frac{k}{2}\|U^m d_t U^m\|_{L^2}^2 + \frac{k}{2}\|d_t U^m\|_{L^2}^2.$$

Adding (3.50) and (3.51), using (3.52) and (3.53), and applying the operator $k\sum_{m=1}^{\ell}$, we get the conclusion for the case when $f^m = (U^m)^3 - U^{m-1}$.

Using the equality below, we get the conclusion for the case when $f^m = (U^m)^3 - U^m$.

$$(U^m)^3 - U^m = \Big[(U^m)^3 - U^{m-1}\Big] - \Big[U^m - U^{m-1}\Big]. \tag{3.54}$$

□

Corollary 3.3.2. *Let $\sigma_*^0 > 0$ be a sufficiently large constant. Suppose that $\sigma_e^0 > \sigma_*^0$ for all $e \in \mathcal{E}_h$. Then scheme (3.34)–(3.35) is stable for all $h, k > 0$ when $f^m = (U^m)^3 - U^{m-1}$ and is stable for $h > 0$ and $k = O(\epsilon^3)$ when $f^m = (U^m)^3 - U^m$.*

Proof. The first case holds trivially from (3.48). In the second case, the "bad term" $\|d_t U^m\|_{L^2}$ can be controlled by the "good terms" $\|U^m\|_{-1,h}^2$ and $|||U^m|||_a^2$ by using the norm interpolation inequality (3.81) provided that $k = O(\epsilon^3)$. □

Theorem 3.3.3. *Suppose that $\sigma_e^0 > \sigma_*^0$ for all $e \in \mathcal{E}_h$. Then scheme (3.34)–(3.35) has a unique solution (U^m, W^m) at each time step for for all $h, k > 0$ in the case $f^m = (U^m)^3 - U^{m-1}$ and for $h > 0$ and $k = O(\epsilon^3)$ in the case $f^m = (U^m)^3 - U^m$.*

Proof. Setting $\eta = -\Delta_h^{-1} v$ in (3.34) we get

$$\big(d_t U^m, v\big)_{-1,h} + \big(W^m, v\big) = 0.$$

Adding the above equation to (3.35) yields

$$\big(d_t U^m, v\big)_{-1,h} + \epsilon a_h\big(U^m, v\big) + \frac{1}{\epsilon}\big(f^m, v\big) = 0.$$

Hence, U^m satisfies

$$\big(U^m, v\big)_{-1,h} + k\epsilon a_h\big(U^m, v\big) + \frac{k}{\epsilon}\big(f^m, v\big) = \big(U^{m-1}, v\big)_{-1,h}. \tag{3.55}$$

In the case $f^m = (U^m)^3 - U^{m-1}$ it is easy to check that (3.55) can be recast as a convex minimization problem (cf. [1, 40]) whose well-posedness holds for all $h, k > 0$. Hence, in this case there is a unique solution U^m to (3.34)–(3.35). On the

other hand, when $f^m = (U^m)^3 - U^m$, there is an extra term $-k\epsilon^{-1}(U^m, v)$ comes out from the nonlinear term in (3.55). This extra term contributes a "bad term" $-k\epsilon^{-1}\|U^m\|_{L^2}^2$ to the functional of the minimization problem. Again, this term can be controlled by the "good terms" $\|U^m\|_{-1,h}^2$ and $|||U^m|||_a^2$ in the functional by using the norm interpolation inequality (3.81), provided that $k = O(\epsilon^3)$. Hence, in the case $f^m = (U^m)^3 - U^m$, there is a unique solution U^m to (3.34)–(3.35) for all $h > 0$ and $k = O(\epsilon^3)$. The proof is complete. □

3.3.3 Discrete spectrum estimate on the discontinuous Galerkin space

In this subsection, we shall establish a discrete spectrum estimate for the linearized Cahn-Hilliard operator on the DG space, which plays a vital role in our error estimates.

To the end, we first state a slightly modified version of a discrete spectrum estimate for the linearized Cahn-Hilliard operator on the continuous finite element space first proved in [46, 47].

Lemma 3.3.4. *Suppose the assumptions of Lemma 3.2.2 hold, and C_0 is the same as in (3.31). C_1 and C_2 are defined by*

$$C_1 := \max_{|\xi|\le 2C_0} |f''(\xi)|, \tag{3.56}$$

$$\|u - \widehat{P}_h u\|_{L^\infty((0,T);L^\infty)} \le C_2 h^{2-\frac{d}{2}} \epsilon^{\min\{-\sigma_1-\frac{5}{2},-\sigma_3-1\}}. \tag{3.57}$$

Then there exists $0 < \epsilon_1 << 1$ such that, for any $\epsilon \in (0, \epsilon_1)$, there holds

$$\lambda_{CH}^{FE} \equiv \inf_{0\neq\psi_h\in L_0^2(\mathcal{D})\cap S_h} \frac{\epsilon\|\nabla\psi_h\|_{L^2}^2 + \frac{2-\epsilon^3}{2\epsilon}\big(f'(\widehat{P}_h u)\psi_h, \psi_h\big)}{\|\nabla\Delta^{-1}\psi_h\|_{L^2}^2} \ge -(C_0+1), \tag{3.58}$$

provided that h satisfies

$$h^{2-\frac{d}{2}} \leq (C_1 C_2)^{-1} \epsilon^{\max\{\sigma_1+\frac{11}{2}, \sigma_3+4\}}. \tag{3.59}$$

Here $\Delta^{-1} : L_0^2(\mathcal{D}) \to H^1(\mathcal{D}) \cap L_0^2(\mathcal{D})$ denotes the inverse Laplace operator.

Proof. Following Theorem 2.3 in [3], we can know that

$$\|u\|_{L^\infty(\mathcal{D}_T)} \leq \frac{3}{2} C_0. \tag{3.60}$$

By using (3.57), (3.59) and (3.60), when ϵ is small enough, we have

$$\|\widehat{P}_r^h u\|_{L^\infty(J;L^\infty)} \leq \|u\|_{L^\infty(J;L^\infty)} + \|\widehat{P}_r^h u - u\|_{L^\infty(J;L^\infty)} \leq 2C_0. \tag{3.61}$$

It then follows from the Mean Value Theorem that

$$\begin{aligned} \|f'(\widehat{P}_r^h u) - f'(u)\|_{L^\infty(J;L^\infty)} \leq & \max_{|\xi| \leq 2C_0} |f''(\xi)| \|\widehat{P}_r^h u - u\|_{L^\infty(J;L^\infty)} \\ \leq & C_1 C_2 h^{2-\frac{d}{2}} \epsilon^{\min\{-\sigma_1-\frac{5}{2}, -\sigma_3-1\}} \\ \leq & \epsilon^3, \end{aligned} \tag{3.62}$$

where the last inequality comes from the assumption (3.59). Therefore,

$$f'(\widehat{P}_r^h u) \geq f'(u) - \epsilon^3, \tag{3.63}$$

and we have

$$\begin{aligned}\lambda_{CH}^h \geq & \inf_{0\neq\psi\in L_0^2(\mathcal{D})\cap H^1(\mathcal{D})} \Bigg\{(1-\frac{\epsilon^3}{2})\frac{\epsilon\|\nabla\psi\|_{L^2}^2+\frac{1}{\epsilon}(f'(u)\psi,\psi)}{\|\nabla\Delta^{-1}\psi\|_{L^2}^2} \\ & + \frac{\frac{\epsilon^4}{2}\|\nabla\psi\|_{L^2}^2-\epsilon^2(1-\frac{\epsilon^3}{2})\|\psi\|_{L^2}^2}{\|\nabla\Delta^{-1}\psi\|_{L^2}^2}\Bigg\} \\ \geq & -(1-\frac{\epsilon^3}{2})C_0-\frac{(1-\frac{\epsilon^3}{2})^2}{2} \\ \geq & -(C_0+1).\end{aligned} \tag{3.64}$$

□

We are now ready to state the discrete spectrum estimate on the DG space.

Proposition 3.3.5. *Suppose the assumptions of Lemma 3.2.2 hold. Let u be the solution of (1.30)–(1.33) and $P_h u$ denote its DG elliptic projection. Assume*

$$\operatorname*{ess\,sup}_{t\in[0,\infty)}\|u\|_{W^{1+r,\infty}}\leq C\epsilon^{-\gamma}, \tag{3.65}$$

for a constant γ, then there exists $0<\epsilon_2<<1$ and an ϵ-independent and h-independent constant $c_0>0$, such that for any $\epsilon\in(0,\epsilon_2)$, there holds

$$\lambda_{CH}^{DG}=\inf_{0\neq\Phi_h\in L_0^2(\mathcal{D})\bigcap V_h}\frac{\epsilon a_h(\Phi_h,\Phi_h)+\frac{1-\epsilon^3}{\epsilon}(f'(P_hu)\Phi_h,\Phi_h)}{\|\nabla\Delta^{-1}\Phi_h\|_{L^2}^2}\geq -c_0, \tag{3.66}$$

provided that h satisfies the constraints

$$h^{2-\frac{d}{2}}\leq(C_1C_2)^{-1}\epsilon^{\max\{\sigma_1+\frac{11}{2},\sigma_3+4\}}, \tag{3.67}$$

$$h^{1+r}|\ln h|^{\bar{r}}\leq(C_1C_3)^{-1}\epsilon^{\gamma+3}, \tag{3.68}$$

where C_1 and C_2 are same as in Lemma 3.3.4, $\bar{r}$ and C_3 are defined by

$$\bar{r} = \min\{1, r\} - \min\{1, r-1\},$$
$$\|u - P_h u\|_{L^\infty((0,T);L^\infty)} \le C_3 h^{1+r} |\ln h|^{\bar{r}} \epsilon^{-\gamma}.$$

Proof. By Proposition 2 in [46], under the mesh constraint (3.67), we have

$$\|f'(\widehat{P}_h u) - f'(u)\|_{L^\infty((0,T);L^\infty)} \le \epsilon^3. \tag{3.69}$$

Similarly, under the mesh condition (3.68), we can show that for any $\epsilon > 0$, there holds

$$\|f'(P_h u) - f'(u)\|_{L^\infty((0,T);L^\infty)} \le \epsilon^3. \tag{3.70}$$

It follows from (3.69) and (3.70) that

$$\|f'(P_h u) - f'(\widehat{P}_h u)\|_{L^\infty((0,T);L^\infty)} \le 2\epsilon^3 \quad \text{and} \quad f'(P_h u) \ge f'(\widehat{P}_h u) - 2\epsilon^3. \tag{3.71}$$

Therefore,

$$\begin{aligned}
&\epsilon a_h(\Phi_h, \Phi_h) + \frac{1-\epsilon^3}{\epsilon}\big(f'(P_h u)\Phi_h, \Phi_h\big) \qquad (3.72)\\
&\ge \epsilon a_h(\Phi_h, \Phi_h) + \frac{1-\epsilon^3}{\epsilon}\big(f'(\widehat{P}_h u)\Phi_h, \Phi_h\big) - 2\epsilon^2(1-\epsilon^3)\|\Phi_h\|_{L^2}^2\\
&= \epsilon\frac{1-\epsilon^3}{1-\frac{\epsilon^3}{2}} a_h(\Phi_h, \Phi_h) + \frac{1-\epsilon^3}{\epsilon}\big(f'(\widehat{P}_h u)\Phi_h, \Phi_h\big)\\
&\quad - 2\epsilon^2(1-\epsilon^3)\|\Phi_h\|_{L^2}^2 + \frac{\epsilon^4}{2-\epsilon^3} a_h(\Phi_h, \Phi_h)\\
&= \epsilon\frac{1-\epsilon^3}{1-\frac{\epsilon^3}{2}} a_h(\Phi_h, \Phi_h) + \frac{1-\epsilon^3}{\epsilon}\int_{\mathcal{D}} f'(\widehat{P}_h u)\Big((\Phi_h)^2 - (\Phi_h^{FE})^2\Big)\, dx\\
&\quad + \frac{1-\epsilon^3}{\epsilon}\int_{\mathcal{D}} f'(\widehat{P}_h u)(\Phi_h^{FE})^2 dx - 2\epsilon^2(1-\epsilon^3)\|\Phi_h\|_{L^2}^2 + \frac{\epsilon^4}{2-\epsilon^3} a_h(\Phi_h, \Phi_h).
\end{aligned}$$

Next, we derive a lower bound for each of the first two terms on the right-hand side of (3.72). Notice that the first term can be rewritten as

$$\begin{aligned} a_h(\Phi_h,\Phi_h) &= a_h(\Phi_h-\Phi_h^{FE},\Phi_h-\Phi_h^{FE}) + 2a_h(\Phi_h,\Phi_h^{FE}) - a_h(\Phi_h^{FE},\Phi_h^{FE}) \\ &= a_h(\Phi_h-\Phi_h^{FE},\Phi_h-\Phi_h^{FE}) + \|\nabla\Phi_h^{FE}\|_{L^2}^2 + 2\alpha\|\Phi_h^{FE}-\Phi_h\|_{L^2}^2 \\ &\quad + 2\alpha\big(\Phi_h^{FE}-\Phi_h,\Phi_h\big). \end{aligned} \tag{3.73}$$

To bound $\|\Phi_h-\Phi_h^{FE}\|_{L^2}$ from above, we consider the following auxiliary problem:

$$\widetilde{a}_h(\phi,\chi) = \big(\Phi_h-\Phi_h^{FE},\chi\big) \qquad \forall\,\chi\in H^1(\mathcal{D}).$$

For $\sigma_e^0>\sigma_*^0$ for all $e\in\mathcal{E}_h$, the above problem has a unique solution $\phi\in H^{1+\theta}(\mathcal{D})$ for $0<\theta\le 1$ such that

$$\|\phi\|_{H^{1+\theta}(\mathcal{D})} \le C\|\Phi_h-\Phi_h^{FE}\|_{L^2} \qquad \text{for } \theta\in(0,1]. \tag{3.74}$$

By the definition of Φ_h^{FE}, we immediately get the following Galerkin orthogonality:

$$\widetilde{a}_h\big(\Phi_h-\Phi_h^{FE},\chi_h\big) = 0 \qquad \forall\chi_h\in S_h.$$

It follows from the duality argument (cf. [82, Theorem 2.14]) that

$$\begin{aligned} \|\Phi_h-\Phi_h^{FE}\|_{L^2}^2 &\le Ch^{2\theta}\widetilde{a}_h(\Phi_h-\Phi_h^{FE},\Phi_h-\Phi_h^{FE}) \\ &\le Ch^{2\theta}a_h(\Phi_h-\Phi_h^{FE},\Phi_h-\Phi_h^{FE}) + Ch^{2\theta}\alpha\|\Phi_h-\Phi_h^{FE}\|_{L^2}^2. \end{aligned} \tag{3.75}$$

For all h satisfying $Ch^{2\theta}\alpha<1$, we get

$$\|\Phi_h-\Phi_h^{FE}\|_{L^2}^2 \le \frac{Ch^{2\theta}}{1-Ch^{2\theta}\alpha}a_h(\Phi_h-\Phi_h^{FE},\Phi_h-\Phi_h^{FE}). \tag{3.76}$$

Now the last term on the right-hand side of (3.73) can be bounded as follows:

$$
\begin{aligned}
2\alpha\big(\Phi_h^{FE}-\Phi_h,\Phi_h\big) &\geq -2\alpha\|\Phi_h^{FE}-\Phi_h\|_{L^2}\|\Phi_h\|_{L^2} \\
&\geq -2\alpha\sqrt{\frac{Ch^{2\theta}a_h(\Phi_h-\Phi_h^{FE},\Phi_h-\Phi_h^{FE})}{1-Ch^{2\theta}\alpha}}\,\|\Phi_h\|_{L^2} \\
&\geq -\frac{1}{2}a_h(\Phi_h-\Phi_h^{FE},\Phi_h-\Phi_h^{FE})-\frac{2C\alpha^2h^{2\theta}}{1-Ch^{2\theta}\alpha}\|\Phi_h\|_{L^2}^2.
\end{aligned}
\tag{3.77}
$$

The second term on the right-hand side of (3.72) can be bounded by

$$
\begin{aligned}
&\int_{\mathcal{D}} f'(\widehat{P}_hu)\big((\Phi_h)^2-(\Phi_h^{FE})^2\big)\,dx \geq -C\int_{\mathcal{D}}\big|(\Phi_h)^2-(\Phi_h^{FE})^2\big|\,dx \\
&\quad = -C\int_{\mathcal{D}}\Big|-\big(\Phi_h-\Phi_h^{FE}\big)^2+2\Phi_h\big(\Phi_h-\Phi_h^{FE}\big)\Big|\,dx \\
&\quad \geq -C\|\Phi_h-\Phi_h^{FE}\|_{L^2}^2-\frac{\epsilon^3(1-\epsilon^3)}{1-\frac{\epsilon^3}{2}}\|\Phi_h\|_{L^2}^2-C\frac{1-\frac{\epsilon^3}{2}}{\epsilon^3(1-\epsilon^3)}\|\Phi_h-\Phi_h^{FE}\|_{L^2}^2.
\end{aligned}
\tag{3.78}
$$

Here we have used the facts that

$$
\|u\|_{L^\infty((0,T);L^\infty)}\leq C,\qquad |f'(\widehat{P}_hu)|\leq |f'(u)|+\epsilon^3\leq C. \tag{3.79}
$$

Substituting (3.76) into (3.78) yields

$$
\begin{aligned}
&\frac{1-\epsilon^3}{\epsilon}\int_{\mathcal{D}} f'(\widehat{P}_hu)\big((\Phi_h)^2-(\Phi_h^{FE})^2\big)\,dx \\
&\quad \geq -\gamma_3\frac{\epsilon(1-\epsilon^3)}{1-\frac{\epsilon^3}{2}}a_h(\Phi_h-\Phi_h^{FE},\Phi_h-\Phi_h^{FE})-\frac{\epsilon^2(1-\epsilon^3)}{1-\frac{\epsilon^3}{2}}\|\Phi_h\|_{L^2}^2,
\end{aligned}
\tag{3.80}
$$

where

$$
\gamma_3\geq\frac{Ch^{2\theta}}{1-Ch^{2\theta}\alpha}\cdot 2C\frac{1-\frac{\epsilon^3}{2}}{\epsilon(1-\epsilon^3)}\Big(1+\frac{1-\frac{\epsilon^3}{2}}{\epsilon^3(1-\epsilon^3)}\Big),
$$

and h is chosen small enough such that $\gamma_3<1/4$.

The term $\|\Phi_h\|_{L^2}^2$ can be bounded by

$$\begin{aligned}\|\Phi_h\|_{L^2}^2 &= (\Phi_h,\Phi_h) = a_h(\Delta_h^{-1}\Phi_h,\Phi_h) \le a_h(\Delta_h^{-1}\Phi_h,\Delta_h^{-1}\Phi_h)^{\frac{1}{2}} a_h(\Phi_h,\Phi_h)^{\frac{1}{2}} \\ &\le \frac{\rho}{2} a_h(\Delta_h^{-1}\Phi_h,\Delta_h^{-1}\Phi_h) + \frac{1}{2\rho} a_h(\Phi_h,\Phi_h)\end{aligned} \tag{3.81}$$

for any constant $\rho > 0$.

Adding the fifth term on the right-hand side of (3.72), the last term on the right-hand side of (3.77) and that of (3.80), we get for all h satisfying $2C\alpha^2 h^{2\theta}/(1 - Ch^{2\theta}\alpha) \le \epsilon$

$$\begin{aligned}-\left(\frac{\epsilon(1-\epsilon^3)}{1-\frac{\epsilon^3}{2}}\frac{2C\alpha^2h^{2\theta}}{1-Ch^{2\theta}\alpha} + \frac{3\epsilon^2(1-\epsilon^3)}{1-\frac{\epsilon^3}{2}}\right)\|\Phi_h\|_{L^2}^2 &\ge -\frac{4\epsilon^2(1-\epsilon^3)}{1-\frac{\epsilon^3}{2}}\|\Phi_h\|_{L^2}^2 \\ &\ge -\frac{\epsilon^4}{2(2-\epsilon^3)} a_h(\Phi_h,\Phi_h) - Ca_h(\Delta_h^{-1}\Phi_h,\Delta_h^{-1}\Phi_h).\end{aligned} \tag{3.82}$$

Combining (3.73), (3.77), (3.80) and (3.82) with (3.72), we have

$$\begin{aligned}&\epsilon a_h(\Phi_h,\Phi_h) + \frac{1-\epsilon^3}{\epsilon}\int_{\mathcal{D}} f'(P_hu)(\Phi_h)^2\,dx \\ &\ge \frac{\epsilon(1-\epsilon^3)}{4-2\epsilon^3} a_h(\Phi_h-\Phi_h^{FE},\Phi_h-\Phi_h^{FE}) + \frac{2\alpha\epsilon(1-\epsilon^3)}{1-\frac{\epsilon^3}{2}}\|\Phi_h^{FE}-\Phi_h\|_{L^2}^2 \\ &+ \frac{\epsilon(1-\epsilon^3)}{1-\frac{\epsilon^3}{2}}\|\nabla\Phi_h^{FE}\|_{L^2}^2 - Ca_h(\Delta_h^{-1}\Phi_h,\Delta_h^{-1}\Phi_h) \\ &+ \frac{1-\epsilon^3}{\epsilon}\int_{\mathcal{D}} f'(\widehat{P}_hu)(\Phi_h^{FE})^2\,dx + \frac{\epsilon^4}{2(2-\epsilon^3)} a_h(\Phi_h,\Phi_h).\end{aligned} \tag{3.83}$$

Applying the spectrum estimate (3.58), we get

$$\begin{aligned}&\epsilon\frac{1-\epsilon^3}{1-\frac{\epsilon^3}{2}}\|\nabla\Phi_h^{FE}\|_{L^2}^2 + \frac{1-\epsilon^3}{\epsilon}\int_{\mathcal{D}} f'(\widehat{P}_hu)(\Phi_h^{FE})^2\,dx \\ &= \frac{1-\epsilon^3}{1-\frac{\epsilon^3}{2}}\left(\epsilon\|\nabla\Phi_h^{FE}\|_{L^2}^2 + \frac{1-\frac{\epsilon^3}{2}}{\epsilon}\int_{\mathcal{D}} f'(\widehat{P}_hu)(\Phi_h^{FE})^2 dx\right) \\ &\ge -\frac{1-\epsilon^3}{1-\frac{\epsilon^3}{2}}(C_0+1)\|\nabla\Delta^{-1}\Phi_h^{FE}\|_{L^2}^2,\end{aligned}$$

which together with (3.83) implies that

$$\epsilon a_h(\Phi_h,\Phi_h)+\frac{1-\epsilon^3}{\epsilon}\int_{\mathcal{D}} f'(P_hu)(\Phi_h)^2\,dx \tag{3.84}$$
$$\geq -Ca_h(\Delta_h^{-1}\Phi_h,\Delta_h^{-1}\Phi_h)-C\|\nabla\Delta^{-1}\Phi_h^{FE}\|_{L^2}^2+\frac{2\alpha\epsilon(1-\epsilon^3)}{1-\frac{\epsilon^3}{2}}\|\Phi_h^{FE}-\Phi_h\|_{L^2}^2.$$

By the stability of Δ^{-1}, we have

$$\|\nabla\Delta^{-1}(\Phi_h-\Phi_h^{FE})\|_{L^2}^2\leq\widehat{C}\|\Phi_h-\Phi_h^{FE}\|_{L^2}^2,$$

which together with the triangle inequality yields

$$\|\nabla\Delta^{-1}\Phi_h^{FE}\|_{L^2}^2\leq 2\|\nabla\Delta^{-1}\Phi_h\|_{L^2}^2+2\widehat{C}\|\Phi_h-\Phi_h^{FE}\|_{L^2}^2.$$

Similarly, since $\Delta_h^{-1}\Phi_h$ is the elliptic projection of $\Delta^{-1}\Phi_h$, there holds

$$a_h(\Delta_h^{-1}\Phi_h,\Delta_h^{-1}\Phi_h)\leq C\|\nabla\Delta^{-1}\Phi_h\|_{L^2}^2.$$

Therefore, choosing $\alpha=O(\widehat{C}\epsilon^{-1})$, (3.84) can be further reduced into

$$\epsilon a_h(\Phi_h,\Phi_h)+\frac{1-\epsilon^3}{\epsilon}\int_{\mathcal{D}} f'(P_hu)(\Phi_h)^2\,dx\geq -c_0\|\nabla\Delta^{-1}\Phi_h\|_{L^2}^2$$

for some $c_0>0$. This proves (3.66), and the proof is complete. □

3.3.4 Error analysis

In this subsection, we shall derive some optimal error estimates for the proposed MIP-DG schemes (3.34)–(3.35), in which the constants in the error bounds depend on ϵ^{-1} only in low polynomial orders, instead of exponential orders. The key to obtaining such refined error bounds is to use the discrete spectrum estimate (3.66). In addition, the nonlinear Gronwall inequality presented in Lemma 2.2.4 also plays an important

role in the proof. To ease the presentation, we set $r = 1$ in this subsection and section 3.4, and generalization to $r > 1$ can be proven similarly.

The main results of this subsection are stated in the following theorem.

Theorem 3.3.6. *Let $\{(U^m, W^m)\}_{m=0}^M$ be the solution of scheme (3.34)–(3.35) with $r = 1$. Suppose that (GA) holds and $\sigma_e^0 > \sigma_*^0$ for all $e \in \mathcal{E}_h$, and define*

$$\rho_3(\epsilon) := \epsilon^{-\max\{2\sigma_1+\frac{13}{2}, 2\sigma_3+\frac{7}{2}, 2\sigma_2+4, 2\sigma_4\}-4}, \tag{3.85}$$

$$r(h, k; \epsilon, d, \sigma_i) := k^2\rho_1(\epsilon; d) + h^6\rho_3(\epsilon). \tag{3.86}$$

Then, under the following mesh and starting value conditions:

$$h^{2-\frac{d}{2}} \leq (C_1C_2)^{-1}\epsilon^{\max\{\sigma_1+\frac{11}{2}, \sigma_3+4\}}, \tag{3.87}$$

$$h^{1+r}|\ln h|^{\bar{r}} \leq (C_1C_3)^{-1}\epsilon^{\gamma+3}, \tag{3.88}$$

$$k \leq \epsilon^3 \qquad \text{when } f^m = (U^m)^3 - U^m, \tag{3.89}$$

$$h^{2\theta} \leq C\frac{\epsilon(1-\epsilon^3)}{8-4\epsilon^3}, \tag{3.90}$$

$$k \leq C\epsilon^{\frac{4(6+d)}{4-d}+(4d-2)\sigma_1}, \tag{3.91}$$

$$(U^0, 1) = (u_0, 1), \tag{3.92}$$

$$\|u_0 - U^0\|_{H^{-1}} \leq Ch^3\|u_0\|_{H^2}, \tag{3.93}$$

there hold the error estimates

$$\max_{0\le m\le M}\|u(t_m)-U^m\|_{-1,h}+\Big(\sum_{m=1}^{M}k^2\|d_t(u(t_m)-U^m)\|_{-1,h}^2\Big)^{\frac{1}{2}} \tag{3.94}$$
$$\le Cr(h,k;\epsilon,d,\sigma_i)^{\frac{1}{2}},$$
$$\Big(k\sum_{m=1}^{M}\|u(t_m)-U^m\|_{L^2}^2\Big)^{\frac{1}{2}} \tag{3.95}$$
$$\le C\Big(h^2\epsilon^{-\max\{\sigma_1+\frac{5}{2},\sigma_3+1\}}+\epsilon^{-2}r(h,k;\epsilon,d,\sigma_i)^{\frac{1}{2}}\Big),$$
$$\Big(k\sum_{m=1}^{M}\|\nabla\big(u(t_m)-U^m\big)\|_{L^2}^2\Big)^{\frac{1}{2}} \tag{3.96}$$
$$\le C\Big(h\epsilon^{-\max\{\sigma_1+\frac{5}{2},\sigma_3+1\}}+\epsilon^{-2}r(h,k;\epsilon,d,\sigma_i)^{\frac{1}{2}}\Big).$$

Moreover, if the starting value U^0 satisfies

$$\|u_0-U^0\|_{L^2}\le Ch^2\|u_0\|_{H^2}, \tag{3.97}$$

then there hold

$$\max_{0\le m\le M}\|u(t_m)-U^m\|_{L^2}+\Big(k\sum_{m=1}^{M}k\|d_t(u(t_m)-U^m)\|_{L^2}^2\Big)^{\frac{1}{2}} \tag{3.98}$$
$$+\Big(\frac{k}{\epsilon}\sum_{m=1}^{M}\|w(t_m)-W^m\|_{L^2}^2\Big)^{\frac{1}{2}}$$
$$\le C\Big(h^2\rho_3(\epsilon)^{\frac{1}{2}}+\epsilon^{-\frac{7}{2}}r(h,k;\epsilon,d,\sigma_i)^{\frac{1}{2}}\Big),$$
$$\max_{0\le m\le M}\|u(t_m)-U^m\|_{L^\infty} \tag{3.99}$$
$$\le C\Big(h^2|\ln h|\epsilon^{-\gamma}+h^{-\frac{d}{2}}\epsilon^{-\frac{7}{2}}r(h,k;\epsilon,d,\sigma_i)^{\frac{1}{2}}\Big).$$

Furthermore, suppose that the starting value W^0 satisfies

$$\|P_hw_0-W^0\|_{L^2}\le Ch^\beta \tag{3.100}$$

for some $\beta > 1$, and there exists a constant γ' such that

$$\operatorname*{ess\,sup}_{t\in[0,\infty)} \|w\|_{W^{2,\infty}} \le C\epsilon^{-\gamma'}, \tag{3.101}$$

then we have

$$\begin{aligned}\max_{0\le m\le M} \|w(t_m) - W^m\|_{L^2} \le C\Big(&h^2\rho_3(\epsilon) + h^\beta \\ &+ k^{-\frac{1}{2}}\epsilon^{-3} r(h,k;\epsilon,d,\sigma_i)^{\frac{1}{2}}\Big),\end{aligned} \tag{3.102}$$

$$\begin{aligned}\max_{0\le m\le M} \|w(t_m) - W^m\|_{L^\infty} \le C\Big(&h^{-\frac{d}{2}}\Big(k^{-\frac{1}{2}}\epsilon^{-3} r(h,k;\epsilon,d,\sigma_i)^{\frac{1}{2}} + h^\beta\Big) \\ &+ h^2|\ln h|\epsilon^{-\gamma'}\Big).\end{aligned} \tag{3.103}$$

Proof. In the following, we only give a proof for the convex splitting scheme corresponding to $f^m = (u^m)^3 - u^{m-1}$ in (3.44) because the proof for the fully implicit scheme with $f^m = (u^m)^3 - u^m$ is almost same. Since the proof is long, we divide it into four steps.

Step 1: It is obvious that equations (1.30)–(1.33) imply that

$$\big(u_t(t_m), \eta_h\big) + a_h(w(t_m), \eta_h) = 0 \qquad \forall \eta_h \in V_h, \tag{3.104}$$

$$\epsilon a_h(u(t_m), v_h) + \frac{1}{\epsilon}\big(f(u(t_m)), v_h\big) = \big(w(t_m), v_h\big) \qquad \forall v_h \in V_h. \tag{3.105}$$

Define error functions $E^m := u(t_m) - U^m$ and $G^m := w(t_m) - W^m$. Subtracting (3.34) from (3.104) and (3.35) from (3.105) yield the following error equations:

$$\big(d_t E^m, \eta_h\big) + a_h(G^m, \eta_h) = \big(R(u_{tt}, m), \eta_h\big) \qquad \forall \eta_h \in V_h, \tag{3.106}$$

$$\epsilon a_h(E^m, v_h) + \frac{1}{\epsilon}\big(f(u(t_m)) - f(U^m), v_h\big) = \big(G^m, v_h\big) \qquad \forall v_h \in V_h, \tag{3.107}$$

where

$$R(u_{tt}; m) := \frac{1}{k}\int_{t_{m-1}}^{t_m} (s - t_{m-1}) u_{tt}(s)\, ds.$$

It follows from (xiv) in Proposition 3.2.1 that

$$k\sum_{m=1}^{M}\|R(u_{tt};m)\|_{H^{-1}}^2 \le \frac{1}{k}\sum_{m=1}^{M}\Big(\int_{t_{m-1}}^{t_m}(s-t_{m-1})^2\,ds\Big)\Big(\int_{t_{m-1}}^{t_m}\|u_{tt}(s)\|_{H^{-1}}^2\,ds\Big)$$
$$\le Ck^2\rho_1(\epsilon,d).$$

Introduce the error decompositions

$$E^m=\Theta^m+\Phi^m,\qquad G^m=\Lambda^m+\Psi^m, \tag{3.108}$$

where

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

Using the definition of the operator P_h in (3.37), (3.106)–(3.107) can be rewritten as

$$\big(d_t\Phi^m,\eta_h\big)+a_h(\Psi^m,\eta_h)=-\big(d_t\Theta^m,\eta_h\big)+\big(R(u_{tt},m),\eta_h\big)\qquad\forall\eta_h\in V_h, \tag{3.109}$$
$$\epsilon a_h(\Phi^m,v_h)+\frac{1}{\epsilon}\big(f(u(t_m))-f^m,v_h\big)=\big(\Psi^m,v_h\big)+\big(\Lambda^m,v_h\big)\qquad\forall v_h\in V_h. \tag{3.110}$$

Setting $\eta_h=-\Delta_h^{-1}\Phi^m$ in (3.109) and $v_h=\Phi^m$ in (3.110), adding the resulting equations and summing over m from 1 to ℓ, we get

$$a_h(\Delta_h^{-1}\Phi^\ell,\Delta_h^{-1}\Phi^\ell)+\sum_{m=1}^{\ell}a_h(\Delta_h^{-1}\Phi^m-\Delta_h^{-1}\Phi^{m-1},\Delta_h^{-1}\Phi^m-\Delta_h^{-1}\Phi^{m-1}) \tag{3.111}$$
$$+2k\sum_{m=1}^{\ell}\epsilon a_h(\Phi^m,\Phi^m)+2k\sum_{m=1}^{\ell}\frac{1}{\epsilon}\big(f(u(t_m))-f^m,\Phi^m\big)$$
$$=2k\sum_{m=1}^{\ell}\Big(\big(R(u_{tt},m),-\Delta_h^{-1}\Phi^m\big)-\big(d_t\Theta^m,-\Delta_h^{-1}\Phi^m\big)+\big(\Lambda^m,\Phi^m\big)\Big)$$
$$+a_h(\Delta_h^{-1}\Phi^0,\Delta_h^{-1}\Phi^0).$$

Step 2: For $\sigma_e^0 > \sigma_*^e$ for all $e \in \mathcal{E}_h$, the first long term on the right-hand side of (3.111) can be bounded as follows

$$
\begin{aligned}
&2k\sum_{m=1}^{\ell}\Big(\big(R(u_{tt},m),-\Delta_h^{-1}\Phi^m\big)+\big(d_t\Theta^m,-\Delta_h^{-1}\Phi^m\big)+\big(\Lambda^m,\Phi^m\big)\Big) \qquad (3.112)\\
&\quad\le Ck\sum_{m=1}^{\ell}\Big(\|R(u_{tt};m)\|_{H^{-1}}^2+\|d_t\Theta^m\|_{H^{-1}}^2+(1-\epsilon^3)\epsilon^{-4}\|\Lambda^m\|_{H^{-1}}^2\Big)\\
&\qquad+k\sum_{m=1}^{\ell}\Big(a_h(\Delta_h^{-1}\Phi^m,\Delta_h^{-1}\Phi^m)+\frac{\epsilon^4}{1-\epsilon^3}a_h(\Phi^m,\Phi^m)\Big)\\
&\quad\le k\sum_{m=1}^{\ell}\Big(a_h(\Delta_h^{-1}\Phi^m,\Delta_h^{-1}\Phi^m)+\frac{\epsilon^4}{1-\epsilon^3}a_h(\Phi^m,\Phi^m)\Big)\\
&\qquad+C\Big(k^2\rho_1(\epsilon,d)+h^6\rho_3(\epsilon)\Big),
\end{aligned}
$$

where we have used (xi) in Proposition 3.2.1 and the following facts [29]:

$$
\|u-P_hu\|_{H^{-1}}\le Ch^3\|u\|_{H^2},\qquad \|w-P_hw\|_{H^{-1}}\le Ch^3\|w\|_{H^2}.
$$

We now bound the last term on the left-hand side of (3.111). By the definition of f^m, we have

$$
\begin{aligned}
f(u(t_m))-f^m&=f(u(t_m))-f\big(P_hu(t_m)\big)+f\big(P_hu(t_m)\big)-f^m\\
&\ge-\big|f(u(t_m))-f\big(P_hu(t_m)\big)\big|+\big(P_hu(t_m)\big)^3-P_hu(t_m)-(U^m)^3+U^{m-1}\\
&\ge-C|\Theta^m|+\Big(\big(P_hu(t_m)\big)^2+P_hu(t_m)\,U^m+(U^m)^2\Big)\Phi^m-\Phi^m-kd_tU^m\\
&\ge-C|\Theta^m|+f'\big(P_hu(t_m)\big)\,\Phi^m-3P_hu(t_m)\,(\Phi^m)^2+(\Phi^m)^3-kd_tU^m.
\end{aligned}
$$

By the discrete energy law (3.48), (3.44) and (3.81), we obtain for any $1 \le \ell \le M$

$$
\begin{aligned}
&2k\sum_{m=1}^{\ell}\frac{1}{\epsilon}\big(f(u(t_m))-f^m,\Phi^m\big) \qquad (3.113)\\
&\ge -\frac{Ck}{\epsilon}\sum_{m=1}^{\ell}\|\Theta^m\|_{H^{-1}(\mathcal{T}_h)}\|\Phi^m\|_{H^1(\mathcal{T}_h)}+2k\sum_{m=1}^{\ell}\frac{1}{\epsilon}\Big(f'\big(P_hu(t_m)\big),(\Phi^m)^2\Big)\\
&\quad-\frac{Ck}{\epsilon}\sum_{m=1}^{\ell}\|\Phi^m\|_{L^3}^3+\frac{2k}{\epsilon}\sum_{m=1}^{\ell}\|\Phi^m\|_{L^4}^4-\frac{2k}{\epsilon}\sum_{m=1}^{\ell}k\|d_tU^m\|_{-1,h}\,\|\Phi^m\|_{\alpha}\\
&\ge 2k\sum_{m=1}^{\ell}\frac{1}{\epsilon}\Big(f'\big(P_hu(t_m)\big),(\Phi^m)^2\Big)+\frac{2k}{\epsilon}\sum_{m=1}^{\ell}\|\Phi^m\|_{L^4}^4-\frac{Ck}{\epsilon}\sum_{m=1}^{\ell}\|\Phi^m\|_{L^3}^3\\
&\quad-k\epsilon^4\sum_{m=1}^{\ell}\|\Phi^m\|_a^2-C\Big(h^6\epsilon^{-6}\|u\|^2_{L^2((0,T);H^s(\mathcal{D}))}+k^2\epsilon^{-6}E_h(u_h^0)\Big)\\
&\ge 2k\sum_{m=1}^{\ell}\frac{1}{\epsilon}\Big(f'\big(P_hu(t_m)\big),(\Phi^m)^2\Big)+\frac{2k}{\epsilon}\sum_{m=1}^{\ell}\|\Phi^m\|_{L^4}^4-\frac{Ck}{\epsilon}\sum_{m=1}^{\ell}\|\Phi^m\|_{L^3}^3\\
&\quad-k\frac{\epsilon^4}{1-\epsilon^3}\sum_{m=1}^{\ell}a_h(\Phi^m,\Phi^m)-C\Big(h^6\epsilon^{-6}\|u\|^2_{L^2((0,T);H^s(\mathcal{D}))}+k^2\epsilon^{-6}E_h(U^0)\Big).
\end{aligned}
$$

Substituting (3.112) and (3.113) into (3.111) we get

$$
\begin{aligned}
&a_h(\Delta_h^{-1}\Phi^\ell,\Delta_h^{-1}\Phi^\ell)+\sum_{m=1}^{\ell}a_h(\Delta_h^{-1}\Phi^m-\Delta_h^{-1}\Phi^{m-1},\Delta_h^{-1}\Phi^m-\Delta_h^{-1}\Phi^{m-1}) \qquad (3.114)\\
&\qquad+\frac{2k(1-5\epsilon^3)}{1-\epsilon^3}\sum_{m=1}^{\ell}\Big(\epsilon a_h(\Phi^m,\Phi^m)+\frac{1-\epsilon^3}{\epsilon}\big(f'(P_hu(t_m))\Phi^m,\Phi^m\big)\Big)\\
&\qquad+\frac{6\epsilon^4}{1-\epsilon^3}k\sum_{m=1}^{\ell}a_h(\Phi^m,\Phi^m)+\frac{2k}{\epsilon}\sum_{m=1}^{\ell}\|\Phi^m\|_{L^4}^4\\
&\le Ck\sum_{m=1}^{\ell}a_h(\Delta_h^{-1}\Phi^m,\Delta_h^{-1}\Phi^m)+\frac{Ck}{\epsilon}\sum_{m=1}^{\ell}\|\Phi^m\|_{L^3}^3\\
&\qquad-10k\epsilon^2\sum_{m=1}^{\ell}\big(f'(P_hu(t_m))\Phi^m,\Phi^m\big)+C\big(k^2\rho_1(\epsilon;d)+h^6\rho_3(\epsilon)\big)\\
&\qquad+C\Big(h^6\epsilon^{-6}\|u\|^2_{L^2((0,T);H^s(\mathcal{D})}+k^2\epsilon^{-6}E_h(U^0)\Big).
\end{aligned}
$$

Step 3: To control the second term on the right-hand side of (3.114), we appeal to the following Gagliardo-Nirenberg inequality [2]:

$$\|v\|_{L^3(K)}^3 \le C\Big(\|\nabla v\|_{L^2(K)}^{\frac{d}{2}}\big\|v\big\|_{L^2(K)}^{\frac{6-d}{2}} + \|v\|_{L^2(K)}^3\Big) \qquad \forall K \in \mathcal{T}_h.$$

Thus we get

$$\begin{aligned}\frac{Ck}{\epsilon}\sum_{m=1}^{\ell}\|\Phi^m\|_{L^3}^3 \le& \epsilon^4 k\sum_{m=1}^{\ell}\|\nabla\Phi^m\|_{L^2(\mathcal{T}_h)}^2 + \frac{Ck}{\epsilon}\sum_{m=1}^{\ell}\|\Phi^m\|_{L^2}^3 \\ &+ C\epsilon^{-\frac{4(1+d)}{4-d}}k\sum_{m=1}^{\ell}\big\|\Phi^m\big\|_{L^2}^{\frac{2(6-d)}{4-d}} \\ \le& \frac{\epsilon^4}{1-\epsilon^3}k\sum_{m=1}^{\ell}a_h(\Phi^m,\Phi^m) + \frac{Ck}{\epsilon}\sum_{m=1}^{\ell}\|\Phi^m\|_{L^2}^3 \\ &+ C\epsilon^{-\frac{4(1+d)}{4-d}}k\sum_{m=1}^{\ell}\big\|\Phi^m\big\|_{L^2}^{\frac{2(6-d)}{4-d}}.\end{aligned} \tag{3.115}$$

The third item on the right-hand side of (3.114) can be bounded by

$$\begin{aligned}&-10k\epsilon^2(f'(P_hu(t_m))\Phi^m,\Phi^m) \\ &\qquad\le k\frac{\epsilon^4}{1-\epsilon^3}a_h(\Phi^m,\Phi^m) + kCa_h(\Delta_h^{-1}\Phi^m,\Delta_h^{-1}\Phi^m).\end{aligned} \tag{3.116}$$

Again, here we have used (3.81).

Finally, for the third term on the left-hand side of (3.114), we utilize the discrete spectrum estimate (3.66) to bound it from below as follows:

$$\epsilon a_h(\Phi^m,\Phi^m) + \frac{1-\epsilon^3}{\epsilon}\big(f'(P_hu(t_m))\Phi^m,\Phi^m\big) \ge -c_0\|\nabla\Delta^{-1}\Phi^m\|_{L^2}^2. \tag{3.117}$$

By the stability of Δ^{-1} and (3.81), we also have

$$c_0\|\nabla\Delta^{-1}\Phi^m\|_{L^2}^2 \le C\|\Phi^m\|_{L^2}^2 \le \frac{\epsilon^4}{1-\epsilon^3}a_h(\Phi^m,\Phi^m) + Ca_h(\Delta_h^{-1}\Phi^m,\Delta_h^{-1}\Phi^m). \tag{3.118}$$

Step 4: Substituting (3.115), (3.116), (3.117), (3.118) into (3.114), we get

$$
\begin{aligned}
&a_h(\Delta_h^{-1}\Phi^\ell, \Delta_h^{-1}\Phi^\ell) + \sum_{m=1}^{\ell} a_h(\Delta_h^{-1}\Phi^m - \Delta_h^{-1}\Phi^{m-1}, \Delta_h^{-1}\Phi^m - \Delta_h^{-1}\Phi^{m-1}) \qquad (3.119)\\
&\qquad + \frac{2\epsilon^4 k}{1-\epsilon^3}\sum_{m=1}^{\ell} a_h(\Phi^m, \Phi^m) + \frac{2k}{\epsilon}\sum_{m=1}^{\ell}\|\Phi^m\|_{L^4}^4\\
&\le Ck\sum_{m=1}^{\ell} a_h(\Delta_h^{-1}\Phi^m, \Delta_h^{-1}\Phi^m) + \frac{Ck}{\epsilon}\sum_{m=1}^{\ell}\|\Phi^m\|_{L^2}^3\\
&\quad + C\epsilon^{-\frac{4(1+d)}{4-d}} k\sum_{m=1}^{\ell}\left\|\Phi^m\right\|_{L^2}^{\frac{2(6-d)}{4-d}} + C\big(k^2\rho_1(\epsilon;d) + h^6\rho_3(\epsilon)\big)\\
&\quad + C\Big(h^6\epsilon^{-6}\|u\|^2_{L^2((0,T);H^s(\mathcal{D})} + k^2\epsilon^{-6}E_h(U^0)\Big).
\end{aligned}
$$

By discrete energy law (3.48), General Assumption (3.2), H^1 stability of elliptic projection, L^∞ stability(or L^∞ error estimate and triangle inequality) of elliptic projection, we can get for any $0 \le \ell \le M$

$$
\|U^\ell\|_{L^2} \le k\sum_{m=1}^{\ell}\|d_t U^m\|_{L^2} + \|U^0\|_{L^2} \le C\epsilon^{-\sigma_1}.
$$

Since the projection of u is bounded, then for any $0 \le \ell \le M$

$$
\|\Phi^\ell\|_{L^2} \le C\epsilon^{-\sigma_1}. \qquad (3.120)
$$

We point out that the exponent for $\left\|\Phi^m\right\|_{L^2}$ is $\frac{2(6-d)}{4-d}$, which is bigger than 3 for $d = 2, 3$. By (3.120) we have

$$
\left\|\Phi^m\right\|_{L^2}^4 \le C\epsilon^{-\sigma_1}\left\|\Phi^m\right\|_{L^2}^3, \qquad \left\|\Phi^m\right\|_{L^2}^6 \le C\epsilon^{-3\sigma_1}\left\|\Phi^m\right\|_{L^2}^3.
$$

Using the Schwarz and Young's inequalities, we have

$$\begin{aligned}\left\|\Phi^m\right\|_{L^2}^3 &= \left(\left\|\Phi^m\right\|_{L^2}^2\right)^{\frac{3}{2}} = a_h(-\Delta_h^{-1}\Phi^m,\Phi^m)^{\frac{3}{2}} \\ &\le a_h(\Delta_h^{-1}\Phi^m,\Delta_h^{-1}\Phi^m)^{\frac{3}{4}}\, a_h(\Phi^m,\Phi^m)^{\frac{3}{4}} \\ &\le \epsilon^{\frac{4(1+d)}{4-d}+\sigma_1+2(d-2)\sigma_1}\frac{\epsilon^4}{1-\epsilon^3}a_h(\Phi^m,\Phi^m) \\ &\quad + C\epsilon^{-4}\epsilon^{-\frac{4(1+d)}{4-d}-\sigma_1-2(d-2)\sigma_1}a_h(\Delta_h^{-1}\Phi^m,\Delta_h^{-1}\Phi^m)^3.\end{aligned} \tag{3.121}$$

Therefore, (3.119) becomes

$$\begin{aligned}&a_h(\Delta_h^{-1}\Phi^\ell,\Delta_h^{-1}\Phi^\ell) + \sum_{m=1}^{\ell} a_h(\Delta_h^{-1}\Phi^m-\Delta_h^{-1}\Phi^{m-1},\Delta_h^{-1}\Phi^m-\Delta_h^{-1}\Phi^{m-1}) \\ &\qquad + \frac{\epsilon^4 k}{1-\epsilon^3}\sum_{m=1}^{\ell} a_h(\Phi^m,\Phi^m) + \frac{2k}{\epsilon}\sum_{m=1}^{\ell}\|\Phi^m\|_{L^4}^4 \\ &\le Ck\sum_{m=1}^{\ell} a_h(\Delta_h^{-1}\Phi^m,\Delta_h^{-1}\Phi^m) \\ &\quad + Ck\epsilon^{-\frac{4(6+d)}{4-d}-2\sigma_1-4(d-2)\sigma_1}\sum_{m=1}^{\ell} a_h(\Delta_h^{-1}\Phi^m,\Delta_h^{-1}\Phi^m)^3 \\ &\quad + C\big(k^2\rho_1(\epsilon;d)+h^6\rho_3(\epsilon)\big) + C\Big(h^6\epsilon^{-6}\|u\|^2_{L^2((0,T);H^s(\mathcal{D})}+k^2\epsilon^{-6}E_h(U^0)\Big) \\ &\le Ck\sum_{m=1}^{\ell} a_h(\Delta_h^{-1}\Phi^m,\Delta_h^{-1}\Phi^m) \\ &\quad + Ck\epsilon^{-\frac{4(6+d)}{4-d}-2\sigma_1-4(d-2)\sigma_1}\sum_{m=1}^{\ell} a_h(\Delta_h^{-1}\Phi^m,\Delta_h^{-1}\Phi^m)^3 \\ &\quad + C\big(k^2\rho_1(\epsilon;d)+h^6\rho_3(\epsilon)\big).\end{aligned} \tag{3.122}$$

On noting that U^m can be written as

$$U^\ell = k\sum_{m=1}^{\ell} d_t U^m + U^0, \tag{3.123}$$

then by (3.2) and (3.48), we get

$$\|U^\ell\|_{-1,h} \le k\sum_{m=1}^{\ell}\|d_tU^m\|_{-1,h} + \|U^0\|_{-1,h} \le C\epsilon^{-\sigma_1}. \tag{3.124}$$

Using the boundedness of the projection, we have

$$\|\Phi^\ell\|^2_{-1,h} \le C\epsilon^{-2\sigma_1}. \tag{3.125}$$

Also, (3.122) can be written in the following equivalent form

$$\begin{aligned} &a_h(\Delta_h^{-1}\Phi^\ell,\Delta_h^{-1}\Phi^\ell) + \sum_{m=1}^{\ell} a_h(\Delta_h^{-1}\Phi^m - \Delta_h^{-1}\Phi^{m-1}, \Delta_h^{-1}\Phi^m - \Delta_h^{-1}\Phi^{m-1}) \\ &\quad + \frac{\epsilon^4 k}{1-\epsilon^3}\sum_{m=1}^{\ell} a_h(\Phi^m,\Phi^m) + \frac{2k}{\epsilon}\sum_{m=1}^{\ell}\|\Phi^m\|^4_{L^4} \le M_1 + M_2, \end{aligned} \tag{3.126}$$

where

$$\begin{aligned} M_1 :=\ & Ck\sum_{m=1}^{\ell-1} a_h(\Delta_h^{-1}\Phi^m,\Delta_h^{-1}\Phi^m) \\ &+ Ck\epsilon^{-\frac{4(6+d)}{4-d}-2\sigma_1-4(d-2)\sigma_1}\sum_{m=1}^{\ell-1} a_h(\Delta_h^{-1}\Phi^m,\Delta_h^{-1}\Phi^m)^3 \\ &+ C\big(k^2\rho_1(\epsilon;d) + h^6\rho_3(\epsilon)\big), \end{aligned} \tag{3.127}$$

$$\begin{aligned} M_2 :=\ & Cka_h(\Delta_h^{-1}\Phi^\ell,\Delta_h^{-1}\Phi^\ell) \\ &+ Ck\epsilon^{-\frac{4(6+d)}{4-d}-2\sigma_1-4(d-2)\sigma_1} a_h(\Delta_h^{-1}\Phi^\ell,\Delta_h^{-1}\Phi^\ell)^3. \end{aligned} \tag{3.128}$$

It is easy to check that

$$M_2 \le \frac{1}{2}\|\Phi^\ell\|^2_{-1,h} \qquad \text{provided that} \quad k \le C\epsilon^{\frac{4(6+d)}{4-d}+(4d-2)\sigma_1}. \tag{3.129}$$

Under this restriction, we have

$$
\begin{aligned}
&a_h(\Delta_h^{-1}\Phi^\ell, \Delta_h^{-1}\Phi^\ell) + 2\sum_{m=1}^{\ell} a_h(\Delta_h^{-1}\Phi^m - \Delta_h^{-1}\Phi^{m-1}, \Delta_h^{-1}\Phi^m - \Delta_h^{-1}\Phi^{m-1}) \qquad (3.130)\\
&\qquad\qquad + \frac{2\epsilon^4 k}{1-\epsilon^3}\sum_{m=1}^{\ell} a_h(\Phi^m, \Phi^m) + \frac{4k}{\epsilon}\sum_{m=1}^{\ell} \|\Phi^m\|_{L^4}^4\\
&\le 2Ck\sum_{m=1}^{\ell-1} a_h(\Delta_h^{-1}\Phi^m, \Delta_h^{-1}\Phi^m) + 2C\big(k^2\rho_1(\epsilon;d) + h^6\rho_3(\epsilon)\big)\\
&\quad + 2Ck\epsilon^{-\frac{4(6+d)}{4-d}-2\sigma_1-4(d-2)\sigma_1}\sum_{m=1}^{\ell-1} a_h(\Delta_h^{-1}\Phi^m, \Delta_h^{-1}\Phi^m)^3\\
&\le Ck\sum_{m=1}^{\ell-1} a_h(\Delta_h^{-1}\Phi^m, \Delta_h^{-1}\Phi^m) + C\big(k^2\rho_1(\epsilon;d) + h^6\rho_3(\epsilon)\big)\\
&\quad + Ck\epsilon^{-\frac{4(6+d)}{4-d}-2\sigma_1-4(d-2)\sigma_1}\sum_{m=1}^{\ell-1} a_h(\Delta_h^{-1}\Phi^m, \Delta_h^{-1}\Phi^m)^3.
\end{aligned}
$$

Define the slack variable $d_\ell \ge 0$ such that

$$
\begin{aligned}
&a_h(\Delta_h^{-1}\Phi^\ell, \Delta_h^{-1}\Phi^\ell) + 2\sum_{m=1}^{\ell} a_h(\Delta_h^{-1}\Phi^m - \Delta_h^{-1}\Phi^{m-1}, \Delta_h^{-1}\Phi^m - \Delta_h^{-1}\Phi^{m-1}) \qquad (3.131)\\
&\qquad\qquad + \frac{2\epsilon^4 k}{1-\epsilon^3}\sum_{m=1}^{\ell} a_h(\Phi^m, \Phi^m) + \frac{4k}{\epsilon}\sum_{m=1}^{\ell} \|\Phi^m\|_{L^4)}^4 + d_\ell\\
&= Ck\sum_{m=1}^{\ell-1} a_h(\Delta_h^{-1}\Phi^m, \Delta_h^{-1}\Phi^m) + C\big(k^2\rho_1(\epsilon;d) + h^6\rho_3(\epsilon)\big)\\
&\quad + Ck\epsilon^{-\frac{4(6+d)}{4-d}-2\sigma_1-4(d-2)\sigma_1}\sum_{m=1}^{\ell-1} a_h(\Delta_h^{-1}\Phi^m, \Delta_h^{-1}\Phi^m)^3.
\end{aligned}
$$

We also define $\{S_\ell\}_{\ell\ge 1}$ by

$$
\begin{aligned}
S_\ell = d_\ell &+ 2\sum_{m=1}^{\ell} a_h(\Delta_h^{-1}\Phi^m - \Delta_h^{-1}\Phi^{m-1}, \Delta_h^{-1}\Phi^m - \Delta_h^{-1}\Phi^{m-1}) \qquad (3.132)\\
&+ a_h(\Delta_h^{-1}\Phi^\ell, \Delta_h^{-1}\Phi^\ell) + \frac{2\epsilon^4 k}{1-\epsilon^3}\sum_{m=1}^{\ell} a_h(\Phi^m, \Phi^m) + \frac{4k}{\epsilon}\sum_{m=1}^{\ell} \|\Phi^m\|_{L^4}^4,
\end{aligned}
$$

and equation (3.131) shows that

$$S_1 = C\big(k^2\rho_1(\epsilon;d) + h^6\rho_3(\epsilon)\big).$$

Then

$$S_{\ell+1} - S_\ell \le CkS_\ell + Ck\epsilon^{-\frac{4(6+d)}{4-d}-2\sigma_1-4(d-2)\sigma_1}S_\ell^3 \qquad \forall \ell \ge 1. \tag{3.133}$$

Applying Lemma 2.2.4 ([40, 42, 73])to $\{S_\ell\}_{\ell\ge 1}$ defined above, we obtain $\forall \ell \ge 1$,

$$S_\ell \le a_\ell^{-1}\left\{S_1^{-2} - 2C\epsilon^{-\frac{4(6+d)}{4-d}-2\sigma_1-4(d-2)\sigma_1}k\sum_{s=1}^{\ell-1}a_{s+1}^{-2}\right\}^{-\frac{1}{2}} \tag{3.134}$$

provided that

$$S_1^{-2} - 2C\epsilon^{-\frac{4(6+d)}{4-d}-2\sigma_1-4(d-2)\sigma_1}k\sum_{s=1}^{\ell-1}a_{s+1}^{-2} > 0. \tag{3.135}$$

We note that $a_s\,(1 \le s \le \ell)$ are all bounded as $k \to 0$, therefore, (3.135) holds under the mesh constraint stated in the theorem. It follows from (3.92) and (3.93) that

$$S_\ell \le 2a_\ell^{-1}S_1 \le C\big(k^2\rho_1(\epsilon;d) + h^6\rho_3(\epsilon)\big). \tag{3.136}$$

Then (3.94) follows from the triangle inequality on $E^m = \Theta^m + \Phi^m$. (3.96) is obtained by taking the test function $\eta_h = \Phi^m$ in (3.109) and $v_h = \Phi^m$ in (3.110), and (3.95) is a consequence of the Poincarè inequality.

Now setting $\eta_h = \Phi^m$ in (3.109) and $v_h = -\frac{1}{\epsilon}\Psi^m$ in (3.110), and adding the resulting equations yield

$$\begin{aligned}\frac{1}{2}d_t\|\Phi^m\|_{L^2}^2 + \frac{k}{2}\|d_t\Phi^m\|_{L^2}^2 + \frac{1}{\epsilon}\|\Psi^m\|_{L^2}^2 = \frac{1}{\epsilon^2}\big(f(u(t_m)) - f(U^m), \Psi^m\big) \\ + \big(R(u_{tt};m), \Phi^m\big) - \big(d_t\Theta^m, \Phi^m\big) - \frac{1}{\epsilon}\big(\Lambda^m, \Psi^m\big).\end{aligned} \tag{3.137}$$

The last three terms on the right-hand side of (3.137) can be bounded in the same way as in (3.112), and the first term can be controlled as

$$\begin{aligned}\frac{1}{\epsilon^2}\big(f(u(t_m))-f(U^m),\Psi^m\big) &= \frac{1}{\epsilon^2}\big(f'(\xi)E^m,\Psi^m\big)\\ &\le \frac{1}{2\epsilon}\|\Psi^m\|_{L^2}^2+\frac{C}{\epsilon^3}\|E^m\|_{L^2}^2.\end{aligned} \tag{3.138}$$

Multiplying both sides of (3.137) by k and summing over m from 1 to M yield the desired estimate (3.98). Estimate (3.99) follows from an applications of the following inverse inequality:

$$\|\Phi^m\|_{L^\infty}\le h^{-\frac{d}{2}}\|\Phi^m\|_{L^2}, \tag{3.139}$$

and the following L^∞ estimate for the elliptic projection:

$$\|u-P_hu\|_{L^\infty}\le Ch^2|\ln h|\|u\|_{W^{s,\infty}}\qquad \forall u\in H^2(\mathcal{D}). \tag{3.140}$$

Finally, it is well known that there holds the following estimate for the elliptic projection operator:

$$\max_{0\le m\le M}\|\Lambda^m\|_{L^2}+\left(k\sum_{m=0}^{M}k\|d_t\Lambda^m\|_{L^2}^2\right)^{\frac{1}{2}}\le Ch^2\rho_2(\epsilon). \tag{3.141}$$

Using the identity

$$\big(d_t\Phi^m,\Phi^m\big)=\frac{1}{2}d_t\|\Phi^m\|_{L^2}^2+\frac{k}{2}\|d_t\Phi^m\|_{L^2}^2, \tag{3.142}$$

we get

$$\begin{aligned}\frac{1}{2}\|\Psi^M\|_{L^2}^2+k\sum_{m=1}^{M}\frac{k}{2}\|d_t\Psi^m\|_{L^2}^2 &= k\sum_{m=1}^{M}\big(d_t\Psi^m,\Psi^m\big)+\frac{1}{2}\|\Psi^0\|_{L^2}^2\\ &\le k\sum_{m=1}^{M}\left(\frac{k}{4}\|d_t\Psi^m\|_{L^2}^2+\frac{1}{k}\|\Psi^m\|_{L^2}^2\right)+\frac{1}{2}\|\Psi^0\|_{L^2}^2.\end{aligned} \tag{3.143}$$

The first term on the right hand side of (3.143) can be absorbed by the second term on the left hand side of (3.143). The second tern on the right hand side of (3.143) has been obtained in (3.98). Estimate (3.102) for W^m then follows from (3.141) and (3.143). (3.103) follows from an application of the triangle inequality, the inverse inequality, and (3.140). This completes the proof. □

3.4 Convergence of numerical interfaces

In this section, we prove that the numerical interface defined as the zero level set of the finite element interpolation of the solution U^m converges to the moving interface of the Hele-Shaw problem under the assumption that the Hele-Shaw problem has a unique global (in time) classical solution. To the end, we first cite the following PDE convergence result proved in [3].

Theorem 3.4.1. *Let $\mathcal{D}$ be a given smooth domain and Γ_{00} be a smooth closed hypersurface in $\mathcal{D}$. Suppose that the Hele-Shaw problem starting from Γ_{00} has a unique smooth solution $\big(w, \Gamma := \bigcup_{0\le t\le T}(\Gamma_t \times \{t\})\big)$ in the time interval $[0,T]$ such that $\Gamma_t \subseteq \mathcal{D}$ for all $t \in [0,T]$. Then there exists a family of smooth functions $\{u_0^\epsilon\}_{0<\epsilon\le 1}$ which are uniformly bounded in $\epsilon \in (0,1]$ and $(x,t) \in \overline{\mathcal{D}}_T$, such that if u^ϵ solves the Cahn-Hilliard problem (1.30)–(1.33), then*

(i) $\lim\limits_{\epsilon\to 0} u^\epsilon(x,t) = \begin{cases} 1 & \text{if } (x,t) \in \mathcal{O} \\ -1 & \text{if } (x,t) \in \mathcal{I} \end{cases}$ *uniformly on compact subsets, where $\mathcal{I}$ and $\mathcal{O}$ stand for the "inside" and "outside" of Γ;*

(ii) $\lim\limits_{\epsilon\to 0}\big(\epsilon^{-1} f(u^\epsilon) - \epsilon\Delta u^\epsilon\big)(x,t) = -w(x,t)$ *uniformly on $\overline{\mathcal{D}}_T$.*

We note that since U^m is multi-valued on the edges of the mesh $\mathcal{T}_h$, its zero-level set is not well defined. To avoid this technicality, we use a continuous finite element interpolation of U^m to define the numerical interface. Let $\widehat{U}^m \in S_h$ denote the finite element approximation of U^m which is defined using the averaged degrees of freedom

of U^m as the degrees of freedom for determining $\widehat{U}^m$ (cf. [58]). By the construction, $\widehat{U}^m$ is expected to be very close to U^m, hence, $\widehat{U}^m$ should also be very close to $u(t_m)$. This is indeed the case as stated in the following theorem, which says that Theorem 3.3.6 also hold for $\widehat{U}^m$.

Theorem 3.4.2. *Let U^m denote the solution of scheme (3.32)–(3.45) and $\widehat{U}^m$ denote its finite element approximation as defined above. Then under the assumptions of Theorem 3.3.6 the error estimates for U^m given in Theorem 3.3.6 are still valid for $\widehat{U}^m$, in particular, there holds*

$$\max_{0\le m\le M} \|u(t_m) - \widehat{U}^m\|_{L^\infty(\mathcal{T}_h)} \le C\Big(h^2|\ln h|\epsilon^{-\gamma} + h^{-\frac{d}{2}}\epsilon^{-\frac{7}{2}} r(h,k;\epsilon,d,\sigma_i)^{\frac{1}{2}}\Big). \tag{3.144}$$

We omit the proof because it is essentially the same as the proof of Theorem 2.4.1.

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

Theorem 3.4.3. *Let $\{\Gamma_t\}_{t\ge 0}$ denote the zero level set of the Hele-Shaw problem and $\big(U_{\epsilon,h,k}(x,t), W_{\epsilon,h,k}(x,t)\big)$ denote the piecewise linear interpolation in time of the finite element interpolation $\{(\widehat{U}^m, \widehat{W}^m)\}$ of the DG solution $\{(U^m, W^m)\}$, namely,*

$$U_{\epsilon,h,k}(x,t) := \frac{t-t_{m-1}}{k}\widehat{U}^m(x) + \frac{t_m - t}{k}\widehat{U}^{m-1}(x), \tag{3.145}$$

$$W_{\epsilon,h,k}(x,t) := \frac{t-t_{m-1}}{k}W^m(x) + \frac{t_m - t}{k}W^{m-1}(x), \tag{3.146}$$

for $t_{m-1} \le t \le t_m$ and $1 \le m \le M$. Then, under the mesh and starting value constraints of Theorem 3.3.6 and $k = O(h^{2-\gamma})$ with $\gamma > 0$, we have

(i) *$U_{\epsilon,h,k}(x,t) \overset{\epsilon\searrow 0}{\longrightarrow} 1$ uniformly on compact subset of $\mathcal{O}$,*

(ii) *$U_{\epsilon,h,k}(x,t) \overset{\epsilon\searrow 0}{\longrightarrow} -1$ uniformly on compact subset of $\mathcal{I}$.*

(iii) *Moreover, in the case that dimension $d = 2$, when $k = O(h^3)$, suppose that W^0 satisfies $\|w_0^\epsilon - W^0\|_{L^2} \le Ch^\beta$ for some $\beta > \frac{3}{2}$, then we have $W_{\epsilon,h,k}(x,t) \overset{\epsilon\searrow 0}{\longrightarrow} -w(x,t)$ uniformly on $\overline{\mathcal{D}}_T$.*

Proof. For any compact set $A \subset \mathcal{O}$ and for any $(x,t) \in A$, we have

$$\begin{aligned} |U_{\epsilon,h,k} - 1| &\leq |U_{\epsilon,h,k} - u^{\epsilon}(x,t)| + |u^{\epsilon}(x,t) - 1| \\ &\leq |U_{\epsilon,h,k} - u^{\epsilon}(x,t)|_{L^{\infty}(\mathcal{D}_T)} + |u^{\epsilon}(x,t) - 1|. \end{aligned} \tag{3.147}$$

Equation (3.99) of Theorem 3.3.6 infers that there exists a constant $0 < \alpha < \frac{4-d}{2}$ such that

$$|U_{\epsilon,h,k} - u^{\epsilon}(x,t)|_{L^{\infty}(\mathcal{D}_T)} \leq Ch^{\alpha}. \tag{3.148}$$

The first term on the right-hand side of (3.147) tends to 0 when $\epsilon \searrow 0$ (note that $h, k \searrow 0$, too). The second term converges uniformly to 0 on the compact set A, which is ensured by (i) of Theorem 3.4.1. Hence, the assertion (i) holds.

To show (ii), we only need to replace $\mathcal{O}$ by $\mathcal{I}$ and 1 by -1 in the above proof. To prove (iii), under the assumptions $k = O(h^3)$, (3.103) in Theorem 3.3.6 implies that there exists a positive constant $0 < \zeta < \frac{4-d}{2}$ such that

$$\|W_{\epsilon,h,k} - w^{\epsilon}\|_{L^{\infty}(\mathcal{D}_T)} \leq Ch^{\zeta}. \tag{3.149}$$

Then by the triangle inequality we obtain for any $(x,t) \in \overline{\mathcal{D}}_T$,

$$\begin{aligned} |W_{\epsilon,h,k}(x,t) - (-w)| &\leq |W_{\epsilon,h,k}(x,t) - w^{\epsilon}(x,t)| + |w^{\epsilon}(x,t) - (-w)|, \\ &\leq \|W_{\epsilon,h,k}(x,t) - w^{\epsilon}(x,t)\|_{L^{\infty}(\mathcal{D}_T)} + |w^{\epsilon}(x,t) - (-w)|. \end{aligned} \tag{3.150}$$

The first term on the right-hand side of (3.150) tends to 0 when $\epsilon \searrow 0$ (note that $h, k \searrow 0$, too). The second term converges uniformly to 0 in $\overline{\mathcal{D}}_T$, which is ensured by (ii) of Theorem 3.4.1. Thus the assertion (iii) is proved. The proof is complete. □

The second main theorem of this section which is given below addresses the convergence of numerical interfaces.

Theorem 3.4.4. *Let* $\Gamma_t^{\epsilon,h,k} := \{x \in \mathcal{D};\, U_{\epsilon,h,k}(x,t) = 0\}$ *be the zero level set of* $U_{\epsilon,h,k}(x,t)$*, then under the assumptions of Theorem 3.4.3, we have*

$$\sup_{x\in\Gamma_t^{\epsilon,h,k}} \operatorname{dist}(x,\Gamma_t) \xrightarrow{\epsilon\searrow 0} 0 \quad \textit{uniformly on } [0,T].$$

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

$$\mathcal{N}_\eta := \{(x,t) \in \mathcal{D}_T;\, \operatorname{dist}(x,\Gamma_t) < \eta\}. \tag{3.151}$$

Let A and B denote the complements of the neighborhood $\mathcal{N}_\eta$ in $\mathcal{O}$ and $\mathcal{I}$, respectively, i.e.

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

Note that A is a compact subset outside Γ_t and B is a compact subset inside Γ_t, then there exists $\epsilon_3 > 0$, which only depends on η, such that for any $\epsilon \in (0,\epsilon_3)$

$$|U_{\epsilon,h,k}(x,t) - 1| \le \eta \quad \forall (x,t) \in A, \tag{3.152}$$

$$|U_{\epsilon,h,k}(x,t) + 1| \le \eta \quad \forall (x,t) \in B. \tag{3.153}$$

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

$$|U_{\epsilon,h,k}(x,t) - 1| = 1 \qquad \forall (x,t) \in A, \tag{3.154}$$

$$|U_{\epsilon,h,k}(x,t) + 1| = 1 \qquad \forall (x,t) \in B. \tag{3.155}$$

(3.152) and (3.154) imply that (x,t) is not in A, and (3.153) and (3.155) imply that (x,t) is not in B, then (x,t) must lie in the tubular neighborhood $\mathcal{N}_\eta$. Therefore, for any $\epsilon \in (0,\epsilon_3)$,

$$\sup_{x\in\Gamma_t^{\epsilon,h,k}} \operatorname{dist}(x,\Gamma_t) \le \eta \qquad \text{uniformly on } [0,T]. \tag{3.156}$$

The proof is complete. □

3.5 Numerical experiments

In this section, we present three two-dimensional numerical tests to gauge the performance of the proposed fully discrete MIP-DG methods using the linear element (i.e., $r = 1$). The square domain $\mathcal{D} = [-1, 1]^2$ is used in all three tests and the initial condition is chosen to have the form $u_0 = \tanh\big(\frac{d_0(x)}{\sqrt{2}\epsilon}\big)$, where $d_0(x)$ denotes the signed distance from x to the initial interface Γ_0.

Our first test uses a smooth initial condition to satisfy the requirement for u_0, consequently, the theoretical results established in this chapter apply to this test problem. On the other hand, non-smooth initial conditions are used in the second and third tests, hence, the theoretical results of this chapter may not apply. But we still use our MIP-DG methods to compute the error order, energy decay and the evolution of the numerical interfaces. Our numerical results suggest that the proposed DG schemes work well, even a convergence theory is missing for them.

Test 1. Consider the Cahn-Hilliard problem (1.30)-(1.33) with the following initial condition:

$$u_0(x) = \tanh\Big(\frac{d_0(x)}{\sqrt{2}\epsilon}\Big),$$

where $\tanh(t) = (e^t - e^{-t})/(e^t + e^{-t})$, and $d_0(x)$ represents the signed distance function to the ellipse:

$$\frac{x_1^2}{0.36} + \frac{x_2^2}{0.04} = 1.$$

Hence, u_0 has the desired form as stated in Proposition 3.3.5.

Table 3.1 shows the spatial L^2 and H^1-norm errors and convergence rates, which are consistent with what are proved for the linear element in the convergence theorem. $\epsilon = 0.1$ is used to generate the table.

Figure 3.1 plots the change of the discrete energy $E_h(U^\ell)$ in time, which should decrease according to (3.48). This graph clearly confirms this decay property. Figure 3.2 displays four snapshots at four fixed time points of the numerical interface with four different ϵ. They clearly indicate that at each time point the numerical interface

Table 3.1: Spatial errors and convergence rates of Test 1 with $\epsilon = 0.1$.

	$L^\infty(L^2)$ error	$L^\infty(L^2)$ order	$L^2(H^1)$ error	$L^2(H^1)$ order
$h = 0.4\sqrt{2}$	0.53325		0.84260	
$h = 0.2\sqrt{2}$	0.21280	1.3253	0.64843	0.3779
$h = 0.1\sqrt{2}$	0.07164	1.5707	0.43273	0.5835
$h = 0.05\sqrt{2}$	0.01779	2.0097	0.21411	1.0151
$h = 0.025\sqrt{2}$	0.00454	1.9703	0.10890	0.9753

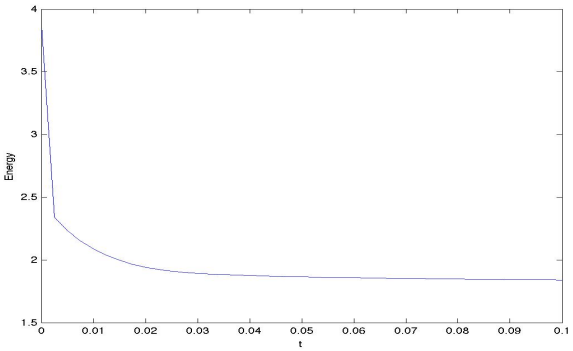

Figure 3.1: Decay of the numerical energy $E_h(U^\ell)$ of Test 1.

converges to the sharp interface Γ_t of the Hele-Shaw flow as ϵ tends to zero. It also shows that the numerical interface evolves faster in time for larger ϵ and confirms the mass conservation property of the Cahn-Hilliard problem as the total mass does not change in time, which approximates a constant 3.064.

Test 2. Consider the Cahn-Hilliard problem (1.30)-(1.33) with the following initial condition:

$$u_0(x) = \tanh\Big(\frac{1}{\sqrt{2}\epsilon}\big(\min\big\{\sqrt{(x_1+0.3)^2+x_2^2}-0.3, \sqrt{(x_1-0.3)^2+x_2^2}-0.25\big\}\big)\Big).$$

We note that u_0 can be written as

$$u_0(x) = \tanh\Big(\frac{d_0(x)}{\sqrt{2}\epsilon}\Big).$$

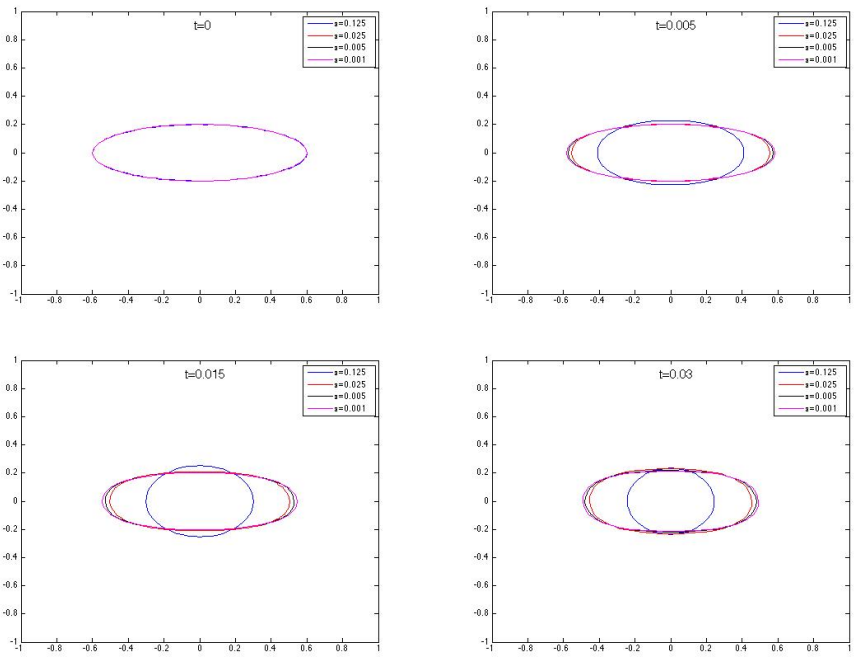

Figure 3.2: Test 1: Snapshots of the zero-level set of $u^{\epsilon,h,k}$ at time $t = 0, 0.005, 0.015, 0.03$ and $\epsilon = 0.125, 0.025, 0.005, 0.001$.

Here $d_0(x)$ represents the signed distance function. We note that u_0 does not have the desired form as stated in Proposition 3.3.5.

Table 3.2 shows the spatial L^2 and H^1-norm errors and convergence rates, which are consistent with what are proved for the linear element in the convergence theorem. $\epsilon = 0.1$ is used to generate the table. Figure 3.3 plots the change of the discrete energy

Table 3.2: Spatial errors and convergence rates of Test 2 with $\epsilon = 0.1$.

	$L^\infty(L^2)$ error	$L^\infty(L^2)$ order	$L^2(H^1)$ error	$L^2(H^1)$ order
$h = 0.4\sqrt{2}$	0.26713		0.35714	
$h = 0.2\sqrt{2}$	0.07161	1.8993	0.18411	0.9559
$h = 0.1\sqrt{2}$	0.01833	1.9660	0.09620	0.9365
$h = 0.05\sqrt{2}$	0.00476	1.9452	0.04928	0.9650
$h = 0.025\sqrt{2}$	0.00121	1.9760	0.02497	0.9808

$E_h(U^\ell)$ in time, which should decrease according to (3.48). This graph clearly confirms this decay property. Figure 3.4 displays four snapshots at four fixed time points of

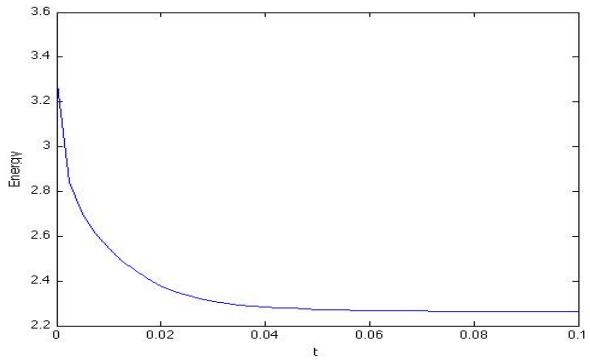

Figure 3.3: Decay of the numerical energy $E_h(U^\ell)$ of Test 2.

the numerical interface with four different ϵ. They clearly indicate that at each time point the numerical interface converges to the sharp interface Γ_t of the Hele-Shaw flow as ϵ tends to zero. It again shows that the numerical interface evolves faster in time for larger ϵ and confirms the mass conservation property of the Cahn-Hilliard problem as the total mass does not change in time, which approximates a constant 3.032.

Test 3. Consider the Cahn-Hilliard problem (1.30)–(1.33) with the following initial condition:

$$u_0(x) = \tanh\Big(\frac{1}{\sqrt{2}\epsilon}\big(\min\big\{\sqrt{(x_1+0.3)^2+x_2^2}-0.2, \sqrt{(x_1-0.3)^2+x_2^2}-0.2, \\ \sqrt{x_1^2+(x_2+0.3)^2}-0.2, \sqrt{x_1^2+(x_2-0.3)^2}-0.2\big\}\big)\Big).$$

Notice that the above u_0 does not have the desired form as stated in Proposition 3.3.5.

Table 3.3 shows the spatial L^2 and H^1-norm errors and convergence rates with $\epsilon = 0.1$, which are consistent with what are proved for the linear element in the convergence theorem. Figure 3.5 plots the change of the discrete energy $E_h(U^\ell)$ in time, which again decreases as predicted by (3.48). Figure 3.6 displays four snapshots at four fixed time points of the numerical interface with four different ϵ. Once again, we observe that at each time point the numerical interface converges to the sharp

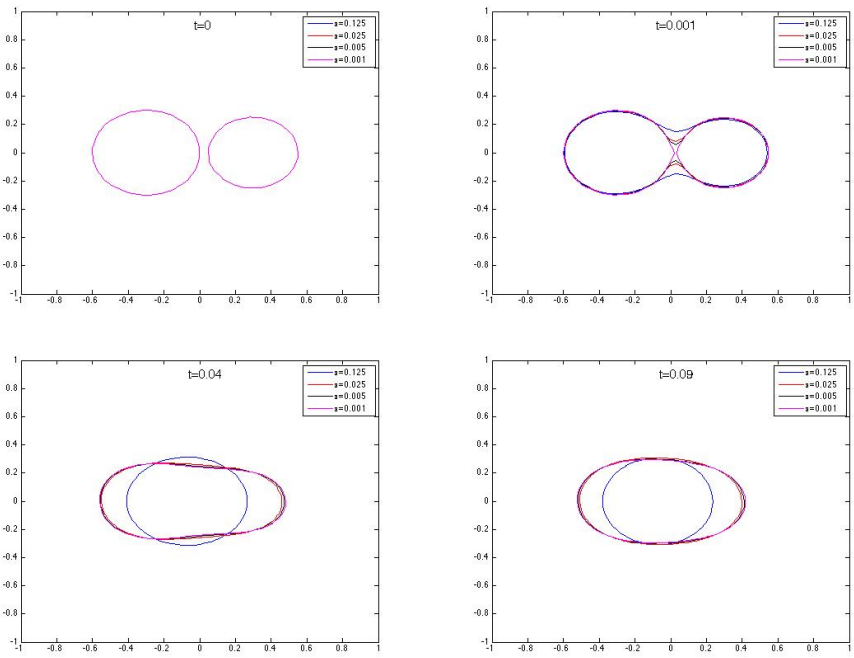

Figure 3.4: Test 2: Snapshots of the zero-level set of $u^{\epsilon,h,k}$ at time $t = 0, 0.001, 0.04, 0.09$ and $\epsilon = 0.125, 0.025, 0.005, 0.001$.

interface Γ_t of the Hele-Shaw flow as ϵ tends to zero, the interface evolves faster in time for larger ϵ and the mass conservation property is preserved. The total mass approximates a constant 2.989.

Table 3.3: Spatial errors and convergence rates of Test 3 with $\epsilon = 0.1$.

	$L^\infty(L^2)$ error	$L^\infty(L^2)$ order	$L^2(H^1)$ error	$L^2(H^1)$ order
$h = 0.4\sqrt{2}$	0.38576		0.84157	
$h = 0.2\sqrt{2}$	0.12347	1.6435	0.55082	0.6115
$h = 0.1\sqrt{2}$	0.03599	1.7785	0.31149	0.8224
$h = 0.05\sqrt{2}$	0.00965	1.8990	0.16199	0.9433
$h = 0.025\sqrt{2}$	0.00247	1.9660	0.08218	0.9790

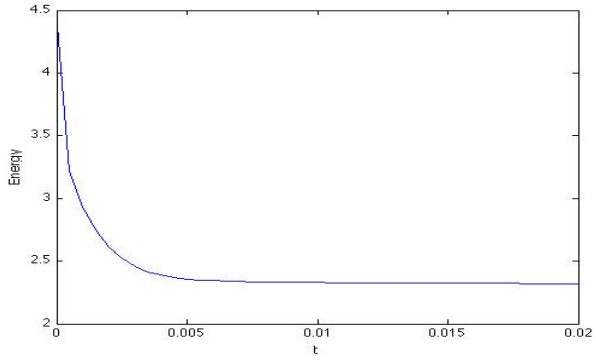

Figure 3.5: Decay of the numerical energy $E_h(U^\ell)$ of Test 3.

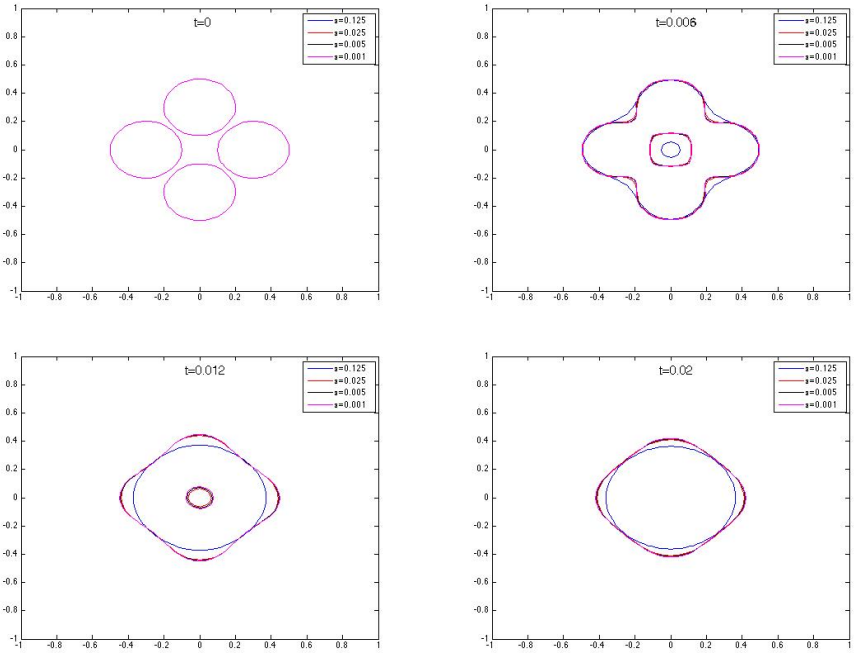

Figure 3.6: Test 3: Snapshots of the zero-level set of $u^{\epsilon,h,k}$ at time $t = 0, 0.006, 0.012, 0.02$ and $\epsilon = 0.125, 0.025, 0.005, 0.001$.

Chapter 4

Finite Element Methods for the Stochastic Mean Curvature Flow of Planer Curves of Graphs

4.1 Introduction

It is easy to check that (cf. [94, 32]) the level set formulation of (1.35) is given by the following nonlinear parabolic stochastic partial differential equation (SPDE):

$$df = |\nabla_{x'} f| \, \mathrm{div}_{x'} \Big(\frac{\nabla_{x'} f}{|\nabla_{x'} f|} \Big) \, dt + \epsilon |\nabla_{x'} f| \circ dW_t, \tag{4.1}$$

where $f = f(x', t)$ with $x' = (x, x_{d+1})$ denotes the level set function so that Γ_t is represented by the zero level set of f, and '$\circ$' refers to the Stratonovich interpretation of the stochastic integral. Again, stochastic effects are modeled by a standard $\mathbb{R}$-valued Wiener process $W \equiv \{W_t; \, t \geq 0\}$ which is defined on a given filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t; \, t \geq 0\}, \mathbb{P})$.

In the case that f is a d-dimensional graph, that is, $f(x',t) = x_{d+1} - u(x,t)$, equation (4.1) reduces to

$$du = \sqrt{1+|\nabla_x u|^2}\,\mathrm{div}_x\Big(\frac{\nabla_x u}{\sqrt{1+|\nabla_x u|^2}}\Big)\,dt + \epsilon\sqrt{1+|\nabla_x u|^2}\circ dW_t. \tag{4.2}$$

To the best of our knowledge, a comprehensive PDE theory for the SPDE (4.2) is still missing in the literature. For the case $d = 1$, (4.2) reduces to the following one-dimensional nonlinear parabolic SPDE:

$$\begin{aligned} du &= \frac{\partial_x^2 u}{1+|\partial_x u|^2}dt + \epsilon\sqrt{1+|\partial_x u|^2}\circ dW_t \\ &= \partial_x\big(\arctan(\partial_x u)\big)dt + \epsilon\sqrt{1+|\partial_x u|^2}\circ dW_t. \end{aligned} \tag{4.3}$$

Here $\partial_x u$ stands for the derivative of u with respect to x. This Stratonovich SPDE can be equivalently converted into the following Itô SPDE:

$$\begin{aligned} du &= \Big[\frac{\epsilon^2}{2}\partial_x^2 u + \big(1-\frac{\epsilon^2}{2}\big)\frac{\partial_x^2 u}{1+|\partial_x u|^2}\Big]dt + \epsilon\sqrt{1+|\partial_x u|^2}\,dW_t \\ &= \partial_x\Big(\frac{\varepsilon^2}{2}\partial_x u + (1-\frac{\varepsilon^2}{2})\arctan(\partial_x u)\Big)dt + \epsilon\sqrt{1+|\partial_x u|^2}\,dW_t. \end{aligned} \tag{4.4}$$

As is evident from (4.3), (4.4), the stochastic mean curvature flow (4.2) for $d = 1$ may be interpreted as a gradient flow with multiplicative noise. Recently, Es-Sarhir and von Renesse [32] proved existence and uniqueness of (stochastically) strong solutions for (4.3) by a variational method, based on the Lyapunov structure of the problem (cf. [32, property (H3)]) which replaces the standard coercivity assumption (cf. [32, property (A)]). As is pointed out in [32], mild solutions for (4.3) may not be expected due to its quasilinear character.

The primary goal of this chapter is to develop and analyze by a variational method some semi-discrete and fully discrete finite element methods for approximating (with rates) the strong solution of the Itô form (4.4) of the stochastic MCF. The error analysis presented in this chapter differs from most existing works on the numerical

analysis of SPDEs, where mild solutions are mostly approximated with the help of corresponding discrete semi-groups (see [61] and the references therein). We also note that the error estimates derived in [55] which hold for general quasilinear SPDEs do not apply to (4.4) because the structural assumptions, such as the coercivity assumption [55, cf. Assumption 2.1, (ii)] and the strong monotonicity assumption [55, cf. Assumption 2.2, (i)] fail to hold for (4.4), and also the regularity assumptions [55, cf. Assumption 2.3] are not known to hold in the present case. In this chapter, we use a variational approach similar to [55, 15, 17] to analyze the convergence of our finite element methods. One main difficulty for approximating the strong solution of (4.4) with certain rates is caused by the low regularity of the solution. To circumvent this difficulty, we first regularize the SPDE (4.4) by adding an additional linear diffusion term $\delta\partial_x^2 u$ to the drift coefficient of (4.4); as a consequence the related drift operator in (4.7) becomes strongly monotone, and the corresponding solution process u^δ is then H^2-valued in space. However, it is due to the 'gradient-type' noise that a relevant Hölder estimate in the H^1-norm for the solution u^δ seems not available, which is necessary to properly control time-discretization errors. In order to circumvent this problematic issue, we proceed first with the spatial discretization (4.12); we may then use an inverse finite element estimate, and the weaker Hölder estimate (4.27) for the process u_h^δ to control time-discretization errors. We remark that addressing space discretization errors first requires to efficiently cope with the limited regularity of Lagrange finite element functions in the context of required higher norm estimates, which is overcome by a perturbation argument (cf. Proposition 4.3.4).

The remainder of this chapter consists of three additional sections. In section 4.2 we first recall some relevant facts about the solution of (4.4) from [32]; we then present an analysis for the regularized problem. The main result of this section is to prove an error bound for $u^\delta - u$ in powers of δ. In section 4.3 we propose a semi-discrete (in space) and a fully discrete finite element method for the regularized equation (4.7) of the SPDE (4.4). The main result of this section is the strong L^2-error estimate for the finite element solution. Finally, in section 4.4 we present several computational

results to validate the theoretical error estimate, and to study relative effects due to geometric evolution and gradient-type noises.

4.2 Preliminaries and error estimates for a partial differential equation regularization

The standard function and space notation will be adapted in this chapter. For example, $H^2(I)$ denotes the Sobolev space $W^{2,2}(I)$ on the interval $I = (0,1)$, and $H^0(I) = L^2(I)$. We also use $H^m_p(I)$ to denote the subspace of $H^m(I)$ which consists of all periodic functions in $H^m(I)$. Let $(\cdot,\cdot)_I$ denote the L^2-inner product on I. The quadruple $(\Omega, \mathcal{F}, \{\mathcal{F}_t; t \geq 0\}, \mathbb{P})$ stands for a given probability space, on which an $\mathbb{R}$-valued Wiener process W is given. For a random variable X, we denote by $\mathbb{E}[X]$ the expected value of X.

We first quote the following existence and uniqueness result from [32] for the SPDE (4.4) with periodic boundary conditions. In this context, we refer to the $\{\mathcal{F}_t\}$-adapted process $u : I \times [0,T] \times \Omega \to \mathbb{R}$ as a (stochastically) strong solution in case it satisfies $\mathbb{P}$-a.s. (4.4) in an analytically weak sense, i.e., tested with deterministic functions.

Theorem 4.2.1. *Suppose that $u_0 \in H^1_p(I)$ and fix $T > 0$. Let $\epsilon \leq \sqrt{2}$. There exists a unique strong solution to SPDE (4.3) with periodic boundary conditions and attaining the initial condition $u(0) = u_0$, that is, there exists a unique H^1_p-valued $\{\mathcal{F}_t\}_{t\in[0,T]}$-adapted process $u \equiv \{u(t);\, t \in [0,T]\}$ such that $\mathbb{P}$-almost surely*

$$\begin{aligned}\big(u(t),\varphi\big)_I \;=\; & \big(u_0,\varphi\big)_I - \frac{\epsilon^2}{2}\int_0^t \big(\partial_x u, \partial_x\varphi\big)_I\, ds \\ & -\big(1-\frac{\epsilon^2}{2}\big)\int_0^t \Big(\big(\arctan(\partial_x u), \partial_x\varphi\big)_I\Big] ds \\ & +\epsilon\int_0^t \Big(\sqrt{1+|\partial_x u|^2},\varphi\Big)_I dW_s \qquad \forall\varphi\in H^1(I) \quad \forall t\in[0,T].\end{aligned} \tag{4.5}$$

Moreover, u satisfies for some $C > 0$ independent of $T > 0$,

$$\sup_{t\in[0,T]} \mathbb{E}\big[\|u(t)\|^2_{H^1(I)}\big] \le C. \tag{4.6}$$

It is not clear if such a regularity can be improved from the analysis of [32] because of the difficulty caused by the gradient-type noise. In particular, H^2-regularity in space, which would be desirable in order to derive some rates of convergence for finite element methods, seems not clear. To overcome this difficulty, we introduce the following simple regularization of (4.4):

$$du^\delta = \Big[\big(\delta + \frac{\epsilon^2}{2}\big)\partial_x^2 u^\delta + \big(1 - \frac{\epsilon^2}{2}\big)\frac{\partial_x^2 u^\delta}{1 + |\partial_x u^\delta|^2}\Big]dt + \epsilon\sqrt{1 + |\partial_x u^\delta|^2}\, dW_t. \tag{4.7}$$

To make this indirect approach successful, we need to address the well-posedness and regularity issues for (4.7) and to estimate the difference between the strong solutions u^δ of (4.7) and u of (4.4).

Theorem 4.2.2. *Suppose that $u_0^\delta \in H_p^1(I)$ and $\|u_0^\delta\|_{H^1(I)} \le C_0$, where $C_0 > 0$ is independent of δ. Let $\epsilon \le \sqrt{2(1+\delta)}$. Then there exists a unique strong solution to SPDE (4.7) with periodic boundary conditions and initial condition $u^\delta(0) = u_0^\delta$, that is, there exists a unique H_p^1-valued $\{\mathcal{F}_t\}_{t\in[0,T]}$-adapted process $u^\delta \equiv \{u^\delta(t);\, t \in [0,T]\}$ such that there holds $\mathbb{P}$-almost surely*

$$\begin{aligned}\big(u^\delta(t), \varphi\big)_I = \big(u_0^\delta, \varphi\big)_I &- \big(\delta + \frac{\epsilon^2}{2}\big)\int_0^t \big(\partial_x u^\delta, \partial_x \varphi\big)_I\, ds \\ &- \big(1 - \frac{\epsilon^2}{2}\big)\int_0^t \big(\arctan(\partial_x u^\delta), \partial_x \varphi\big)_I\, ds \\ &+ \epsilon\int_0^t \Big(\sqrt{1 + |\partial_x u^\delta|^2}, \varphi\Big)_I dW_s \quad \forall \varphi \in H^1(I) \quad \forall\, t \in [0,T].\end{aligned} \tag{4.8}$$

Moreover, u^δ satisfies

$$\sup_{t\in[0,T]} \mathbb{E}\Big[\frac{1}{2}\|\partial_x u^\delta(t)\|^2_{L^2(I)}\Big] + \delta\,\mathbb{E}\Big[\int_0^T \|\partial_x^2 u^\delta(s)\|^2_{L^2(I)}\, ds\Big] \le \mathbb{E}\Big[\frac{1}{2}\|\partial_x u_0^\delta\|^2_{L^2(I)}\Big]. \tag{4.9}$$

Proof. Existence of u^δ can be shown in the same way as done in Theorem 4.2.1 (cf. [32]). To verify (4.9), we proceed formally and apply Ito's formula (cf. e.g. [62], or [78]) with $f(\cdot) = \frac{1}{2}\|\partial_x \cdot\|^2_{L^2(I)}$ to (a Galerkin approximation of) the solution u^δ to get

$$\begin{aligned}
&\frac{1}{2}\|\partial_x u^\delta(t)\|^2_{L^2(I)} + \int_0^t \Big[(\frac{\varepsilon^2}{2}+\delta)\|\partial_x^2 u^\delta\|^2_{L^2(I)} + (1-\frac{\varepsilon^2}{2})\Big\|\frac{\partial_x^2 u^\delta}{\sqrt{1+|\partial_x u^\delta|^2}}\Big\|^2_{L^2}\Big]\,ds \\
&\quad = \frac{1}{2}\|\partial_x u_0^\delta\|^2_{L^2(I)} + \frac{\varepsilon^2}{2}\int_0^t \Big\|\partial_x\sqrt{1+|\partial_x u^\delta|^2}\Big\|^2_{L^2}\,ds + M_t \\
&\quad = \frac{1}{2}\|\partial_x u_0^\delta\|^2_{L^2(I)} + \frac{\varepsilon^2}{2}\int_0^t \Big\|\frac{\partial_x u^\delta\cdot\partial_x^2 u^\delta}{\sqrt{1+|\partial_x u^\delta|^2}}\Big\|^2_{L^2(I)}\,ds + M_t \qquad \forall\, t\in[0,T],
\end{aligned}$$

where

$$M_t := \epsilon\int_0^t \Big(\partial_x\sqrt{1+|\partial_x u^\delta(s)|^2}, \partial_x u^\delta\Big)_I\, dW_s$$

is a martingale. Taking expectation yields

$$\begin{aligned}
\mathbb{E}\Big[\frac{1}{2}\|\partial_x u^\delta(t)\|^2_{L^2(I)} + \int_0^t \Big[\delta\|\partial_x^2 u^\delta\|^2_{L^2} + (1-\frac{\varepsilon^2}{2})\Big\|\frac{\partial_x^2 u^\delta}{\sqrt{1+|\partial_x u^\delta|^2}}\Big\|^2_{L^2}\Big]\,ds\Big] \\
\le \mathbb{E}\big[\frac{1}{2}\|\partial_x u_0^\delta\|^2_{L^2}\big].
\end{aligned}$$

Hence, (4.9) hold. The proof is complete. □

Next, we shall derive an upper bound for the error $u^\delta - u$ as a low order power function of δ.

Theorem 4.2.3. *Suppose that $u_0^\delta \equiv u_0$. Let u and u^δ denote respectively the strong solutions of the initial-boundary value problems (4.4) and (4.7) as stated in Theorems 4.2.1 and 4.2.2. Then there holds the following error estimate:*

$$\sup_{t\in[0,T]}\mathbb{E}\Big[\|u^\delta(t)-u(t)\|^2_{L^2(I)}\Big] + \delta\,\mathbb{E}\Big[\int_0^T \|\partial_x\big(u^\delta(s)-u(s)\big)\|^2_{L^2(I)}\,ds\Big] \le CT\delta. \quad (4.10)$$

Proof. Let $e^\delta := u^\delta - u$. Subtracting (4.5) from (4.8) we get that $\mathbb{P}$-a.s.

$$\begin{aligned}\big(e^\delta(t),\varphi\big)_I = -\int_0^t \Big[\delta\big(\partial_x u, \partial_x\varphi\big)_I + \big(\delta + \frac{\epsilon^2}{2}\big)\big(\partial_x e^\delta, \partial_x\varphi\big)_I \\ + \big(1 - \frac{\epsilon^2}{2}\big)\big(\arctan(\partial_x u^\delta) - \arctan(\partial_x u), \partial_x\varphi\big)_I\Big] ds + M_t\end{aligned}$$

for all $\varphi \in H^1(I)$ and $t \in [0,T]$, with the martingale

$$M_t := \epsilon \int_0^t \big(\sqrt{1+|\partial_x u^\delta|^2} - \sqrt{1+|\partial_x u|^2}, \varphi\big)_I dW_s.$$

By Itô's formula (cf. [62]) we get

$$\begin{aligned}\|e^\delta(t)\|^2_{L^2(I)} = -2\int_0^t \Big[\delta\big(\partial_x u, \partial_x e^\delta\big)_I + \big(\delta + \frac{\epsilon^2}{2}\big)\|\partial_x e^\delta\|^2_{L^2(I)} \\ + \big(1 - \frac{\epsilon^2}{2}\big)\big(\arctan(\partial_x u^\delta) - \arctan(\partial_x u), \partial_x e^\delta\big)_I\Big] ds \\ + \epsilon^2 \int_0^t \Big\|\sqrt{1+|\partial_x u^\delta|^2} - \sqrt{1+|\partial_x u|^2}\Big\|^2_{L^2(I)} ds \\ + 2\epsilon \int_0^t \Big(\sqrt{1+|\partial_x u^\delta|^2} - \sqrt{1+|\partial_x u|^2}, e^\delta\Big)_I dW_s.\end{aligned} \tag{4.11}$$

Taking expectations on both sides, and using the monotonicity property of the arctan function and the inequality $\big(\sqrt{1+x^2} - \sqrt{1+y^2}\big)^2 \le |x-y|^2$ yield

$$\begin{aligned}\mathbb{E}\big[\|e^\delta(t)\|^2_{L^2(I)}\big] + 2\delta\,\mathbb{E}\Big[\int_0^t \|\partial_x e^\delta\|^2_{L^2(I)}\, ds\Big] &\le -2\delta\,\mathbb{E}\Big[\int_0^t \big(\partial_x u, \partial_x e^\delta\big)_I ds\Big] \\ &\le \delta\,\mathbb{E}\Big[\int_0^T \Big[\|\partial_x u\|^2_{L^2(I)} + \|\partial_x e^\delta\|^2_{L^2(I)}\Big]\, ds\Big],\end{aligned}$$

which and (4.6) imply that

$$\begin{aligned}\mathbb{E}\big[\|e^\delta(t)\|^2_{L^2(I)}\big] + \delta\,\mathbb{E}\Big[\int_0^t \|\partial_x e^\delta\|^2_{L^2(I)}\, ds\Big] &\le \delta\,\mathbb{E}\Big[\int_0^T \|\partial_x u\|^2_{L^2(I)}\, ds\Big] \\ &\le (CT)\delta.\end{aligned}$$

The desired estimate (4.10) follows immediately. The proof is complete. □

4.3 Finite element methods

In this section we propose a fully discrete finite element method to solve the regularized SPDE (4.7) and to derive an error estimate for the finite element solution. This goal will be achieved in two steps. We first present and study a semi-discrete in space finite element method and then discretize it in time to obtain our fully discrete finite element method.

4.3.1 Semi-discretization in space

Let $0 = x_0 < x_1 < \cdots < x_{J+1} = 1$ be a quasiuniform partition of $I = (0,1)$. Define $h_j := x_{j+1} - x_j$ and $h := \max_{0\le j\le J} h_j$. Introduce the finite element spaces

$$V_r^h := \left\{ v_h \in C^0(\bar{I});\, v_h|_{[x_j,x_{j+1}]} \in P_r([x_j,x_{j+1}]),\, j = 0,1,\cdots,J \right\} \cap H_p^1(I),$$

where $P_r([x_j,x_{j+1}]$ denotes the space of all polynomials of degree not exceeding $r(\ge 0)$ on $[x_j,x_{j+1}]$. We note that functions in V_r^h are piecewise continuous periodic functions. Our semi-discrete finite element method for SPDE (4.7) is defined by seeking $u_h(\cdot,t,\omega) : [0,T]\times\Omega \to V_r^h$ such that $\mathbb{P}$-almost surely

$$\begin{aligned} \left(u_h^\delta(t), v_h\right)_I = \left(u_h^\delta(0), v_h\right)_I &- \left(\delta + \frac{\epsilon^2}{2}\right) \int_0^t \left(\partial_x u_h^\delta, \partial_x v_h\right)_I ds \\ &- \left(1 - \frac{\epsilon^2}{2}\right) \int_0^t \left(\arctan(\partial_x u_h^\delta), \partial_x v_h\right)_I ds \\ &+ \epsilon \int_0^t \left(\sqrt{1+|\partial_x u_h^\delta|^2}, v_h\right)_I dW_s \qquad \forall v_h \in V_r^h \quad \forall t \in [0,T], \end{aligned} \tag{4.12}$$

where $u_h^\delta(0) = P_h^r u_0^\delta$, and P_h^r denotes the L^2-projection operator from $L^2(I)$ to V_r^h which is defined by

$$\left(P_h^r w, v_h\right)_I = \left(w, v_h\right)_I \qquad \forall v_h \in V_r^h.$$

To derive an SDE for u_h^δ from the above weak formulation, we introduce the discrete (nonlinear) operator $A_h^\delta : V_r^h \to V_r^h$ by

$$\begin{aligned}\big(A_h^\delta w_h, v_h\big)_I &:= \big(\delta + \frac{\epsilon^2}{2}\big)\big(\partial_x w_h, \partial_x v_h\big)_I \\ &\quad + \big(1 - \frac{\epsilon^2}{2}\big)\big(\arctan(\partial_x w_h), \partial_x v_h\big)_I \quad \forall w_h, v_h \in V_r^h.\end{aligned} \tag{4.13}$$

Then (4.12) can be equivalently written as

$$du_h^\delta(t) = -A_h^\delta u_h(t)\, dt + \epsilon P_h\Big(\sqrt{1 + |\partial_x u_h^\delta(t)|^2}\Big)\, dW_t. \tag{4.14}$$

Proposition 4.3.1. *For $\epsilon \le \sqrt{2(1+\delta)}$, there is a unique solution $u_h^\delta \in C\big([0,T]; L^2(\Omega; V_r^h)\big)$ to scheme (4.12). Moreover, there holds*

$$\begin{aligned}\sup_{0\le t\le T} \mathbb{E}\Big[\frac{1}{2}\big\|u_h^\delta(t)\big\|_{L^2(I)}^2\Big] + \delta\,\mathbb{E}\Big[\int_0^T \big\|\partial_x u_h^\delta(s)\big\|_{L^2(I)}^2\, ds\Big] \\ \le \mathbb{E}\Big[\frac{1}{2}\big\|u_h^\delta(0)\big\|_{L^2(I)}^2\Big] + \epsilon^2 T.\end{aligned} \tag{4.15}$$

Proof. Well-posedness of (4.14) follows from the standard theory for stochastic ODEs with Lipschitz drift and diffusion. To verify (4.15), applying Itô's formula (cf. [62]) to $f(u_h^\delta) = \|u_h^\delta\|_{L^2(I)}^2$ and using (4.14) we get

$$\begin{aligned}\|u_h^\delta(t)\|_{L^2(I)}^2 &= \|u_h^\delta(0)\|_{L^2(I)}^2 - 2\int_0^t \Big(A_h^\delta u_h^\delta(s), u_h^\delta(s)\Big)_I\, ds \\ &\quad + \epsilon^2 \int_0^t \Big\|P_h^r \sqrt{1 + |\partial_x u_h^\delta(s)|^2}\Big\|_{L^2(I)}^2\, ds \\ &\quad + 2\epsilon \int_0^t \Big(P_h^r \sqrt{1 + |\partial_x u_h^\delta(s)|^2}, u_h^\delta\Big)_I\, dW_s.\end{aligned} \tag{4.16}$$

It follows from the definitions of A_h^δ and P_h^r that

$$\begin{aligned}\|u_h^\delta(t)\|_{L^2(I)}^2 \le \|u_h^\delta(0)\|_{L^2(I)}^2 &- (2\delta+\epsilon^2)\int_0^t \big\|\partial_x u_h^\delta(s)\big\|_{L^2(I)}^2\,ds \\ &- (2-\epsilon^2)\int_0^t \Big(\arctan(\partial_x u_h^\delta(s)), \partial_x u_h^\delta(s)\Big)_I\,ds \\ &+ \epsilon^2\int_0^t \Big[1+\big\|\partial_x u_h^\delta(s)\big\|_{L^2(I)}^2\Big]\,ds \\ &+ 2\epsilon\int_0^t \Big(\sqrt{1+|\partial_x u_h^\delta(s)|^2}, u_h^\delta\Big)_I\,dW_s.\end{aligned} \tag{4.17}$$

Then (4.15) follows from applying expectation to (4.17), and using the coercivity of arctan. The proof is complete. □

An a priori estimate for u_h^δ in stronger norms is more difficult to obtain, which is due to low global smoothness and local nature of finite element functions. We shall derive some of these estimates in Proposition 4.3.4 using a perturbation argument after establishing error estimates for u_h^δ.

To derive error estimates for u_h^δ, we introduce the elliptic H^1-projection $R_h^r : H^1(I) \to V_r^h$, i.e., for any $w \in H^1(I)$, $R_h^r w \in V_r^h$ is defined by

$$\big(\partial_x[R_h^r w - w], \partial_x v_h\big)_I + \big(R_h^r w - w, v_h\big)_I = 0 \qquad \forall v_h \in V_r^h. \tag{4.18}$$

The following error bounds are well-known (cf. [13]),

$$\big\|w - R_h^r w\big\|_{L^2(I)} + h\big\|w - R_h^r w\big\|_{H^1(I)} \le Ch^2\|w\|_{H^2(I)} \qquad \forall w \in H_p^2(I). \tag{4.19}$$

Theorem 4.3.2. *Let $\epsilon \le \sqrt{2(1+\delta)}$. Then there holds*

$$\begin{aligned}\sup_{t\in[0,T]} \mathbb{E}\Big[\big\|u^\delta(t) - u_h^\delta(t)\big\|_{L^2(I)}^2\Big] + \delta\,\mathbb{E}\Big[\int_0^T \big\|\partial_x[u^\delta(s) - u_h^\delta(s)]\big\|_{L^2(I)}^2\,ds\Big] \\ \le Ch^2\big(1+\delta^{-2}\big).\end{aligned} \tag{4.20}$$

Proof. Let

$$e^\delta(t) := u^\delta(t) - u_h^\delta(t), \quad \eta^\delta := u^\delta(t) - R_h^r u^\delta(t), \quad \xi^\delta := R_h^r u^\delta(t) - u_h^\delta(t).$$

Then $e^\delta = \eta^\delta + \xi^\delta$. Subtracting (4.12) from (4.8) we obtain the following error equation which holds $\mathbb{P}$-almost surely:

$$\begin{aligned} &\big(e^\delta(t), v_h\big)_I + \big(\delta + \frac{\epsilon^2}{2}\big) \int_0^t \big(\partial_x e^\delta(s), \partial_x v_h\big)_I ds \\ &= -\big(1 - \frac{\epsilon^2}{2}\big) \int_0^t \Big(\arctan(\partial_x u^\delta(s)) - \arctan(\partial_x u_h^\delta(s)), \partial_x v_h\Big)_I ds \\ &\quad + \epsilon \int_0^t \Big(\sqrt{1 + |\partial_x u^\delta(s)|^2} - \sqrt{1 + |\partial_x u_h^\delta(s)|^2}, v_h\Big)_I dW_s + \big(e^\delta(0), v_h\big)_I \end{aligned} \tag{4.21}$$

for all $v_h \in V_r^h$. Substituting $e^\delta = \eta^\delta + \xi^\delta$ and rearranging terms leads to

$$\begin{aligned} &\big(\xi^\delta(t), v_h\big)_I + \big(\delta + \frac{\epsilon^2}{2}\big) \int_0^t \big(\partial_x \xi^\delta(s), \partial_x v_h\big)_I ds \\ &\quad + \big(1 - \frac{\epsilon^2}{2}\big) \int_0^t \Big(\arctan\big(\partial_x u^\delta(s)\big) - \arctan\big(\partial_x u_h^\delta(s)\big), \partial_x v_h\Big)_I ds \\ &= \epsilon \int_0^t \Big(\sqrt{1 + |\partial_x u^\delta(s)|^2} - \sqrt{1 + |\partial_x u_h^\delta(s)|^2}, v_h\Big)_I dW_s \\ &\quad - \big(\delta + \frac{\epsilon^2}{2}\big) \int_0^t \big(\eta^\delta(s), v_h\big)_I ds - \big(\eta^\delta(t), v_h\big)_I + \big(e^\delta(0), v_h\big)_I. \end{aligned} \tag{4.22}$$

Applying Itô's formula (cf. [62]) with $f(\xi^\delta) = \|\xi^\delta\|^2_{L^2(I)}$, and using (4.22) and (4.13) we obtain

$$
\begin{aligned}
&\big\|\xi^\delta(t)\big\|^2_{L^2(I)} + \big(2\delta+\epsilon^2\big)\int_0^t \big\|\partial_x\xi^\delta(s)\big\|^2_{L^2(I)}\,ds \\
&\quad + (2-\epsilon^2)\int_0^t \Big(\arctan\big(\partial_x R_h^r u^\delta(s)\big) - \arctan\big(\partial_x u_h^\delta(s)\big), \partial_x\xi^\delta(s)\Big)_I ds \\
&= -\big(2-\epsilon^2\big)\int_0^t \Big(\arctan\big(\partial_x u^\delta(s)\big) - \arctan\big(\partial_x R_h^r u^\delta(s)\big), \partial_x\xi^\delta(s)\Big)_I ds \\
&\quad + \epsilon^2\int_0^t \Big\|\sqrt{1+|\partial_x u^\delta(s)|^2} - \sqrt{1+|\partial_x u_h^\delta(s)|^2}\Big\|^2_{L^2(I)} ds \\
&\quad + 2\epsilon\int_0^t \Big(\sqrt{1+|\partial_x u^\delta(s)|^2} - \sqrt{1+|\partial_x u_h^\delta(s)|^2}, \xi^\delta(s)\Big)_I dW_s \\
&\quad - \big(2\delta+\epsilon^2\big)\int_0^t \big(\partial_x\eta^\delta(s), \partial_x\xi_h(s)\big)_I\,ds - 2\big(\eta^\delta(t), \xi^\delta(t)\big)_I \\
&\quad + 2\big(\eta^\delta(0), \xi^\delta(t)\big)_I + \big(\xi^\delta(0), \xi^\delta(0)\big)_I.
\end{aligned}
\tag{4.23}
$$

By the monotonicity of arctan, (4.19), (4.9), and the inequality $\big(\sqrt{1+x^2} - \sqrt{1+y^2}\big)^2 \le |x-y|^2$ we have

$$
\begin{aligned}
&\mathbb{E}\Big[\int_0^t \Big(\arctan(\partial_x R_h^r u^\delta(s)) - \arctan(\partial_x u_h^\delta(s)), \partial_x\xi^\delta(s)\Big)_I ds\Big] \ge 0, \\
&(2-\epsilon^2)\mathbb{E}\Big[\int_0^t \Big(\arctan\big(\partial_x u^\delta(s)\big) - \arctan\big(\partial_x R_h^r u^\delta(s)\big), \partial_x\xi^\delta(s)\Big)_I ds\Big] \\
&\qquad \le \mathbb{E}\Big[\int_0^t \Big(\frac{\delta}{4}\|\partial_x\xi^\delta(s)\|^2_{L^2(I)} + 4\delta^{-1}\big\|\partial_x u^\delta(s) - \partial_x R_h^r u^\delta(s)\big\|^2_{L^2(I)}\Big) ds\Big] \\
&\qquad \le \frac{\delta}{4}\mathbb{E}\Big[\int_0^t \|\partial_x\xi^\delta(s)\|^2_{L^2(I)}\,ds\Big] + Ch^2\delta^{-2},
\end{aligned}
$$

$$
\begin{aligned}
&\mathbb{E}\Big[\epsilon^2 \int_0^t \Big\| \sqrt{1+|\partial_x u^\delta(s)|^2} - \sqrt{1+|\partial_x u_h^\delta(s)|^2} \Big\|_{L^2(I)}^2 ds\Big] \\
&\qquad \le \mathbb{E}\Big[\big(\epsilon^2 + \frac{\delta}{4}\big) \int_0^t \|\partial_x \xi^\delta(s)\|_{L^2(I)}^2 \, ds\Big] + C\delta^{-1}\mathbb{E}\Big[\int_0^t \|\partial_x \eta^\delta\|_{L^2(I)}^2 \, ds\Big] \\
&\qquad \le \big(\epsilon^2 + \frac{\delta}{4}\big)\mathbb{E}\Big[\int_0^t \|\partial_x \xi^\delta(s)\|_{L^2(I)}^2 \, ds\Big] + Ch^2\delta^{-2}, \\
&\mathbb{E}\Big[\big(\eta^\delta(t), \xi^\delta(t)\big)_I\Big] \le \mathbb{E}\Big[\frac{1}{4}\|\xi^\delta(t)\|_{L^2(I)}^2 + \|\eta^\delta(t)\|_{L^2(I)}^2\Big] \le \frac{1}{4}\mathbb{E}\Big[\|\xi^\delta(t)\|_{L^2(I)}^2\Big] + Ch^2, \\
&\mathbb{E}\Big[\big(e^\delta(0), \xi^\delta(t)\big)_I\Big] \le \frac{1}{4}\mathbb{E}\Big[\|\xi^\delta(t)\|_{L^2(I)}^2\Big] + \mathbb{E}\Big[\|e^\delta(0)\|_{L^2(I)}^2\Big] \le \frac{1}{4}\|\xi^\delta(t)\|_{L^2(I)}^2 + Ch^2.
\end{aligned}
$$

Taking the expectation in (4.23) and using the above estimates then yields

$$
\sup_{t\in[0,T]} \mathbb{E}\Big[\big\|\xi^\delta(t)\big\|_{L^2(I)}^2\Big] + 3\delta\,\mathbb{E}\Big[\int_0^T \big\|\partial_x \xi^\delta(s)\big\|_{L^2(I)}^2 \, ds\Big] \le Ch^2\big(1+\delta^{-2}\big). \tag{4.24}
$$

Finally, (4.20) follows from the triangle inequality, (4.19), and (4.24). The proof is complete. □

Remark 4.3.3. *(a) Estimate* (4.20) *is optimal in the H^1-norm, but suboptimal in the L^2-norm. The suboptimal rate for the L^2-error is caused by the stochastic effect, i.e., the second term on the right-hand side of* (4.23)*, and it is also caused by the lack of the space-time regularity in $L^\infty((0,T);H^2(I))$ for u^δ.*

(b) The proof still holds if the elliptic projection R_h^r is replaced by the L^2-projection P_h^r.

We now use estimate (4.24) to derive some stronger norm estimates for u_h^δ. To this end, we define the *discrete Laplacian* $\partial_h^2 : V_r^h \to V_r^h$ by

$$
(\partial_h^2 w_h, v_h)_I = -(\partial_x w_h, \partial_x v_h)_I \qquad \forall w_h, v_h \in V_r^h. \tag{4.25}
$$

Proposition 4.3.4. *For $\varepsilon \le \sqrt{2(1+\delta)}$ there hold the following estimates for the solution u_h^δ of scheme (4.12):*

$$\sup_{0\le t\le T} \mathbb{E}\Big[\|\partial_x u_h^\delta(t)\|_{L^2(I)}^2\Big] + \delta\,\mathbb{E}\Big[\int_0^T \|\partial_h^2 u_h^\delta(s)\|_{L^2(I)}^2\,ds\Big] \le C\big(1+\delta^{-2}\big), \tag{4.26}$$

$$\begin{aligned}\mathbb{E}\Big[\|u_h^\delta(t)-u_h^\delta(s)\|_{L^2(I)}^2 + \frac{\delta}{2}\int_s^t \|\partial_x[u_h^\delta(\zeta)-u_h^\delta(s)]\|_{L^2(I)}^2\,d\zeta\Big] \\ \le C\big(1+\delta^{-3}\big)|t-s| \qquad \forall\, 0\le s\le t\le T.\end{aligned} \tag{4.27}$$

Proof. Notice that $u_h^\delta = \xi^\delta + R_h^r u^\delta$ with $\xi^\delta = u_h^\delta - R_h^r u^\delta \in V_r^h$. By the H^1-stability of R_h^r, the following inverse inequality for piecewise polynomial function ξ^δ (cf. [13]),

$$\|\partial_x \xi^\delta(t)\|_{L^2(I)} \le Ch^{-1}\|\xi^\delta(t)\|_{L^2(I)},$$

(4.9), and (4.24) we get

$$\begin{aligned}\sup_{t\in[0,T]} \mathbb{E}\Big[\|\partial_x u_h^\delta(t)\|_{L^2(I)}^2\Big] &\le 2\sup_{t\in[0,T]} \mathbb{E}\Big[\|\partial_x R_h^r u^\delta(t)\|_{L^2(I)}^2\Big] + 2\sup_{t\in[0,T]} \mathbb{E}\Big[\|\partial_x \xi^\delta(t)\|_{L^2(I)}^2\Big] \\ &\le C\sup_{t\in[0,T]} \mathbb{E}\Big[\|\partial_x u^\delta(t)\|_{L^2(I)}^2\Big] + \frac{C}{h^2}\sup_{t\in[0,T]} \mathbb{E}\Big[\|\xi^\delta(t)\|_{L^2(I)}^2\Big] \\ &\le C(1+\delta^{-2}).\end{aligned}$$

It follows from (4.25) and (4.18) that

$$\begin{aligned}\|\partial_h^2 R_h^r w\|_{L^2(I)}^2 &= -\big(\partial_x\partial_h^2 R_h^r w, \partial_x R_h^r w\big)_I \\ &= \big(\partial_h^2 R_h^r w, \partial_x^2 w\big)_I + \big(w - R_h^r w, \partial_h^2 R_h^r w\big)_I \qquad \forall w\in H^2(I),\end{aligned}$$

and hence

$$\|\partial_h^2 R_h^r w\|_{L^2(I)} \le \|\partial_x^2 w\|_{L^2(I)} + \|w - R_h^r w\|_{L^2(I)} \le (1+Ch^2)\|w\|_{H^2(I)}. \tag{4.28}$$

By an inverse estimate, (4.28), and (4.24) we have

$$\begin{aligned}\mathbb{E}\Big[\int_0^T \|\partial_h^2 u_h^\delta(s)\|_{L^2(I)}^2\,ds\Big] &\le 2\mathbb{E}\Big[\int_0^T \Big(\|\partial_h^2\xi^\delta(s)\|_{L^2(I)}^2 + \|\partial_h^2 R_h^r u^\delta(s)\|_{L^2(I)}^2\Big)ds\Big]\\ &\le 2\mathbb{E}\Big[\int_0^T \Big(Ch^{-2}\|\partial_x\xi^\delta(s)\|_{L^2(I)}^2 + C\|\partial_x^2 u^\delta(s)\|_{L^2(I)}^2\Big)ds\Big]\\ &\le C\delta^{-1}(1+\delta^{-2}) + C\mathbb{E}\Big[\int_0^T \|\partial_x^2 u^\delta(s)\|_{L^2(I)}^2\,ds\Big],\end{aligned}$$

which and (4.9) give the desired bound in (4.26).

To show (4.27), we fix $s \ge 0$ and apply Ito's formula (cf. [62]) to $f(u_h^\delta) = \|u_h^\delta(t) - u_h^\delta(s)\|_{L^2(I)}^2$ to get that

$$\begin{aligned}&\|u_h^\delta(t) - u_h^\delta(s)\|_{L^2(I)}^2 \qquad (4.29)\\ &= -(\epsilon^2+2\delta)\int_s^t \Big(\partial_x[u_h^\delta(\zeta) \pm u_h^\delta(s)], \partial_x[u_h^\delta(\zeta) - u_h^\delta(s)]\Big)_I\,d\zeta\\ &\quad - (2-\epsilon^2)\int_s^t \Big(\arctan\big(\partial_x u_h^\delta(\zeta)\big) \pm \arctan\big(\partial_x u_h^\delta(s)\big), \partial_x[u_h^\delta(\zeta) - u_h^\delta(s)]\Big)_I\,d\zeta\\ &\quad + \epsilon^2\int_s^t \Big\|P_h^r\sqrt{1+|\partial_x\big[u_h^\delta(\zeta) \pm u_h^\delta(s)\big]|^2}\Big\|_{L^2(I)}^2\,d\zeta + M_t,\end{aligned}$$

where

$$M_t := \epsilon\int_s^t \Big(\sqrt{1+|\partial_x u_h^\delta(\zeta)|^2}, u_h^\delta(\zeta) - u_h^\delta(s)\Big)_I\,dW_\zeta,$$

which is an $\{\mathcal{F}_t; t\in[s,T]\}$-martingale.

By the L^2-stability of P_h^r, the triangle and Young's inequality, and the properties of the square root function, we can bound the third term on the right-hand side as

follows:

$$\begin{aligned}
&\epsilon^2 \int_s^t \Big\| P_h^r \sqrt{1 + |\partial_x \big[u_h^\delta(\zeta) \pm u_h^\delta(s)\big]|^2} \Big\|_{L^2(I)}^2 \, d\zeta \\
&\qquad \le \epsilon^2 \int_s^t \Big(\big\| \partial_x [u_h^\delta(\zeta) - u_h^\delta(s)] \big\|_{L^2(I)} + \big\| 1 + |\partial_x u_h^\delta(s)| \big\|_{L^2(I)} \Big)^2 \, d\zeta \\
&\qquad \le \epsilon^2 (1+\delta) \int_s^t \big\| \partial_x [u_h^\delta(\zeta) - u_h^\delta(s)] \big\|_{L^2(I)}^2 \, d\zeta \\
&\qquad\qquad + \epsilon^2 (4 + \delta^{-1}) \Big(1 + \|\partial_x u_h^\delta(s)\|_{L^2(I)}^2 \Big) \, |t-s|.
\end{aligned}$$

Also

$$\begin{aligned}
&(\epsilon^2 + 2\delta) \int_s^t \Big(\partial_x u_h^\delta(s), \partial_x [u_h^\delta(\zeta) - u_h^\partial(s)] \Big)_I \, d\zeta \\
&\le \frac{\delta}{4} \int_s^t \|\partial_x [u_h^\delta(\zeta) - u_h^\delta(s)]\|_{L^2(I)}^2 \, d\zeta + (\epsilon^2 + 2\delta)^2 \delta^{-1} \, |t-s| \|\partial_x u_h^\delta(s)\|_{L^2(I)}^2, \\
&(2 - \epsilon^2) \int_s^t \Big(\arctan\big(\partial_x u_h^\delta(s)\big), \partial_x [u_h^\delta(\zeta) - u_h^\partial(s)] \Big)_I \, d\zeta \\
&\le \frac{\delta}{4} \int_s^t \|\partial_x [u_h^\delta(\zeta) - u_h^\delta(s)]\|_{L^2(I)}^2 \, d\zeta + 4(2-\epsilon^2)^2 \delta^{-1} \, |t-s|.
\end{aligned}$$

Substituting the above estimates into (4.29) yields

$$\begin{aligned}
&\|u_h^\delta(t) - u_h^\delta(s)\|_{L^2(I)}^2 + \frac{\delta}{2} \int_s^t \|\partial_x [u_h^\delta(\zeta) - u_h^\delta(s)]\|^2 \, d\zeta \\
&\qquad \le C\delta^{-1} \Big((2-\epsilon^2)^2 + \epsilon^4 + \delta^2 \Big) \Big(1 + \|\partial_x u_h^\delta(s)\|_{L^2(I)}^2 \Big) \, |t-s| + M_t.
\end{aligned}$$

Finally, (4.27) follows from applying the expectation to the above inequality and using (4.26) as well as the fact that $\mathbb{E}[M_t] = 0$. □

4.3.2 Full discretization in space and in time

Let $t_n = n\tau$ for $n = 0, 1, \cdots, N$ be a uniform partition of $[0, T]$ with $\tau = T/N$. Our fully discrete finite element method for SPDE (4.7) is defined by seeking an $\{\mathcal{F}_{t_n}; n = 0, 1, \cdots, N\}$-adapted V_r^h-valued process $\{u_h^n;\, n = 0, 1, \cdots, N\}$ such that

$\mathbb{P}$-almost surely

$$\begin{aligned}
&\big(u_h^{\delta,n+1}, v_h\big)_I + \tau\big(\delta + \frac{\epsilon^2}{2}\big)\big(\partial_x u_h^{\delta,n+1}, \partial_x v_h\big)_I \\
&\qquad + \tau\big(1 - \frac{\epsilon^2}{2}\big)\big(\arctan(\partial_x u_h^{\delta,n+1}), \partial_x v_h\big)_I \\
&= \big(u_h^{\delta,n}, v_h\big)_I + \epsilon\Big(\sqrt{1 + |\partial_x u_h^{\delta,n}|^2}, v_h\Big)_I \Delta W_{n+1} \qquad \forall v_h \in V_r^h,
\end{aligned} \tag{4.30}$$

where $\Delta W_{n+1} := W(t_{n+1}) - W(t_n) \sim \mathcal{N}(0, \tau)$.

We first establish the following stability estimate for $u_h^{\delta,n}$.

Proposition 4.3.5. *Let $\epsilon \le \sqrt{2(1+\delta)}$. For each $n = 0, 1, \cdots, N$, there is a V_r^h-valued discrete process $\{u_h^{\delta,n+1};\ 0 \le n \le N-1\}$ which solves scheme (4.30). Moreover, there holds*

$$\max_{0\le n\le N} \mathbb{E}\Big[\big\|u_h^{\delta,n}\big\|_{L^2(I)}^2\Big] + 2\delta \sum_{n=0}^{N} \tau \mathbb{E}\Big[\big\|\partial_x u_h^{\delta,n}\big\|_{L^2(I)}^2\Big] \le \mathbb{E}\Big[\big\|u_h^{\delta,0}\big\|_{L^2(I)}^2\Big] + \epsilon^2 T. \tag{4.31}$$

Proof. The existence of solutions to scheme (4.30) for $\tau, h > 0$ can be proved by Brouwer's fixed-point theorem, which uses the coercivity of the operator $I + \tau A_h^\delta$ (see (4.13)).

To show (4.31), we choose $v_h = u_h^{\delta,n+1}(\omega)$ in (4.30) to find $\mathbb{P}$-almost surely

$$\begin{aligned}
&\frac{1}{2}\Big[\big\|u_h^{\delta,n+1}\big\|_{L^2(I)}^2 - \big\|u_h^{\delta,n}\big\|_{L^2(I)}^2\Big] + \frac{1}{2}\big\|u_h^{\delta,n+1} - u_h^{\delta,n}\big\|_{L^2(I)}^2 \\
&\quad + \tau\big(\delta + \frac{\epsilon^2}{2}\big)\big\|\partial_x u_h^{\delta,n+1}\big\|_{L^2(I)}^2 + \tau\big(1 - \frac{\epsilon^2}{2}\big)\big(\arctan(\partial_x u_h^{\delta,n+1}), \partial_x u_h^{\delta,n+1}\big)_I \\
&= \epsilon\Big(\sqrt{1 + |\partial_x u_h^{\delta,n}|^2}, u_h^{\delta,n} + u_h^{\delta,n+1} - u_h^{\delta,n}\Big)_I \Delta W_{n+1}.
\end{aligned} \tag{4.32}$$

We compute

$$\begin{aligned}
&\big(\arctan(\partial_x u_h^{\delta,n+1}), \partial_x u_h^{\delta,n+1}\big)_I \geq 0,\\
&\epsilon\Big(\sqrt{1+|\partial_x u_h^{\delta,n}|^2}, u_h^{\delta,n+1}-u_h^{\delta,n}\Big)_I \Delta W_{n+1}\\
&\qquad \leq \frac{1}{2}\big\|u_h^{\delta,n+1}-u_h^{\delta,n}\big\|_{L^2(I)}^2 + \frac{\epsilon^2}{2}\Big\|\sqrt{1+|\partial_x u_h^{\delta,n}|^2}\Big\|_{L^2(I)}^2 |\Delta W_{n+1}|^2.
\end{aligned}$$

The last estimate controls one part of the stochastic term in (4.32), while the expectation of the remaining part vanishes. By the tower property for expectations, there holds

$$\frac{\epsilon^2}{2}\mathbb{E}\Big[\Big\|\sqrt{1+|\partial_x u_h^{\delta,n}|^2}\Big\|_{L^2(I)}^2 \mathbb{E}\big[|\Delta W_{n+1}|^2|\mathcal{F}_{t_n}\big]\Big] = \frac{\epsilon^2}{2}\tau\,\mathbb{E}\Big[1+\|\partial_x u^{\delta,n}\|_{L^2(I)}^2\Big],$$

such that we get

$$\begin{aligned}
&\frac{1}{2}\mathbb{E}\left[\big\|u_h^{\delta,n+1}\big\|_{L^2(I)}^2 - \big\|u_h^{\delta,n}\big\|_{L^2(I)}^2\right] + \tau\delta\,\mathbb{E}\Big[\big\|\partial_x u_h^{\delta,n+1}\big\|_{L^2(I)}^2\Big] \\
&\qquad + \frac{\epsilon^2}{2}\tau\,\mathbb{E}\Big[\big\|\partial_x u_h^{\delta,n+1}\big\|_{L^2(I)}^2 - \big\|\partial_x u_h^{\delta,n}\big\|_{L^2(I)}^2\Big] \leq \epsilon^2\tau.
\end{aligned} \tag{4.33}$$

After summation, we arrive at

$$\max_{0\leq n\leq N}\mathbb{E}\big[\big\|u_h^{\delta,n}\big\|_{L^2(I)}^2\big] + 2\delta\tau\sum_{n=0}^{N}\mathbb{E}\Big[\big\|\partial_x u_h^{\delta,n}\big\|_{L^2(I)}^2\Big] \leq \mathbb{E}\Big[\big\|u_h^{\delta,0}\big\|_{L^2(I)}^2\Big] + \epsilon^2 T.$$

So (4.31) holds. The proof is complete. □

Next, we derive an error bound for $u_h^{\delta}(t_n) - u_h^{\delta,n}$.

Theorem 4.3.6. *Let $r = 1$. There holds the following error estimate:*

$$\begin{aligned}
&\sup_{0\leq n\leq N}\mathbb{E}\Big[\big\|u_h^{\delta}(t_n) - u_h^{\delta,n}\big\|_{L^2(I)}^2\Big] + \delta\,\mathbb{E}\Big[\sum_{n=0}^{N}\tau\big\|\partial_x u_h^{\delta}(t_n) - \partial_x u_h^{\delta,n}\big\|_{L^2(I)}^2\Big] \\
&\qquad\qquad \leq CT\big(1+\delta^{-2}\big)h^{-2}\tau.
\end{aligned} \tag{4.34}$$

Proof. Let $e^{\delta,n} := u_h^\delta(t_n) - u_h^{\delta,n}$. It follows from (4.12) that for all $\{t_n; n \geq 0\}$ there holds $\mathbb{P}$-almost surely

$$\begin{aligned}
&\big(u_h^\delta(t_{n+1}), v_h\big)_I - \big(u_h^\delta(t_n), v_h\big)_I \qquad (4.35)\\
&\quad = -\big(\delta + \frac{\epsilon^2}{2}\big) \int_{t_n}^{t_{n+1}} \big(\partial_x u_h^\delta(s), \partial_x v_h\big)_I ds\\
&\qquad - \big(1 - \frac{\epsilon^2}{2}\big) \int_{t_n}^{t_{n+1}} \big(\arctan(\partial_x u_h^\delta(s)), \partial_x v_h\big)_I ds\\
&\qquad + \epsilon \int_{t_n}^{t_{n+1}} \Big(\sqrt{1 + |\partial_x u_h^\delta(s)|^2}, v_h\Big)_I dW_s \qquad \forall v_h \in V_r^h.
\end{aligned}$$

Subtracting (4.30) from (4.35) yields the following error equation:

$$\begin{aligned}
&\big(e^{\delta,n+1}, v_h\big)_I - \big(e^{\delta,n}, v_h\big)_I \qquad (4.36)\\
&= -\big(\delta + \frac{\epsilon^2}{2}\big) \int_{t_n}^{t_{n+1}} \big(\partial_x u_h^\delta(s) - \partial_x u_h^{\delta,n+1}, \partial_x v_h\big)_I ds\\
&\quad - \big(1 - \frac{\epsilon^2}{2}\big) \int_{t_n}^{t_{n+1}} \Big(\arctan\big(\partial_x u_h^\delta(s)\big) - \arctan(\partial_x u_h^{\delta,n+1}), \partial_x v_h\Big)_I ds\\
&\quad + \epsilon \int_{t_n}^{t_{n+1}} \Big(\sqrt{1 + |\partial_x u_h^\delta(s)|^2} - \sqrt{1 + |\partial_x u_h^{\delta,n}|^2}, v_h\Big)_I dW_s.
\end{aligned}$$

Choosing $v_h = e^{\delta,n+1}(\omega)$ in (4.36) leads to $\mathbb{P}$-almost surely

$$\begin{aligned}
&\frac{1}{2}\big[\|e^{\delta,n+1}\|_{L^2(I)}^2 - \|e^{\delta,n}\|_{L^2(I)}^2\big] + \frac{1}{2}\big\|e^{\delta,n+1} - e^{\delta,n}\big\|_{L^2(I)}^2 \qquad (4.37)\\
&\qquad + \big(\frac{\epsilon^2}{2} + \delta\big)\tau \,\|\partial_x e^{\delta,n+1}\|_{L^2(I)}^2\\
&= -\big(\frac{\epsilon^2}{2} + \delta\big) \int_{t_n}^{t_{n+1}} \Big(\partial_x u_h^\delta(s) - \partial_x u_h^\delta(t_{n+1}), \partial_x e^{\delta,n+1}\Big)_I ds\\
&\quad - \big(1 - \frac{\epsilon^2}{2}\big) \int_{t_n}^{t_{n+1}} \Big(\arctan\big(\partial_x u_h^\delta(s)\big) - \arctan\big(\partial_x u_h^\delta(t_{n+1})\big)\\
&\quad + \arctan\big(\partial_x u_h^\delta(t_{n+1})\big) - \arctan(\partial_x u_h^{\delta,n+1}), \partial_x e^{\delta,n+1}\Big)_I ds\\
&\quad + \epsilon \int_{t_n}^{t_{n+1}} \Big(\sqrt{1 + |\partial_x u_h^\delta(s)|^2} - \sqrt{1 + |\partial_x u_h^{\delta,n}|^2}, e^{\delta,n+1}\Big)_I dW_s.
\end{aligned}$$

We now bound each term on the right-hand side. First, since $\mathbb{E}[\Delta W_{n+1}] = 0$, by Ito's isometry, the inequality $\left(\sqrt{1+x^2} - \sqrt{1+y^2}\right)^2 \leq (x-y)^2$, and the inverse inequality we get

$$
\begin{aligned}
&\mathbb{E}\Big[\epsilon \int_{t_n}^{t_{n+1}} \Big(\sqrt{1+|\partial_x u_h^\delta(s)|^2} - \sqrt{1+|\partial_x u_h^{\delta,n}|^2}, e^{\delta,n+1} \pm e^{\delta,n}\Big)_I dW_s\Big] \\
&\leq \mathbb{E}\Big[\frac{1}{2}\|e^{\delta,n+1} - e^{\delta,n}\|_{L^2(I)}^2\Big] + \frac{\epsilon^2}{2}\int_{t_n}^{t_{n+1}} \|\partial_x[u_h^\delta(s) - u_h^{\delta,n} \pm u_h^\delta(t_n)]\|_{L^2(I)}^2 \, ds\Big] \\
&\leq \frac{1}{2}\mathbb{E}\Big[\|e^{\delta,n+1} - e^{\delta,n}\|_{L^2(I)}^2\Big] + \mathbb{E}\Big[(\frac{\epsilon^2}{2} + \frac{\delta}{2})\tau \, \|\partial_x e^{\delta,n}\|_{L^2(I)}^2 \\
&\qquad + (\frac{\epsilon^2}{2} + \frac{2}{\delta})\int_{t_n}^{t_{n+1}} \|\partial_x[u_h^\delta(s) - u_h^\delta(t_n)]\|_{L^2(I)}^2 \, ds\Big] \\
&\leq \frac{1}{2}\mathbb{E}\Big[\|e^{\delta,n+1} - e^{\delta,n}\|_{L^2(I)}^2\Big] + (\frac{\epsilon^2}{2} + \frac{\delta}{2})\tau \, \mathbb{E}\Big[\|\partial_x e^{\delta,n}\|_{L^2(I)}^2\Big] \\
&\qquad + C\big(1+\delta^{-1}\big)h^{-2}\,\mathbb{E}\Big[\int_{t_n}^{t_{n+1}} \|u_h^\delta(s) - u_h^\delta(t_n)\|_{L^2(I)}^2 \, ds\Big]. \qquad (4.38)
\end{aligned}
$$

An elementary calculation and an application of an inverse inequality yield

$$
\begin{aligned}
&(\frac{\epsilon^2}{2} + \delta)\int_{t_n}^{t_{n+1}} \Big(\partial_x u_h^\delta(s) - \partial_x u_h^\delta(t_{n+1}), \partial_x e^{\delta,n+1}\Big)_I ds \\
&\leq \frac{\delta}{8}\tau \, \|\partial_x e^{\delta,n+1}\|_{L^2(I)}^2 + \frac{2(\frac{\epsilon^2}{2} + \delta)^2}{\delta}\int_{t_n}^{t_{n+1}} \|\partial_x[u_h^\delta(s) - u_h^\delta(t_{n+1})]\|_{L^2(I)}^2 \, ds \\
&\leq \frac{\delta}{8}\tau \, \|\partial_x e^{\delta,n+1}\|_{L^2(I)}^2 + (\epsilon^2 + 2\delta)^2\delta^{-1}h^{-2}\int_{t_n}^{t_{n+1}} \|u_h^\delta(s) - u_h^\delta(t_n)\|_{L^2(I)}^2 \, ds. \qquad (4.39)
\end{aligned}
$$

By the monotonicity of arctan we get

$$
\begin{aligned}
&-\big(1-\frac{\epsilon^2}{2}\big)\int_{t_n}^{t_{n+1}}\Big(\arctan\big(\partial_x u_h^\delta(s)\big)-\arctan\big(\partial_x u_h^\delta(t_{n+1})\big)\\
&+\arctan\big(\partial_x u_h^\delta(t_{n+1})\big)-\arctan(\partial_x u_h^{\delta,n+1}),\partial_x e^{\delta,n+1}\Big)_I ds\\
&\le -\big(1-\frac{\epsilon^2}{2}\big)\int_{t_n}^{t_{n+1}}\Big(\arctan\big(\partial_x u_h^\delta(s)\big)-\arctan\big(\partial_x u_h^\delta(t_{n+1})\big),\partial_x e^{\delta,n+1}\Big)\,ds\\
&\le \frac{\delta}{8}\tau\,\|\partial_x e^{\delta,n+1}\|^2_{L^2(I)}+\frac{2}{\delta}(1-\frac{\epsilon^2}{2})^2\int_{t_n}^{t_{n+1}}\|\partial_x[u_h^\delta(s)-u_h^\delta(t_{n+1})]\|^2_{L^2(I)}\,ds\\
&\le \frac{\delta}{8}\tau\,\|\partial_x e^{\delta,n+1}\|^2_{L^2(I)}+(2-\epsilon^2)^2\delta^{-1}h^{-2}\int_{t_n}^{t_{n+1}}\|u_h^\delta(s)-u_h^\delta(t_n)\|^2_{L^2(I)}\,ds. \qquad (4.40)
\end{aligned}
$$

Finally, substituting the above estimates into (4.37), summing over $n = 0,1,2,\cdots,N-1$, and using (4.27) and the fact that $e^{\delta,0}=0$ we get

$$
\begin{aligned}
&\sup_{0\le n\le N}\mathbb{E}\Big[\|e^{\delta,n}\|^2_{L^2(I)}\Big]+\frac{\delta}{2}\mathbb{E}\Big[\tau\sum_{n=0}^{N}\|\partial_x e^{\delta,n+1}\|^2_{L^2(I)}\Big]\\
&\qquad\le C\big(1+\delta^{-1}\big)h^{-2}\sum_{n=0}^{N}\int_{t_n}^{t_{n+1}}\mathbb{E}\Big[\sup_{s\in[t_n,t_{n+1}]}\|u_h^\delta(s)-u_h^\delta(t_n)\|^2_{L^2(I)}\Big]\,ds\\
&\qquad\le CT(1+\delta^{-1})^2h^{-2}\tau,
\end{aligned}
$$

which infers (4.34). The proof is complete. □

Remark 4.3.7. *Due to the lack of a Hölder continuity (in time) estimate for $\partial_x u_h^\delta$ in L^2-norm, we have used the inverse inequality to get inequalities (4.38)–(4.40), which leads to a restrictive coupling of the spatial and time mesh parameters h and τ in the error bound.*

We conclude this section by stating the following error estimates for the fully discrete finite element solution $u_h^{\delta,n}$ as an approximation to the solution of the original mean curvature flow equation (4.4).

Theorem 4.3.8. *Let u and $u_h^{\delta,n}$ denote respectively the solutions of SPDE (4.4) and scheme (4.30). Under assumptions of Theorems 4.2.1, 4.3.2, and 4.3.6, there holds*

the following error estimate:

$$\sup_{0\le n\le N}\mathbb{E}\Big[\big\|u(t_n)-u_h^{\delta,n}\big\|_{L^2(I)}^2\Big]+\delta\,\mathbb{E}\Big[\sum_{n=0}^{N}\tau\big\|\partial_x u(t_n)-\partial_x u_h^{\delta,n}\big\|_{L^2(I)}^2\Big] \tag{4.41}$$
$$\le CT\delta+C\big(1+\delta^{-2}\big)h^2+CT\big(1+\delta^{-2}\big)h^{-2}\tau.$$

Inequality (4.41) follows immediately from Theorems 4.2.3, 4.3.2 and 4.3.6, and an application of the triangle inequality.

Remark 4.3.9. *Again, we note that the main reason to have a restrictive coupling between numerical parameters in* (4.41) *is due to the lack of Hölder continuity (in time) estimate for* $\partial_x u_h^\delta$ *in* L^2*-norm. On the other hand, it can be shown that, under a stronger regularity assumption, the estimate* (4.41) *can be improved to*

$$\sup_{0\le n\le N}\mathbb{E}\Big[\big\|u(t_n)-u_h^{\delta,n}\big\|_{L^2(I)}^2\Big] \tag{4.42}$$
$$+\,\delta\,\mathbb{E}\Big[\sum_{n=0}^{N}\tau\big\|\partial_x u(t_n)-\partial_x u_h^{\delta,n}\big\|_{L^2(I)}^2\Big]\le C\big(h^2+\tau+\delta\big)\,.$$

This is because we no longer need to use the inverse inequality to get (4.38)–(4.40)*, and* (4.42) *can be obtained by starting with a control of the time discretization first.*

4.4 Numerical experiments

In this section we shall first present some numerical experiments to gauge the performance of the proposed fully discrete finite element method and to examine the effect of the noise for long-time dynamics of the stochastic MCF of planar graphs, and we then present a numerical study of the stochastic MCF driven by both colored and space-time white noises where no theoretical result is known so far in the literature. We like to note that all our numerical experiments are done in Matlab. At each time step, a nonlinear algebraic system must be solved, which is done by using Matlab's

built-in Newton solver in all our numerical tests. In addition, all space norms are computed approximately using sufficiently high order numerical quadrature formulas.

4.4.1 Verifying the rate of convergence of time discretization

To verify the rate of convergence of the time discretization obtained in Theorem 4.3.8, in this first test we use the following parameters $\epsilon = 1$, $\delta = 10^{-5}$, and $T = 0.1$. In order to computationally generate a driving reference $\mathbb{R}$-valued Wiener process, we use the smaller time step $\tau = 10^{-5}$. The initial condition is set to be $u_0(x) = \sin(\pi x)$. To calculate the rate, we compute the solution $u_h^{\delta,n}$ for varying $\tau = 0.0005, 0.001, 0.002, 0.004$. We take 500 stochastic samples at each time step t_n in order to compute the expected values of the $L^\infty(L^2)$-norm of the error. The computed errors along with the computed convergence rates are exhibited in Table 4.1 and Fig. 4.1. The numerical results confirm the theoretical result of Theorem 4.3.6. In Fig. 4.2, we plot the errors of the computed solution with regularization (i.e., $\delta > 0$) and without regularization (i.e., $\delta = 0$). The comparison shows that without the regularization term our numerical methods still compute correct solutions for some problems although our convergence theory requires that $\delta > 0$.

Table 4.1: Computed time discretization errors and convergence rates.

	Expected values of error	order of convergence
dt=0.004	0.41965657	–
dt=0.002	0.27206448	0.62526
dt=0.001	0.18136210	0.58508
dt=0.0005	0.12373884	0.55157

4.4.2 Dynamics of the stochastic mean curvature flow

We shall perform several numerical tests to demonstrate the dynamics of the stochastic MCF with different magnitudes of noise (i.e., different sizes of the parameter ϵ).

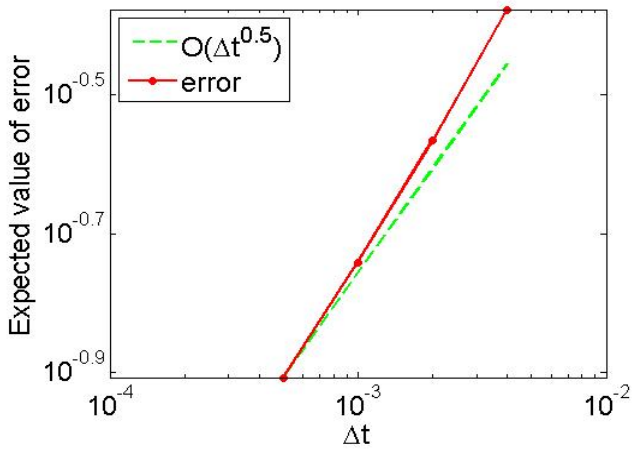

Figure 4.1: Plot of the errors in Table 4.1.

Figure 4.3 shows the surface plots of the computed solution $u_h^{\delta,n}$ at one stochastic sample over the space-time domains $(0,1)\times(0,0.1)$ (left) and $(0,1)\times(0,2^8\times10^{-5})$ (right) with the initial value $u_0(x)=\sin(\pi x)$ and the noise intensity parameter $\epsilon=0.1$. The test shows that the solution converges to a steady state solution at the end.

Figures 4.4–4.6 are the counterparts of Figure 4.3 with noise intensity parameter $\epsilon=1,\sqrt{2},5$, respectively. We note that the error estimate of Theorem 4.3.8 does not apply to the latter case because the condition $\epsilon\le\sqrt{2(1+\delta)}$ is violated. However, the computation result suggests that the stochastic MCF also converges to the steady state solution at the end although the paths to reach the steady state are different for different noise intensity parameter ϵ.

We then repeat the above four tests after replacing the smooth initial function u_0 by the following non-smooth initial function:

$$u_0(x)=\begin{cases}10x, & \text{if } x\le 0.25,\\ 5-10x, & \text{if } 0.25<x\le 0.5,\\ 10x-5, & \text{if } 0.5<x\le 0.75,\\ 10-10x, & \text{if } 0.75<x\le 1.\end{cases}\tag{4.43}$$

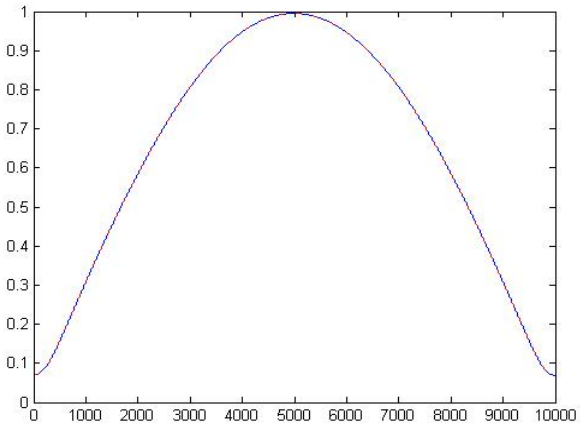

Figure 4.2: Comparison of the computed solution with (blue line) and without (red line) the regularization term (Color figure online).

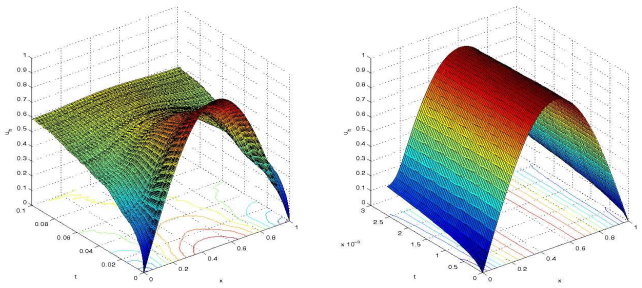

Figure 4.3: Surface plots of computed solution at a fixed stochastic sample on the space time domains $(0,1) \times (0,0.1)$ (left) and $(0,1) \times (0, 2^8 \times 10^{-5})$ (right). $u_0(x) = \sin(\pi x)$ and $\epsilon = 0.1$.

The surface plots of the computed solutions are shown in Figures 4.7–4.10, respectively. Again, the numerical results suggest that the solution of the stochastic MCF converges to the steady state solution at the end although the paths to reach the steady state are different for different noise intensity parameter ϵ. As expected, the geometric evolution dominates for small ϵ, but the noise dominates the geometric evolution for large ϵ.

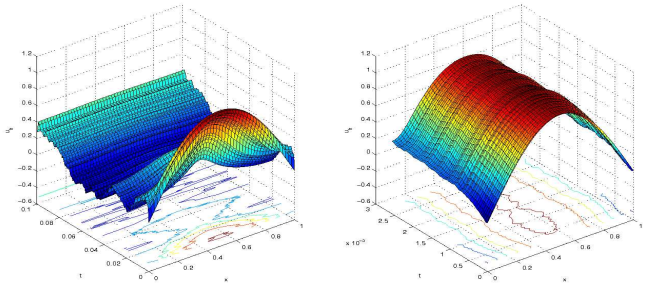

Figure 4.4: Surface plots of computed solution at a fixed stochastic sample on the space time domains $(0,1)\times(0,0.1)$ (left) and $(0,1)\times(0,2^8\times10^{-5})$ (right). $u_0(x)=\sin(\pi x)$ and $\epsilon=1$.

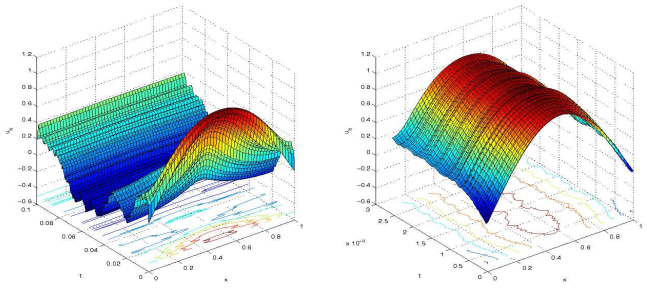

Figure 4.5: Surface plots of computed solution at a fixed stochastic sample on the space time domains $(0,1)\times(0,0.1)$ (left) and $(0,1)\times(0,2^8\times10^{-5})$ (right). $u_0(x)=\sin(\pi x)$ and $\epsilon=\sqrt{2}$.

4.4.3 Verifying energy dissipation

It follows from (4.9) that the "energy" $J(t) := \frac{1}{2}\mathbb{E}\big[\|\partial_x u^\delta(t)\|^2_{L^2(I)}\big]$ decreases monotonically in time. In the following we verify this fact numerically. Again, we consider the case with the initial function $u_0(x)=\sin(\pi x)$ and the noise intensity parameter $\epsilon=1$. It is not hard to prove that $J(t)$ converges to zero as $t\to\infty$. Figure 4.11 plots the computed $J(t)$ as a function of t. The numerical result suggests that $J(t)$ does not change anymore for $t\geq 0.1$.

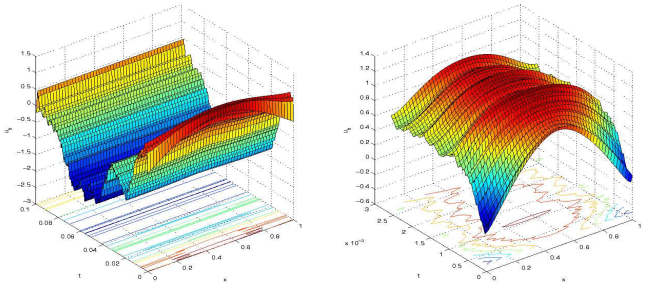

Figure 4.6: Surface plots of computed solution at a fixed stochastic sample on the space time domains $(0,1)\times(0,0.1)$ (left) and $(0,1)\times(0,2^8\times10^{-5})$ (right). $u_0(x) = \sin(\pi x)$ and $\epsilon = 5$.

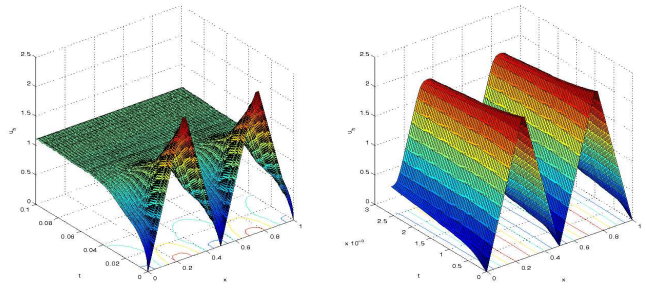

Figure 4.7: Surface plots of computed solution at a fixed stochastic sample on the space time domains $(0,1)\times(0,0.1)$ (left) and $(0,1)\times(0,2^8\times10^{-5})$ (right). u_0 is given in (4.43) and $\epsilon = 0.1$.

4.4.4 Thresholding for colored noise

In this subsection we present a computational study of the interplay of noise and geometric evolution in (4.4), which is beyond our theoretical results in section 4.3.1 and 4.3.2. For this purpose, we use driving colored noise represented by the Q-Wiener process ($J \in \mathbb{N}$)

$$W_t = \sum_{j=1}^{J} q_j^{\frac{1}{2}} \beta_j(t) e_j \,, \tag{4.44}$$

where $\{\beta_j(t);\, t \geq 0\}_{j\geq 1}$ denotes a family of real-valued independent Wiener processes on $\big(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P}\big)$, and $\{(q_j, e_j)\}_{j=1}^{J}$ is an eigen-system of the symmetric, non-negative trace-class operator $Q : L^2(I) \to L^2(I)$, with $e_j = \sqrt{2}\sin(j\pi x)$. In particular, we like to numerically address the following questions:

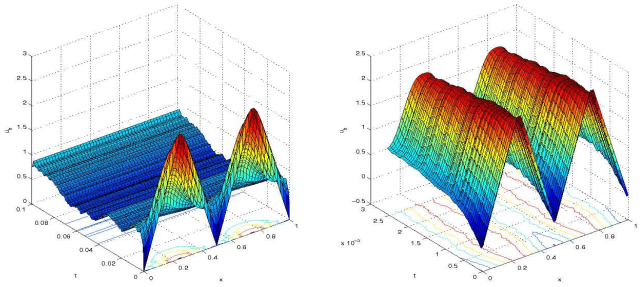

Figure 4.8: Surface plots of computed solution at a fixed stochastic sample on the space time domains $(0,1)\times(0,0.1)$ (left) and $(0,1)\times(0,2^8\times10^{-5})$ (right). u_0 is given in (4.43) and $\epsilon = 1$.

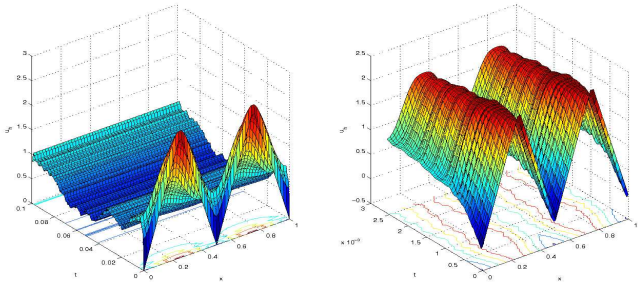

Figure 4.9: Surface plots of computed solution at a fixed stochastic sample on the space time domains $(0,1)\times(0,0.1)$ (left) and $(0,1)\times(0,2^8\times10^{-5})$ (right). u_0 is given in (4.43) and $\epsilon = \sqrt{2}$.

(A) *Thresholding*: By Theorem 4.2.1, strong solutions of (4.4) exist for $\varepsilon \leq \sqrt{2}$, and a similar result can be shown for the PDE problem with the noise (4.44). What are admissible intensities of the noise suggested by computations? Moreover, what do the computations suggest about the stochastic MCF in the case of spatially white noise (i.e., $q_j \equiv 1, J = \infty$) where no theoretical result is available so far?

(B) *General initial profiles*: The deterministic evolution of Lipschitz initial graphs is well-understood. For example, the (upper) graph of two touching spheres may trigger non-uniqueness. What are the regularization and the noise excitation effects in the case of the initial data with infinite energy and using different noises?

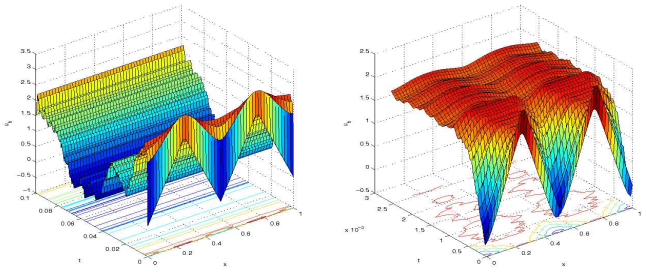

Figure 4.10: Surface plots of computed solution at a fixed stochastic sample on the space time domains $(0,1)\times(0,0.1)$ (left) and $(0,1)\times(0,2^8\times10^{-5})$ (right). u_0 is given in (4.43) and $\epsilon = 5$.

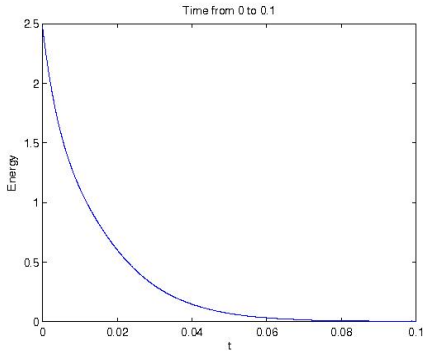

Figure 4.11: Decay of the energy $J(t)$ on the interval $(0, 0.1)$.

Recall that the estimate in Proposition 4.3.5 for V_r^h-valued solution $u_h^{\delta,n}$ suggests that $\varepsilon > 0$ ought be sufficiently small to ensure the existence. In our test, we employ the colored noise (4.44) with $q_j^{\frac{1}{2}} = j^{-0.6}$, $J = 20$, and the following non-Lipschitz initial data:

$$u_0(x) = |0.5 - x|^\kappa \qquad \forall\, x \in (0,1)\,, \tag{4.45}$$

where $\kappa = 0.1$. In addition, we set $(\tau, h) = (0.01, 0.02)$ and $T = \frac{1}{2}$. Figure 4.12 shows the single trajectory of the stochastic MCF plotted as graphs over the space-time domain with, respectively, $\epsilon = 0.1, 0.5, \sqrt{2}$. The results indicate thresholding, namely, the trajectories grow rapidly in time for sufficiently large values ε, and the noise effect dominates the geometric evolution. The excitation effect of the noise on the geometric evolution is illustrated by corresponding plots for the evolution of the

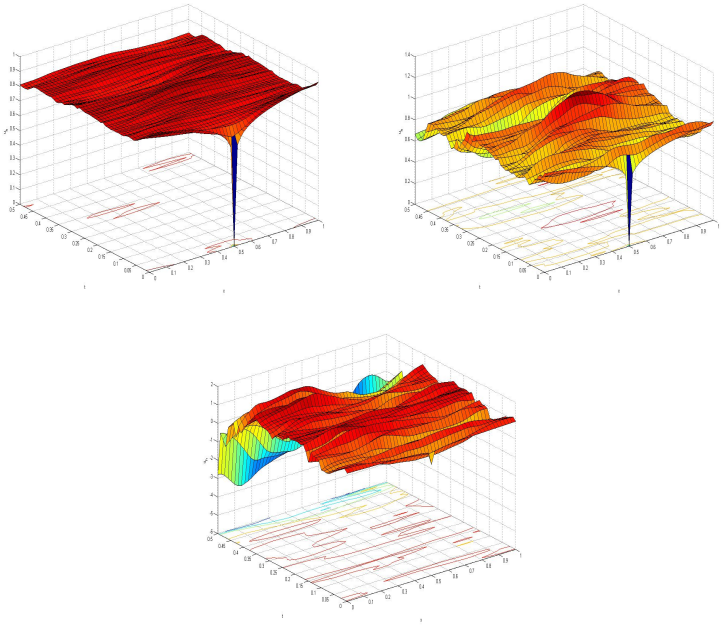

Figure 4.12: Thresholding for colored noise: Trajectories for $\varepsilon = 0.1$ (top left), $\varepsilon = 0.5$ (top right), $\varepsilon = \sqrt{2}$ (bottom).

functional $n \mapsto \|\partial_x u_h^{\delta,n}(\omega)\|_{L^2}^2$ vs its expectation $n \mapsto \mathbb{E}\big[\|\partial_x u_h^{\delta,n}\|_{L^2}^2\big]$ in Figure 4.13 and 4.14. We observe that the geometric evolution dominates for small values of ε, while the noise evolution takes over for large values of ε.

4.4.5 Thresholding for white noise

We now consider the case of white noise in (4.30), that is, $q_j \equiv 1$ in (4.44) and $J = \infty$, for which the solvability of (4.2) is not known. Figure 4.15 shows the single trajectory of the stochastic MCF (with the same data as in section 4.4.4) plotted as graphs over the space-time domain with, respectively, $\epsilon = 0.1, 0.5, \sqrt{2}$. We observe a very rapid growth of trajectories (numerical values range between 10^{14} and 10^{21}) even for small values of $\varepsilon > 0$. These numerical results suggest either a rapid growth or a finite time explosion for the stochastic MCF in the case of white noise.

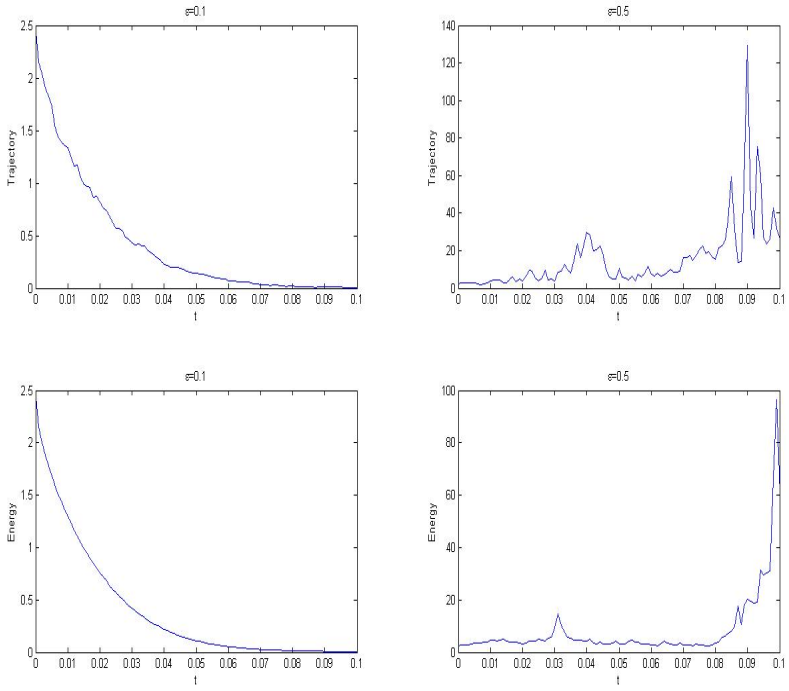

Figure 4.13: Geometric evolution vs colored noise evolution ($q_j^{\frac{1}{2}} = j^{-0.6}$, $J = 20$): 1st row: single trajectory for $n \mapsto \|\partial_x u_h^{\delta,n}(\omega)\|_{L^2}^2$ and $\varepsilon = 0.1$ (left), $\varepsilon = 0.5$ (right); 2nd row: $n \mapsto \mathbb{E}\big[\|\partial_x u_h^{\delta,n}\|_{L^2}^2\big]$ for $\varepsilon = 0.1$ (left), $\varepsilon = 0.5$ (right).

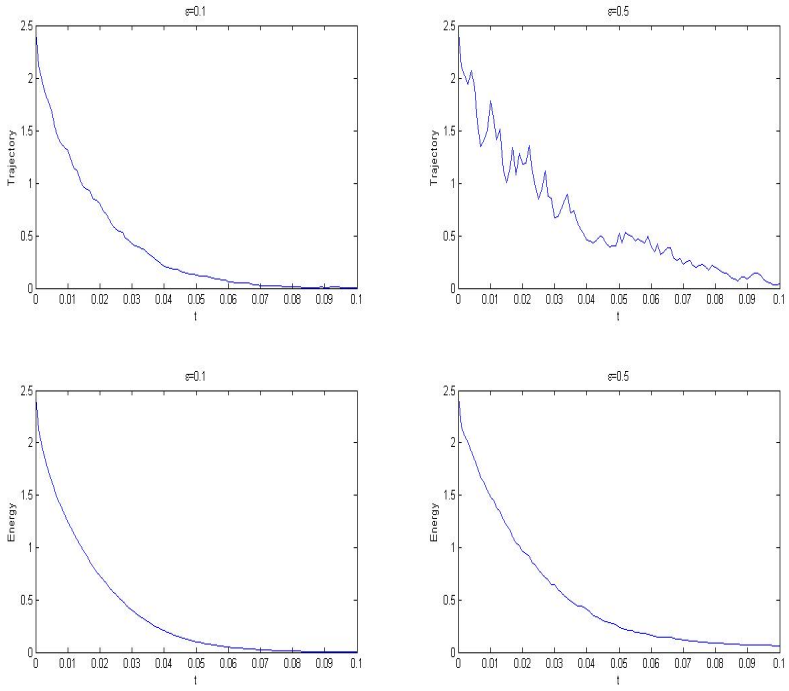

Figure 4.14: Geometric evolution vs colored noise evolution ($q_j^{\frac{1}{2}} = j^{-1}$, $J = 50$): 1st row: single trajectory for $n \mapsto \|\partial_x u_h^{\delta,n}(\omega)\|_{L^2}^2$ and $\varepsilon = 0.1$ (left), $\varepsilon = 0.5$ (right); 2nd row: $n \mapsto \mathbb{E}\big[\|\partial_x u_h^{\delta,n}\|_{L^2}^2\big]$ for $\varepsilon = 0.1$ (left), $\varepsilon = 0.5$ (right).

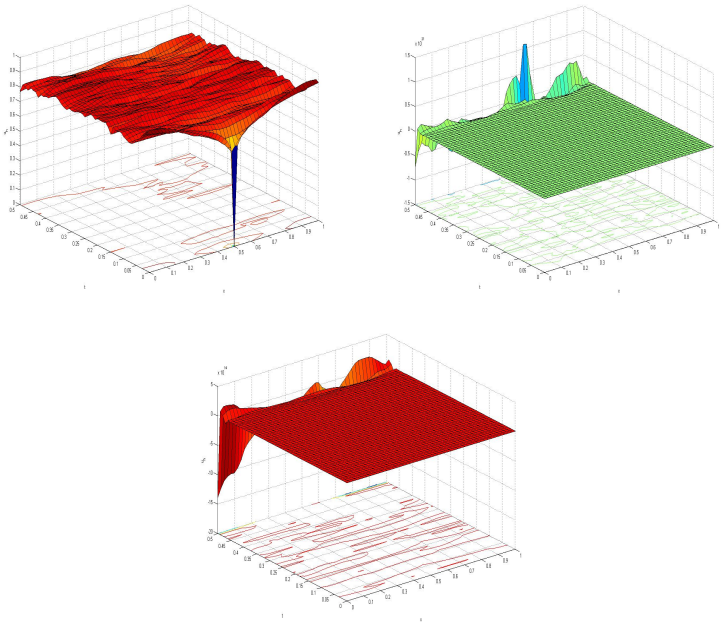

Figure 4.15: Thresholding and white noise: $\varepsilon = 0.1$ (top left), $\varepsilon = 0.5$ (top right) $\varepsilon = \sqrt{2}$ (bottom).

Chapter 5

Finite Element Methods for the Stochastic Allen-Cahn Equation with Gradient Multiplicative Noises

5.1 Introduction

In this chapter, we consider a stochastic version of mean curvature flow which incorporates the influence of noises for (1.4). The uncertainty may arise from thermal fluctuation, impurities of the materials, and the intrinsic instabilities of the deterministic evolutions. A simple model is given by [59, 83]:

$$V_n(x,t) = -H(x,t) + \delta \overset{\circ}{\mathbb{X}}(x,t) \cdot n \qquad x \in \Gamma_t, \tag{5.1}$$

where $\overset{\circ}{\mathbb{X}}$: $\mathbb{R}^d \times [0,T] \longrightarrow \mathbb{R}^d$ indicates the Stratonovich derivative of $\mathbb{X}$ and $\delta > 0$ is the noise intensity that controls the strength of the noise. Formally, the phase field formulation of (5.1) is given by the following Stratonovich stochastic partial

differential equation (SPDE):

$$\frac{du}{dt} = \Delta u - \frac{1}{\epsilon^2} f(u) + \delta \nabla u \cdot \overset{\circ}{\mathbb{X}} . \tag{5.2}$$

More specifically, we assume $\mathbb{X}$ is a vector-field Brownian motion that is white in time and smooth in space, i.e.,

$$\mathbb{X}(x,t) = X(x)W(t), \tag{5.3}$$

where $X : \mathbb{R}^d \longrightarrow \mathbb{R}^d$ is a time independent deterministic smooth vector field with compact support in $\mathcal{D}$ and $W(t)$ is a standard $\mathbb{R}$-valued Brownian motion (Wiener process) on a given filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t : t \geq 0\}, \mathbb{P})$. Thus the Stratonovich SPDE (5.2) becomes

$$du = \left[\Delta u - \frac{1}{\epsilon^2} f(u)\right] dt + \delta \nabla u \cdot X \circ dW(t), \tag{5.4}$$

and the corresponding Itô SPDE is given by

$$\begin{aligned} du &= \left[\Delta u - \frac{1}{\epsilon^2} f(u) + \frac{\delta^2}{2} \nabla(\nabla u \cdot X) \cdot X\right] dt + \delta \nabla u \cdot X \, dW(t) \\ &= \left[\Delta u - \frac{1}{\epsilon^2} f(u) + \frac{\delta^2}{2} (B : D^2 u + b \cdot \nabla u)\right] dt + \delta \nabla u \cdot X \, dW(t) \\ &= \left[\Delta u - \frac{1}{\epsilon^2} f(u) + \frac{\delta^2}{2} \big(\mathrm{div}(B \nabla u) + (b - \mathrm{div} B) \cdot \nabla u\big)\right] dt \\ &\quad + \delta \nabla u \cdot X \, dW(t), \end{aligned} \tag{5.5}$$

where $B = X \otimes X \in \mathbb{R}^{d \times d}$ with $B_{ij} = X_i X_j$, $b = (\nabla X) X \in \mathbb{R}^d$ with $b_j = (\partial_i X_j) X_i$ and $\mathrm{div} B - b = (\mathrm{div} X) X$. For convenience, here we suppress the summation notation for repeated indices. Note that the Itô SPDE (5.5) has two correction terms which is hidden in the Stratonovich SPDE (5.4).

Remark 5.1.1. *In general, we can also consider the following vector field [83]*

$$\mathbb{X}(x,t) = \sum_{k=1}^{K} X^k(x) W_k(t),$$

where $K < \infty$ and $W_k(t)$ are independent $\mathbb{R}$-valued standard Brownian motions. In this case, the corresponding B and b in the SPDE (5.5) are given by

$$B = \sum_{k=1}^{K} X^k \otimes X^k \qquad \text{and} \qquad b = \sum_{k=1}^{K} (\nabla X^k) X^k.$$

It is relatively straightforward to generalize the results in this chapter for this $\mathbb{X}$.

We consider the SPDE (5.5) with certain boundary and initial conditions (1.26)–(1.27) under the following assumptions:

$$u_0 \in C^{\infty}(\bar{\mathcal{D}}), \tag{5.6}$$

$$X \in C_0^{4,\beta_0}(\mathcal{D}), \tag{5.7}$$

for some $\beta_0 \in (0,1]$. According to the existence and regularity results in [83, Theorem 4.1], if the domain $\mathcal{D}$ is smooth and the assumptions (5.6)–(5.7) hold, there exists a unique strong solution $u(\cdot,t)$ such that

$$\begin{aligned} u(x,t) = u_0(x) + \int_0^t \Big[\Delta u(x,s) - \frac{1}{\epsilon^2} f(u(x,s)) + \frac{\delta^2}{2} \Big(B : D^2 u(x,s) \\ + b \cdot \nabla u(x,s) \Big) \Big] \, ds + \int_0^t \delta \nabla u(x,s) \cdot X(x) \, dW(s), \end{aligned} \tag{5.8}$$

$$\frac{\partial u}{\partial n} = 0 \qquad \text{on } \partial \mathcal{D}_T, \tag{5.9}$$

hold $\mathbb{P}$-almost surely, and $u(\cdot,t)$ is a continuous $C^{3,\beta}(\bar{\mathcal{D}})$-semimartingale for any $0 < \beta < \beta_0$. Furthermore, for any multi-index $|\sigma| \le 3$ and $p \ge 1$, there exists a positive constant $C_0 = C(p,\delta,\epsilon)$ such that the following bound for the p-th moment of the

spatial derivatives holds

$$\sup_{t\in[0,T]} \mathbb{E}\left[\sup_{x\in\mathcal{D}} |\partial^\sigma u(x,t)|^p\right] < C_0. \tag{5.10}$$

Note that there are also other solution concepts for the SPDE (5.5) such as mild solution (semigroup approach) [78] and variational solution [63] (variational approach). For the analysis, we assume the strong solution of the SPDE (5.5) exists and satisfies (5.8)–(5.10). Then it is straightforward to see that the solution $u(\cdot,t)$ is an adapted $H^1(\mathcal{D})$-valued process such that for any $t \in (0,T]$, the following holds $\mathbb{P}$-almost surely:

$$\begin{aligned}(u(t),\phi) = (u_0,\phi) &- \int_0^t \left(\left(I + \frac{\delta^2}{2}B\right)\nabla u(s), \nabla\phi\right) ds \\ &- \frac{1}{\epsilon^2}\int_0^t (f(u(s)),\phi)\, ds - \frac{\delta^2}{2}\int_0^t \left((\mathrm{div}B - b)\cdot\nabla u(s),\phi\right) ds \\ &+ \delta\int_0^t (\nabla u(s)\cdot X,\phi)\, dW(s) \qquad \forall\,\phi\in H^1(\mathcal{D}),\end{aligned} \tag{5.11}$$

where $(\cdot,\cdot)$ is the inner product on $L^2(\mathcal{D})$.

The primary goal of this chapter is to develop and analyze some fully discrete finite element methods for approximating the solution of the SPDE (5.5) with initial and boundary conditions (1.26)–(1.27). In order to derive the strong convergence with rates for finite element methods, we will assume $u(\cdot,t) \in W^{s,\infty}(\mathcal{D})$ $(s \geq 3)$ for any $t \in [0,T]$ and has the following regularity estimate:

$$\sup_{t\in[0,T]} \mathbb{E}\left[\|u\|^p_{W^{s,\infty}(\mathcal{D})}\right] \leq C_0 = C(p,\delta,\epsilon) \qquad \forall p \geq 1. \tag{5.12}$$

Note that the assumption (5.12) is reasonable in view of (5.10).

The SPDE (5.5) was proposed in [83] (cf. also [94]) in which the tightness of solutions in the sharp interface limit was shown. However, the rigorous justification

of the convergence of (5.5) to (5.1) is still missing. In [84, 51], the stochastic Allen-Cahn equation in one dimension with additive space-time white noise was studied and the sharp interface limit was obtained. Also finite element methods of this model was proposed in [66]. But in higher dimensions, it requires spatial correlations for the noise otherwise the space-time white noise is too rough which prevents the existence of solution in reasonable function spaces. In [41], finite element methods for the stochastic mean curvature flow of planar curves of graphs were investigated, where the SPDE is quasilinear and arises from the level set formulation of the mean curvature flow. However, the results in [41] only holds for $d = 1$. In [79], finite element methods for a stochastic Allen-Cahn equation with multiplicative noise were proposed, where the strong convergence with rates are obtained. Due to the limited regularity in space, the spatial error estimate in [79] is not optimal.

In this chapter, we consider fully discrete finite element approximations of the SPDE (5.5). Since we are interested in the case where ϵ is small, it is important to see how estimates depend on ϵ. It is well-known that the nonlinear term in the deterministic Allen-Cahn equation needs to be controlled properly in order to obtain a reasonable estimate [45, 40, 69]. However, it is unclear whether the techniques used in [45, 40, 69] can be extended to the stochastic case. Therefore we apply the standard div's inequality to bound the nonlinear term and thus the estimate has an exponential dependence on $\frac{1}{\epsilon}$. Moreover, the gradient-type noises in the SPDE (5.5) brings new challenges as it introduces additional diffusion and convection terms. The additional diffusion contribution in the SPDE (5.5) automatically controls the gradient-depended noise term so that the stochastic parabolicity condition always holds [50, 66]. In addition, we assume δ is not too large to prevent the convection-dominance in the SPDE (5.5).

The rest of the chapter is organized as follows. We present in Section 5.2 some technical lemmas which are crucial for the convergence analysis of finite element methods. In Section 5.3, two fully discrete finite element schemes are proposed. The main result of this section is the error estimate for the finite element approximations.

Finally, in Section 5.4 we present several computational results to illustrate the performances of the methods and to verify the theoretical error estimates obtained in Section 5.3.

5.2 Preliminary results

In this section, we present some lemmas that will be used in Section 5.3. Throughout the chapter we will use C to denote a generic positive constant independent of ϵ, δ, space and time step sizes h and τ, which can take different values at different occurrences.

We begin with the following uniform estimate for the expectation of the p-th moment of Cahn-Hilliard energy functional

$$J(v) := \int_{\mathcal{D}} \left(\frac{1}{2}|\nabla v|^2 + \frac{1}{\epsilon^2} F(v) \right) dx.$$

A similar result can be found in [83] without tracking parameters in the estimate. For convenience of the reader, we include a proof here.

Lemma 5.2.1. *Let $u(t)$ be the strong solution to (5.5). For any $p \geq 1$ we have*

$$\sup_{t\in[0,T]} \mathbb{E}\left[J(u(t))^p\right] + \mathbb{E}\left[\int_0^t pJ(u(s))^{p-1} \left\|w(s)\right\|^2_{L^2(\mathcal{D})} \, ds\right] \leq C_1, \tag{5.13}$$

where $w(\cdot) := -\Delta u(\cdot) + \frac{1}{\epsilon^2} F'(u(\cdot))$ and $C_1 = e^{C\delta^2 p^2 \|X\|_{C^2(\bar{\mathcal{D}})} T} \left[J(u_0)\right]^p$.

Proof. We have

$$\begin{aligned}
D(J(u)^p) &= pJ(u)^{p-1} \left(-\Delta u + \frac{1}{\epsilon^2} F'(u) \right), \\
D^2(J(u)^p) &= p(p-1)J(u)^{p-2} \left(-\Delta u + \frac{1}{\epsilon^2} F'(u) \right) \otimes \left(-\Delta u + \frac{1}{\epsilon^2} F'(u) \right) \\
&\quad + pJ(u)^{p-1} \left(-\Delta + \frac{1}{\epsilon^2} F''(u) \right),
\end{aligned}$$

and apply Itô's formula to the functional $\Phi(u(\cdot)) := J(u(\cdot))^p$ and integration by parts to obtain

$$
\begin{aligned}
J(u(t))^p &= J(u_0)^p - \int_0^t pJ(u(s))^{p-1} \left\|w(s)\right\|^2_{L^2(\mathcal{D})}\, ds \\
&+ \int_0^t pJ(u(s))^{p-1} \left(w(s), \frac{\delta^2}{2}\big(B : D^2u(s) + b\cdot\nabla u(s)\big)\right) ds + M_t \\
&+ \frac{\delta^2}{2}\int_0^t p(p-1)J(u(s))^{p-2} \left(w(s), \nabla u(s)\cdot X\right)^2 ds \\
&+ \frac{\delta^2}{2}\int_0^t pJ(u(s))^{p-1} \int_{\mathcal{D}} \left[|\nabla(\nabla u(s)\cdot X)|^2 + \frac{1}{\epsilon^2}F''(u(s))(\nabla u(s)\cdot X)^2\right] dx\, ds,
\end{aligned}
\tag{5.14}
$$

where M_t is the martingale given by

$$
M_t = \delta \int_0^t pJ(u(s))^{p-1} \left(w(s), \nabla u(s)\cdot X\right)\, dW(s). \tag{5.15}
$$

By integration by parts and a direct calculation, we have

$$
\begin{aligned}
&\frac{\delta^2}{2}\int_{\mathcal{D}} \left[|\nabla(\nabla u(s)\cdot X)|^2 + \frac{1}{\epsilon^2}F''(u(s))(\nabla\cdot X)^2\right] dx \\
&\qquad + \left(w(s), \frac{\delta^2}{2}\big(B : D^2u(s) + b\cdot\nabla u(s)\big)\right) \\
&= \frac{\delta^2}{2}\int_0^t \int_{\mathcal{D}} \left(G(x) : (\nabla u(s)\otimes\nabla u(s)) + \frac{1}{\epsilon^2}g(x)F(u(s))\right) dx\, ds,
\end{aligned}
\tag{5.16}
$$

where $G(\cdot) : \mathcal{D} \longrightarrow \mathbb{R}^{d\times d}$ and $g(\cdot) : \mathcal{D} \longrightarrow \mathbb{R}$ are defined by

$$
G_{ij} = [\partial_k(X_k\partial_l X_l)]\delta_{ij} + \partial_i[X_k\partial_k X_j] - 2\partial_k[X_k\partial_i X_j] + (\partial_k X_i)(\partial_k X_j), \tag{5.17}
$$

$$
g = \partial_k[X_k\partial_l X_l]. \tag{5.18}
$$

Taking expectation on both sides of (5.14), by (5.15) and the fact that $\mathbb{E}[M_t] = 0$, we have the energy equation

$$\begin{aligned}\mathbb{E}\left[J(u(t))^p\right] = J(u_0)^p - \mathbb{E}\left[\int_0^t pJ(u(s))^{p-1}\left\|w(s)\right\|^2_{L^2(\mathcal{D})}\,ds\right] \\ + \frac{\delta^2}{2}\mathbb{E}\Bigg[\int_0^t pJ(u(s))^{p-1}\int_{\mathcal{D}}\Big(G(x) : (\nabla u(s)\otimes\nabla u(s)) \\ + \frac{1}{\epsilon^2}g(x)F(u(s))\Big)\,dx\,ds\Bigg] \\ + \frac{\delta^2}{2}\mathbb{E}\left[\int_0^t p(p-1)J(u(s))^{p-2}\left(w(s),\nabla u(s)\cdot X\right)^2\,ds\right].\end{aligned} \tag{5.19}$$

Now it remains to estimate the third and fourth terms on the right-hand side of (5.19). Observe that

$$\|G\|_{C(\bar{\mathcal{D}})} + \|g\|_{C(\bar{\mathcal{D}})} \le C\|X\|^2_{C^2(\bar{\mathcal{D}})},$$

the third term can be estimated by

$$\begin{aligned}&\frac{\delta^2}{2}\mathbb{E}\left[\int_0^t pJ(u(s))^{p-1}\int_{\mathcal{D}}\left(G(x) : (\nabla u(s)\otimes\nabla u(s)) + \frac{1}{\epsilon^2}g(x)F(u(s))\right)\,dx\,ds\right] \\ &\quad\le Cp\delta^2\|X\|^2_{C^2(\bar{\mathcal{D}})}\int_0^t \mathbb{E}\left[J(u(s))^p\right]\,ds.\end{aligned} \tag{5.20}$$

For the fourth term, by integration by parts and the fact that

$$\left(-\Delta u + \frac{1}{\epsilon^2}F'(u)\right)\nabla u = -\nabla\cdot(\nabla u\otimes\nabla u) + \nabla\left(\frac{1}{2}|\nabla u|^2 + \frac{1}{\epsilon^2}F(u)\right),$$

we have

$$\begin{aligned}&\frac{\delta^2}{2}\mathbb{E}\left[\int_0^t p(p-1)J(u(s))^{p-2}\left(w(s),\nabla u(s)\cdot X\right)^2\,ds\right] \\ &\quad\le Cp(p-1)\delta^2\|X\|^2_{C^1(\bar{\mathcal{D}})}\int_0^t \mathbb{E}\left[J(u(s))^p\right]\,ds.\end{aligned} \tag{5.21}$$

Finally, combine (5.19)–(5.21) and apply the Gronwall's inequality, we obtain the estimate (5.13). □

Remark 5.2.2. *In the case of $p > 1$ in Theorem 5.2.1, the fourth term on the right-hand side of (5.19) can also be bounded by*

$$\begin{aligned} &\frac{\delta^2}{2}\mathbb{E}\left[\int_0^t p(p-1)J(u(s))^{p-2}\left(w(s),\nabla u(s)\cdot X\right)^2\,ds\right] \\ &\qquad \le (p-1)\delta^2\|X\|^2_{C(\bar{\mathcal{D}})}\mathbb{E}\left[\int_0^t pJ(u(s))^{p-1}\left\|w(s)\right\|^2_{L^2(\mathcal{D})}\,ds\right]. \end{aligned} \tag{5.22}$$

Therefore the estimate (5.13) is replaced by

$$\sup_{t\in[0,T]}\mathbb{E}\left[J(u(t))^p\right] + C_p\mathbb{E}\left[\int_0^t pJ(u(s))^{p-1}\left\|w(s)\right\|^2_{L^2(\mathcal{D})}\,ds\right] \le \tilde{C}_1, \tag{5.23}$$

where

$$C_p = 1-(p-1)\delta^2\|X\|^2_{C(\bar{\mathcal{D}})} \qquad \text{and} \qquad \tilde{C}_1 = e^{C\delta^2 p\|X\|_{C^2(\bar{\mathcal{D}})}T}\left[J(u_0)\right]^p.$$

However, (5.23) implicitly requires the noise intensity to be sufficiently small, i.e., $\delta \le 1/(\sqrt{p-1}\|X\|_{C(\bar{\mathcal{D}})})$.

Remark 5.2.3. *Note that in case of $p = 1$ and $\delta = 0$, we recover the deterministic energy law:*

$$\sup_{t\in[0,T]} J(u(t)) + \int_0^t \left\|w(s)\right\|_{L^2(\mathcal{D})}\,ds \le J(u_0).$$

Next we derive estimates of Hölder continuity (in time) for the solution $u(t)$ in L^2-norm and H^1-seminorm that will be key ingredients for the error analysis in Section 5.3. Clearly, the constants in these estimates depend on ϵ^{-1} in some polynomial orders.

Lemma 5.2.4. *Let $u(t)$ be the strong solution to (5.5). For any $s, t \in [0,T]$ with $s < t$ and $p \geq 2$, we have*

$$\begin{aligned}
&\mathbb{E}\left[\|u(t)-u(s)\|^p_{L^2(\mathcal{D})}\right] \\
&\quad + \frac{p}{2}\mathbb{E}\left[\int_s^t \|u(\zeta)-u(s)\|^{p-2}_{L^2(\mathcal{D})}\|\nabla(u(\zeta)-u(s))\|^2_{L^2(\mathcal{D})}\,d\zeta\right] \\
&\quad + \frac{\delta^2 p}{4}\mathbb{E}\left[\int_s^t \|u(\zeta)-u(s)\|^{p-2}_{L^2(\mathcal{D})}\|\nabla(u(\zeta)-u(s))\cdot X\|^2_{L^2(\mathcal{D})}\,d\zeta\right] \\
&\leq C_2(t-s),
\end{aligned} \tag{5.24}$$

where $C_2 = C(p,\delta,\|X\|_{C^2(\bar{\mathcal{D}})},T)\epsilon^{-p/2}\left([J(u_0)]^{p/2}+C\right)$.

Proof. We apply Itô's formula to the functional $\Phi(u(\cdot)) := \|u(\cdot)-u(s)\|^p_{L^2(\mathcal{D})}$ with fixed $s \in [0,T)$ and integration by parts to obtain

$$\begin{aligned}
&\|u(t)-u(s)\|^p_{L^2(\mathcal{D})} \\
&= p\int_s^t \|u(\zeta)-u(s)\|^{p-2}_{L^2(\mathcal{D})}\left(u(\zeta)-u(s), \Delta u(\zeta) - \frac{1}{\epsilon^2}f(u(\zeta))\right)\,d\zeta \\
&\quad + \frac{\delta^2 p}{2}\int_s^t \|u(\zeta)-u(s)\|^{p-2}_{L^2(\mathcal{D})}\left(u(\zeta)-u(s), B:D^2u(\zeta)+b\cdot\nabla u(\zeta)\right)\,d\zeta \\
&\quad + p\int_s^t \|u(\zeta)-u(s)\|^{p-2}_{L^2(\mathcal{D})}\left(u(\zeta)-u(s), \delta\nabla u(\zeta)\cdot X\right)\,dW(\zeta) \\
&\quad + \frac{\delta^2 p}{2}\int_s^t (p-2)\|u(\zeta)-u(s)\|^{p-4}_{L^2(\mathcal{D})}(u(\zeta)-u(s), \nabla u(\zeta)\cdot X)^2\,d\zeta \\
&\quad + \frac{\delta^2 p}{2}\int_s^t \|u(\zeta)-u(s)\|^{p-2}_{L^2(\mathcal{D})}\|\nabla u(\zeta)\cdot X\|^2_{L^2(\mathcal{D})}\,d\zeta \\
&= I + II + III + IV + V.
\end{aligned} \tag{5.25}$$

For simplicity, we assume $p > 2$ in the following proof since $p = 2$ case is easier to prove. By Young's inequality, we have

$$
\begin{aligned}
&\mathbb{E}\left[p\int_s^t \|u(\zeta)-u(s)\|_{L^2(\mathcal{D})}^{p-2}\left(u(\zeta)-u(s),\Delta u(\zeta)\right)\,d\zeta\right] \qquad (5.26)\\
&\quad=-\mathbb{E}\left[p\int_s^t \|u(\zeta)-u(s)\|_{L^2(\mathcal{D})}^{p-2}\|\nabla u(\zeta)-\nabla u(s)\|_{L^2(\mathcal{D})}^2\,d\zeta\right]\\
&\qquad-\mathbb{E}\left[p\int_s^t \|u(\zeta)-u(s)\|_{L^2(\mathcal{D})}^{p-2}\left(\nabla u(\zeta)-\nabla u(s),\nabla u(s)\right)\,d\zeta\right]\\
&\quad\le-\frac{p}{2}\mathbb{E}\left[\int_s^t \|u(\zeta)-u(s)\|_{L^2(\mathcal{D})}^{p-2}\|\nabla u(\zeta)-\nabla u(s)\|_{L^2(\mathcal{D})}^2\,d\zeta\right]\\
&\qquad+\frac{p}{2}\mathbb{E}\left[\int_s^t \|u(\zeta)-u(s)\|_{L^2(\mathcal{D})}^{p-2}\|\nabla u(s)\|_{L^2(\mathcal{D})}^2\,d\zeta\right]\\
&\quad\le-\frac{p}{2}\mathbb{E}\left[\int_s^t \|u(\zeta)-u(s)\|_{L^2(\mathcal{D})}^{p-2}\|\nabla u(\zeta)-\nabla u(s)\|_{L^2(\mathcal{D})}^2\,d\zeta\right]\\
&\qquad+\frac{p-2}{2}\int_s^t \mathbb{E}\left[\|u(\zeta)-u(s)\|_{L^2(\mathcal{D})}^{p}\right]\,d\zeta+\mathbb{E}\left[\|\nabla u(s)\|_{L^2(\mathcal{D})}^{p}\right](t-s).
\end{aligned}
$$

Similarly, we have

$$
\begin{aligned}
&\mathbb{E}\left[p\int_s^t \|u(\zeta)-u(s)\|_{L^2(\mathcal{D})}^{p-2}\left(u(\zeta)-u(s),-\frac{1}{\epsilon^2}f(u(\zeta))\right)\,d\zeta\right] \qquad (5.27)\\
&\le\mathbb{E}\Bigg[p\int_s^t \|u(\zeta)-u(s)\|_{L^2(\mathcal{D})}^{p-2}\\
&\qquad\times\left(\frac{1}{2\epsilon}\|u(\zeta)(u(\zeta)-u(s))\|_{L^2(\mathcal{D})}^2+\frac{2}{\epsilon}\left\|\frac{1}{\epsilon^2}F(u(\zeta))\right\|_{L^1(\mathcal{D})}\right)\,d\zeta\Bigg]\\
&\le 2(p-2)\int_s^t \mathbb{E}\left[\|u(\zeta)-u(s)\|_{L^2(\mathcal{D})}^{p}\right]\,d\zeta\\
&+C_p\epsilon^{-p/2}\left(\sup_{s\le\zeta\le t}\mathbb{E}\left[\|u(\zeta)\|_{L^4(\mathcal{D})}^{2p}\right]+\sup_{s\le\zeta\le t}\mathbb{E}\left[\left\|\frac{1}{\epsilon^2}F(u(\zeta))\right\|_{L^1(\mathcal{D})}^{p/2}\right]\right)(t-s).
\end{aligned}
$$

From (5.26)–(5.27), we can estimate the first term on the right-hand side of (5.25):

$$\begin{aligned}\mathbb{E}[I] \leq & -\frac{p}{2}\mathbb{E}\left[\int_s^t \|u(\zeta)-u(s)\|_{L^2(\mathcal{D})}^{p-2}\|\nabla u(\zeta)-\nabla u(s)\|_{L^2(\mathcal{D})}^2\,d\zeta\right] \\ & + \frac{5(p-2)}{2}\int_s^t \mathbb{E}\left[\|u(\zeta)-u(s)\|_{L^2(\mathcal{D})}^p\right]d\zeta + \mathbb{E}\left[\|\nabla u(s)\|_{L^2(\mathcal{D})}^p\right](t-s) \\ & + C_p\epsilon^{-p/2}\left(\sup_{s\leq\zeta\leq t}\mathbb{E}\left[\|u(\zeta)\|_{L^4(\mathcal{D})}^{2p}\right] + \sup_{s\leq\zeta\leq t}\mathbb{E}\left[\left\|\frac{1}{\epsilon^2}F(u(\zeta))\right\|_{L^1(\mathcal{D})}^{p/2}\right]\right)(t-s).\end{aligned} \tag{5.28}$$

The second term on the right-hand side of (5.25) can also be estimated by Young's inequality and integration by parts in the following

$$\begin{aligned}\mathbb{E}[II] \leq & -\frac{\delta^2 p}{4}\mathbb{E}\left[\int_s^t \|u(\zeta)-u(s)\|_{L^2(\mathcal{D})}^{p-2}\|\nabla u(\zeta)\cdot X-\nabla u(s)\cdot X\|_{L^2(\mathcal{D})}^2\,d\zeta\right] \\ & + \frac{\delta^2}{2}(p-2)\int_s^t \mathbb{E}\left[\|u(\zeta)-u(s)\|_{L^2(\mathcal{D})}^p\right]d\zeta \\ & + \frac{\delta^2}{2}(\|X\|_{C(\bar{\mathcal{D}})}^p + \|X\|_{C^1(\bar{\mathcal{D}})}^{2p})\sup_{s\leq\zeta\leq t}\mathbb{E}\left[\|\nabla u(\zeta)\|_{L^2(\mathcal{D})}^p\right](t-s).\end{aligned} \tag{5.29}$$

Next, we bound the fourth and fifth terms on the right-hand side of (5.25) by Young's inequality

$$\begin{aligned}\mathbb{E}\left[IV+V\right] &\leq \mathbb{E}\left[\frac{\delta^2}{2}p(p-1)\int_s^t \|u(\zeta)-u(s)\|_{L^2(\mathcal{D})}^{p-2}\|\nabla u(\zeta)\cdot X\|_{L^2(\mathcal{D})}^2\,d\zeta\right] \\ &\leq \frac{\delta^2}{2}(p-1)(p-2)\int_s^t \mathbb{E}\left[\|u(\zeta)-u(s)\|_{L^2(\mathcal{D})}^p\right]d\zeta \\ &\quad + (p-1)\delta^2\|X\|_{C(\bar{\mathcal{D}})}^p\sup_{s\leq\zeta\leq t}\mathbb{E}\left[\|\nabla u(\zeta)\|_{L^2(\mathcal{D})}^p\right](t-s).\end{aligned} \tag{5.30}$$

Combining (5.25), (5.28)–(5.30) and using the facts $\|u\|_{L^4(\mathcal{D})}^4 \le 8\epsilon^2 J(u) + C$ and $E[III] = 0$, we have

$$\begin{aligned}
&\mathbb{E}\left[\|u(t) - u(s)\|_{L^2(\mathcal{D})}^p\right] \\
&\quad + \frac{p}{2}\mathbb{E}\left[\int_s^t \|u(\zeta) - u(s)\|_{L^2(\mathcal{D})}^{p-2} \|\nabla(u(\zeta) - u(s))\|_{L^2(\mathcal{D})}^2 \, d\zeta\right] \\
&\quad + \frac{\delta^2 p}{4}\mathbb{E}\left[\int_s^t \|u(\zeta) - u(s)\|_{L^2(\mathcal{D})}^{p-2} \|\nabla(u(\zeta) - u(s)) \cdot X\|_{L^2(\mathcal{D})}^2 \, d\zeta\right] \\
&\le \frac{p-2}{2}(p\delta^2 + 5)\int_s^t \mathbb{E}\left[\|u(\zeta) - u(s)\|_{L^2(\mathcal{D})}^p\right] d\zeta \\
&\quad + C(p, \delta^2, \|X\|_{C^1(\bar{\mathcal{D}})})\epsilon^{-p/2}\left(\sup_{s\le\zeta\le t}\mathbb{E}\left[J(u(\zeta)^{p/2})\right] + C\right)(t-s).
\end{aligned} \tag{5.31}$$

Finally, the estimate (5.24) follows from (5.31), Gronwall's inequality and Lemma 5.2.1. □

Lemma 5.2.5. *Let $u(t)$ be the strong solution to (5.5). For any $s, t \in [0, T]$ with $s < t$, we have*

$$\mathbb{E}\left[\|\nabla(u(t) - u(s))\|_{L^2(\mathcal{D})}^2\right] + \frac{1}{2}\mathbb{E}\left[\int_s^t \|\Delta(u(\zeta) - u(s))\|_{L^2(\mathcal{D})}^2 \, d\zeta\right] \le C_3(t-s), \tag{5.32}$$

where

$$C_3 = C\left(\left(\delta^2\|X\|_{C^1(\bar{\mathcal{D}})}^2 + \delta^4\|X\|_{C^1(\bar{\mathcal{D}})}^4 + \frac{1}{\epsilon^2} + 1\right)\left(C_1 + \sup_{s\le\zeta\le t}\mathbb{E}\left[\|\Delta u(\zeta)\|_{L^2(\mathcal{D})}^2\right]\right)\right.$$

Proof. Applying Itô's formula to the functional $\Phi(u(\cdot)) := \|\nabla u(\cdot) - \nabla u(s)\|^2_{L^2(\mathcal{D})}$ with fixed $s \in [0, T)$ and integration by parts, we have

$$\begin{aligned}
\|\nabla u(t) - \nabla u(s)\|^2_{L^2(\mathcal{D})} &= -2\int_s^t (\Delta u(\zeta) - \Delta u(s), \Delta u(\zeta))\, d\zeta \\
&+ 2\int_s^t \left(\Delta u(\zeta) - \Delta u(s), \frac{1}{\epsilon^2} f(u(\zeta))\right) d\zeta \\
&- \delta^2 \int_s^t \left(\Delta u(\zeta) - \Delta u(s), B : D^2 u(\zeta) + b \cdot \nabla u(\zeta)\right) d\zeta \\
&- 2\delta \int_s^t (\Delta u(\zeta) - \Delta u(s), \nabla u(\zeta) \cdot X)\, dW(\zeta) \\
&+ \delta^2 \int_s^t \int_{\mathcal{D}} |\nabla(\nabla u(\zeta) \cdot X)|^2\, dx\, d\zeta.
\end{aligned} \tag{5.33}$$

The first term on the right-hand side of (5.33) can be estimated by Cauchy-Schwarz inequality as follows

$$\begin{aligned}
&- 2\mathbb{E}\left[\int_s^t (\Delta u(\zeta) - \Delta u(s), \Delta u(\zeta))\, d\zeta\right] \\
&= -2\mathbb{E}\left[\int_s^t \|\Delta u(\zeta) - \Delta u(s)\|^2_{L^2(\mathcal{D})}\, d\zeta\right] \\
&\quad - 2\mathbb{E}\left[\int_s^t (\Delta u(\zeta) - \Delta u(s), \Delta u(s))\, d\zeta\right] \\
&\leq -\mathbb{E}\left[\int_s^t \|\Delta u(\zeta) - \Delta u(s)\|^2_{L^2(\mathcal{D})}\, d\zeta\right] + \|\Delta u(s)\|^2_{L^2(\mathcal{D})}(t - s).
\end{aligned} \tag{5.34}$$

Using the Sobolev embedding $H^1(\mathcal{D}) \subset L^6(\mathcal{D})$ for $d \leq 3$, the second term on the right-hand side of (5.33) can be written as

$$\begin{aligned}
&2\mathbb{E}\left[\int_s^t \left(\Delta u(\zeta) - \Delta u(s), \frac{1}{\epsilon^2} f(u(\zeta))\right) d\zeta\right] \\
&\quad \leq \frac{2}{\epsilon^2}\mathbb{E}\left[\int_s^t \left(\|\Delta u(\zeta) - \Delta u(s)\|^2_{L^2(\mathcal{D})} + \|f(u(\zeta))\|^2_{L^2(\mathcal{D})}\right) d\zeta\right] \\
&\quad \leq \frac{C}{\epsilon^2}\left(\sup_{s\leq\zeta\leq t} \mathbb{E}\left[\|\Delta u(\zeta)\|^2_{L^2(\mathcal{D})}\right] + \sup_{s\leq\zeta\leq t} \mathbb{E}\left[J(u(\zeta))^3\right]\right)(t - s).
\end{aligned} \tag{5.35}$$

Next we bound the third and fifth terms on the right-hand side of (5.33) in the following:

$$\begin{aligned}
&\mathbb{E}\left[-\delta^2\int_s^t\Big(\Delta u(\zeta)-\Delta u(s), B:D^2u(\zeta)+b\cdot\nabla u(\zeta)\Big)\,d\zeta\right] \qquad (5.36)\\
&\quad\le\frac{1}{2}\mathbb{E}\left[\int_s^t\|\Delta u(\zeta)-\Delta u(s)\|^2_{L^2(\mathcal{D})}\,d\zeta\right]+\frac{1}{2}\delta^4\|X\|^4_{C^1(\bar{\mathcal{D}})}\\
&\qquad\times\left(\sup_{s\le\zeta\le t}\mathbb{E}\left[\|\Delta u(\zeta)\|^2_{L^2(\mathcal{D})}\right]+\sup_{s\le\zeta\le t}\mathbb{E}\left[\|\nabla u(\zeta)\|^2_{L^2(\mathcal{D})}\right]\right)(t-s),
\end{aligned}$$

and

$$\begin{aligned}
&\mathbb{E}\left[\delta^2\int_s^t\int_{\mathcal{D}}|\nabla(\nabla u(\zeta)\cdot X)|^2\,dx\,d\zeta\right]\le C\delta^2\|X\|^2_{C^1(\bar{\mathcal{D}})} \qquad (5.37)\\
&\qquad\times\Big(\sup_{s\le\zeta\le t}\mathbb{E}\left[\|\Delta u(\zeta)\|^2_{L^2(\mathcal{D})}\right]+\sup_{s\le\zeta\le t}\mathbb{E}\left[\|\nabla u(\zeta)\|^2_{L^2(\mathcal{D})}\right]\Big)(t-s).
\end{aligned}$$

The estimate (5.32) now follows from (5.33)–(5.37), Lemma 5.2.1 and the fact that the expectation of the fourth term on the right-hand side of (5.34) is 0. □

Finally, we include a Hölder continuity estimate for the nonlinear increment $f(u(t))-f(u(s)$ in the L^2 norm which will be useful to control the nonlinear term in the error analysis.

Lemma 5.2.6. *Let $u(t)$ be the strong solution to (5.5). For any $s,t\in[0,T]$ with $s<t$, we have*

$$\mathbb{E}\big[\|f(u(t))-f(u(s))\|^2_{L^2(\mathcal{D})}\big]\le C_4(t-s), \qquad (5.38)$$

where

$$C_4=C(\delta,\epsilon,\|X\|_{C^1(\bar{\mathcal{D}})},C_0,C_1). \qquad (5.39)$$

Proof. Consider the functional $\Phi(u(t)):=\|f(u(t))-f(u(s))\|^2_{L^2(\mathcal{D})}$ with fixed $s\in[0,T)$, and it's not hard to prove that its first order and second order Frechet

derivatives are

$$\Phi'(u(t))(v(t)) = 2\int_{\mathcal{D}}(f(u(t)) - f(u(s)))(3(u(t))^2 - 1)v(t)dx, \tag{5.40}$$

$$\begin{aligned}\Phi''(u(t))(v(t), m(t)) = 12\int_{\mathcal{D}}(f(u(t)) - f(u(s)))u(t)v(t)m(t)dx \\ + 2\int_{\mathcal{D}}(3(u(t))^2 - 1)^2 v(t)m(t)dx.\end{aligned} \tag{5.41}$$

Apply Itô's formula to the functional $\Phi(u(t)) := \|f(u(t)) - f(u(s))\|^2_{L^2(\mathcal{D})}$, we get

$$\begin{aligned}\|f(u(t)) - f(u(s))\|^2_{L^2(\mathcal{D})} &= 2\int_s^t\int_{\mathcal{D}}(f(u(\eta)) - f(u(s)))(3(u(\eta))^2 - 1) \\ &\times \left[\Delta u(\eta) - \frac{1}{\epsilon^2}f(u(\eta)) + \frac{\delta^2}{2}\mathrm{div}\,(B\nabla u(\eta)) + \frac{\delta^2}{2}(b - \mathrm{div}\,B)\cdot\nabla u(\eta)\right]dxd\eta \\ &+ 2\int_s^t\int_{\mathcal{D}}(f(u(\eta)) - f(u(s)))(3u(\eta)^2 - 1)\delta\nabla u(\eta)\cdot X dx dW_\eta \\ &+ 6\int_s^t\int_{\mathcal{D}}(f(u(\eta)) - f(u(s)))u(\eta)|\delta\nabla u(\eta)\cdot X|^2 dxd\eta \\ &+ \int_s^t\int_{\mathcal{D}}(3u(\eta)^2 - 1)^2|\delta\nabla u(\eta)\cdot X|^2 dxd\eta.\end{aligned} \tag{5.42}$$

Using integration by part and Young's inequality, we have

$$\begin{aligned}\|f(u(t)) - f(u(s))\|^2_{L^2(\mathcal{D})} &\le C\int_s^t(\|u(s)\|^4_{L^\infty(\mathcal{D})} + \|u(\eta)\|^4_{L^\infty(\mathcal{D})})J(u(\eta)) \\ &+ C\|X\|^2_{C^1(\mathcal{D})}\int_s^t(\|u(s)\|^4_{L^\infty(\mathcal{D})} + \|u(\eta)\|^4_{L^\infty(\mathcal{D})})(J(u(s)) + J(u(\eta))) \\ &+ C(t-s) + 2\int_s^t\int_{\mathcal{D}}(f(u(\eta)) - f(u(s)))(3u(\eta)^2 - 1)\delta\nabla u(\eta)\cdot X dx dW_\eta.\end{aligned} \tag{5.43}$$

Taking the expectation on both sides, using Lemma 5.2.1, and using Young's inequality again, we get

$$\|f(u(t)) - f(u(s))\|^2_{L^2(\mathcal{D})} \le C_4(t-s). \tag{5.44}$$

Here

$$C_4 = C\bigg(\delta, \epsilon, \sup_{s\le\eta\le t}\mathbb{E}\left[\|u(\eta)\|^6_{L^\infty(\mathcal{D})}\right], \sup_{s\le\eta\le t}\mathbb{E}\left[J(u(\eta))^3\right], \|X\|_{C^1(\bar{\mathcal{D}})}\bigg), \tag{5.45}$$

which can be written as (5.39) by regularity result (5.10) and energy law (5.13). □

5.3 Finite element methods

In this section we propose two fully discrete finite element schemes to solve (5.5) and derive optimal order error estimates for both finite element methods.

Let $t_n = n\tau$ $(n = 0, 1, \ldots, N)$ be a uniform partition of $[0, T]$ with $\tau = T/N$ and $\mathcal{T}_h$ be a quasi-uniform triangulation of $\mathcal{D}$. We consider the finite element space

$$V_h^r := \{v_h \in H^1(\mathcal{D}) : v_h|_K \in \mathcal{P}_r(K) \quad \forall\, K \in \mathcal{T}_h\},$$

where $\mathcal{P}_r(K)$ denotes the space of all polynomials of degree not exceeding a given positive integer r on $K \in \mathcal{T}_h$. The fully discrete finite element method for SPDE (5.5) is defined by seeking an $\mathcal{F}_{t_n}$-adapted V_h^r-valued process $\{u_h^n\}$ $(n = 0, 1, \ldots, N)$ such that $\mathbb{P}$-almost surely

$$\begin{aligned}(u_h^{n+1}, v_h) + \tau\left(\big(I + \frac{\delta^2}{2}B\big)\nabla u_h^{n+1}, \nabla v_h\right) + \tau\frac{1}{\epsilon^2}(f^{n+1}, v_h) \\ = (u_h^n, v_h) - \tau\frac{\delta^2}{2}\left((\mathrm{div}B - b)\cdot\nabla u_h^n, v_h\right) \\ + \delta(\nabla u_h^n \cdot X, v_h)\,\Delta W_{n+1} \qquad \forall\, v_h \in V_h^r,\end{aligned} \tag{5.46}$$

where $\Delta W_{n+1} := W(t_{n+1}) - W(t_n) \sim \mathcal{N}(0, \tau)$ and

$$f^{n+1} := (u_h^{n+1})^3 - u_h^{n+1} \qquad \text{or} \qquad f^{n+1} := (u_h^{n+1})^3 - u_h^n. \tag{5.47}$$

Note that in the deterministic case ($\delta = 0$), (5.47) corresponds to a fully implicit scheme or a convex-splitting scheme. We choose $u_h(0) = P_h u_0$ to complement (5.46) where $P_h : L^2(\mathcal{D}) \longrightarrow V_h^r$ is the L^2-projection operator defined by

$$(P_h w, v_h) = (w, v_h) \qquad v_h \in V_h^r.$$

The following error estimate results are well-known [13, 23]:

$$\|w - P_h w\|_{L^2(\mathcal{D})} + h\|\nabla(w - P_h w)\|_{L^2(\mathcal{D})} \leq Ch^{\min(r+1,s)}\|w\|_{H^s(\mathcal{D})}, \tag{5.48}$$

$$\|w - P_h w\|_{L^\infty(\mathcal{D})} \leq Ch^{2-\frac{d}{2}}\|w\|_{H^2(\mathcal{D})}, . \tag{5.49}$$

for any $w \in H^s(\mathcal{D})$.

We consider a convex decomposition $F(v) = F^+(v) - F^-(v)$ where

$$F^+(v) = \frac{1}{4}(v^4 + 1) \qquad \text{and} \qquad F^-(v) = \frac{1}{2}v^2. \tag{5.50}$$

Now define

$$\begin{aligned} G(v) := {} & \frac{1}{2}(v, v) + \frac{\tau}{2}(\nabla v, \nabla v) + \frac{\tau\delta^2}{2}(\nabla v \cdot X, \nabla v \cdot X) + \frac{\tau}{\epsilon^2}(F(v), 1) \\ & - (u_h^n, v) + \frac{\tau\delta^2}{2}\left((\mathrm{div}B - b) \cdot \nabla u_h^n, v\right) - \delta((\nabla u_h^n \cdot X)\Delta W_{n+1}, v), \end{aligned} \tag{5.51}$$

$$\begin{aligned} H(v) := {} & \frac{1}{2}(v, v) + \frac{\tau}{2}(\nabla v, \nabla v) + \frac{\tau\delta^2}{2}(\nabla v \cdot X, \nabla v \cdot X) + \frac{\tau}{\epsilon^2}(F^+(v), 1) \\ & - \left(1 + \frac{\tau}{\epsilon^2}\right)(u_h^n, v) + \frac{\tau\delta^2}{2}\left((\mathrm{div}B - b) \cdot \nabla u_h^n, v\right) - \delta((\nabla u_h^n \cdot X)\Delta W_{n+1}, v). \end{aligned} \tag{5.52}$$

It is clear that $H(v)$ is strictly convex for all $h, \tau > 0$ and $G(v)$ is strictly convex when $\tau \leq \epsilon^2$. Then a straightforward calculation implies that the discrete problem (5.46)

is equivalent to the following finite-dimensional convex minimization problems:

$$u_h^{n+1} = \underset{v_h \in V_h^r}{\operatorname{argmin}}\, G(v_h), \qquad \text{when} \quad f^{n+1} = (u_h^{n+1})^3 - u_h^{n+1} \quad \text{and} \quad \tau \le \epsilon^2 \tag{5.53}$$

$$u_h^{n+1} = \underset{v_h \in V_h^r}{\operatorname{argmin}}\, H(v_h), \qquad \text{when} \quad f^{n+1} = (u_h^{n+1})^3 - u_h^{n}. \tag{5.54}$$

Therefore, the existence and uniqueness of the solution to (5.46) is guaranteed for all $h, \tau > 0$ when $f^{n+1} = (u_h^{n+1})^3 - u_h^n$ and for all $\tau \le \epsilon^2$ when $f^{n+1} = (u_h^{n+1})^3 - u_h^{n+1}$.

Remark 5.3.1. *We can also consider a modified scheme*

$$\begin{aligned}(u_h^{n+1}, v_h) + \tau \left(\left(I + \frac{\delta^2}{2} B\right) \nabla u_h^{n+1}, \nabla v_h \right) + \tau \frac{1}{\epsilon^2} (f^{n+1}, v_h) \\ + \tau \frac{\delta^2}{2} \left((\operatorname{div} B - b) \cdot \nabla u_h^{n+1}, v_h \right) \\ = (u_h^n, v_h) + \delta (\nabla u_h^n \cdot X, v_h)\, \Delta W_{n+1} \qquad \forall\, v_h \in V_h,\end{aligned} \tag{5.55}$$

where we replace the term $\tau \frac{\delta^2}{2} ((\operatorname{div} B - b) \cdot \nabla u_h^n, v_h)$ *in* (5.46) *by* $\tau \frac{\delta^2}{2} \big((\operatorname{div} B - b) \cdot \nabla u_h^{n+1}, v_h \big)$ *in* (5.55). *Clearly,* (5.46) *has a simpler form and the stiffness matrix for* (5.46) *is symmetric. On the other hand,* (5.55) *has one more advective contribution making the stiffness to be non-symmetric.*

Remark 5.3.2. *Due to the time discretization, it is unclear whether the discrete analog of energy bound in Lemma 5.2.1 is valid or not. However, the error estimate below does not require any discrete estimate.*

Let $e^n := u(t_n) - u_h^n$ $(n = 0, 1, 2, ..., N)$, we now derive an error estimate for e^n.

Theorem 5.3.3. *Let* u *and* $\{u_h^n\}_{n=1}^N$ *denote respectively the solutions of problems* (5.5) *and* (5.46), *and* $u(\cdot)$ *satisfies* (5.12). *In the case of* $f^{n+1} = (u_h^{n+1})^3 - u_h^{n+1}$, *under the following mesh constraint:*

$$\tau \le C \left(\epsilon^{-2} + \delta^4 \right)^{-1}, \tag{5.56}$$

we have

$$\begin{aligned}
&\sup_{0\le n\le N}\mathbb{E}\left[\|e^n\|^2_{L^2(\mathcal{D})}\right]+\mathbb{E}\left[\sum_{n=1}^{N}\tau\|\nabla e^n\|^2_{L^2(\mathcal{D})}\right] \qquad (5.57)\\
&\le T\Bigg[\Big(C_3(1+\delta^2+\delta^4)+\frac{C_4}{\epsilon^2}\Big)\tau\\
&\qquad +C_0(1+\delta^2+\delta^4)h^{2\min(r,s-1)}+\Big(C_0+\frac{C_0^2}{\epsilon^2}\Big)h^{2\min(r+1,s)}\Bigg]e^{C(\delta^4+\frac{1}{\epsilon^2})T}\\
&\le C(\delta,\epsilon,T,\|X\|,C_0,C_2,C_3,C_4)\big[C\tau+h^{2\min(r,s-1)}\big].
\end{aligned}$$

In the case of $f^{n+1}=(u_h^{n+1})^3-u_h^n$*, under the following mesh constraint:*

$$\tau\le C(1+\delta^4)^{-1}, \qquad (5.58)$$

we have

$$\begin{aligned}
&\sup_{0\le n\le N}\mathbb{E}\left[\|e^n\|^2_{L^2(\mathcal{D})}\right]+\mathbb{E}\left[\sum_{n=1}^{N}\tau\|\nabla e^n\|^2_{L^2(\mathcal{D})}\right] \qquad (5.59)\\
&\le Te^{C(1+\delta^4+\frac{1}{\epsilon^4})T}\Bigg[\Big(C_3(1+\delta^2+\delta^4)+\frac{C_2+C_4}{\epsilon^4}\Big)\tau\\
&\qquad +C_0(1+\delta^2+\delta^4)h^{2\min(r,s-1)}+\Big(C_0+\frac{C_0^2+C_0}{\epsilon^4}\Big)h^{2\min(r+1,s)}\Bigg]\\
&\le C(\delta,\epsilon,T,\|X\|,C_0,C_2,C_3,C_4)\big[C\tau+h^{2\min(r,s-1)}\big].
\end{aligned}$$

Proof. We write $e^n=\eta^n+\xi^n$ where

$$\eta^n:=u(t_n)-P_hu(t_n)\qquad\text{and}\qquad \xi^n:=P_hu(t_n)-u_h^n,\qquad n=0,1,2,...,N.$$

It follows from (5.11) that for all t_n ($n \geq 0$) there holds $\mathbb{P}$-almost surely

$$
\begin{aligned}
&(u(t_{n+1}), v_h) - (u(t_n), v_h) + \int_{t_n}^{t_{n+1}} (\nabla u(s), \nabla v_h)\, ds \\
&\quad + \frac{1}{\epsilon^2} \int_{t_n}^{t_{n+1}} (f(u(s)), v_h)\, ds + \frac{\delta^2}{2} \int_{t_n}^{t_{n+1}} (B \nabla u(s), \nabla v_h)\, ds \\
&\quad + \frac{\delta^2}{2} \int_{t_n}^{t_{n+1}} ((\mathrm{div} B - b) \cdot \nabla u(s), v_h)\ ds \\
&= \delta \int_{t_n}^{t_{n+1}} (\nabla u(s) \cdot X, v_h)\, dW(s) \qquad\qquad \forall\, v_h \in V_h.
\end{aligned} \tag{5.60}
$$

Subtracting (5.46) from (5.60) and using the decomposition of e^{n+1}, we obtain the following error equation:

$$
\begin{aligned}
(\xi^{n+1} - \xi^n, v_h) &= -(\eta^{n+1} - \eta^n, v_h) - \int_{t_n}^{t_{n+1}} (\nabla u(s) - \nabla u_h^{n+1}, \nabla v_h)\, ds \\
&\quad - \frac{1}{\epsilon^2} \int_{t_n}^{t_{n+1}} (f(u(s)) - f^{n+1}, v_h)\, ds \\
&\quad - \frac{\delta^2}{2} \int_{t_n}^{t_{n+1}} \left(B(\nabla u(s) - \nabla u_h^{n+1}), \nabla v_h\right)\ ds \\
&\quad - \frac{\delta^2}{2} \int_{t_n}^{t_{n+1}} ((\mathrm{div} B - b) \cdot (\nabla u(s) - \nabla u_h^n), v_h)\ ds \\
&\quad + \delta \int_{t_n}^{t_{n+1}} ((\nabla u(s) - \nabla u_h^n) \cdot X, v_h))\ dW(s) \qquad\qquad \forall\, v_h \in V_h^r.
\end{aligned} \tag{5.61}
$$

Testing (5.61) with $v_h = \xi^{n+1}(\omega)$, we have $\mathbb{P}$-almost surely

$$
\begin{aligned}
(\xi^{n+1} - \xi^n, \xi^{n+1}) = &-(\eta^{n+1} - \eta^n, \xi^{n+1}) - \int_{t_n}^{t_{n+1}} (\nabla u(s) - \nabla u_h^{n+1}, \nabla \xi^{n+1})\, ds \\
&- \frac{1}{\epsilon^2} \int_{t_n}^{t_{n+1}} (f(u(s)) - f^{n+1}, \xi^{n+1})\, ds \\
&- \frac{\delta^2}{2} \int_{t_n}^{t_{n+1}} \left(B(\nabla u(s) - \nabla u_h^{n+1}), \nabla \xi^{n+1}\right)\, ds \\
&- \frac{\delta^2}{2} \int_{t_n}^{t_{n+1}} \left((\mathrm{div} B - b) \cdot (\nabla u(s) - \nabla u_h^n), \xi^{n+1}\right)\, ds \\
&+ \delta \int_{t_n}^{t_{n+1}} \left((\nabla u(s) - \nabla u_h^n) \cdot X, \xi^{n+1}\right)\, dW(s), \\
:= &\, T_1 + T_2 + T_3 + T_4 + T_5 + T_6.
\end{aligned} \tag{5.62}
$$

It is easy to obtain from an elementary identity $(a-b)a = \frac{1}{2}(a^2 - b^2) + \frac{1}{2}(a-b)^2$ that

$$
\begin{aligned}
\mathbb{E}\left[(\xi^{n+1} - \xi^n, \xi^{n+1})\right] = &\frac{1}{2}\mathbb{E}\left[\|\xi^{n+1}\|_{L^2(\mathcal{D})}^2 - \|\xi^n\|_{L^2(\mathcal{D})}^2\right] \\
&+ \frac{1}{2}\mathbb{E}\left[\|\xi^{n+1} - \xi^n\|_{L^2(\mathcal{D})}^2\right].
\end{aligned} \tag{5.63}
$$

Next, we estimate the right-hand side of (5.62). First, since P_h is the L^2-projection operator, we have $\mathbb{E}[T_1] = 0$. For the second term on the right-hand side of (5.62),

we have by Lemma 5.2.5 and Young's inequality that

$$\begin{aligned}\mathbb{E}\left[T_2\right] &= -\mathbb{E}\left[\int_{t_n}^{t_{n+1}} (\nabla u(s) - \nabla u(t_{n+1}), \nabla \xi^{n+1})\, ds\right] \\ &\quad - \mathbb{E}\left[\int_{t_n}^{t_{n+1}} (\nabla \eta^{n+1} + \nabla \xi^{n+1}, \nabla \xi^{n+1})\, ds\right] \\ &\leq C\mathbb{E}\left[\int_{t_n}^{t_{n+1}} \|\nabla u(s) - \nabla u(t_{n+1})\|_{L^2(\mathcal{D})}^2\, ds\right] + \frac{3}{16}\mathbb{E}\left[\|\nabla \xi^{n+1}\|_{L^2(\mathcal{D})}^2\right]\tau \\ &\quad + \mathbb{E}\left[\|\nabla \eta^{n+1}\|_{L^2(\mathcal{D})}^2\right]\tau + \frac{1}{4}\mathbb{E}\left[\|\nabla \xi^{n+1}\|_{L^2(\mathcal{D})}^2\right]\tau \\ &\quad - \mathbb{E}\left[\|\nabla \xi^{n+1}\|_{L^2(\mathcal{D})}^2\right]\tau \\ &\leq C_3\tau^2 + \mathbb{E}\left[\|\nabla \eta^{n+1}\|_{L^2(\mathcal{D})}^2\right]\tau - \frac{9}{16}\mathbb{E}\left[\|\nabla \xi^{n+1}\|_{L^2(\mathcal{D})}^2\right]\tau.\end{aligned} \tag{5.64}$$

In order to estimate the third term on the right-hand side of (5.62), we write

$$\begin{aligned}-\left(f(u(s)) - f^{n+1}, \xi^{n+1}\right) &= -\left(f(u(s)) - f(u(t_{n+1})), \xi^{n+1}\right) \\ &\quad - \left(f(u(t_{n+1}) - f(P_h u(t_{n+1})), \xi^{n+1}\right) \\ &\quad - \left(f(P_h u(t_{n+1})) - f^{n+1}, \xi^{n+1}\right),\end{aligned} \tag{5.65}$$

By Lemma 5.2.6, we obtain

$$\begin{aligned}&-\mathbb{E}\left[\left(f(u(s)) - f(u(t_{n+1})), \xi^{n+1}\right)\right] \\ &\quad \leq \frac{1}{2C_\dagger}\mathbb{E}\left[\|f(u(s)) - f(u(t_{n+1}))\|_{L^2(\mathcal{D})}^2\right] + \frac{C_\dagger}{2}\mathbb{E}\left[\|\xi^{n+1}\|_{L^2(\mathcal{D})}^2\right] \\ &\quad \leq \frac{C_4}{2C_\dagger}\tau + \frac{C_\dagger}{2}\mathbb{E}\left[\|\xi^{n+1}\|_{L^2(\mathcal{D})}^2\right].\end{aligned} \tag{5.66}$$

Next, by (5.10) and (5.48)–(5.49), we have

$$
\begin{aligned}
&-\mathbb{E}\left[\left(f(u(t_{n+1})-f(P_hu(t_{n+1})),\xi^{n+1}\right)\right] \qquad (5.67)\\
&\quad=-\mathbb{E}\left[\left(\eta^{n+1}(u(t_{n+1})^2+u(t_{n+1})P_hu(t_{n+1})+P_hu(t_{n+1})^2-1),\xi^{n+1}\right)\right]\\
&\quad\le\frac{1}{4C_\dagger}\mathbb{E}\Bigg[\|u(t_{n+1})^2+u(t_{n+1})P_hu(t_{n+1})+P_hu(t_{n+1})^2-1\|^2_{L^\infty(\mathcal{D})}\\
&\qquad\times\|\eta^{n+1}\|^2_{L^2(\mathcal{D})}\Bigg]+C_\dagger\mathbb{E}\left[\|\xi^{n+1}\|^2_{L^2(\mathcal{D})}\right]\\
&\quad\le\frac{1}{4C_\dagger}\left(\mathbb{E}\left[\|u(t_{n+1})^2+u(t_{n+1})P_hu(t_{n+1})+P_hu(t_{n+1})^2-1\|^3_{L^\infty(\mathcal{D})}\right]\right)^{\frac{2}{3}}\\
&\qquad\times\left(\mathbb{E}\left[\|\eta^{n+1}\|^6_{L^2(\mathcal{D})}\right]\right)^{\frac{1}{3}}+C_\dagger\mathbb{E}\left[\|\xi^{n+1}\|^2_{L^2(\mathcal{D})}\right]\\
&\quad\le\frac{C}{4C_\dagger}\left(\mathbb{E}\left[\|P_hu(t_{n+1})\|^6_{L^\infty(\mathcal{D})}+\|u(t_{n+1})\|^6_{L^\infty(\mathcal{D})}+|D|^3\right]\right)^{\frac{2}{3}}\\
&\qquad\times\left(\mathbb{E}\left[\|\eta^{n+1}\|^6_{L^2(\mathcal{D})}\right]\right)^{\frac{1}{3}}+C_\dagger\mathbb{E}\left[\|\xi^{n+1}\|^2_{L^2(\mathcal{D})}\right]\\
&\quad\le\frac{C_0}{4C_\dagger}\left(\mathbb{E}\left[\|\eta^{n+1}\|^6_{L^2(\mathcal{D})}\right]\right)^{\frac{1}{3}}+C_\dagger\mathbb{E}\left[\|\xi^{n+1}\|^2_{L^2(\mathcal{D})}\right].
\end{aligned}
$$

When $f^{n+1}=(u_h^{n+1})^3-u_h^{n+1}$, the last term on the right-hand side of (5.65) can be bounded by the monotonicity property

$$
-\mathbb{E}\big[\left(f(P_hu(t_{n+1}))-f^{n+1},\xi^{n+1}\right)\big]\le\mathbb{E}\big[\|\xi^{n+1}\|^2_{L^2(\mathcal{D})}\big]. \qquad (5.68)
$$

When $f^{n+1}=(u_h^{n+1})^3-u_h^n$, we have an extra term $\left(u_h^{n+1}-u_h^n,\xi^{n+1}\right)$ adding to (5.67) and we can control it by Lemma 5.2.4 as follows

$$
\begin{aligned}
\mathbb{E}\left[\left(u_h^{n+1}-u_h^n,\xi^{n+1}\right)\right]&=-\mathbb{E}\left[\left((\eta^{n+1}-\eta^n)+(\xi^{n+1}-\xi^n),\xi^{n+1}\right)\right] \qquad (5.69)\\
&\quad+\mathbb{E}\left[\left(u(t_{n+1})-u(t_n),\xi^{n+1}\right)\right]\\
&\le\frac{1}{2C_\dagger}\mathbb{E}\left[\|\eta^{n+1}-\eta^n\|^2_{L^2(\mathcal{D})}\right]+C_\dagger\mathbb{E}\left[\|\xi^{n+1}\|^2_{L^2(\mathcal{D})}\right]+\frac{C_2}{C_\dagger}\tau\\
&\quad-\mathbb{E}\left[\|\xi^{n+1}\|^2_{L^2(\mathcal{D})}\right]+\frac{1}{4C_\dagger}\mathbb{E}\left[\|\xi^n\|^2_{L^2(\mathcal{D})}\right]+C_\dagger\mathbb{E}\left[\|\xi^{n+1}\|^2_{L^2(\mathcal{D})}\right].
\end{aligned}
$$

Combining (5.65)–(5.69), we choose $C_\dagger = \frac{1}{2}$ whenever $f^{n+1} = (u_h^{n+1})^3 - u_h^{n+1}$ to obtain

$$\mathbb{E}\left[T_3\right] \leq \frac{C_4}{\epsilon^2}\tau^2 + \frac{1}{\epsilon^2}\mathbb{E}\left[\|\xi^{n+1}\|_{L^2(\mathcal{D})}^2\right]\tau + \frac{C_0}{\epsilon^2}\left(\mathbb{E}\left[\|\eta^{n+1}\|_{L^2(\mathcal{D})}^6\right]\right)^{\frac{1}{3}}\tau, \tag{5.70}$$

and choose $C_\dagger = \epsilon^2$ whenever $f^{n+1} = (u_h^{n+1})^3 - u_h^n$ to obtain

$$\begin{aligned}\mathbb{E}\left[T_3\right] \leq{}& \frac{C_2 + C_4}{\epsilon^4}\tau^2 + \mathbb{E}\left[\|\xi^{n+1}\|_{L^2(\mathcal{D})}^2\right]\tau + \frac{1}{\epsilon^4}\mathbb{E}\left[\|\xi^{n}\|_{L^2(\mathcal{D})}^2\right]\tau \\ &+ \frac{C_0}{\epsilon^4}\left(\mathbb{E}\left[\|\eta^{n+1}\|_{L^2(\mathcal{D})}^6\right]\right)^{\frac{1}{3}}\tau + \frac{1}{\epsilon^4}\mathbb{E}\left[\|\eta^{n+1}\|_{L^2(\mathcal{D})}^2 + \|\eta^{n}\|_{L^2(\mathcal{D})}^2\right]\tau.\end{aligned} \tag{5.71}$$

Similar to the estimate of T_2, the fourth and fifth terms on the right-hand side of (5.62) can also be estimated by Lemma 5.2.5 and Young's inequality:

$$\begin{aligned}\mathbb{E}\left[T_4\right] ={}& -\frac{\delta^2}{2}\mathbb{E}\left[\int_{t_n}^{t_{n+1}}\left((\nabla u(s) - \nabla u(t_{n+1})\cdot X, \nabla\xi^{n+1}\cdot X\right)\,ds\right] \\ &- \frac{\delta^2}{2}\mathbb{E}\left[\int_{t_n}^{t_{n+1}}\left((\nabla\eta^{n+1} + \nabla\xi^{n+1})\cdot X, \nabla\xi^{n+1}\cdot X\right)\,ds\right] \\ \leq{}& \delta^4\|X\|_{C(\bar{\mathcal{D}})}^4 C_3\tau^2 + \frac{2}{16}\mathbb{E}\left[\|\nabla\xi^{n+1}\|_{L^2(\mathcal{D})}^2\right]\tau \\ &+ \delta^4\|X\|_{C(\bar{\mathcal{D}})}^4\mathbb{E}\left[\|\nabla\eta^{n+1}\|_{L^2(\mathcal{D})}^2\right]\tau - \frac{\delta^2}{2}\mathbb{E}\left[\|\nabla\xi^{n+1}\cdot X\|_{L^2(\mathcal{D})}^2\right]\tau,\end{aligned} \tag{5.72}$$

and

$$\begin{aligned}\mathbb{E}\left[T_5\right] ={}& -\frac{\delta^2}{2}\mathbb{E}\left[\int_{t_n}^{t_{n+1}}\left((\mathrm{div}B - b)\cdot(\nabla u(s) - \nabla u(t_n)), \xi^{n+1}\right)\,ds\right] \\ &- \frac{\delta^2}{2}\mathbb{E}\left[\int_{t_n}^{t_{n+1}}\left((\mathrm{div}B - b)\cdot(\nabla\eta^{n} + \nabla\xi^{n}), \xi^{n+1}\right)\,ds\right] \\ \leq{}& \|X\|_{C^1(\bar{\mathcal{D}})}^4 C_3\tau^2 + (1 + \|X\|_{C^1(\bar{\mathcal{D}})}^4)\delta^4\mathbb{E}\left[\|\xi^{n+1}\|_{L^2(\mathcal{D})}^2\right]\tau \\ &+ \|X\|_{C^1(\bar{\mathcal{D}})}^4\mathbb{E}\left[\|\nabla\eta^{n}\|_{L^2(\mathcal{D})}^2\right]\tau + \frac{1}{16}\mathbb{E}\left[\|\nabla\xi^{n}\|_{L^2(\mathcal{D})}^2\right]\tau.\end{aligned} \tag{5.73}$$

By the martingale property and Itô's isometry and Lemma 5.2.5, we have

$$\begin{aligned}
\mathbb{E}[T_6] &\le \frac{1}{2}\mathbb{E}\left[\|\xi^{n+1}-\xi^n\|^2_{L^2(\mathcal{D})}\right] \\
&\quad + \frac{\delta^2}{2}\mathbb{E}\left[\int_{t_n}^{t_{n+1}} \|(\nabla u(s)-\nabla u_h^n)\cdot X\|^2_{L^2(\mathcal{D})}\,ds\right] \\
&\le \frac{1}{2}\mathbb{E}\left[\|\xi^{n+1}-\xi^n\|^2_{L^2(\mathcal{D})}\right] \\
&\quad + \frac{\delta^2}{2}(1+C')\mathbb{E}\left[\int_{t_n}^{t_{n+1}} \|(\nabla u(s)-\nabla u(t_n))\cdot X\|^2_{L^2(\mathcal{D})}\,ds\right] \\
&\quad + \frac{\delta^2}{2}\mathbb{E}\left[\|(\nabla\eta^n+\nabla\xi^n)\cdot X\|^2_{L^2(\mathcal{D})}\right]\tau + \frac{\delta^2}{2C'}\mathbb{E}\left[\|(\nabla\eta^n+\nabla\xi^n)\cdot X\|^2_{L^2(\mathcal{D})}\right]\tau \\
&\le \frac{1}{2}\mathbb{E}\left[\|\xi^{n+1}-\xi^n\|^2_{L^2(\mathcal{D})}\right] + \frac{\delta^2}{2}(1+C')\|X\|^2_{C(\bar{\mathcal{D}})}C_3\tau^2 \\
&\quad + \delta^2\left(\frac{1+C''}{2}+\frac{1}{C'}\right)\|X\|^2_{C(\bar{\mathcal{D}})}\mathbb{E}\left[\|\nabla\eta^n\|^2_{L^2(\mathcal{D})}\right]\tau \\
&\quad + \frac{\delta^2}{2}\mathbb{E}\left[\|\nabla\xi^n\cdot X\|^2_{L^2(\mathcal{D})}\right]\tau + \delta^2\left(\frac{1}{2C''}+\frac{1}{C'}\right)\|X\|^2_{C(\bar{\mathcal{D}})}\mathbb{E}\left[\|\nabla\xi^n\|^2_{L^2(\mathcal{D})}\right]\tau.
\end{aligned} \tag{5.74}$$

Now taking $C' = 16\delta^2\|X\|^2_{C(\bar{\mathcal{D}})}$ and $C'' = 8\delta^2\|X\|^2_{C(\bar{\mathcal{D}})}$ in (5.74), we obtain an estimate for the last term on the right-hand side of (5.62):

$$\begin{aligned}
\mathbb{E}[T_6] &\le \frac{1}{2}\mathbb{E}\left[\|\xi^{n+1}-\xi^n\|^2_{L^2(\mathcal{D})}\right] + \frac{\delta^2}{2}(1+16\delta^2\|X\|^2_{C(\bar{\mathcal{D}})})\|X\|^2_{C(\bar{\mathcal{D}})}C_3\tau^2 \\
&\quad + \delta^2\left(\frac{1+8\delta^2\|X\|^2_{C(\bar{\mathcal{D}})}}{2}+\frac{1}{16\delta^2\|X\|^2_{C(\bar{\mathcal{D}})}}\right)\|X\|^2_{C(\bar{\mathcal{D}})}\mathbb{E}\left[\|\nabla\eta^n\|^2_{L^2(\mathcal{D})}\right]\tau \\
&\quad + \frac{\delta^2}{2}\mathbb{E}\left[\|\nabla\xi^n\cdot X\|^2_{L^2(\mathcal{D})}\right]\tau + \frac{2}{16}\mathbb{E}\left[\|\nabla\xi^n\|^2_{L^2(\mathcal{D})}\right]\tau.
\end{aligned} \tag{5.75}$$

Taking expectation on (5.62) and combining estimates (5.63)–(5.64), (5.70)/(5.71)–(5.73) and (5.75), summing over $n = 0, 1, 2, ..., l-1$ with $1 \le l \le N$, and using the

properties of L^2-projection and the regularity assumption (5.12), we obtain

$$\begin{aligned}
&\left[\frac{1}{2}-\left(\frac{1}{\epsilon^2}+\delta^4\right)\tau\right]\mathbb{E}\left[\|\xi^l\|^2_{L^2(\mathcal{D})}\right]+\frac{1}{4}\mathbb{E}\left[\tau\sum_{n=1}^{l}\|\nabla\xi^n\|^2_{L^2(\mathcal{D})}\right] \qquad (5.76)\\
&\quad\le\frac{1}{2}\mathbb{E}\left[\|\xi^0\|^2_{L^2(\mathcal{D})}\right]+\frac{\delta}{2}\tau\mathbb{E}\left[\|\nabla\xi^0\cdot X\|^2_{L^2(\mathcal{D})}\right]+\frac{3}{16}\tau\mathbb{E}\left[\|\nabla\xi^0\|^2_{L^2(\mathcal{D})}\right]\\
&\quad\quad+\left[C_3(1+\delta^2+\delta^4)+\frac{C_4}{\epsilon^2}\right]T\tau\\
&\quad\quad+C_0(1+\delta^2+\delta^4)Th^{2\min(r,s-1)}+\frac{C_0^2}{\epsilon^2}Th^{2\min(r+1,s)}\\
&\quad\quad+\left(\frac{1}{\epsilon^2}+\delta^4\right)\mathbb{E}\left[\tau\sum_{n=0}^{l-1}\|\xi^n\|^2_{L^2(\mathcal{D})}\right],
\end{aligned}$$

whenever $f^{n+1}=(u_h^{n+1})^3-u_h^{n+1}$, and

$$\begin{aligned}
&\left[\frac{1}{2}-\left(1+\delta^4\right)\tau\right]\mathbb{E}\left[\|\xi^l\|^2_{L^2(\mathcal{D})}\right]+\frac{1}{4}\mathbb{E}\left[\tau\sum_{n=1}^{l}\|\nabla\xi^n\|^2_{L^2(\mathcal{D})}\right] \qquad (5.77)\\
&\quad\le\frac{1}{2}\mathbb{E}\left[\|\xi^0\|^2_{L^2(\mathcal{D})}\right]+\frac{\delta}{2}\tau\mathbb{E}\left[\|\nabla\xi^0\cdot X\|^2_{L^2(\mathcal{D})}\right]+\frac{3}{16}\tau\mathbb{E}\left[\|\nabla\xi^0\|^2_{L^2(\mathcal{D})}\right]\\
&\quad\quad+\left(C_3(1+\delta^2+\delta^4)+\frac{C_2+C_4}{\epsilon^4}\right)T\tau\\
&\quad\quad+C_0(1+\delta^2+\delta^4)Th^{2\min(r,s-1)}+\frac{C_0^2+C_0}{\epsilon^4}Th^{2\min(r+1,s)}\\
&\quad\quad+\left(1+\frac{1}{\epsilon^4}+\delta^4\right)\mathbb{E}\left[\tau\sum_{n=0}^{l-1}\|\xi^n\|^2_{L^2(\mathcal{D})}\right],
\end{aligned}$$

whenever $f^{n+1}=(u_h^{n+1})^3-u_h^n$. Here we have not explicitly tracked the constant $\|X\|_{C^1(\bar{\mathcal{D}})}$ or $\|X\|_{C(\bar{\mathcal{D}})}$ and some other constants in (5.76) and (5.77).

Finally, estimates (5.57) and (5.59) follow from (5.76)–(5.77), the discrete Gronwall's inequality, the L^2-projection properties and the fact $\xi^0=0$ and the triangle inequality. □

Remark 5.3.4. *Error estimates in Theorem 5.3.3 remain unchanged if we consider the modified schemes (5.55). In fact, we only need to check the fifth term on the*

right-hand side of (5.62)*:*

$$\begin{aligned}\mathbb{E}\left[T_5\right] &= -\frac{\delta^2}{2}\mathbb{E}\left[\int_{t_n}^{t_{n+1}}\left((\operatorname{div} B - b)\cdot(\nabla u(s) - \nabla u(t_{n+1})), \xi^{n+1}\right) ds\right] \\ &\quad - \frac{\delta^2}{2}\mathbb{E}\left[\int_{t_n}^{t_{n+1}}\left((\operatorname{div} B - b)\cdot(\nabla \eta^{n+1} + \nabla \xi^{n+1}), \xi^{n+1}\right) ds\right] \\ &\le \|X\|^4_{C^1(\bar{\mathcal{D}})} C_3 \tau^2 + (1 + \|X\|^4_{C^1(\bar{\mathcal{D}})})\delta^4 \mathbb{E}\left[\|\xi^{n+1}\|^2_{L^2(\mathcal{D})}\right]\tau \\ &\quad + \|X\|^4_{C^1(\bar{\mathcal{D}})}\mathbb{E}\left[\|\nabla \eta^{n+1}\|^2_{L^2(\mathcal{D})}\right]\tau + \frac{1}{16}\mathbb{E}\left[\|\nabla \xi^{n+1}\|^2_{L^2(\mathcal{D})}\right]\tau.\end{aligned} \tag{5.78}$$

Hence the third term $\frac{3}{16}\tau\mathbb{E}\left[\|\nabla\xi^0\|^2_{L^2(\mathcal{D})}\right]$ *on the right-hand side of* (5.76)/(5.77) *is replaced by* $\frac{1}{8}\tau\mathbb{E}\left[\|\nabla\xi^0\|^2_{L^2(\mathcal{D})}\right]$.

Remark 5.3.5. *Spatial estimates* (5.57) *and* (5.59) *are optimal in the* H^1*-norm, but suboptimal in the* L^2*-norm. From the proof of Theorem 5.3.3, the suboptimal estimate in the* L^2*-norm is caused by gradient type of noise, i.e., the existence of* T_4*,* T_5 *and* T_6 *on the right-hand side of* (5.62)*. The proof in Theorem 5.3.3 relies on the strong p-th moment estimate* (5.12) *for the solution to* (5.5)*. Otherwise, we may lose some order of h if a weaker regularity result is used (cf. the derivation of* (5.67)*). Note also the estimate depends exponentially on* $1/\epsilon^2$ *(or* $1/\epsilon^4$*), which seems to be pessimistic. However, this is the case even in the deterministic case (i.e.,* $\delta = 0$*) unless the standard error estimate technique is replaced by a much involved nonstandard technique as used in [45]. We intend to address this issue in a future work.*

5.4 Numerical results

In this section we present some two-dimensional numerical experiments to gauge the performance of the proposed fully discrete finite element methods with $r = 1$, i.e., V_h is the linear finite element space. We also numerically examine the influence of noises on the dynamics of the numerical interfaces.

We consider the SPDE (5.5) on the square domain $\mathcal{D} = [-0.5, 0.5]^2$ with two different initial conditions, and in both tests we define $X(x) = \varphi(x)[x_1+x_2,\ x_1-x_2]^T$, where

$$\varphi(x) := \begin{cases} e^{-\frac{0.001}{0.09^2-|x|^2}}, & \text{if } |x| < 0.3, \\ 0, & \text{if } |x| \geq 0.3. \end{cases}$$

For both tests, we take the Brownian motion step to be 1×10^{-4} and examine the interplay of the geometric evolution and gradient type-noises for $M = 500$ Monte Carlo realizations. In this section, we denote the numerical solution by $u^{\delta,\epsilon,h,\tau}(\omega)$ where $\delta > 0$ denotes the noise intensity, $\epsilon > 0$ is the interface width, h is the spatial mesh size, τ is the time step size and ω is a particular sample.

Next, we give a brief description of the algorithm that we use to solve the discrete problem (5.46). Let $N_h = \dim V_h$ and $\{\psi_i\}_{i=1}^{N_h}$ be the nodal basis of V_h. Denote by $\mathbf{u^{n+1}}$ the coefficient vector corresponding to the discrete solution $u_h^{n+1} = \sum_{i=0}^{N_h} u_i^{n+1}\psi_i$ at time $t_{n+1} = (n+1)\tau$. Suppose $f^{n+1} = (u_h^{n+1})^3 - u_h^n$ (the other case is similar), (5.46) is then equivalent to

$$\begin{aligned} &\left[\mathbf{M} + \tau\left(\mathbf{A} + \frac{\delta^2}{2}\mathbf{A_X}\right)\right]\mathbf{u^{n+1}} + \frac{\tau}{\epsilon^2}\mathbf{N}(\mathbf{u^{n+1}}) \\ &\quad = \left(1 + \frac{\tau}{\epsilon^2}\right)\mathbf{M}\mathbf{u^n} - \frac{\tau\delta^2}{2}\mathbf{C_1}\mathbf{u^n} + \delta\mathbf{\Delta W_{n+1}}\mathbf{C_2}\mathbf{u^n}, \end{aligned}$$

where $\mathbf{M}, \mathbf{A}$ are the standard mass and stiffness matrices, respectively, $\mathbf{A_X}$ is the weighted stiffness matrix whose (i,j) component is $(\nabla\psi_j \cdot X, \nabla\psi_i \cdot X)$, $\mathbf{N}(\mathbf{u^{n+1}})$ is the nonlinear term, $(\mathbf{C_1})_{\mathbf{ij}} = ((\mathrm{div}\mathbf{B} - \mathbf{b}) \cdot \nabla\psi_\mathbf{j}, \psi_\mathbf{i})$, $(\mathbf{C_2})_{\mathbf{ij}} = (\nabla\psi_\mathbf{j}, \psi_\mathbf{i})$ and $\mathbf{W}$ is the discrete Brownian motion with $\Delta\mathbf{W_{n+1}} = \mathbf{W_{n+1}} - \mathbf{W_n}$. Since we want to generate as many samples as possible in order to recover the statistical properties, it will be expensive to use the Newton method to solve this nonlinear system. In fact, we can solve it by a cheaper fixed-point iteration so that in each iteration. Although in one simulation it converges slower than Newton's method, we can store the Cholesky

factorization

$$\mathbf{M} + \tau \left(\mathbf{A} + \frac{\delta^2}{2} \mathbf{A_X} \right) = \mathbf{L}\mathbf{L}',$$

and use it for every Monte Carlo realization and at every time step for fixed τ. From this point of view, the fixed-point iteration outperforms Newton's method, especially for large sample size M. In fact, we observed in our experiment that the fixed-point iteration is more efficient than Newton's method.

Test1. The initial condition u_0 is set to be

$$u_0(x) = \tanh\left(\frac{d(x)}{\sqrt{2}\epsilon}\right),$$

where $\tanh(x) = (e^x - e^{-x})/(e^x + e^{-x})$ and $d(x)$ represents the signed distance between the point x and the ellipse

$$\frac{x_1^2}{0.04} + \frac{x_2^2}{0.01} = 1.$$

In Table 5.1, we record the expected values of the $L^\infty([0,T], ||\cdot||_{L_2})$ -norm of the errors and rates of convergence of the time discretization for varying τ with the fixed parameters $\delta = 1$, $\epsilon = 0.1$, and $T = 0.016$. The numerical results confirm the theoretical result of Theorem 5.3.3.

Table 5.1: Computed time discretization errors and convergence rates.

	Expected values of error	Order of convergence
τ=0.008	0.09895	
τ=0.004	0.06557	0.5937
τ=0.002	0.04472	0.5521
τ=0.001	0.03136	0.5120

In Figure 5.1–5.3, we display some snapshots of the zero-level set of the averaged numerical solution

$$\bar{u}^{\delta,\epsilon,h,\tau} = \frac{1}{M} \sum_{i=1}^{M} u^{\delta,\epsilon,h,\tau}(\omega_i)$$

at several time points with $\epsilon = 0.01$, and three different noise intensity parameters $\delta = 0.1, 1, 10$. We observe that the shape of the zero-level set of the expected value of the numerical solution undergoes more changes as δ increases.

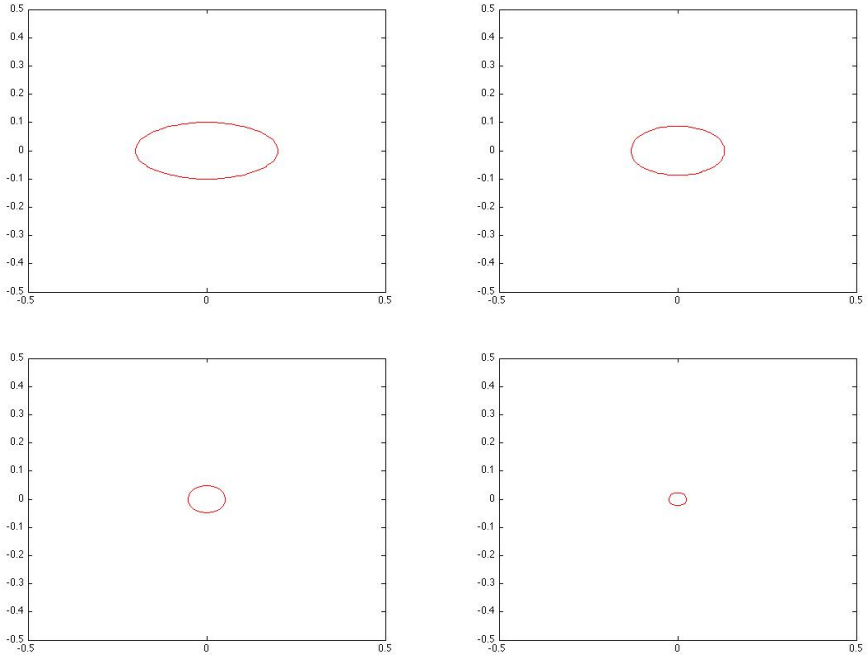

Figure 5.1: Snapshots of the zero-level set of $\bar{u}^{\delta,\epsilon,h,\tau}$ at time $t = 0, 0.020, 0.040, 0.043$ with $\delta = 0.1$ and $\epsilon = 0.01$.

Next, we study the effect of ϵ on the evolution process. For this aim, we fix $\delta = 0.1$. In Figure 5.4, we depict four snapshots at four fixed time points of the zero-level set of the averaged numerical solution $\bar{u}^{\delta,\epsilon,h,\tau}$ with three different $\epsilon = 0.01, 0.011, 0.02$. We observe that at each time step the zero-level set converges to the stochastic mean curvature flow as $\epsilon \to 0$, and furthermore it evolves faster in time for larger ϵ.

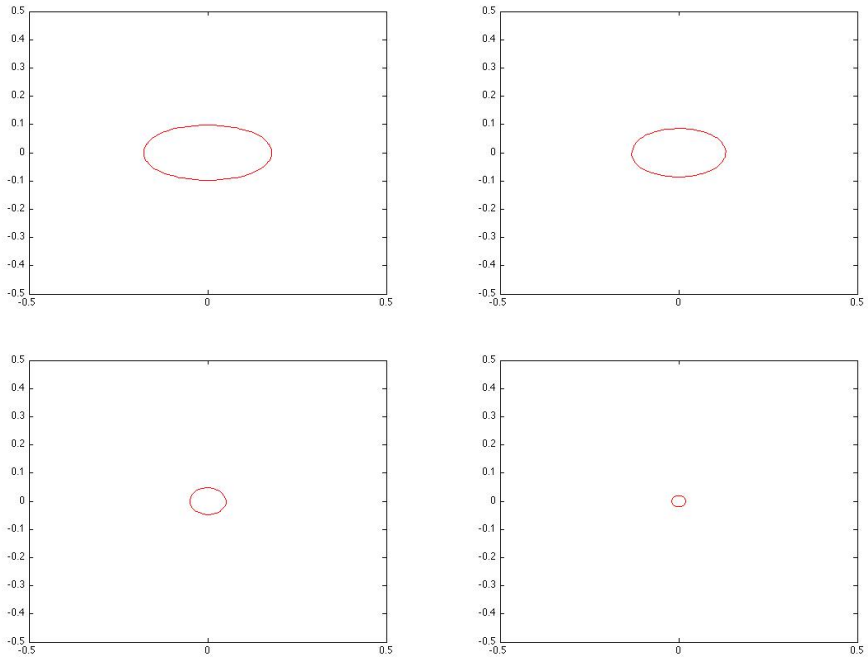

Figure 5.2: Snapshots of the zero-level set of $\bar{u}^{\delta,\epsilon,h,\tau}$ at time $t = 0.005, 0.020, 0.040, 0.043$ with $\delta = 1$ and $\epsilon = 0.01$.

Test2. First, we define

$$\tanh(y) := \frac{e^y - e^{-y}}{e^y + e^{-y}}, \qquad \psi_1(x_2) := \frac{-1 + \sqrt{0.8x_2 + 0.04}}{2},$$
$$\psi_2(x_2) := \frac{1 - \sqrt{1.92x_2 + 0.2304}}{2}, \qquad \psi_3(x_2) := \frac{-1 + \sqrt{-0.8x_2 + 0.04}}{2},$$
$$\psi_4(x_2) := \frac{1 - \sqrt{-1.92x_2 + 0.2304}}{2}, \qquad \psi_5(x_2) := -\sqrt{\frac{1 - 0.2451x_2^2}{0.0049}}.$$

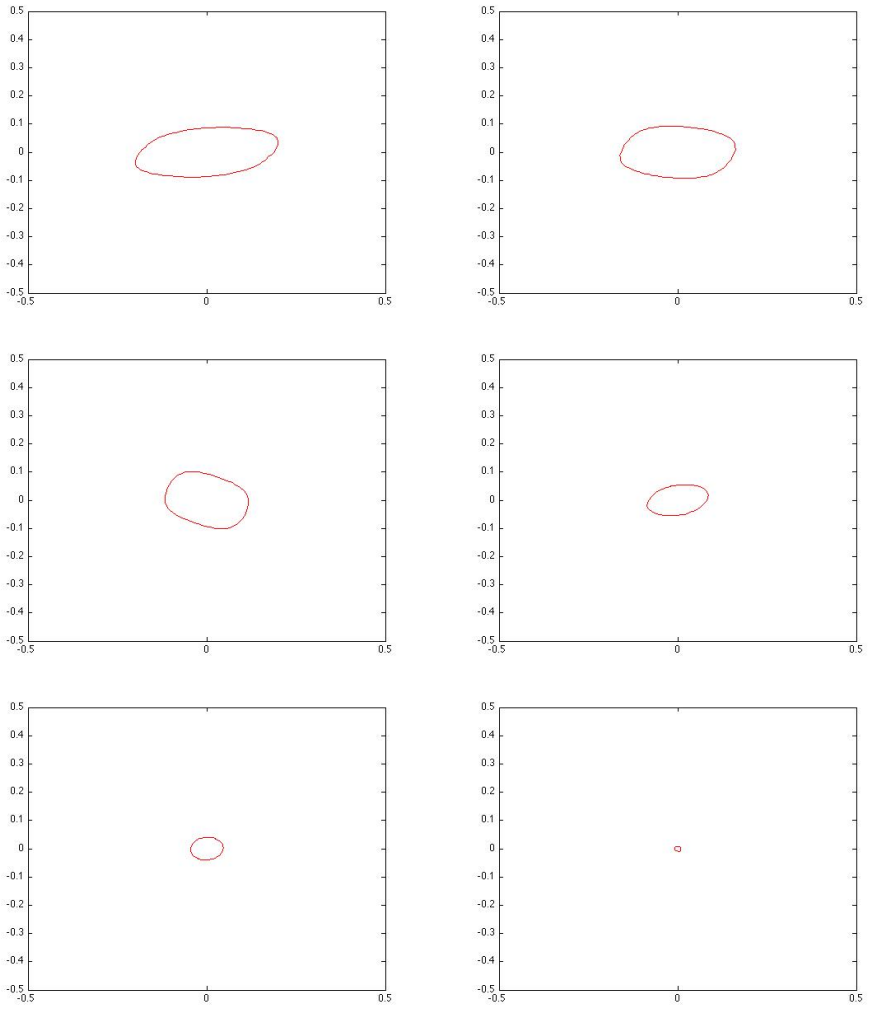

Figure 5.3: Snapshots of the zero-level set of $\bar{u}^{\delta,\epsilon,h,\tau}$ at time $t = 0.0025, 0.0050, 0.0100, 0.0200, 0.0250, 0.0280$ with $\delta = 10$ and $\epsilon = 0.01$.

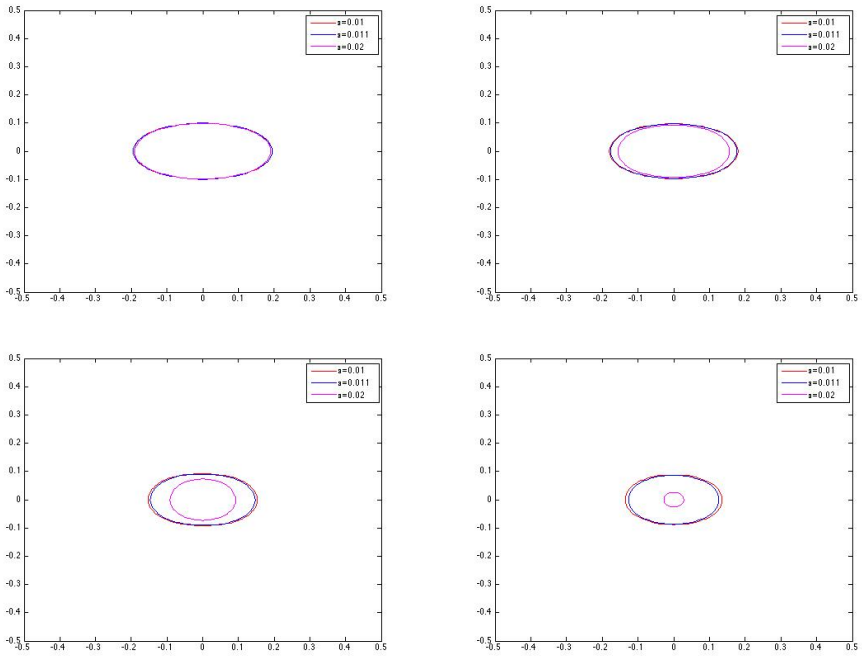

Figure 5.4: Snapshots of the zero-level set of $\bar{u}^{\delta,\epsilon,h,\tau}$ at time $t = 0.0010, 0.0050, 0.0125, 0.0175$ with $\delta = 0.1$ and $\epsilon = 0.01, 0.011, 0.02$.

and consider the initial condition $u_0(x_1, x_2) = u_1(3x_1, 3x_2)$, where $u_1(x_1, x_2)$ is defined by (cf. Test 2 in [40]):

$u_1(x_1, x_2)$

$$
= \begin{cases}
\tanh(\frac{1}{\sqrt{2}\epsilon}(-\sqrt{(x_1-0.14)^2+(x_2-0.15)^2})), & \text{if } x_1 > 0.14, 0 \le x_2 < -\frac{5}{12}(x_1-0.5), \\
\tanh(\frac{1}{\sqrt{2}\epsilon}(-\sqrt{(x_1-0.14)^2+(x_2+0.15)^2})), & \text{if } x_1 > 0.14, \frac{5}{12}(x_1-0.5) < x_2 < 0, \\
\tanh(\frac{1}{\sqrt{2}\epsilon}(-\sqrt{(x_1+0.3)^2+(x_2-0.15)^2})), & \text{if } x_1 < -0.3, 0 \le x_2 < \frac{3}{4}(x_1+0.5), \\
\tanh(\frac{1}{\sqrt{2}\epsilon}(-\sqrt{(x_1+0.3)^2+(x_2+0.15)^2})), & \text{if } x_1 < -0.3, -\frac{3}{4}(x_1+0.5) < x_2 < 0, \\
\tanh(\frac{1}{\sqrt{2}\epsilon}(\sqrt{(x_1-0.5)^2+x_2^2}-0.39)), & \text{if } x_1 > 0.14, x_2 \ge -\frac{5}{12}(x_1-0.5) \\
 & \text{or } x_2 \le \frac{5}{12}(x_1-0.5), \\
\tanh(\frac{1}{\sqrt{2}\epsilon}(\sqrt{(x_1+0.5)^2+x_2^2}-0.25)), & \text{if } x_1 < -0.3, x_2 \ge -\frac{3}{4}(x_1+0.5) \\
 & \text{or } x_2 \le -\frac{3}{4}(x_1+0.5), \\
\tanh(\frac{1}{\sqrt{2}\epsilon}(|x_2|-0.15)), & \text{if } -0.3 \le x_1 \le 0.14, \\
 & \psi_1(x_2) \le x_1 \le \psi_2(x_2) \\
 & \text{and } \psi_3(x_2) \le x_1 \le \psi_4(x_2), \\
\tanh(\frac{1}{\sqrt{2}\epsilon}(\sqrt{(x_1-0.5)^2+x_2^2}-0.39)), & \text{if } -0.3 \le x_1 \le 0.14,\ x_1 \ge \psi_2(x_2) \\
 & \text{and } x_1 \ge \psi_5(x_2), \\
\tanh(\frac{1}{\sqrt{2}\epsilon}(\sqrt{(x_1-0.5)^2+x_2^2}-0.39)), & \text{if } -0.3 \le x_1 \le 0.14,\ x_1 \ge \psi_4(x_2) \\
 & \text{and } x_1 \ge \psi_5(x_2), \\
\tanh(\frac{1}{\sqrt{2}\epsilon}(\sqrt{(x_1+0.5)^2+x_2^2}-0.25)), & \text{if } -0.3 \le x_1 \le 0.14,\ x_1 \le \psi_1(x_2) \\
 & \text{and } x_1 \le \psi_5(x_2), \\
\tanh(\frac{1}{\sqrt{2}\epsilon}(\sqrt{(x_1+0.5)^2+x_2^2}-0.25)), & \text{if } -0.3 \le x_1 \le 0.14,\ x_1 \le \psi_3(x_2) \\
 & \text{and } x_1 \le \psi_5(x_2).
\end{cases}
$$

Note that the initial condition is not smooth due to the dumbbell shape of the zero-level set. Nevertheless, we study the effects of δ and ϵ on the evolution process. Figure 5.5–5.7 display a a few snapshots of the zero-level set of the averaged numerical

solution $\bar{u}^{\delta,\epsilon,h,\tau}$ at several time steps with $\epsilon = 0.01$ and three different noise intensity $\delta = 0.1, 1, 10$. Similar to Test 1, the zero-level set evolves faster for larger δ and the shape is more irregular. Figure 5.8 plots snapshots at four fixed time steps of the zero-level set of $\bar{u}^{\delta,\epsilon,h,\tau}$ with $\delta = 0.1$ and $\epsilon = 0.01, 0.011, 0.02$. Again, it suggests the convergence of the zero-level set at each time step to the stochastic mean curvature flow as $\epsilon \to 0$.

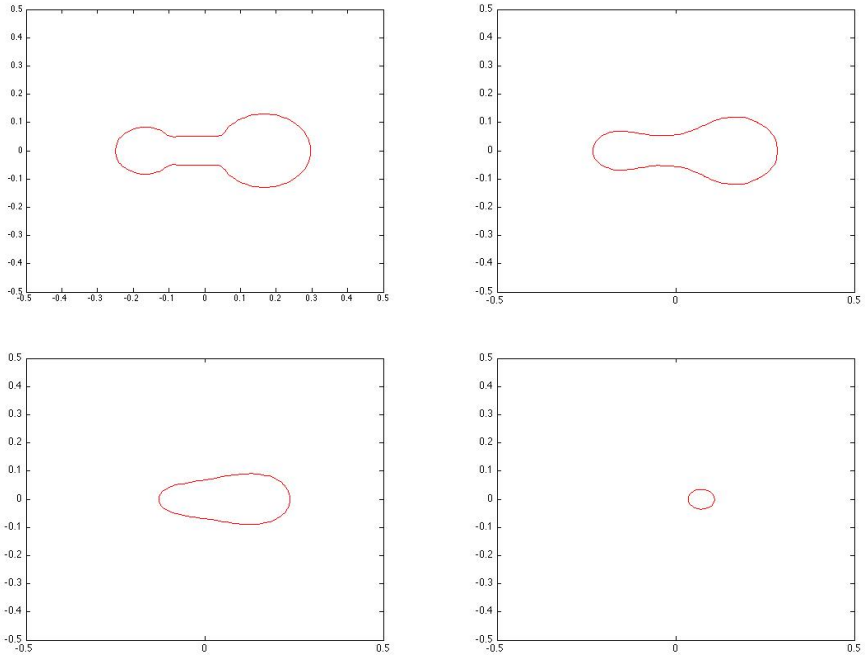

Figure 5.5: Snapshots of the zero-level set of $\bar{u}^{\delta,\epsilon,h,\tau}$ at time $t = 0, 0.040, 0.200, 0.456$ with $\delta = 0.1$ and $\epsilon = 0.01$.

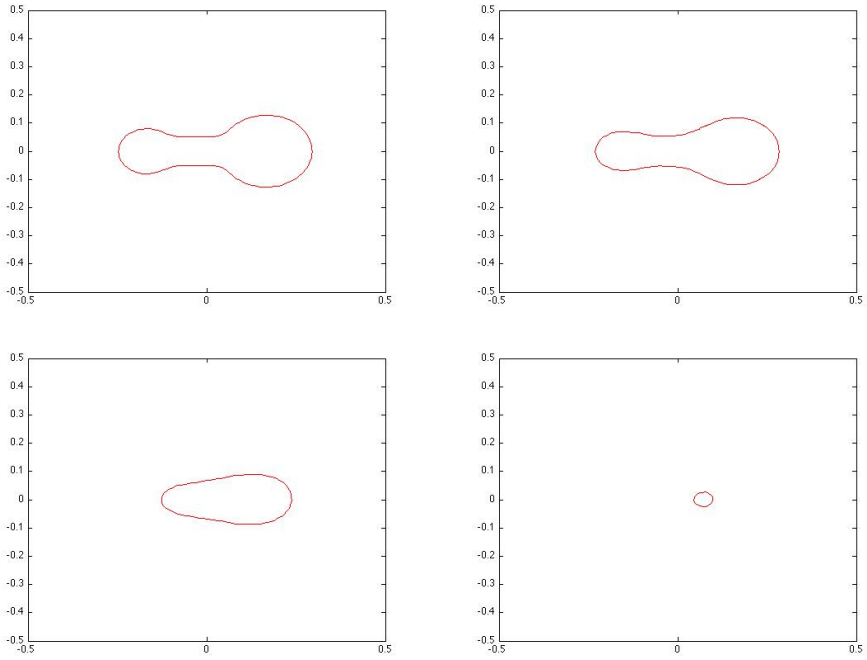

Figure 5.6: Snapshots of the zero-level set of $\bar{u}^{\delta,\epsilon,h,\tau}$ at time $t = 0.004, 0.040, 0.200, 0.456$ with $\delta = 1$ and $\epsilon = 0.01$.

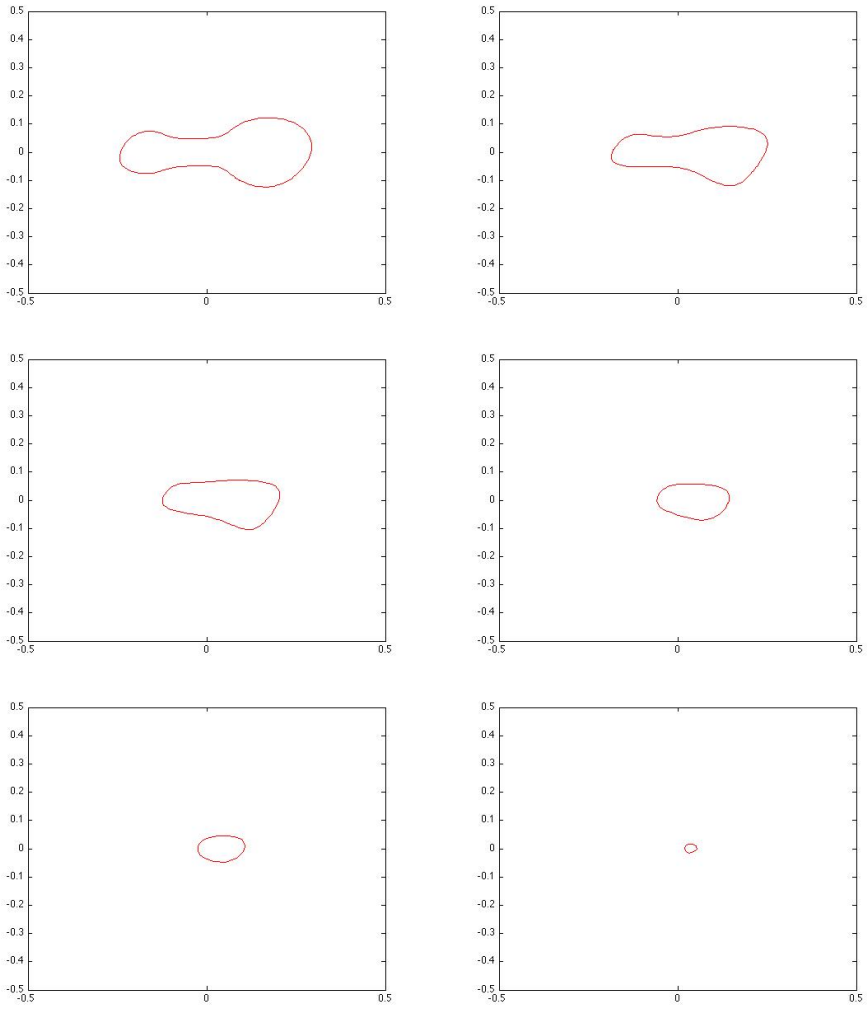

Figure 5.7: Snapshots of the zero-level set of $\bar{u}^{\delta,\epsilon,h,\tau}$ at time $t = 0.004, 0.040, 0.080, 0.140, 0.180, 0.216$ with $\delta = 10$ and $\epsilon = 0.01$.

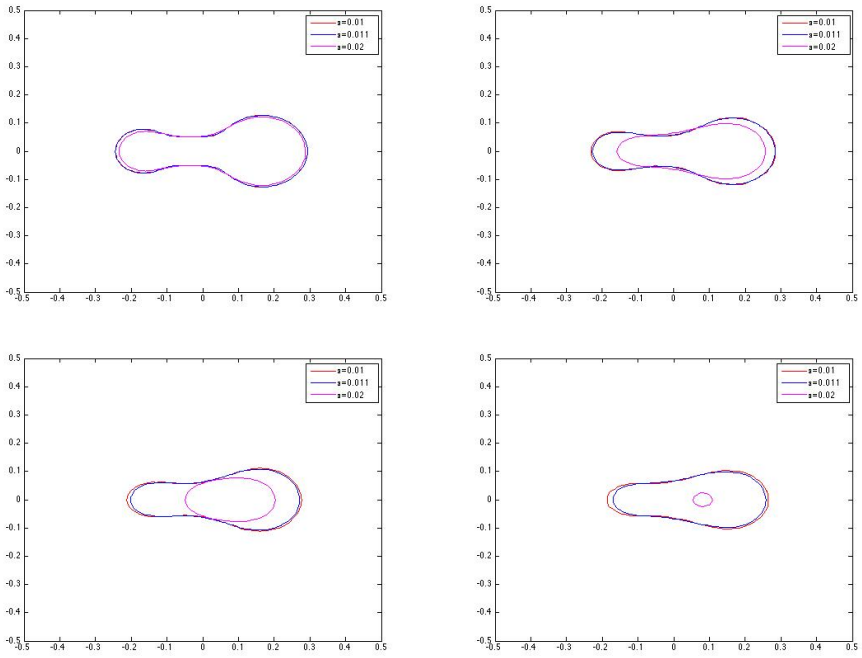

Figure 5.8: Snapshots of the zero-level set of $\bar{u}^{\delta,\epsilon,h,\tau}$ at time $t = 0.008, 0.040, 0.080, 0.120$.

Chapter 6

Multiphysics Finite Element Methods for a Quasi-static Poroelasticity Model

6.1 Introduction

This chapter considers a general quasi-static model of linear poroelasticity which is broad enough to contain the well-known Biot's consolidation model from soil mechanics (cf. [68, 86]) and the Doi's model for polymer gels (cf. [38, 93]). The quasi-static feature is due to the assumption that the acceleration of the solid (described by the second order time derivative of the displacement vector field) is assumed to be negligible. We refer the reader to [24, 75, 90] for a derivation of the model and to [85] for its mathematical analysis. When the parameter c_0, called the constrained specific storage coefficient, vanishes in the model, it reduces into the above mentioned Boit's model and Doi's model arising from two distinct applications. Their mathematical analysis can be found in [38] and their finite element numerical approximations based on two very different approaches were carried out in [68, 38], respectively. In [75, 76] the authors proposed and analyzed a semi-discrete and a fully discrete mixed finite

element method which simultaneously approximate the pressure and its gradient along with the displacement vector field. Since the implicit Euler scheme is used for the time discretization, a combined linear system must be solved at each time step. It is observed in the numerical tests that the proposed fully discrete mixed finite method may exhibit a "locking phenomenon" in the sense that the computed pressure oscillates and its accuracy deteriorate when a rapidly changed initial pressure is given, as explained in the [77] that such a "locking phenomenon" is caused by the difficulty of satisfying the nearly divergence-free condition of $\mathbf{u}$ for very small time $t > 0$.

The goal of this chapter is to present a multiphysics approach for approximating the poroelasticity model. A key idea [37] of this approach is to derive a multiphysics reformulation for the original model which clearly reveals the underlying multiple physics process (i.e., the deformation and diffusion) of the pore-scale fluid-solid interaction system. To the end, two pseudo-pressures are introduced, one of them is shown to satisfy a diffusion equation, while the displacement vector field along with the other pseudo-pressure variable is shown to satisfy a generalized Stokes system. It should be noted that the original pressure is eliminated in the reformulation, thus, it is not approximated as a primary (unknown) variable, instead, it is computed as a linear combination of the two pseudo-pressures. Based on this multiphysics reformulation we then propose a time-stepping algorithm which decouples (or couples) the reformulated PDE problem at each time step into two sub-problems, a generalized Stokes problem for the displacement vector field along with a pseudo-pressures and a diffusion problem for another pseudo-pressure field. To make this multiphysics approach feasible numerically, two critical issue must be resolved: the first one is the uniqueness of the generalized Stokes problem and the other is to find a good boundary condition for the diffusion equation so that it also becomes uniquely solvable.We also demonstrate that, regardless the choice of discretization methods, the proposed formulation has a built-in mechanism to overcome the "locking phenomenon" associated with numerical approximations of the poroelasticity model.

The remainder of this chapter is organized as follows. In Section 6.2 we present a complete PDE analysis of the poroelasticity model which emphasizes the energy law of the underlying model. Several conserved quantities are derived for the PDE solution. Moreover, it is proved that the poroelasticity model converges to the Biot's consolidation model as the constrained specific storage coefficient $c_0 \to 0$. In Section 6.3 we propose and analyze some fully discrete finite element methods based on the above mentioned multiphysics reformulation. Both coupled and decoupled time-stepping are considered and compared. The Taylor-Hood mixed finite element method combined with the P_1-conforming finite element method is chosen as an example for spatial discretization. It is proved that the solutions of the fully discrete finite element methods fulfill a discrete energy law which mimics the differential energy law satisfied by the PDE solution. Optimal order error estimates in the energy norm are also established. Finally, in Section 6.4, several benchmark numerical experiments are provided to show the performance of the proposed approach and methods, and to demonstrate the absence of "locking phenomenon" in our numerical experiments.

6.2 Partial differential equation model and its analysis

6.2.1 Preliminaries

$\mathcal{D} \subset \mathbb{R}^d\,(d = 1, 2, 3)$ denotes a bounded polygonal domain with the boundary $\partial\mathcal{D}$. The standard function space notation is adopted in this chapter, their precise definitions can be found in [13, 23, 89]. In particular, $(\cdot,\cdot)$ and $\langle\cdot,\cdot\rangle$ denote respectively the standard $L^2(\mathcal{D})$ and $L^2(\partial\mathcal{D})$ inner products. For any Banach space B, we let $\mathbf{B} = [B]^d$, and use $\mathbf{B}'$ to denote its dual space. In particular, we use $(\cdot,\cdot)_{\text{dual}}$ to denote the dual product on $(\mathbf{H}^1(\mathcal{D}))' \times \mathbf{H}^1(\mathcal{D})$, and $\|\cdot\|_{L^p(B)}$ is a shorthand notation for $\|\cdot\|_{L^p((0,T);B)}$.

We also introduce the function spaces

$$L_0^2(\mathcal{D}) := \{q \in L^2(\mathcal{D});\ (q,1) = 0\}, \qquad \mathbf{X} := \mathbf{H}^1(\mathcal{D}).$$

It is well known [89] that the following so-called inf-sup condition holds in the space $\mathbf{X} \times L_0^2(\mathcal{D})$:

$$\sup_{\mathbf{v}\in\mathbf{X}} \frac{(\operatorname{div}\mathbf{v}, \varphi)}{\| \nabla\mathbf{v} \|_{L^2(\mathcal{D})}} \geq \alpha_0 \| \varphi \|_{L^2(\mathcal{D})} \qquad \forall \varphi \in L_0^2(\mathcal{D}), \quad \alpha_0 > 0. \tag{6.1}$$

Let

$$\mathbf{RM} := \{\mathbf{r} := \mathbf{a} + \mathbf{b} \times x;\ \mathbf{a}, \mathbf{b}, x \in \mathbb{R}^d\}$$

denote the space of infinitesimal rigid motions. It is well known [13, 23, 89] that $\mathbf{RM}$ is the kernel of the strain operator ε, that is, $\mathbf{r} \in \mathbf{RM}$ if and only if $\varepsilon(\mathbf{r}) = 0$. Hence, we have

$$\varepsilon(\mathbf{r}) = 0, \quad \operatorname{div}\mathbf{r} = 0 \qquad \forall \mathbf{r} \in \mathbf{RM}. \tag{6.2}$$

Let $\mathbf{L}_\perp^2(\partial\mathcal{D})$ and $\mathbf{H}_\perp^1(\mathcal{D})$ denote respectively the subspaces of $\mathbf{L}^2(\partial\mathcal{D})$ and $\mathbf{H}^1(\mathcal{D})$ which are orthogonal to $\mathbf{RM}$, that is,

$$\mathbf{H}_\perp^1(\mathcal{D}) := \{\mathbf{v} \in \mathbf{H}^1(\mathcal{D});\ (\mathbf{v}, \mathbf{r}) = 0\ \forall \mathbf{r} \in \mathbf{RM}\},$$
$$\mathbf{L}_\perp^2(\partial\mathcal{D}) := \{\mathbf{g} \in \mathbf{L}^2(\partial\mathcal{D});\ \langle \mathbf{g}, \mathbf{r} \rangle = 0\ \forall \mathbf{r} \in \mathbf{RM}\}.$$

It is well known [26] that there exists a constant $c_1 > 0$ such that

$$\inf_{\mathbf{r}\in\mathbf{RM}} \|\mathbf{v} + \mathbf{r}\|_{L^2(\mathcal{D})} \leq c_1 \|\varepsilon(\mathbf{v})\|_{L^2(\mathcal{D})} \qquad \forall \mathbf{v} \in \mathbf{H}^1(\mathcal{D}).$$

Hence, for each $\mathbf{v} \in \mathbf{H}^1_\perp(\mathcal{D})$ there holds

$$\|\mathbf{v}\|_{L^2(\mathcal{D})} = \inf_{\mathbf{r}\in\mathbf{RM}} \|\mathbf{v}+\mathbf{r}\|_{L^2(\mathcal{D})} \le c_1 \|\varepsilon(\mathbf{v})\|_{L^2(\mathcal{D})}, \tag{6.3}$$

which and the well-known Korn's inequality [26] yield that for some $c_2 > 0$

$$\begin{aligned} \|\mathbf{v}\|_{H^1(\mathcal{D})} &\le c_2[\|\mathbf{v}\|_{L^2(\mathcal{D})} + \|\varepsilon(\mathbf{v})\|_{L^2(\mathcal{D})}] \\ &\le c_2(1+c_1)\|\varepsilon(\mathbf{v})\|_{L^2(\mathcal{D})} \qquad \forall \mathbf{v} \in \mathbf{H}^1_\perp(\mathcal{D}). \end{aligned} \tag{6.4}$$

By Lemma 2.1 of [12] we know that for any $q \in L^2(\mathcal{D})$, there exists $\mathbf{v} \in \mathbf{H}^1_\perp(\mathcal{D})$ such that $\operatorname{div}\mathbf{v} = q$ and $\|\mathbf{v}\|_{H^1(\mathcal{D})} \le C\|q\|_{L^2(\mathcal{D})}$. An immediate consequence of this lemma is that there holds the following alternative version of the inf-sup condition:

$$\sup_{\mathbf{v}\in\mathbf{H}^1_\perp(\mathcal{D})} \frac{(\operatorname{div}\mathbf{v}, \varphi)}{\|\nabla\mathbf{v}\|_{L^2(\mathcal{D})}} \ge \alpha_1 \|\varphi\|_{L^2(\mathcal{D})} \qquad \forall \varphi \in L^2_0(\mathcal{D}), \quad \alpha_1 > 0. \tag{6.5}$$

Throughout the chapter, we assume $\mathcal{D} \subset \mathbb{R}^d$ is a bounded polygonal domain such that $\Delta : H^1_0(\mathcal{D}) \cap H^2(\mathcal{D}) \to L^2(\mathcal{D})$ is an isomorphism (cf. [25, 54]). In addition, C is used to denote a generic positive (pure) constant which may be different in different places.

6.2.2 Partial differential equation model and its multiphysics reformulation

The quasi-static poroelasticity model to be studied in this chapter is given by (cf. [75])

$$-\operatorname{div}\sigma(\mathbf{u}) + \alpha\nabla p = \mathbf{f} \qquad \text{in } \mathcal{D}_T := \mathcal{D}\times(0,T) \subset \mathbb{R}^d \times (0,T), \tag{6.6}$$

$$(c_0 p + \alpha \operatorname{div}\mathbf{u})_t + \operatorname{div}\mathbf{v}_f = \phi \qquad \text{in } \mathcal{D}_T, \tag{6.7}$$

where

$$\sigma(\mathbf{u}) := \mu\varepsilon(\mathbf{u}) + \lambda \mathrm{div}\, \mathbf{u} I, \qquad \varepsilon(\mathbf{u}) := \frac{1}{2}\big(\nabla \mathbf{u} + \nabla \mathbf{u}^T\big), \tag{6.8}$$

$$\mathbf{v}_f := -\frac{K}{\mu_f}\big(\nabla p - \rho_f \mathbf{g}\big). \tag{6.9}$$

Where $\mathbf{u}$ denotes the displacement vector of the solid and p denotes the pressure of the solvent. $\mathbf{f}$ is the body force. I denotes the $d \times d$ identity matrix and $\varepsilon(\mathbf{u})$ is known as the strain tensor. The parameters in the model are Lamé constants λ and μ, the (symmetric) permeability tensor K, the solvent viscosity μ_f, Biot-Willis constant α, and the constrained specific storage coefficient c_0. In addition, $\sigma(\mathbf{u})$ is called the (effective) stress tensor. $\widehat{\sigma}(\mathbf{u}, p) := \sigma(\mathbf{u}) - \alpha p I$ is the total stress tensor. $\mathbf{v}_f$ is the volumetric solvent flux and (6.9) is the well-known Darcy's law. We assume that $\rho_f \not\equiv 0$, which is a realistic assumption.

To close the above system, suitable boundary and initial conditions must also be prescribed. The following set of boundary and initial conditions will be considered in this chapter:

$$\widehat{\sigma}(\mathbf{u}, p)\mathbf{n} = \sigma(\mathbf{u})\mathbf{n} - \alpha p \mathbf{n} = \mathbf{f}_1 \qquad \text{on } \partial\mathcal{D}_T := \partial\mathcal{D} \times (0, T), \tag{6.10}$$

$$\mathbf{v}_f \cdot \mathbf{n} = -\frac{K}{\mu_f}\big(\nabla p - \rho_f \mathbf{g}\big) \cdot \mathbf{n} = \phi_1 \qquad \text{on } \partial\mathcal{D}_T, \tag{6.11}$$

$$\mathbf{u} = \mathbf{u}_0, \qquad p = p_0 \qquad \text{in } \mathcal{D} \times \{t = 0\}. \tag{6.12}$$

We note that in some engineering literature the second Lamé constant μ is also called the *shear modulus* and denoted by G, and $B := \lambda + \frac{2}{3}G$ is called the *bulk modulus*. λ, μ and B are computed from the *Young's modulus* E and the *Poisson ratio* ν by the following formulas:

$$\lambda = \frac{E\nu}{(1+\nu)(1-2\nu)}, \qquad \mu = G = \frac{E}{2(1+\nu)}, \qquad B = \frac{E}{3(1-2\nu)}.$$

Unlike the existing approaches in the literature [75, 68], in this chapter we will not approximate the above original model directly, instead, we first derive a (multiphysics) reformulation for the model, we then approximate the reformulated model. This is a key idea of this chapter and it will be seen in the later sections that this new approach is advantageous. To the end, we introduce new variables

$$q := \operatorname{div} \mathbf{u}, \qquad \eta := c_0 p + \alpha q, \qquad \xi := \alpha p - \lambda q.$$

It is easy to check that

$$p = \kappa_1 \xi + \kappa_2 \eta, \qquad q = \kappa_1 \eta - \kappa_3 \xi, \tag{6.13}$$

where

$$\kappa_1 := \frac{\alpha}{\alpha^2 + \lambda c_0}, \quad \kappa_2 := \frac{\lambda}{\alpha^2 + \lambda c_0}, \quad \kappa_3 := \frac{c_0}{\alpha^2 + \lambda c_0}. \tag{6.14}$$

Then (6.6)–(6.9) can be written as

$$-\mu \operatorname{div} \varepsilon(\mathbf{u}) + \nabla \xi = \mathbf{f} \qquad \text{in } \mathcal{D}_T, \tag{6.15}$$

$$\kappa_3 \xi + \operatorname{div} \mathbf{u} = \kappa_1 \eta \qquad \text{in } \mathcal{D}_T, \tag{6.16}$$

$$\eta_t - \frac{1}{\mu_f} \operatorname{div} \left[K(\nabla(\kappa_1 \xi + \kappa_2 \eta) - \rho_f \mathbf{g}) \right] = \phi \qquad \text{in } \mathcal{D}_T, \tag{6.17}$$

where p and q are related to ξ and η through the algebraic equations in (6.13).

It is now clear that $(\mathbf{u}, \xi)$ satisfies a generalized Stokes problem and η satisfies a diffusion problem. This new formulation reveals the underlying deformation and diffusion multiphysics process which occurs in the poroelastic material. In particular, the diffusion part of the process is hidden in the original formulation but is apparent in the new formulation. To make use the above reformulation for computation, we need to address a crucial issue of the uniqueness for the generalized Stokes problem and the diffusion problem after they are decoupled.

6.2.3 Analysis of the partial differential equation model

We start this section with a definition of weak solutions to problem (6.6)–(6.12). For convenience, we assume that $\mathbf{f}, \mathbf{f}_1, \phi$ and ϕ_1 all are independent of t in the remaining of the chapter. We note that all the results of this chapter can be easily extended to the time-dependent cases.

Definition 6.2.1. *Let* $\mathbf{u}_0 \in \mathbf{H}^1(\mathcal{D}), \mathbf{f} \in \mathbf{L}^2(\mathcal{D}), \mathbf{f}_1 \in \mathbf{L}^2(\partial\mathcal{D}), p_0 \in L^2(\mathcal{D}), \phi \in L^2(\mathcal{D})$, *and* $\phi_1 \in L^2(\partial\mathcal{D})$. *Assume* $(\mathbf{f}, \mathbf{v}) + \langle \mathbf{f}_1, \mathbf{v}\rangle = 0$ *for any* $\mathbf{v} \in \mathbf{RM}$. *Given* $T > 0$, *a tuple* $(\mathbf{u}, p)$ *with*

$$\begin{aligned} &\mathbf{u} \in L^\infty\big(0, T; \mathbf{H}^1_\perp(\mathcal{D})\big), && p \in L^2\big(0, T; H^1(\mathcal{D})\big), \\ &(c_0 p + \alpha \operatorname{div} \mathbf{u})_t \in L^2(0, T; H^{-1}(\mathcal{D})), && c_0^{\frac{1}{2}} p \in L^\infty\big(0, T; L^2(\mathcal{D})\big), \end{aligned}$$

is called a weak solution to (6.6)–(6.12), *if there hold for almost every* $t \in [0, T]$

$$\begin{aligned} &\mu\big(\varepsilon(\mathbf{u}), \varepsilon(\mathbf{v})\big) + \lambda\big(\operatorname{div} \mathbf{u}, \operatorname{div} \mathbf{v}\big) - \alpha\big(p, \operatorname{div} \mathbf{v}\big) && (6.18) \\ &\qquad = (\mathbf{f}, \mathbf{v}) + \langle \mathbf{f}_1, \mathbf{v}\rangle \qquad \forall \mathbf{v} \in \mathbf{H}^1(\mathcal{D}), \\ &\big((c_0 p + \alpha \operatorname{div} \mathbf{u})_t, \varphi\big)_{\text{dual}} + \frac{1}{\mu_f}\big(K(\nabla p - \rho_f \mathbf{g}), \nabla \varphi\big) && (6.19) \\ &\qquad = \big(\phi, \varphi\big) + \langle \phi_1, \varphi\rangle \qquad \forall \varphi \in H^1(\mathcal{D}), \\ &\mathbf{u}(0) = \mathbf{u}_0, \qquad p(0) = p_0. && (6.20) \end{aligned}$$

Similarly, we can define weak solutions to problem (6.15)–(6.17), (6.10)–(6.12).

Definition 6.2.2. *Let* $\mathbf{u}_0 \in \mathbf{H}^1(\mathcal{D}), \mathbf{f} \in \mathbf{L}^2(\mathcal{D}), \mathbf{f}_1 \in \mathbf{L}^2(\partial\mathcal{D}), p_0 \in L^2(\mathcal{D}), \phi \in L^2(\mathcal{D})$, *and* $\phi_1 \in L^2(\partial\mathcal{D})$. *Assume* $(\mathbf{f}, \mathbf{v}) + \langle \mathbf{f}_1, \mathbf{v}\rangle = 0$ *for any* $\mathbf{v} \in \mathbf{RM}$. *Given* $T > 0$, *a*

5-*tuple* $(\mathbf{u}, \xi, \eta, p, q)$ *with*

$$\begin{aligned}
&\mathbf{u} \in L^\infty\big(0,T;\mathbf{H}^1_\perp(\mathcal{D})\big), && \xi \in L^2\big(0,T;L^2(\mathcal{D})\big),\\
&\eta \in L^\infty\big(0,T;L^2(\mathcal{D})\big) \cap H^1\big(0,T;H^{-1}(\mathcal{D})\big), && q \in L^\infty(0,T;L^2(\mathcal{D})),\\
&p \in L^2\big(0,T;H^1(\mathcal{D})\big),
\end{aligned}$$

is called a weak solution to (6.15)–(6.17), (6.10)–(6.12) *if there hold for almost every* $t \in [0,T]$

$$\mu\big(\varepsilon(\mathbf{u}),\varepsilon(\mathbf{v})\big) - \big(\xi, \operatorname{div}\mathbf{v}\big) = (\mathbf{f},\mathbf{v}) + \langle \mathbf{f}_1, \mathbf{v}\rangle \qquad \forall \mathbf{v} \in \mathbf{H}^1(\mathcal{D}), \tag{6.21}$$

$$\kappa_3\big(\xi,\varphi\big) + \big(\operatorname{div}\mathbf{u},\varphi\big) = \kappa_1\big(\eta,\varphi\big) \qquad \forall \varphi \in L^2(\mathcal{D}), \tag{6.22}$$

$$\begin{aligned}
\big(\eta_t,\psi\big)_{\text{dual}} + \frac{1}{\mu_f}\big(K(\nabla(\kappa_1\xi+\kappa_2\eta) - \rho_f\mathbf{g}), \nabla\psi\big) \\
= (\phi,\psi) + \langle\phi_1,\psi\rangle \quad \forall \psi \in H^1(\mathcal{D}),
\end{aligned} \tag{6.23}$$

$$p := \kappa_1\xi + \kappa_2\eta, \qquad q := \kappa_1\eta - \kappa_3\xi, \tag{6.24}$$

$$\mathbf{u}(0) = \mathbf{u}_0, \qquad p(0) = p_0, \tag{6.25}$$

$$q(0) = q_0 := \operatorname{div}\mathbf{u}_0, \qquad \eta(0) = \eta_0 := c_0 p_0 + \alpha q_0. \tag{6.26}$$

Remark 6.2.3. *(a) After* ξ *and* η *are computed,* p *and* q *are simply updated by their algebraic expressions in* (6.24).

(b) Equation (6.23) *implicitly imposes the following boundary condition for* η*:*

$$\kappa_2 K\frac{\partial\eta}{\partial n} = K\rho_f g\cdot n - \kappa_1 K\frac{\partial\xi}{\partial n}. \tag{6.27}$$

(c) It should be pointed out that the only reason for introducing the space $\mathbf{H}^1_\perp(\mathcal{D})$ *in the above two definitions is that the boundary condition* (6.10) *is a pure "Neumann condition". If it is replaced by a pure Dirichlet condition or by a mixed Dirichlet-Neumann condition, there is no need to introduce this space. This fact will be used in*

our numerical experiments in Section 6.4. We also note that from the analysis point of view, the pure Neumann condition case is the most difficult case.

Lemma 6.2.4. *Every weak solution* $(\mathbf{u}, p)$ *of problem* (6.18)–(6.20) *satisfies the following energy law:*

$$E(t) + \frac{1}{\mu_f}\int_0^t \big(K(\nabla p - \rho_f \mathbf{g}), \nabla p\big)\, ds - \int_0^t \big(\phi, p\big)\, ds - \int_0^t \langle \phi_1, p\rangle\, ds = E(0) \tag{6.28}$$

for all $t \in [0, T]$, *where*

$$\begin{aligned} E(t) := \frac{1}{2}\Big[\mu \| \varepsilon(\mathbf{u}(t)) \|^2_{L^2(\mathcal{D})} + \lambda \| \operatorname{div} \mathbf{u}(t) \|^2_{L^2(\mathcal{D})} + c_0 \| p(t) \|^2_{L^2(\mathcal{D})} \\ - 2\big(\mathbf{f}, \mathbf{u}(t)\big) - 2\langle \mathbf{f}_1, \mathbf{u}(t)\rangle\Big]. \end{aligned} \tag{6.29}$$

Moreover,

$$\begin{aligned} \| (c_0 p + \alpha \operatorname{div} \mathbf{u})_t \|_{L^2(0.T; H^{-1}(\mathcal{D}))} \le \frac{K}{\mu_f} \| \nabla p - \rho_f \mathbf{g} \|_{L^2(\mathcal{D}_T)} \\ + \|\phi\|_{L^2(\mathcal{D}_T)} + \|\phi_1\|_{L^2(\partial\mathcal{D}_T)} < \infty. \end{aligned} \tag{6.30}$$

Proof. We only consider the case $\mathbf{u}_t \in \mathbf{L}^2((0,T); \mathbf{L}^2(\mathcal{D}))$, the general case can be converted into this case using the Steklov average technique (cf. [70, Chapter 2]). Setting $\varphi = p$ in (6.19) and $\mathbf{v} = \mathbf{u}_t$ in (6.18) yields for a.e. $t \in [0, T]$

$$\big((c_0 p + \alpha \operatorname{div} \mathbf{u})_t, p(t)\big)_{\text{dual}} + \frac{1}{\mu_f}\big(K(\nabla p - \rho_f \mathbf{g}), \nabla p\big) = \big(\phi, p\big) + \langle \phi_1, p\rangle, \tag{6.31}$$

$$\mu\big(\varepsilon(\mathbf{u}), \varepsilon(\mathbf{u}_t)\big) + \lambda\big(\operatorname{div} \mathbf{u}, \operatorname{div} \mathbf{u}_t\big) - \alpha\big(p, \operatorname{div} \mathbf{u}_t\big) = (\mathbf{f}, \mathbf{u}_t) + \langle \mathbf{f}_1, \mathbf{u}_t\rangle. \tag{6.32}$$

Adding the above two equations and integrating the sum in t over the interval $(0, s)$ for any $s \in (0, T]$ yield

$$E(s) + \frac{1}{\mu_f}\int_0^s \big(K(\nabla p - \rho_f \mathbf{g}), \nabla p\big)\, dt - \int_0^s \big(\phi, p\big)\, dt - \int_0^s \langle \phi_1, p\rangle\, dt = E(0), \tag{6.33}$$

where $E(\cdot)$ is given by (6.29). Here we have used the fact that $\mathbf{f}$ and $\mathbf{f}_1$ are independent of t. Hence, (6.28) holds.

(6.30) follows immediately from (6.19). The proof is complete. □

Likewise, weak solutions of (6.21)–(6.26) satisfy a similar energy law which is a rewritten version of (6.28) in the new variables.

Lemma 6.2.5. *Every weak solution* $(\mathbf{u}, \xi, \eta, p, q)$ *of problem* (6.21)–(6.26) *satisfies the following energy law:*

$$J(t) + \frac{1}{\mu_f}\int_0^t \big(K(\nabla p - \rho_f \mathbf{g}), \nabla p\big)\, ds - \int_0^t \big(\phi, p\big)\, ds - \int_0^t \langle \phi_1, p\rangle\, ds = J(0) \tag{6.34}$$

for all $t \in [0, T]$, *where*

$$\begin{aligned} J(t) := \frac{1}{2}\Big[\mu \| \varepsilon(\mathbf{u}(t)) \|^2_{L^2(\mathcal{D})} + \kappa_2 \| \eta(t) \|^2_{L^2(\mathcal{D})} + \kappa_3 \| \xi(t) \|^2_{L^2(\mathcal{D})} \\ - 2\big(\mathbf{f}, \mathbf{u}(t)\big) - 2\langle \mathbf{f}_1, \mathbf{u}(t)\rangle\Big]. \end{aligned} \tag{6.35}$$

Moreover,

$$\begin{aligned} \| \eta_t \|_{L^2(0.T;H^{-1}(\mathcal{D}))} \le \frac{K}{\mu_f} \| \nabla p - \rho_f \mathbf{g} \|_{L^2(\mathcal{D}_T)} \\ + \|\phi\|_{L^2(\mathcal{D}_T)} + \|\phi_1\|_{L^2(\partial\mathcal{D}_T)} < \infty. \end{aligned} \tag{6.36}$$

Proof. Again, we only consider the case that $\mathbf{u}_t \in L^2(0, T; L^2(\mathcal{D}))$. Setting $\mathbf{v} = \mathbf{u}_t$ in (6.21), differentiating (6.22) with respect to t followed by taking $\varphi = \xi$, and setting $\psi = p = \kappa_1 \xi + \kappa_2 \eta$ in (6.23); adding the resulting equations and integrating in t yield the desired equality (6.34). The inequality (6.36) follows immediately from (6.23). □

The above energy law immediately implies the following solution estimates.

Lemma 6.2.6. *There exists a positive constant* $C_1 = C_1\big(\|\mathbf{u}_0\|_{H^1(\mathcal{D})}, \|p_0\|_{L^2(\mathcal{D})}, \|\mathbf{f}\|_{L^2(\mathcal{D})}, \|\mathbf{f}_1\|_{L^2(\partial\mathcal{D})}, \|\phi\|_{L^2(\mathcal{D})}, \|\phi_1\|_{L^2(\partial\mathcal{D})}\big)$ *such that*

$$\begin{aligned}&\sqrt{\mu}\|\varepsilon(\mathbf{u})\|_{L^\infty(0,T;L^2(\mathcal{D}))} + \sqrt{\kappa_2}\|\eta\|_{L^\infty(0,T;L^2(\mathcal{D}))} \\ &\quad + \sqrt{\kappa_3}\|\xi\|_{L^\infty(0,T;L^2(\mathcal{D}))} + \sqrt{\frac{K}{\mu_f}}\|\nabla p\|_{L^2(0,T;L^2(\mathcal{D}))} \le C_1.\end{aligned} \tag{6.37}$$

$$\|\mathbf{u}\|_{L^\infty(0,T;L^2(\mathcal{D}))} \le C_1, \qquad \|p\|_{L^\infty(0,T;L^2(\mathcal{D}))} \le C_1\Big(1 + \sqrt{\frac{\kappa_3}{\kappa_1}}\Big). \tag{6.38}$$

We note that (6.38) follows from (6.37), (6.3) and the relation $p = \kappa_1\xi + \kappa_2\eta$.

Furthermore, by exploiting the linearity of the PDE system, we have the following a priori estimates for the weak solution.

Theorem 6.2.7. *Suppose that* $\mathbf{u}_0$ *and* p_0 *are sufficiently smooth, then there exists a positive constant* $C_2 = C_2\big(C_1, \|\nabla p_0\|_{L^2(\mathcal{D})}\big)$ *and* $C_3 = C_3\big(C_1, C_2, \|\mathbf{u}_0\|_{H^2(\mathcal{D})}, \|p_0\|_{H^2(\mathcal{D})}\big)$ *such that there hold the following estimates for the solution to problem* (6.15)–(6.17),(6.10)–(6.12)*:*

$$\begin{aligned}&\sqrt{\mu}\|\varepsilon(\mathbf{u}_t)\|_{L^2(0,T;L^2(\mathcal{D}))} + \sqrt{\kappa_2}\|\eta_t\|_{L^2(0,T;L^2(\mathcal{D}))} \\ &\quad + \sqrt{\kappa_3}\|\xi_t\|_{L^2(0,T;L^2(\mathcal{D}))} + \sqrt{\frac{K}{\mu_f}}\|\nabla p\|_{L^\infty(0,T;L^2(\mathcal{D}))} \le C_2.\end{aligned} \tag{6.39}$$

$$\begin{aligned}&\sqrt{\mu}\|\varepsilon(\mathbf{u}_t)\|_{L^\infty(0,T;L^2(\mathcal{D}))} + \sqrt{\kappa_2}\|\eta_t\|_{L^\infty(0,T;L^2(\mathcal{D}))} \\ &\quad + \sqrt{\kappa_3}\|\xi_t\|_{L^\infty(0,T;L^2(\mathcal{D}))} + \sqrt{\frac{K}{\mu_f}}\|\nabla p_t\|_{L^2(0,T;L^2(\mathcal{D}))} \le C_3.\end{aligned} \tag{6.40}$$

$$\|\eta_{tt}\|_{L^2(H^{-1}(\mathcal{D}))} \le \sqrt{\frac{K}{\mu_f}}C_3. \tag{6.41}$$

Proof. On noting that $\mathbf{f}, \mathbf{f}_1, \phi$ and ϕ_1 all are assumed to be independent of t, differentiating (6.21) and (6.22) with respect to t, taking $\mathbf{v} = \mathbf{u}_t$ and $\varphi = \xi_t$ in

(6.21) and (6.22) respectively, and adding the resulting equations yield

$$\mu\|\varepsilon(\mathbf{u}_t)\|^2_{L^2(\mathcal{D})} = \big(q_t, \xi_t\big) = \kappa_1\big(\eta_t, \xi_t\big) - \kappa_3\|\xi_t\|^2_{L^2(\mathcal{D})}. \tag{6.42}$$

Setting $\psi = p_t = \kappa_1\xi_t + \kappa_2\eta_t$ in (6.23) gives

$$\kappa_1\big(\eta_t, \xi_t\big) + \kappa_2\|\eta_t\|^2_{L^2(\mathcal{D})} + \frac{K}{2\mu_f}\frac{d}{dt}\|\nabla p - \rho_f\mathbf{g}\|^2_{L^2(\mathcal{D})} = \frac{d}{dt}\Big[(\phi, p) + \langle\phi_1, p\rangle\Big]. \tag{6.43}$$

Adding (6.42) and (6.43) and integrating in t we get for $t \in [0, T]$

$$\begin{aligned}\frac{K}{2\mu_f}\|\nabla p(t) - \rho_f\mathbf{g}\|^2_{L^2(\mathcal{D})} + \int_0^t \Big[\mu\|\varepsilon(\mathbf{u}_t)\|^2_{L^2(\mathcal{D})} + \kappa_2\|\eta_t\|^2_{L^2(\mathcal{D})} + \kappa_3\|\xi_t\|^2_{L^2(\mathcal{D})}\Big]\, ds \\ = \frac{K}{2\mu_f}\|\nabla p_0 - \rho_f\mathbf{g}\|^2_{L^2(\mathcal{D})} + (\phi, p(t) - p_0) + \langle\phi_1, p(t) - p_0\rangle,\end{aligned}$$

which readily infers (6.39).

To show (6.40), first differentiating (6.21) one time with respect to t and setting $\mathbf{v} = \mathbf{u}_{tt}$, differentiating (6.22) twice with respect to t and setting $\varphi = \xi_t$, and adding the resulting equations we get

$$\frac{\mu}{2}\frac{d}{dt}\|\varepsilon(\mathbf{u}_t)\|^2_{L^2(\mathcal{D})} = \big(q_{tt}, \xi_t\big) = \kappa_1\big(\eta_{tt}, \xi_t\big) - \frac{\kappa_3}{2}\frac{d}{dt}\|\xi_t\|^2_{L^2(\mathcal{D})}. \tag{6.44}$$

Second, differentiating (6.23) with respect t one time and taking $\psi = p_t = \kappa_1\xi_t + \kappa_2\eta_t$ yield

$$\kappa_1\big(\eta_{tt}, \xi_t\big) + \frac{\kappa_2}{2}\frac{d}{dt}\|\eta_t\|^2_{L^2(\mathcal{D})} + \frac{K}{\mu_f}\|\nabla p_t\|^2_{L^2(\mathcal{D})} = 0. \tag{6.45}$$

Finally, adding the above two inequalities and integrating in t give for $t \in [0, T]$

$$\begin{aligned}\mu\|\varepsilon(\mathbf{u}_t(t))\|^2_{L^2(\mathcal{D})} + \kappa_2\|\eta_t(t)\|^2_{L^2(\mathcal{D})} + \kappa_3\|\xi_t(t)\|^2_{L^2(\mathcal{D})} + \frac{2K}{\mu_f}\int_0^t \|\nabla p_t\|^2_{L^2(\mathcal{D})}\, ds \\ = \mu\|\varepsilon(\mathbf{u}_t(0))\|^2_{L^2(\mathcal{D})} + \kappa_2\|\eta_t(0)\|^2_{L^2(\mathcal{D})} + \kappa_3\|\xi_t(0)\|^2_{L^2(\mathcal{D})}.\end{aligned} \tag{6.46}$$

Hence, (6.40) holds. (6.41) follows immediately from the following inequality

$$\big(\eta_{tt},\psi\big)=-\frac{1}{\mu_f}\big(K\nabla p_t,\nabla\psi\big)\le\frac{K}{\mu_f}\|\nabla p_t\|_{L^2(\mathcal{D})}\|\nabla\psi\|_{L^2(\mathcal{D})}\qquad\forall\psi\in H_0^1(\mathcal{D}),$$

(6.40) and the definition of the H^{-1}-norm. The proof is complete. □

Remark 6.2.8. *As expected, the above high order norm solution estimates require* $p_0\in H^1(\mathcal{D}),\mathbf{u}_t(0)\in\mathbf{L}^2(\mathcal{D}),\eta_t(0)\in L^2(\mathcal{D})$ *and* $\xi_t(0)\in L^2(\mathcal{D})$. *The values of* $\mathbf{u}_t(0),\eta_t(0)$ *and* $\xi_t(0)$ *can be computed using the PDEs as follows. It follows from* (6.17) *that* $\eta_t(0)$ *satisfies*

$$\eta_t(0)=\phi+\frac{1}{\mu_f}\mathrm{div}\,[K(\nabla p_0-\rho_f\mathbf{g})].$$

Hence $\eta_t(0)\in L^2(\mathcal{D})$ *provided that* $p_0\in H^2(\mathcal{D})$.

To find $\mathbf{u}_t(0)$ *and* $\xi_t(0)$, *differentiating* (6.15) *and* (6.16) *with respect to* t *and setting* $t=0$ *we get*

$$\begin{aligned}-\mu\mathrm{div}\,\varepsilon(\mathbf{u}_t(0))+\nabla\xi_t(0)&=0&&\text{in }\mathcal{D},\\ \kappa_3\xi_t(0)+\mathrm{div}\,\mathbf{u}_t(0)&=\kappa_1\eta_t(0)&&\text{in }\mathcal{D}.\end{aligned}$$

Hence, $\mathbf{u}_t(0)$ *and* $\xi_t(0)$ *can be determined by solving the above generalized Stokes problem.*

The next lemma shows that weak solutions of problem (6.21)–(6.26) preserve some "invariant" quantities, it turns out that these "invariant" quantities play a vital role in the construction of our time-splitting scheme to be introduced in the next section.

Lemma 6.2.9. *Every weak solution* $(\mathbf{u}, \xi, \eta, p, q)$ *to problem* (6.21)–(6.26) *satisfies the following relations:*

$$C_\eta(t) := \big(\eta(\cdot,t),1\big) = \big(\eta_0,1\big) + \big[(\phi,1) + \langle\phi_1,1\rangle\big]t, \quad t \geq 0. \tag{6.47}$$

$$C_\xi(t) := \big(\xi(\cdot,t),1\big) = \frac{1}{d+\mu\kappa_3}\big[\mu\kappa_1 C_\eta(t) - \big(\mathbf{f},x\big) - \langle\mathbf{f}_1,x\rangle\big]. \tag{6.48}$$

$$C_q(t) := \big(q(\cdot,t),1\big) = \kappa_1 C_\eta(t) - \kappa_3 C_\xi(t). \tag{6.49}$$

$$C_p(t) := \big(p(\cdot,t),1\big) = \kappa_1 C_\xi(t) + \kappa_2 C_\eta(t). \tag{6.50}$$

$$C_{\mathbf{u}}(t) := \big\langle \mathbf{u}(\cdot,t)\cdot\mathbf{n},1\big\rangle = C_q(t). \tag{6.51}$$

Proof. We first notice that equation (6.47) follows immediately from taking $\varphi \equiv 1$ in (6.23), which is a valid test function.

To prove (6.48), taking $\mathbf{v} = x$ in (6.21) and $\varphi = 1$ in (6.22), which are valid test functions, and using the identities $\nabla x = I, \operatorname{div} x = d$, and $\varepsilon(x) = I$, we get

$$\mu\big(\operatorname{div}\mathbf{u},1\big) - d\big(\xi,1\big) = \big(\mathbf{f},x\big) + \langle\mathbf{f}_1,x\rangle, \tag{6.52}$$

$$\big(\operatorname{div}\mathbf{u},1\big) = \kappa_1(\eta,1) - \kappa_3(\xi,1). \tag{6.53}$$

Substituting (6.53) into (6.52) and using (6.47) yield

$$C_\xi(t) := \big(\xi(\cdot,t),1\big) = \frac{1}{d+\mu\kappa_3}\big[\mu\kappa_1 C_\eta(t) - \big(\mathbf{f},x\big) - \langle\mathbf{f}_1,x\rangle\big]. \tag{6.54}$$

Hence (6.48) holds. (6.49) follows immediately from (6.53), (6.47) and (6.48).

Finally, since $p = \kappa_1\xi + \kappa_2\eta$, (6.50) then follows from (6.47) and (6.48). (6.51) is an immediate consequence of $q = \operatorname{div}\mathbf{u}$ and the divergence theorem. The proof is complete. □

Remark 6.2.10. *We note that* C_η, C_ξ, C_q *and* C_p *all are (known) linear functions of* t*, and they become (known) constants when* $\phi \equiv 0$ *and* $\phi_1 \equiv 0$*.*

With the help of the above lemmas, we can show the solvability of problem (6.6)–(6.12).

Theorem 6.2.11. *Let* $\mathbf{u}_0 \in \mathbf{H}^1(\mathcal{D}), \mathbf{f} \in \mathbf{L}^2(\mathcal{D}), \mathbf{f}_1 \in \mathbf{L}^2(\partial\mathcal{D}), p_0 \in L^2(\mathcal{D}), \phi \in L^2(\mathcal{D})$, *and* $\phi_1 \in L^2(\partial\mathcal{D})$. *Suppose* $(\mathbf{f}, \mathbf{v}) + \langle \mathbf{f}_1, \mathbf{v} \rangle = 0$ *for any* $\mathbf{v} \in \mathbf{RM}$. *Then there exists a unique solution to problem* (6.6)–(6.12) *in the sense of Definition 6.2.1, likewise, there exists a unique solution to problem* (6.15)–(6.17),(6.10)–(6.12) *in the sense of Definition 6.2.2.*

Proof. We only outline the main steps of the proof and leave the details to the interested reader.

First, since the PDE system is linear, the existence of weak solution can be proved by the standard Galerkin method and compactness argument (cf. [89]). We note that the energy laws established in Lemmas 6.2.4 and 6.2.5 guarantee the required uniform estimates for the Galerkin approximate solutions.

Second, to show the uniqueness, suppose there are two sets of weak solutions, again by the linearity of the PDE system it is trivial to show that the difference of the solutions satisfy the same PDE system with *zero* initial and boundary data. The energy law immediately implies that the difference must be zero, hence, the uniqueness is verified. □

We conclude this section by establishing a convergence result for the solution of problem (6.15)–(6.17),(6.10)–(6.12) when the constrained specific storage coefficient c_0 tends to 0. Such a convergence result is useful and significant for the following two reasons. First, as mentioned earlier, the poroelasticity model studied in this chapter reduces into the well-known Biot's consolidation model from soil mechanics (cf. [75, 68]) and Doi's model for polymer gels (cf. [38, 93]). Second, it proves that the proposed approach and methods of this chapter are robust under such a limit process.

Theorem 6.2.12. *Let* $\mathbf{u}_0 \in \mathbf{H}^1(\mathcal{D}), \mathbf{f} \in \mathbf{L}^2(\mathcal{D}), \mathbf{f}_1 \in \mathbf{L}^2(\partial\mathcal{D}), p_0 \in L^2(\mathcal{D}), \phi \in L^2(\mathcal{D})$, *and* $\phi_1 \in L^2(\partial\mathcal{D})$. *Suppose* $(\mathbf{f}, \mathbf{v}) + \langle \mathbf{f}_1, \mathbf{v} \rangle = 0$ *for any* $\mathbf{v} \in \mathbf{RM}$. *Let*

$(\mathbf{u}_{c_0}, \eta_{c_0}, \xi_{c_0}, p_{c_0}, q_{c_0})$ *denote the unique weak solution to problem* (6.15)–(6.17),(6.10)–(6.12). *Then there exists* $(\mathbf{u}_*, \eta_*, \xi_*, p_*, q_*) \in \mathbf{L}^\infty(0,T;\mathbf{H}^1_\perp(\mathcal{D})) \times L^\infty(0,T;L^2(\mathcal{D})) \times L^\infty(0,T;L^2(\mathcal{D}))) \times L^\infty(0,T;L^2(\mathcal{D})) \cap L^2(0,T;H^1(\mathcal{D}) \times L^\infty(0,T;L^2(\mathcal{D}))$ *such that* $(\mathbf{u}_{c_0}, \eta_{c_0}, \xi_{c_0}, p_{c_0}, q_{c_0})$ *converges weakly to* $(\mathbf{u}_*, \eta_*, \xi_*, p_*, q_*)$ *in the above product space as* $c_0 \to 0$.

Proof. It follows immediately from (6.36)–(6.38) and Korn's inequality that

- $\mathbf{u}_{c_0}$ is uniformly bounded (in c_0) in $\mathbf{L}^\infty(0,T;\mathbf{H}^1_\perp(\mathcal{D}))$.
- $\sqrt{\kappa_2}\eta_{c_0}$ is uniformly bounded (in c_0) in $L^\infty(0,T;L^2(\mathcal{D})) \cap L^2(0,T;H^{-1}(\mathcal{D}))$.
- $\sqrt{\kappa_3}\xi_{c_0}$ is uniformly bounded (in c_0) in $L^\infty(0,T;L^2(\mathcal{D}))$.
- p_{c_0} is uniformly bounded (in c_0) in $L^\infty(0,T;L^2(\mathcal{D})) \cap L^2(0,T;H^1(\mathcal{D}))$.
- q_{c_0} is uniformly bounded (in c_0) in $L^\infty(0,T;L^2(\mathcal{D}))$.

On noting that $\lim_{c_0\to 0}\kappa_1 = \frac{1}{\alpha}$, $\lim_{c_0\to 0}\kappa_2 = \frac{\lambda}{\alpha^2}$ and $\lim_{c_0\to 0}\kappa_3 = 0$, by the weak compactness of reflexive Banach spaces and Aubin-Lions Lemma [26] we have that there exist $(\mathbf{u}_*, \eta_*, \xi_*, p_*, q_*) \in \mathbf{L}^\infty(0,T;\mathbf{H}^1_\perp(\mathcal{D})) \times L^\infty(0,T;L^2(\mathcal{D})) \times L^\infty(0,T;L^2(\mathcal{D}))) \times L^\infty(0,T;\ L^2(\mathcal{D})) \cap L^2(0,T;H^1(\mathcal{D}) \times L^\infty(0,T;L^2(\mathcal{D}))$ and a subsequence of $(\mathbf{u}_{c_0}, \eta_{c_0}, \xi_{c_0}, p_{c_0}, q_{c_0})$ (still denoted by the same notation) such that as $c_0 \to 0$ (a subsequence of c_0, to be exact)

- $\mathbf{u}_{c_0}$ converges to $\mathbf{u}_*$ weak $*$ in $\mathbf{L}^\infty(0,T;\mathbf{H}^1_\perp(\mathcal{D}))$ and weakly in $\mathbf{L}^2(0,T;\mathbf{H}^1_\perp(\mathcal{D}))$.
- $\sqrt{\kappa_2}\eta_{c_0}$ converges to $\frac{\sqrt{\lambda}}{\alpha}\eta_*$ weak $*$ in $L^\infty(0,T;L^2(\mathcal{D}))$ and weakly in $L^2(\mathcal{D}_T)$.
- $\kappa_3\xi_{c_0}$ converges to 0 weakly in $L^2(\mathcal{D}_T)$.
- p_{c_0} converges to p_* weak $*$ in $L^\infty(0,T;L^2(\mathcal{D}))$ and weakly in $L^2(0,T;H^1(\mathcal{D}))$.
- q_{c_0} converges to p_* weak $*$ in $L^\infty(0,T;L^2(\mathcal{D}))$ and weakly in $L^2(\mathcal{D}_T)$.

Then setting $c_0 \to 0$ in (6.21)–(6.26) yields (note that the dependence of the solution on c_0 is suppressed there)

$$\begin{aligned}
\mu\big(\varepsilon(\mathbf{u}_*),\varepsilon(\mathbf{v})\big) - \big(\xi_*, \operatorname{div}\mathbf{v}\big) &= (\mathbf{f},\mathbf{v}) + \langle \mathbf{f}_1,\mathbf{v}\rangle && \forall \mathbf{v}\in \mathbf{H}^1(\mathcal{D}),\\
\big(\operatorname{div}\mathbf{u}_*,\varphi\big) &= \frac{1}{\alpha}\big(\eta_*,\varphi\big) && \forall \varphi\in L^2(\mathcal{D}),\\
\big((\eta_*)_t,\psi\big)_{\text{dual}} + \frac{1}{\mu_f}\big(K(\nabla p_* - \rho_f\mathbf{g}),\nabla\psi\big)&\\
&= (\phi,\psi) + \langle\phi_1,\psi\rangle && \forall \psi\in H^1(\mathcal{D}),\\
p_* := \frac{1}{\alpha}\xi_* + \frac{\lambda}{\alpha^2}\eta_*, \qquad q_* &:= \frac{1}{\alpha}\eta_*,&&\\
\mathbf{u}_*(0) &= \mathbf{u}_0,&&\\
q_*(0) = q_0 := \operatorname{div}\mathbf{u}_0, \qquad \eta_*(0) &= \eta_0 := \alpha q_0.&&
\end{aligned}$$

Equivalently,

$$\begin{aligned}
\mu\big(\varepsilon(\mathbf{u}_*),\varepsilon(\mathbf{v})\big) - \big(\xi_*, \operatorname{div}\mathbf{v}\big) &= (\mathbf{f},\mathbf{v}) + \langle \mathbf{f}_1,\mathbf{v}\rangle && \forall \mathbf{v}\in \mathbf{H}^1(\mathcal{D}),\\
\big(\operatorname{div}\mathbf{u}_*,\varphi\big) &= \big(q_*,\varphi\big) && \forall \varphi\in L^2(\mathcal{D}),\\
\alpha\big((q_*)_t,\psi\big)_{\text{dual}} + \frac{1}{\mu_f}\big(K(\nabla p_* - \rho_f\mathbf{g}),\nabla\psi\big) &= (\phi,\psi) + \langle\phi_1,\psi\rangle && \forall \psi\in H^1(\mathcal{D}),\\
p_* := \frac{1}{\alpha}\Big(\xi_* + \lambda q_*\Big) \quad \text{or} \quad \xi_* &= \alpha p_* - \lambda q_*,&&\\
\mathbf{u}_*(0) &= \mathbf{u}_0,&&\\
q_*(0) = q_0 &:= \operatorname{div}\mathbf{u}_0.&&
\end{aligned}$$

Hence, $(\mathbf{u}_*, \eta_*, \xi_*, p_*, q_*)$ is a weak solution of Biot's consolidation model (cf. [38, 93]). By the uniqueness of its solutions, we conclude that the whole sequence $(\mathbf{u}_{c_0}, \eta_{c_0}, \xi_{c_0}, p_{c_0}, q_{c_0})$ converges to $(\mathbf{u}_*, \eta_*, \xi_*, p_*, q_*)$ as $c_0 \to 0$ in the above sense. The proof is complete. □

6.3 Fully discrete finite element methods

The goal of this section is to design and analyze some fully discrete finite element methods for the poroelasticity model based on the above new formulation. As the time stepping is vital for the overall methods, we first introduce our time-stepping schemes at the PDE level.

6.3.1 Basic time-stepping algorithm

Based on this new formulation, our multiphysics time-stepping algorithm reads as follows:

Splitting Algorithm (SA):

(i) Set

$$\begin{aligned} &\mathbf{u}^0 = \mathbf{u}_0, && p^0 = p_0, \quad q^0 = q_0 := \operatorname{div} \mathbf{u}_0, \\ &\eta^0 = c_0 p^0 + \alpha q^0, && \xi^0 = \alpha p^0 - \lambda q^0. \end{aligned}$$

(ii) For $n = 0, 1, 2, \cdots$, complete the following three steps:

Step 1: Solve for $(\mathbf{u}^{n+1}, \xi^{n+1})$ such that

$$-\mu \operatorname{div} \varepsilon(\mathbf{u}^{n+1}) + \nabla \xi^{n+1} = \mathbf{f}, \qquad \text{in } \mathcal{D}_T, \tag{6.55}$$

$$\kappa_3 \xi^{n+1} + \operatorname{div} \mathbf{u}^{n+1} = \kappa_1 \eta^{n+\theta} \qquad \text{in } \mathcal{D}_T, \tag{6.56}$$

$$\widetilde{\sigma}(\mathbf{u}^{n+1}, \xi^{n+1})\mathbf{n} = \mathbf{f}_1 \qquad \text{on } \partial\mathcal{D}_T. \tag{6.57}$$

Here $\theta = 0$ or 1.

Step 2: Solve for η^{n+1} such that

$$d_t\eta^{n+1} - \frac{1}{\mu_f}\mathrm{div}\,[K(\nabla(\kappa_2\eta^{n+1} + \kappa_1\xi^{n+1}) - \rho_f\mathbf{g})] = \phi, \qquad \text{in } \mathcal{D}_T, \tag{6.58}$$

$$\frac{1}{\mu_f}K[\nabla(\kappa_2\eta^{n+1} + \kappa_1\xi^{n+1}) - \rho_f\mathbf{g}]\cdot\mathbf{n} = \phi_1 \qquad \text{on } \partial\mathcal{D}_T. \tag{6.59}$$

Step 3: Update p^{n+1} and q^{n+1} by

$$p^{n+1} = \kappa_1\xi^{n+1} + \kappa_2\eta^{n+\theta}, \qquad q^{n+1} = \kappa_1\eta^{n+1} - \kappa_3\xi^{n+1}. \tag{6.60}$$

Where $d_t\eta^{n+1} := (\eta^{n+1} - \eta^n)/\Delta t$, Δt denotes the time step size of a uniform partition of the time interval $[0, T]$, and

$$\widetilde{\sigma}(\mathbf{u}^{n+1}, \xi^{n+1}) := \mu\varepsilon(\mathbf{u}^{n+1}) - \xi^{n+1}I. \tag{6.61}$$

We note that (6.58) is the implicit Euler scheme, which is chosen just for the ease of presentation, it can be replaced by other time-stepping schemes. (6.59) provides a flux boundary condition for η^{n+1}.

Remark 6.3.1. *When $\theta = 0$,* Step 1 *and* Step 2 *are decoupled, hence these two sub-problems can be solved independently. On the other hand, when $\theta = 1$, two sub-problems are coupled, hence, they must be solved together.*

6.3.2 Fully discrete finite element methods

In this section, we consider the space-time discretization which combines the above splitting algorithm with appropriately chosen spatial discretization methods. To the end, we introduce some notation.

Assume $\mathcal{D} \in \mathbb{R}^d (d = 2, 3)$ is a polygonal domain. Let $\mathcal{T}_h$ be a quasi-uniform triangulation or rectangular partition of $\mathcal{D}$ with mesh size h, and $\bar{\mathcal{D}} = \bigcup_{K\in\mathcal{T}_h}\bar{K}$. Also, let $(\mathbf{X}_h, M_h)$ be a stable mixed finite element pair, that is, $\mathbf{X}_h \subset \mathbf{H}^1(\mathcal{D})$ and

$M_h \subset L^2(\mathcal{D})$ satisfy the inf-sup condition

$$\sup_{v_h \in \mathbf{X}_h} \frac{(\mathrm{div} v_h, \varphi_h)}{\|\nabla v_h\|_{L^2(\mathcal{D})}} \geq \beta_0 \|\varphi_h\|_{L^2(\mathcal{D})} \quad \forall \varphi_h \in M_{0h} := M_h \cap L_0^2(\mathcal{D}),\ \beta_0 > 0. \tag{6.62}$$

A number of stable mixed finite element spaces $(\mathbf{X}_h, M_h)$ have been known in the literature [14]. A well-known example is the following so-called Taylor-Hood element (cf. [10, 14]):

$$\mathbf{X}_h = \big\{\mathbf{v}_h \in \mathbf{C}^0(\overline{\mathcal{D}});\ \mathbf{v}_h|_K \in \mathbf{P}_2(K)\ \ \forall K \in \mathcal{T}_h\big\},$$
$$M_h = \big\{\varphi_h \in C^0(\overline{\mathcal{D}});\ \varphi_h|_K \in P_1(K)\ \ \forall K \in \mathcal{T}_h\big\}.$$

In the next subsection, we shall only present the analysis for the Taylor-Hood element, but remark that the analysis can be extended to other stable mixed elements. However, piecewise constant space M_h is not recommended because that would result in no rate of convergence for the approximation of the pressure p (see Section 6.3.4).

Finite element approximation space W_h for η variable can be chosen independently, any piecewise polynomial space is acceptable provided that $W_h \supset M_h$. Especially, $W_h \subset L^2(\mathcal{D})$ can be chosen as a fully discontinuous piecewise polynomial space, although it is more convenient to choose W_h to be a continuous (resp. discontinuous) space if M_h is a continuous (resp. discontinuous) space. The most convenient choice is $W_h = M_h$, which will be adopted in the remainder of this chapter.

Recall that $\mathbf{RM}$ denotes the space of the infinitesimal rigid motions (see Section 6.2), evidently, $\mathbf{RM} \subset \mathbf{X}_h$. We now introduce the L^2-projection $\mathcal{P}_R$ from $\mathbf{L}^2(\mathcal{D})$ to $\mathbf{RM}$. For each $\mathbf{v} \in \mathbf{L}^2(\mathcal{D})$, $\mathcal{P}_R \mathbf{v}_h \in \mathbf{RM}$ is defined by

$$(\mathcal{P}_R \mathbf{v}_h, \mathbf{r}) = (\mathbf{v}_h, \mathbf{r}) \qquad \forall \mathbf{r} \in \mathbf{RM}.$$

Moreover, we define

$$\mathbf{V}_h := (I - \mathcal{P}_R)\mathbf{X}_h = \big\{\mathbf{v}_h \in \mathbf{X}_h;\ (\mathbf{v}_h, \mathbf{r}) = 0\ \forall \mathbf{r} \in \mathbf{RM}\big\}. \tag{6.63}$$

It is easy to check that $\mathbf{X}_h = \mathbf{V}_h \bigoplus \mathbf{RM}$. It was proved in [38] that there holds the following alternative version of the above inf-sup condition:

$$\sup_{\mathbf{v}_h \in \mathbf{V}_h} \frac{(\operatorname{div} \mathbf{v}_h, \varphi_h)}{\| \nabla \mathbf{v}_h \|_{L^2(\mathcal{D})}} \geq \beta_1 \| \varphi_h \|_{L^2(\mathcal{D})} \quad \forall \varphi_h \in M_{0h}, \quad \beta_1 > 0. \tag{6.64}$$

Finite Element Algorithm (FEA):

(i) Compute $\mathbf{u}_h^0 \in \mathbf{V}_h$ and $q_h^0 \in W_h$ by

$$\begin{aligned} &\mathbf{u}_h^0 = \mathcal{R}_h \mathbf{u}_0, \qquad && p_h^0 = \mathcal{Q}_h p_0, \qquad && q_h^0 = \mathcal{Q}_h q_0 \ (q_0 = \operatorname{div} \mathbf{u}_0), \\ &\eta_h^0 = c_0 p_h^0 + \alpha q_h^0, \qquad && \xi_h^0 = \alpha p_h^0 - \lambda q_h^0. \end{aligned}$$

(ii) For $n = 0, 1, 2, \cdots$, do the following three steps.

Step 1: Solve for $(\mathbf{u}_h^{n+1}, \xi_h^{n+1}) \in \mathbf{V}_h \times W_h$ such that

$$\mu\big(\varepsilon(\mathbf{u}_h^{n+1}), \varepsilon(\mathbf{v}_h)\big) - \big(\xi_h^{n+1}, \operatorname{div} \mathbf{v}_h\big) = (\mathbf{f}, \mathbf{v}_h) + \langle \mathbf{f}_1, \mathbf{v}_h \rangle \quad \forall \mathbf{v}_h \in \mathbf{V}_h, \tag{6.65}$$

$$\kappa_3\big(\xi_h^{n+1}, \varphi_h\big) + \big(\operatorname{div} \mathbf{u}_h^{n+1}, \varphi_h\big) = \kappa_1\big(\eta_h^{n+\theta}, \varphi_h\big) \qquad \forall \varphi_h \in M_h. \tag{6.66}$$

Step 2: Solve for $\eta_h^{n+1} \in W_h$ such that

$$\begin{aligned} \big(d_t \eta_h^{n+1}, \psi_h\big) + \frac{1}{\mu_f}\big(K(\nabla(\kappa_1 \xi_h^{n+1} + \kappa_2 \eta_h^{n+1}) - \rho_f \mathbf{g}, \nabla \psi_h\big) \\ = (\phi, \psi_h) + \langle \phi_1, \psi_h \rangle. \end{aligned} \tag{6.67}$$

Step 3: Update p_h^{n+1} and q_h^{n+1} by

$$p_h^{n+1} = \kappa_1 \xi_h^{n+1} + \kappa_2 \eta_h^{n+\theta}, \quad q_h^{n+1} = \kappa_1 \eta_h^{n+1} - \kappa_3 \xi_h^{n+1}. \tag{6.68}$$

Remark 6.3.2. *At each time step, problem (6.65)-(6.66) is a generalized Stokes problem with a mixed boundary condition for $(\mathbf{u}, p)$. The well-posedness of the generalized Stokes problem follows easily with the help of the inf-sup condition.*

6.3.3 Stability analysis of fully discrete finite element methods

The primary goal of this subsection is to derive a discrete energy law which mimics the PDE energy law (6.28). It turns out that such a discrete energy law only holds if h and Δt satisfy the mesh constraint $\Delta t = O(h^2)$ when $\theta = 0$ but for all $h, \Delta t > 0$ when $\theta = 1$.

Before discussing the stability of (FEA), We first show that the numerical solution satisfies all side constraints which are fulfilled by the PDE solution.

Lemma 6.3.3. *Let $\{(\mathbf{u}_h^n, \xi_h^n, \eta_h^n)\}_{n\geq 0}$ be defined by the (FEA), then there hold*

$$(\eta_h^n, 1) = C_\eta(t_n) \qquad \text{for } n = 0, 1, 2, \cdots, \tag{6.69}$$

$$(\xi_h^n, 1) = C_\xi(t_{n-1+\theta}) \qquad \text{for } n = 1-\theta, 1, 2, \cdots, \tag{6.70}$$

$$\langle \mathbf{u}_h^n \cdot \mathbf{n}, 1\rangle = C_{\mathbf{u}}(t_{n-1+\theta}) \qquad \text{for } n = 1-\theta, 1, 2, \cdots. \tag{6.71}$$

Proof. Taking $\psi_h = 1$ in (6.67) yields

$$\big(d_t \eta_h^{n+1}, 1\big) = (\phi, 1) + \langle \phi_1, 1\rangle.$$

Then summing over n from 0 to $\ell\,(\geq 0)$ we get

$$(\eta_h^{\ell+1}, 1) = (\eta_h^0, 1) + \big[(\phi, 1) + \langle \phi_1, 1\rangle\big] t_{\ell+1} = (\eta_0, 1) + \big[(\phi, 1) + \langle \phi_1, 1\rangle\big] t_{\ell+1} = C_\eta(t_{\ell+1})$$

for $\ell = 0, 1, 2, \cdots$. So (6.69) holds.

To prove (6.70), taking $\mathbf{v}_h = \mathbf{x}$ in (6.65) and $\varphi_h = 1$ in (6.66), we get

$$\mu\big(\operatorname{div}\mathbf{u}_h^{n+1}, 1\big) - d\big(\xi_h^{n+1}, 1\big) = \big(\mathbf{f}, x\big) + \langle \mathbf{f}_1, x\rangle, \tag{6.72}$$

$$\kappa_3\big(\xi_h^{n+1}, 1\big) + \big(\operatorname{div}\mathbf{u}_h^{n+1}, 1\big) = \kappa_1 C_\eta(t_{n+\theta}). \tag{6.73}$$

Substituting (6.73) into (6.72) yields

$$(d + \mu\kappa_3)\big(\xi_h^{n+1}, 1\big) = \mu\kappa_1 C_\eta(t_{n+\theta}) - \big(\mathbf{f}, x\big) - \langle \mathbf{f}_1, x\rangle.$$

Hence, by the definition of $C_\xi(t)$ we conclude that (6.70) holds for all $n \geq 1-\theta$.

(6.71) follows from (6.69), (6.70), (6.73), and an application of the Divergence Theorem. The proof is complete. □

The next lemma establishes an identity which mimics the continuous energy law for the solution of (FEA).

Lemma 6.3.4. *Let $\{(\mathbf{u}_h^n, \xi_h^n, \eta_h^n)\}_{n\geq 0}$ be defined by (FEA), then there holds the following identity:*

$$J_{h,\theta}^\ell + S_{h,\theta}^\ell = J_{h,\theta}^0 \qquad \text{for } \ell \geq 1,\ \theta = 0, 1, \tag{6.74}$$

where

$$J_{h,\theta}^\ell := \frac{1}{2}\bigg[\mu\|\varepsilon(\mathbf{u}_h^{\ell+1})\|_{L^2(\mathcal{D})}^2 + \kappa_2\|\eta_h^{\ell+\theta}\|_{L^2(\mathcal{D})}^2 + \kappa_3\|\xi_h^{\ell+1}\|_{L^2(\mathcal{D})}^2 - 2\big(\mathbf{f}, \mathbf{u}_h^{\ell+1}\big) - 2\big\langle \mathbf{f}_1, \mathbf{u}_h^{\ell+1}\big\rangle\bigg],$$

$$\begin{aligned} S_{h,\theta}^\ell := \Delta t \sum_{n=0}^{\ell} \bigg[& \frac{\mu\Delta t}{2}\|d_t\varepsilon(\mathbf{u}_h^{n+1})\|_{L^2(\mathcal{D})}^2 + \frac{K}{\mu_f}\big(\nabla p_h^{n+1} - \rho_f\mathbf{g}, \nabla p_h^{n+1}\big) \\ & + \frac{\kappa_2\Delta t}{2}\|d_t\eta_h^{n+\theta}\|_{L^2(\mathcal{D})}^2 + \frac{\kappa_3\Delta t}{2}\|d_t\xi_h^{n+1}\|_{L^2(\mathcal{D})}^2 - (\phi, p_h^{n+1}) - \langle\phi_1, p_h^{n+1}\rangle \\ & - (1-\theta)\frac{\kappa_1 K\Delta t}{\mu_f}\big(d_t\nabla\xi_h^{n+1}, \nabla p_h^{n+1}\big)\bigg]. \end{aligned}$$

$$p_h^{n+1} := \kappa_1\xi_h^{n+1} + \kappa_2\eta_h^{n+\theta}.$$

Proof. Since the proof for the case $\theta = 1$ is exactly same as that of the PDE energy law, so we omit it and leave it to the interested reader to explore. Here we only consider the case $\theta = 0$. Based on (6.66), we can define η_h^{-1} by

$$\kappa_1\big(\eta_h^{-1}, \varphi_h\big) = \kappa_3\big(\xi_h^0, \varphi_h\big) + \big(\operatorname{div} \mathbf{u}_h^0, \varphi_h\big) \tag{6.75}$$

Setting $\mathbf{v}_h = d_t\mathbf{u}_h^{n+1}$ in (6.65), $\varphi_h = \xi_h^{n+1}$ in (6.66), and $\psi_h = p_h^{n+1}$ in (6.67) after lowing the degree from $n+1$ to n, we get

$$\begin{aligned}
&\frac{\mu}{2} d_t \|\varepsilon(\mathbf{u}_h^{n+1})\|_{L^2(\mathcal{D})}^2 + \frac{\mu}{2}\Delta t \|d_t\varepsilon(\mathbf{u}_h^{n+1})\|_{L^2(\mathcal{D})}^2 \qquad (6.76)\\
&\qquad = d_t(\mathbf{f}, \mathbf{u}_h^{n+1}) + d_t\langle \mathbf{f}_1, \mathbf{u}_h^{n+1}\rangle + (\xi_h^{n+1}, \operatorname{div} d_t\mathbf{u}_h^{n+1}),\\
&\kappa_3\big(d_t\xi_h^{n+1}, \xi_h^{n+1}\big) + \big(\operatorname{div} d_t\mathbf{u}_h^{n+1}, \xi_h^{n+1}\big) = \kappa_1\big(d_t\eta_h^n, \xi_h^{n+1}\big), \qquad (6.77)\\
&\big(d_t\eta_h^n, p_h^{n+1}\big) + \frac{1}{\mu_f}\big(K(\nabla(\kappa_1\xi_h^n + \kappa_2\eta_h^n) - \rho_f\mathbf{g}), \nabla p_h^{n+1}\big) \qquad (6.78)\\
&\qquad = (\phi, p_h^{n+1}) + \langle \phi_1, p_h^{n+1}\rangle.
\end{aligned}$$

The first term on the left-hand side of (6.78) can be rewritten as

$$\begin{aligned}
\big(d_t\eta_h^n, p_h^{n+1}\big) &= \big(d_t\eta_h^n, \kappa_1\xi_h^{n+1} + \kappa_2\eta_h^n\big) \qquad (6.79)\\
&= \kappa_1\big(d_t\eta_h^n, \xi_h^{n+1}\big) + \frac{\kappa_2\Delta t}{2}\|d_t\eta_h^n\|_{L^2(\mathcal{D})}^2 + \frac{\kappa_2}{2} d_t\|\eta_h^n\|_{L^2(\mathcal{D})}^2.
\end{aligned}$$

Moreover,

$$\begin{aligned}
&\frac{K}{\mu_f}\big(\nabla(\kappa_1\xi_h^n + \kappa_2\eta_h^n) - \rho_f\mathbf{g}, \nabla p_h^{n+1}\big) \qquad (6.80)\\
&\qquad = \frac{K}{\mu_f}(\nabla p_h^{n+1} - \rho_f\mathbf{g}, \nabla p_h^{n+1}) - \frac{\kappa_1 K\Delta t}{\mu_f}\big(d_t\nabla\xi_h^{n+1}, \nabla p_h^{n+1}\big).\\
&\kappa_3\big(d_t\xi_h^{n+1}, \xi_h^{n+1}\big) = \frac{\kappa_3}{2} d_t\|\xi_h^{n+1}\|_{L^2(\mathcal{D})}^2 + \frac{\kappa_3\Delta t}{2}\|d_t\xi_h^{n+1}\|_{L^2(\mathcal{D})}^2. \qquad (6.81)
\end{aligned}$$

Adding (6.76)–(6.78), using (6.79)–(6.81) and applying the summation operator $\Delta t\sum_{n=0}^{\ell}$ to the both sides of the resulting equation yield the desired equality (6.74). The proof is complete. □

In the case $\theta = 1$, (6.74) gives the desired solution estimates without any mesh constraint. On the other hand, when $\theta = 0$, since the last term in the expression of $S_{h,\theta}^{\ell}$ does not have a fixed sign, hence, it needs to be controlled in order to ensure the positivity of $S_{h,\theta}^{\ell}$.

Corollary 6.3.5. *Let $\{(\mathbf{u}_h^n, \xi_h^n, \eta_h^n)\}_{n\geq 0}$ be defined by (FEA) with $\theta = 0$, then there holds the following inequality:*

$$J_{h,0}^{\ell} + \widehat{S}_{h,0}^{\ell} \leq J_{h,0}^{0} \qquad \text{for } \ell \geq 1, \tag{6.82}$$

provided that $\Delta t = O(h^2)$. Where

$$\begin{aligned}\widehat{S}_{h,0}^{\ell} := \Delta t \sum_{n=0}^{\ell} \Bigg[&\frac{\mu\Delta t}{4}\|d_t\varepsilon(\mathbf{u}_h^{n+1})\|_{L^2(\mathcal{D})}^2 + \frac{K}{2\mu_f}\|\nabla p_h^{n+1}\|_{L^2(\mathcal{D})}^2 - \frac{K}{\mu_f}\big(\rho_f\mathbf{g}, \nabla p_h^{n+1}\big) \\ &+ \frac{\kappa_2\Delta t}{2}\|d_t\eta_h^n\|_{L^2(\mathcal{D})}^2 + \frac{\kappa_3\Delta t}{2}\|d_t\xi_h^{n+1}\|_{L^2(\mathcal{D})}^2 - (\phi, p_h^{n+1}) - \langle\phi_1, p_h^{n+1}\rangle\Bigg].\end{aligned}$$

Proof. By Schwarz inequality and inverse inequality (6.85), we get

$$\begin{aligned}\frac{\kappa_1 K\Delta t}{\mu_f}\big(d_t\nabla\xi_h^{n+1}, \nabla p_h^{n+1}\big) &\leq \frac{\kappa_1^2 K}{2\mu_f}\|\nabla\xi_h^{n+1} - \nabla\xi_h^n\|_{L^2(\mathcal{D})}^2 + \frac{K}{2\mu_f}\|\nabla p_h^{n+1}\|_{L^2(\mathcal{D})}^2 \\ &\leq \frac{c_1^2\kappa_1^2 K}{2\mu_f h^2}\|\xi_h^{n+1} - \xi_h^n\|_{L(\mathcal{D})2}^2 + \frac{K}{2\mu_f}\|\nabla p_h^{n+1}\|_{L^2(\mathcal{D})}^2.\end{aligned} \tag{6.83}$$

To bound the first term on the right-hand side of (6.83), we appeal to the inf-sup condition and get

$$\begin{aligned}\|\xi_h^{n+1}-\xi_h^n\|_{L^2} &\le \frac{1}{\beta_1}\sup_{v_h\in\mathbf{V}_h}\frac{\big(\operatorname{div}\mathbf{v}_h,\xi_h^{n+1}-\xi_h^n\big)}{\|\nabla\mathbf{v}_h\|_{L^2(\mathcal{D})}}\\ &\le \frac{\mu}{\beta_1}\sup_{v_h\in\mathbf{V}_h}\frac{\big(\varepsilon(\mathbf{u}^{n+1}-\mathbf{u}^n),\varepsilon(\mathbf{v}_h)\big)}{\|\nabla\mathbf{v}_h\|_{L^2(\mathcal{D})}}\\ &\le \frac{\mu}{\beta_1}\Delta t\|d_t\varepsilon(\mathbf{u}_h^{n+1})\|_{L^2}.\end{aligned} \tag{6.84}$$

Substituting (6.84) into (6.83) and combining it with (6.74) imply (6.82) provided that $\Delta t \le (\mu_f\beta_1^2)(2\mu K c_1^2\kappa_1^2)^{-1}h^2$. The proof is complete. □

6.3.4 Convergence analysis

The goal of this section is to analyze the fully discrete finite element algorithm (FEA) proposed in the previous subsection. Precisely, we shall derive optimal order error estimates for (FEA) in both $L^\infty(0,T;L^2(\mathcal{D})$ and $L^2(0,T;H^1(\mathcal{D}))$-norm. To the end, we first list some facts, which are well known in the literature [13, 14], about finite element functions.

We first recall the following inverse inequality for polynomial functions [23]:

$$\|\nabla\varphi_h\|_{L^2(K)} \le c_1 h^{-1}\|\varphi_h\|_{L^2(K)} \qquad \forall\varphi_h\in P_r(K), K\in T_h. \tag{6.85}$$

For any $\varphi\in L^2(\mathcal{D})$, we define its L^2-projection $\mathcal{Q}_h : L^2\to W_h$ as

$$\big(\mathcal{Q}_h\varphi,\psi_h\big) = \big(\varphi,\psi_h\big) \qquad \psi_h\in W_h. \tag{6.86}$$

It is well known that the projection operator $\mathcal{Q}_h : L^2\to W_h$ satisfies (cf [13]), for any $\varphi\in H^s(\mathcal{D})(s>1)$,

$$\|\mathcal{Q}_h\varphi-\varphi\|_{L^2(\mathcal{D})} + h\|\nabla(\mathcal{Q}_h\varphi-\varphi)\|_{L^2(\mathcal{D})} \le Ch^\ell\|\varphi\|_{H^\ell(\mathcal{D})}, \quad \ell=\min\{2,s\}. \tag{6.87}$$

We like to point out that when $W_h \notin H^1(\mathcal{D})$, the second term on the left-hand side of (6.87) has to be replaced by the broken H^1-norm.

Next, for any $\varphi \in H^1(\mathcal{D})$, we define its elliptic projection $\mathcal{S}_h\varphi$ by

$$\big(K\nabla\mathcal{S}_h\varphi, \nabla\varphi_h\big) = \big(K\nabla\varphi, \nabla\varphi_h\big) \quad \varphi_h \in W_h, \tag{6.88}$$

$$\big(\mathcal{S}_h\varphi, 1\big) = \big(\varphi, 1\big). \tag{6.89}$$

It is well known that the projection operator $\mathcal{S}_h : H^1(\mathcal{D}) \to W_h$ satisfies (cf [13]), for any $\varphi \in H^s(\mathcal{D})(s > 1)$,

$$\|\mathcal{S}_h\varphi - \varphi\|_{L^2(\mathcal{D})} + h\|\nabla(\mathcal{S}_h\varphi - \varphi)\|_{L^2(\mathcal{D})} \le Ch^\ell\|\varphi\|_{H^\ell(\mathcal{D})}, \quad \ell = \min\{2, s\}. \tag{6.90}$$

Finally, for any $\mathbf{v} \in \mathbf{H}^1_\perp(\mathcal{D})$, we define its elliptic projection $\mathcal{R}_h\mathbf{v}$ by

$$\big(\varepsilon(\mathcal{R}_h\mathbf{v}), \varepsilon(\mathbf{w}_h)\big) = \big(\varepsilon(\mathbf{v}), \varepsilon(\mathbf{w}_h)\big) \quad \mathbf{w}_h \in \mathbf{V}_h. \tag{6.91}$$

It is easy to show that the projection $\mathcal{R}_h\mathbf{v}$ satisfies (cf [13]), for any $\mathbf{v} \in \mathbf{H}^1_\perp(\mathcal{D}) \cap \mathbf{H}^s(\mathcal{D})(s > 1)$,

$$\|\mathcal{R}_h\mathbf{v} - \mathbf{v}\|_{L^2(\mathcal{D})} + h\|\nabla(\mathcal{R}_h\mathbf{v} - \mathbf{v})\|_{L^2(\mathcal{D})} \le Ch^m\|\mathbf{v}\|_{H^m(\mathcal{D})}, \quad m = \min\{3, s\}. \tag{6.92}$$

To derive error estimates, we introduce the following error notation

$$E^n_{\mathbf{u}} := \mathbf{u}(t_n) - \mathbf{u}^n_h, \qquad E^n_\xi := \xi(t_n) - \xi^n_h, \qquad E^n_\eta := \eta(t_n) - \eta^n_h,$$
$$E^n_p := p(t_n) - p^n_h, \qquad E^n_q := q(t_n) - q^n_h.$$

It is easy to check that

$$E^n_p = \kappa_1 E^n_\xi + \kappa_2 E^n_\eta, \qquad E^n_q = \kappa_3 E^n_\xi + \kappa_1 E^n_\eta. \tag{6.93}$$

Also, we denote

$$
\begin{aligned}
E_{\mathbf{u}}^n &= \mathbf{u}(t_n) - \mathcal{R}_h(\mathbf{u}(t_n)) + \mathcal{R}_h(\mathbf{u}(t_n)) - \mathbf{u}_h^n := \Lambda_{\mathbf{u}}^n + \Theta_{\mathbf{u}}^n,\\
E_{\xi}^n &= \xi(t_n) - \mathcal{Q}_h(\xi(t_n)) + \mathcal{Q}_h(\xi(t_n)) - \xi_h^n := \Lambda_{\xi}^n + \Theta_{\xi}^n,\\
E_{\eta}^n &= \eta(t_n) - \mathcal{Q}_h(\eta(t_n)) + \mathcal{Q}_h(\eta(t_n)) - \eta_h^n := \Lambda_{\eta}^n + \Theta_{\eta}^n,\\
E_{p}^n &= p(t_n) - \mathcal{Q}_h(p(t_n)) + \mathcal{Q}_h(p(t_n)) - p_h^n := \Lambda_{p}^n + \Theta_{p}^n.
\end{aligned}
$$

Lemma 6.3.6. *Let $\{(\mathbf{u}_h^n, \xi_h^n, \eta_h^n)\}_{n\geq 0}$ be generated by the (FEA) and $\Lambda_{\mathbf{u}}^n, \Theta_{\mathbf{u}}^n, \Lambda_{\xi}^n, \Theta_{\xi}^n, \Lambda_{\eta}^n$ and Θ_{η}^n be defined as above. Then there holds the following identity:*

$$
\begin{aligned}
&\mathcal{E}_h^{\ell} + \Delta t \sum_{n=0}^{\ell} \Big[\frac{K}{\mu_f} \big(\nabla \hat{\Theta}_p^{n+1} - \rho_f \mathbf{g}, \hat{\Theta}_p^{n+1} \big) \\
&\qquad + \frac{\Delta t}{2} \Big(\mu \| d_t \varepsilon(\Theta_{\mathbf{u}}^{n+1}) \|_{L^2(\mathcal{D})}^2 + \kappa_2 \| d_t \Theta_{\eta}^{n+\theta} \|_{L^2(\mathcal{D})}^2 + \kappa_3 \| d_t \Theta_{\xi}^{n+1} \|_{L^2(\mathcal{D})}^2 \Big) \Big] \\
&= \mathcal{E}_h^0 + \Delta t \sum_{n=0}^{\ell} \Big[\big(\Lambda_{\xi}^{n+1}, \operatorname{div} d_t \Theta_{\mathbf{u}}^{n+1} \big) - \big(\operatorname{div} d_t \Lambda_{\mathbf{u}}^{n+1}, \Theta_{\xi}^{n+1} \big) \Big] \\
&\qquad + (\Delta t)^2 \sum_{n=0}^{\ell} \big(d_t^2 \eta_h(t_{n+1}), \Theta_{\xi}^{n+1} \big) + \Delta t \sum_{n=0}^{\ell} \big(R_h^{n+1}, \hat{\Theta}_p^{n+1} \big) \\
&\qquad + (1-\theta)(\Delta t)^2 \sum_{n=0}^{\ell} \frac{K \kappa_1}{\mu_f} \big(d_t \nabla \Theta_{\xi}^{n+1}, \nabla \hat{\Theta}_p^{n+1} \big),
\end{aligned}
\tag{6.94}
$$

where

$$
\hat{\Theta}_p^{n+1} := \kappa_1 \nabla \Theta_{\xi}^{n+1} + \kappa_2 \nabla \Theta_{\eta}^{n+\theta} \tag{6.95}
$$

$$
\mathcal{E}_h^{\ell} := \frac{1}{2} \Big[\mu \| \varepsilon(\Theta_{\mathbf{u}}^{\ell+1}) \|_{L^2(\mathcal{D})}^2 + \kappa_2 \| \Theta_{\eta}^{\ell+\theta} \|_{L^2(\mathcal{D})}^2 + \kappa_3 \| \Theta_{\xi}^{\ell+1} \|_{L^2(\mathcal{D})}^2 \Big], \tag{6.96}
$$

$$
R_h^{n+1} := -\frac{1}{\Delta t} \int_{t_n}^{t_{n+1}} (s - t_n) \eta_{tt}(s)\, ds. \tag{6.97}
$$

Proof. Subtracting (6.65) from (6.21), (6.66) from (6.22), (6.67) from (6.23), respectively, we get the following error equations:

$$\mu\big(\varepsilon(\mathbf{E}_{\mathbf{u}}^{n+1}),\varepsilon(\mathbf{v}_h)\big)-\big(E_{\xi}^{n+1},\operatorname{div}\mathbf{v}_h\big)=0 \qquad \forall \mathbf{v}_h\in V_h, \tag{6.98}$$

$$\begin{aligned}\kappa_3\big(E_{\xi}^{n+1},\varphi_h\big)+\big(\operatorname{div}\mathbf{E}_{\mathbf{u}}^{n+1},\varphi_h\big)&\\ =\kappa_1\big(E_{\eta}^{n+\theta},\varphi_h\big)+\Delta t\big(d_t\eta(t_{n+1}),\varphi_h\big) &\qquad \forall\varphi_h\in M_h,\end{aligned} \tag{6.99}$$

$$\begin{aligned}\big(d_tE_{\eta}^{n+1},\psi_h\big)+\frac{K}{\mu_f}\big(\nabla(\kappa_1E_{\xi}^{n+1}+\kappa_2E_{\eta}^{n+1})-\rho_f\mathbf{g},\nabla\psi_h\big)&\\ =(R_h^{n+1},\psi_h) &\qquad \forall\psi_h\in W_h.\end{aligned} \tag{6.100}$$

Using the definition of the projection operators $\mathcal{Q}_h,\mathcal{S}_h,\mathcal{R}_h$, we have

$$\mu\big(\varepsilon(\Theta_{\mathbf{u}}^{n+1}),\varepsilon(\mathbf{v}_h)\big)-\big(\Theta_{\xi}^{n+1},\operatorname{div}\mathbf{v}_h\big)=(\Lambda_{\xi}^{n+1},\operatorname{div}\mathbf{v}_h), \qquad \forall\mathbf{v}_h\in V_h, \tag{6.101}$$

$$\begin{aligned}\kappa_3\big(\Theta_{\xi}^{n+1},\varphi_h\big)+\big(\operatorname{div}\Theta_{\mathbf{u}}^{n+1},\varphi_h\big)=\kappa_1\big(\Theta_{\eta}^{n+\theta},\varphi_h\big)&\\ -\big(\operatorname{div}\Lambda_{\mathbf{u}}^{n+1},\varphi_h\big)+\Delta t\big(d_t\eta(t_{n+1}),\varphi_h\big) &\qquad \forall\varphi_h\in M_h,\end{aligned} \tag{6.102}$$

$$\begin{aligned}\big(d_t\Theta_{\eta}^{n+1},\psi_h\big)+\frac{K}{\mu_f}\big(\nabla(\kappa_1\Theta_{\xi}^{n+1}+\kappa_2\Theta_{\eta}^{n+1})-\rho_f\mathbf{g},\nabla\psi_h\big)&\\ =\big(R_h^{n+1},\psi_h\big) &\qquad \forall\psi_h\in W_h.\end{aligned} \tag{6.103}$$

(6.94) follows from setting $\mathbf{v}_h=d_t\Theta_{\mathbf{u}}^{n+1}$ in (6.101), $\varphi_h=\Theta_{\xi}^{n+1}$ (after applying the difference operator d_t to the equation (6.102)), $\psi_h=\hat{\Theta}_p^{n+1}=\kappa_1\Theta_{\xi}^{n+1}+\kappa_2\Theta_{\eta}^{n+\theta}$ in (6.103), adding the resulting equations, and applying the summation operator $\Delta t\sum_{n=0}^{\ell}$ to both sides. □

Theorem 6.3.7. *Let $\{(u_h^n,\xi_h^n,\eta_h^n)\}_{n\geq 0}$ be defined by (FEA), then there holds the error estimate for $\ell\leq N$*

$$\begin{aligned}\max_{0\leq n\leq\ell}\Big[\sqrt{\mu}\|\varepsilon(\Theta_{\mathbf{u}}^{n+1})\|_{L^2(\mathcal{D})}+\sqrt{\kappa_2}\|\Theta_{\eta}^{n+\theta}\|_{L^2(\mathcal{D})}+\sqrt{\kappa_3}\|\Theta_{\xi}^{n+1}\|_{L^2(\mathcal{D})}\Big]&\\ +\Big[\Delta t\sum_{n=0}^{\ell}\frac{K}{\mu_f}\|\hat{\Theta}_p^{n+1}\|_{L^2}^2\Big]^{\frac{1}{2}}\leq C_1(T)\Delta t+C_2(T)h^2,&\end{aligned} \tag{6.104}$$

provided that $\Delta t = O(h^2)$ *when* $\theta = 0$. *Where*

$$C_1(T) = C\|q_t\|^2_{L^2((0,T);L^2(\mathcal{D}))} + C\|(q)_{tt}\|_{L^2((0,T);H^{-1}(\mathcal{D}))}, \tag{6.105}$$

$$C_2(T) = C\|\xi\|_{L^\infty((0,T);H^2(\mathcal{D}))} + C\|\xi_t\|_{L^2((0,T);H^2(\mathcal{D}))} + C\|\mathrm{div}\,(\mathbf{u})_t\|_{L^2((0,T);H^2(\mathcal{D}))}. \tag{6.106}$$

Proof. To derive the above inequality, we need to bound each term on the right-hand side of (6.94). Using the fact $\Theta^0_{\mathbf{u}} = \mathbf{0}$, $\Theta^0_\xi = 0$ and $\Theta^{-1}_\eta = 0$, we have

$$\begin{aligned}
&\mathcal{E}^\ell_h + \Delta t \sum_{n=0}^{\ell} \Big[\frac{K}{\mu_f}\big(\nabla\hat{\Theta}^{n+1}_p - \rho_f \mathbf{g}, \nabla\hat{\Theta}^{n+1}_p\big) \\
&\quad + \frac{\Delta t}{2}\Big(\mu\|d_t\varepsilon(\Theta^{n+1}_{\mathbf{u}})\|^2_{L^2(\mathcal{D})} + \kappa_2\|d_t\Theta^{n+\theta}_\eta\|^2_{L^2(\mathcal{D})} + \kappa_3\|d_t\Theta^{n+1}_\xi\|^2_{L^2(\mathcal{D})}\Big)\Big] \\
&= (\Delta t)^2 \sum_{n=0}^{\ell}\big(d_t^2\eta(t_{n+1}), \Theta^{n+1}_\xi\big) + \Delta t\sum_{n=0}^{\ell}\big(R^{n+1}_h, \hat{\Theta}^{n+1}_p\big) + \frac{\mu}{2}\|\varepsilon(\Theta^1_{\mathbf{u}})\|^2_{L^2(\mathcal{D})} \\
&\quad + \Delta t\sum_{n=0}^{\ell}\Big[\big(\Lambda^{n+1}_\xi, \mathrm{div}\, d_t\Theta^{n+1}_{\mathbf{u}}\big) - \big(\mathrm{div}\, d_t\Lambda^{n+1}_{\mathbf{u}}, \Theta^{n+1}_\xi\big)\Big] + \frac{\kappa_2}{2}\|\Theta^\theta_\eta\|^2_{L^2(\mathcal{D})} \\
&\quad + (1-\theta)(\Delta t)^2\sum_{n=0}^{\ell}\frac{K\kappa_1}{\mu_f}\big(d_t\nabla\Theta^{n+1}_\xi, \nabla\hat{\Theta}^{n+1}_p\big) + \frac{\kappa_3}{2}\|\Theta^1_\xi\|^2_{L^2(\mathcal{D})}.
\end{aligned} \tag{6.107}$$

We now estimate each term on the right-hand side of (6.107). To bound the first term on the right-hand side of (6.107), we first use the summation by parts formula and $d_t\eta_h(t_0) = 0$ to get

$$\sum_{n=0}^{\ell}\big(d_t^2\eta(t_{n+1}), \Theta^{n+1}_\xi\big) = \frac{1}{\Delta t}\big(d_t\eta(t_{l+1}), \Theta^{l+1}_\xi\big) - \sum_{n=1}^{\ell}\big(d_t\eta(t_n), d_t\Theta^{n+1}_\xi\big). \tag{6.108}$$

Now, we bound each term on the right-hand side of (6.108) as follows:

$$\begin{aligned}
\frac{1}{\Delta t}\big(d_t\eta(t_{\ell+1}),\Theta_\xi^{\ell+1}\big) &\le \frac{1}{\Delta t}\|d_t\eta(t_{\ell+1})\|_{L^2(\mathcal{D})}\|\Theta_\xi^{\ell+1}\|_{L^2(\mathcal{D})} \qquad (6.109)\\
&\le \frac{1}{\Delta t}\|\eta_t\|_{L^2((t_\ell,t_{\ell+1});L^2(\mathcal{D}))}\cdot\frac{1}{\beta_1}\sup_{\mathbf{v}_h\in\mathbf{V}_h}\frac{\mu\big(\varepsilon(\Theta_{\mathbf{u}}^{\ell+1}),\varepsilon(\mathbf{v}_h)\big)-(\Lambda_\xi^{\ell+1},\operatorname{div}\mathbf{v}_h)}{\|\nabla\mathbf{v}_h\|_{L^2(\mathcal{D})}}\\
&\le \frac{C\mu}{\beta_1\Delta t}\|\eta_t\|_{L^2((t_\ell,t_{\ell+1});L^2(\mathcal{D}))}\Big[\|\varepsilon(\Theta_{\mathbf{u}}^{\ell+1})\|_{L^2(\mathcal{D})}+\frac{1}{\mu}\|\Lambda_\xi^{\ell+1}\|_{L^2(\mathcal{D})}\Big]\\
&\le \frac{\mu}{4(\Delta t)^2}\|\varepsilon(\Theta_{\mathbf{u}}^{\ell+1})\|_{L^2(\mathcal{D})}^2+\frac{C\mu}{\beta_1^2}\|\eta_t\|_{L^2((t_\ell,t_{\ell+1});L^2(\mathcal{D}))}^2+\frac{C}{\beta_1^2}\|\Lambda_\xi^{\ell+1}\|_{L^2(\mathcal{D})}^2,\\
\sum_{n=0}^{\ell}\big(d_t\eta(t_n),d_t\Theta_\xi^{n+1}\big) &\le \sum_{n=0}^{\ell}\|d_t\eta(t_n)\|_{L^2(\mathcal{D})}\|d_t\Theta_\xi^{n+1}\|_{L^2(\mathcal{D})} \qquad (6.110)\\
&\le \sum_{n=0}^{\ell}\|d_t\eta(t_n)\|_{L^2(\mathcal{D})}\cdot\frac{1}{\beta_1}\sup_{\mathbf{v}_h\in\mathbf{V}_h}\frac{\mu\big(d_t\varepsilon(\Theta_{\mathbf{u}}^{n+1}),\varepsilon(\mathbf{v}_h)\big)-(d_t\Lambda_\xi^{n+1},\operatorname{div}\mathbf{v}_h)}{\|\nabla\mathbf{v}_h\|_{L^2(\mathcal{D})}}\\
&\le \frac{C\mu}{\beta_1}\sum_{n=0}^{\ell}\|d_t\eta(t_n)\|_{L^2(\mathcal{D})}\Big[\|d_t\varepsilon(\Theta_{\mathbf{u}}^{n+1})\|_{L^2(\mathcal{D})}+\frac{1}{\mu}\|d_t\Lambda_\xi^{n+1}\|_{L^2(\mathcal{D})}\Big]\\
&\le \frac{\mu}{4}\sum_{n=0}^{\ell}\|d_t\varepsilon(\Theta_{\mathbf{u}}^{n+1})\|_{L^2(\mathcal{D})}^2+\frac{C\mu}{\beta_1^2}\|\eta_t\|_{L^2((0,T);L^2(\mathcal{D}))}^2+\frac{C}{\beta_1^2}\|d_t\Lambda_\xi^{n+1}\|_{L^2(\mathcal{D})}.
\end{aligned}$$

The second term on the right-hand side of (6.107) can be bounded as

$$\begin{aligned}
\big|\big(R_h^{n+1},\hat\Theta_p^{n+1}\big)\big| &\le \|R_h^{n+1}\|_{H^{-1}(\mathcal{D})}\|\nabla\hat\Theta_p^{n+1}\|_{L^2(\mathcal{D})} \qquad (6.111)\\
&\le \frac{K}{4\mu_f}\|\nabla\hat\Theta_p^{n+1}\|_{L^2(\mathcal{D})}^2+\frac{\mu_f}{K}\|R_h^{n+1}\|_{H^{-1}(\mathcal{D})}^2\\
&\le \frac{K}{4\mu_f}\|\nabla\hat\Theta_p^{n+1}\|_{L^2(\mathcal{D})}^2+\frac{\mu_f\Delta t}{3K}\|\eta_{tt}\|_{L^2((t_n,t_{n+1});H^{-1}(\mathcal{D}))}^2,
\end{aligned}$$

where we have used the fact that

$$\|R_h^{n+1}\|_{H^{-1}(\mathcal{D})}^2\le\frac{\Delta t}{3}\int_{t_n}^{t_{n+1}}\|\eta_{tt}\|_{H^{-1}(\mathcal{D})}^2\,dt.$$

The fourth term on the right-hand side of (6.107) can be bounded by

$$\begin{aligned}
&\Delta t\sum_{n=0}^{\ell}\Big[\big(\Lambda_\xi^{n+1},\operatorname{div} d_t\Theta_{\mathbf{u}}^{n+1}\big)-\big(\operatorname{div} d_t\Lambda_{\mathbf{u}}^{n+1},\Theta_\xi^{n+1}\big)\Big] \qquad (6.112)\\
&\quad=(\Lambda_\xi^{\ell+1},\operatorname{div}\Theta_{\mathbf{u}}^{\ell+1})-\Delta t\sum_{n=0}^{\ell}\Big[\big(d_t\Lambda_\xi^{n+1},\operatorname{div}\Theta_{\mathbf{u}}^{n}\big)+\big(\operatorname{div} d_t\Lambda_{\mathbf{u}}^{n+1},\Theta_\xi^{n+1}\big)\Big]\\
&\quad\le\frac{1}{2}\|\Lambda_\xi^{\ell+1}\|_{L^2(\mathcal{D})}^2+\frac{1}{2}\|\operatorname{div}\Theta_{\mathbf{u}}^{\ell+1}\|_{L^2(\mathcal{D})}^2+\frac{1}{2}\Delta t\sum_{n=0}^{\ell}\Big[\|d_t\Lambda_\xi^{n+1}\|_{L^2(\mathcal{D})}^2\\
&\qquad+\|\varepsilon(\Theta_{\mathbf{u}}^{n+1})\|_{L^2(\mathcal{D})}^2+\|\operatorname{div} d_t\Lambda_{\mathbf{u}}^{n+1}\|_{L^2(\mathcal{D})}^2+\|\Theta_\xi^{n+1}\|_{L^2(\mathcal{D})}^2\Big].
\end{aligned}$$

When $\theta=0$ we also need to bound the last term on the right-hand side of (6.107), which is carried out below.

$$\begin{aligned}
&\sum_{n=0}^{\ell}\big(d_t\nabla\Theta_\xi^{n+1},\nabla\hat{\Theta}_p^{n+1}\big)\le\sum_{n=0}^{\ell}\|d_t\Theta_\xi^{n+1}\|_{L^2(\mathcal{D})}\|\nabla\hat{\Theta}_p^{n+1}\|_{L^2(\mathcal{D})} \qquad (6.113)\\
&\le\sum_{n=0}^{\ell}\sup_{\mathbf{v}_h\in\mathbf{V}_h}\frac{\mu\big(d_t\varepsilon(\Theta_{\mathbf{u}}^{n+1}),\varepsilon(\mathbf{v}_h)\big)-\big(d_t\Lambda_\xi^{n+1},\operatorname{div}\mathbf{v}_h\big)}{\|\nabla\mathbf{v}_h\|_{L^2(\mathcal{D})}}\|\nabla\hat{\Theta}_p^{n+1}\|_{L^2(\mathcal{D})}\\
&\le\sum_{n=0}^{\ell}\Big[\frac{\mu^2\kappa_1\Delta t}{h^2\beta_1^2}\|d_t\varepsilon(\Theta_{\mathbf{u}}^{n+1})\|_{L^2}^2+\frac{\kappa_1\Delta t}{h^2\beta_1^2}\|d_t\Lambda_\xi^{n+1}\|_{L^2}+\frac{\kappa_1^{-1}}{4\Delta t}\|\nabla\hat{\Theta}_p^{n+1}\|_{L^2(\mathcal{D})}^2\Big].
\end{aligned}$$

Substituting (6.108)–(6.113) into (6.107) and rearranging terms we get

$$\begin{aligned}
&\mu\|\varepsilon(\Theta_{\mathbf{u}}^{\ell+1})\|_{L^2(\mathcal{D})}^2+\kappa_2\|\Theta_\eta^{\ell+\theta}\|_{L^2(\mathcal{D})}^2+\kappa_3\|\Theta_\xi^{\ell+1}\|_{L^2(\mathcal{D})}^2 \qquad (6.114)\\
&\qquad+\Delta t\sum_{n=0}^{\ell}\frac{K}{\mu_f}\|\nabla\hat{\Theta}_p^{n+1}\|_{L^2(\mathcal{D})}^2\\
&\le\frac{4\mu_f(\Delta t)^2}{\kappa}\|\eta_{tt}\|_{L^2((0,T);H^{-1})}^2+\frac{4\mu(\Delta t)^2}{\beta_1^2}\|\eta_t\|_{L^2((0,T);L^2(\mathcal{D}))}^2,\\
&\qquad+\|\Lambda_\xi^{\ell+1}\|_{L^2(\mathcal{D})}^2+\Delta t\sum_{n=0}^{\ell}\|d_t\Lambda_\xi^{n+1}\|_{L^2(\mathcal{D})}^2+\Delta t\sum_{n=0}^{\ell}\|\operatorname{div} d_t\Lambda_{\mathbf{u}}^{n+1}\|_{L^2(\mathcal{D})}^2,
\end{aligned}$$

provide that $\Delta t \leq \mu_f \beta_1^2 (4\mu\kappa_1^2 K)^{-1} h^2$ when $\theta = 0$, but it holds for all $\Delta t > 0$ when $\theta = 1$. Hence, (6.104) follows from using the approximation properties of the projection operators $\mathcal{Q}_h, \mathcal{R}_h$ and $\mathcal{S}_h$. The proof is complete. □

We conclude this section by stating the main theorem of the section.

Theorem 6.3.8. *The solution of the (FEA) satisfies the following error estimates:*

$$\max_{0\leq n\leq N}\Big[\sqrt{\mu}\|\nabla(u(t_n)-u_h^n)\|_{L^2(\mathcal{D})}+\sqrt{\kappa_2}\|\eta(t_n)-\eta_h^n\|_{L^2(\mathcal{D})} + \sqrt{\kappa_3}\|\xi(t_n)-\xi_h^n\|_{L^2(\mathcal{D})}\Big] \leq \widehat{C}_1(T)\Delta t + \widehat{C}_2(T)h^2. \tag{6.115}$$

$$\left[\Delta t\sum_{n=0}^{\ell}\frac{K}{\mu_f}\|\nabla\hat{\Theta}_p^{n+1}\|_{L^2(\mathcal{D})}^2\right]^{\frac{1}{2}} \leq \widehat{C}_1(T)\Delta t + \widehat{C}_2(T)h, \tag{6.116}$$

provided that $\Delta t = O(h^2)$ *when* $\theta = 0$ *and* $\Delta t > 0$ *when* $\theta = 1$. *Where*

$$\widehat{C}_1(T) := C_1(T),$$
$$\widehat{C}_2(T) := C_2(T) + \|\xi\|_{L^\infty((0,T);H^2(\mathcal{D}))} + \|\eta\|_{L^\infty((0,T);H^2(\mathcal{D}))} + \|\nabla\mathbf{u}\|_{L^\infty((0,T);H^2(\mathcal{D}))}.$$

Proof. The above estimates follow immediately from an application of the triangle inequality on

$$\mathbf{u}(t_n)-\mathbf{u}_h^n = \Lambda_{\mathbf{u}}^n + \Theta_{\mathbf{u}}^n, \quad \xi(t_n)-\xi_h^n = \Lambda_\xi^n + \Theta_\xi^n, \quad \eta(t_n)-\eta_h^n = \Lambda_\eta^n + \Theta_\eta^n,$$

and appealing to (6.87), (6.92) and Theorem 6.3.7. □

6.4 Numerical experiments

In this section we shall present three 2-dimensional numerical experiments to validate theoretical results for the proposed numerical methods, to numerically examine the performances of the approach and methods as well as to compare them with

existing methods in the literature on two benchmark problems. One of these two problems was used to demonstrate the "locking phenomenon" in [77]. Our numerical result shows that such a "locking phenomenon" phenomenon does not occur in our numerical methods, it confirms the fact that our approach and methods have a built-in mechanism to prevent the "locking phenomenon".

Test 1. Let $\mathcal{D} = [0,1] \times [0,1]$, $\Gamma_1 = \{(1, x_2); 0 \le x_2 \le 1\}$, $\Gamma_2 = \{(x_1, 0); 0 \le x_1 \le 1\}$, $\Gamma_3 = \{(0, x_2); 0 \le x_2 \le 1\}$, $\Gamma_4 = \{(x_1, 1); 0 \le x_1 \le 1\}$, and $T = 0.001$. We consider problem (6.23)–(6.26) with following source functions:

$$\begin{aligned}
\mathbf{f} &= -(\lambda + \mu)t(1,1)^T + \alpha \cos(x_1 + x_2)e^t(1,1)^T, \\
\phi &= \Big(c_0 + \frac{2\kappa}{\mu_f}\Big) \sin(x_1 + x_2)e^t + \alpha(x_1 + x_2),
\end{aligned}$$

and the following boundary and initial conditions:

$$\begin{aligned}
p &= \sin(x_1 + x_2)e^t && \text{on } \partial\mathcal{D}_T, \\
u_1 &= \frac{1}{2}x_1^2 t && \text{on } \Gamma_j \times (0,T),\ j = 1, 3, \\
u_2 &= \frac{1}{2}x_2^2 t && \text{on } \Gamma_j \times (0,T),\ j =, 2, 4, \\
\sigma\mathbf{n} - \alpha p\mathbf{n} &= \mathbf{f}_1 && \text{on } \partial\mathcal{D}_T, \\
\mathbf{u}(x,0) = \mathbf{0}, \quad p(x,0) &= \sin(x_1 + x_2) && \text{in } \mathcal{D},
\end{aligned}$$

where

$$\mathbf{f}_1(x,t) = \mu(x_1 n_1, x_2 n_2)^T t + \lambda(x_1 + x_2)(n_1, n_2)^T t - \alpha \sin(x_1 + x_2)(n_1, n_2)^T e^t.$$

It is easy to check that the exact solution for this problem is

$$\mathbf{u}(x,t) = \frac{t}{2}\big(x_1^2, x_2^2\big)^T, \qquad p(x,t) = \sin(x_1 + x_2)e^t.$$

We note that the boundary conditions used above are not pure Neumann conditions, instead, they are mixed Dirichlet-Neumann conditions. As pointed out in Remark 6.2.3 (c), the approach and methods of this chapter also apply to this case, the only change is to replace the test and trial space $\mathbf{H}^1_\perp(\mathcal{D})$ by $\mathbf{H}^1(\mathcal{D})$ with some appropriately built-in Dirichlet boundary condition in Definition 6.2.2.

The goal of doing this test problem is to compute the order of the exact errors and to show that the theoretical error bounds proved in the previous section are sharp.

Table 6.1 displays the computed $L^\infty(0,T;L^2(\mathcal{D}))$ and $L^2(0,T;H^1(\mathcal{D}))$-norm errors and the convergence rates with respect to h at the terminal time T. In the test, $\Delta t = 10^{-5}$ is used so that the time error is negligible. Evidently, the spatial rates of convergence are consistent with that proved in the convergence theorem.

Table 6.1: Spatial errors and convergence rates of Test 1.

	$L^\infty(L^2)$ error	$L^\infty(L^2)$ order	$L^2(H^1)$ error	$L^2(H^1)$ order
$h = 0.16$	2.0789e-3		5.5045e-2	
$h = 0.08$	5.9674e-4	1.8006	2.9431e-2	0.9032
$h = 0.04$	1.6227e-4	1.8787	1.5332e-2	0.9408
$h = 0.02$	4.0971e-5	1.9857	7.6968e-3	0.9942

Figures 6.1 and 6.2 show respectively the surface plot of the computed pressure p at the terminal time T and the color plot of both the computed pressure p and displacement $\mathbf{u}$ with mesh parameters $h = 0.02$ and $\Delta t = 10^{-5}$. They coincide with the exact solution on the same space-time resolution.

Test 2. In this test we consider so-called Barry-Mercer's problem, which is a Benchmark test problem for the poroelasticity model (6.23)–(6.26) (cf. [77, 76] and the references therein). Again, $\mathcal{D} = [0,1] \times [0,1]$ but $T = 1$. Barry-Mercer's problem assumes no source, that is, $\mathbf{f} \equiv 0$ and $\phi \equiv 0$, and takes the following boundary

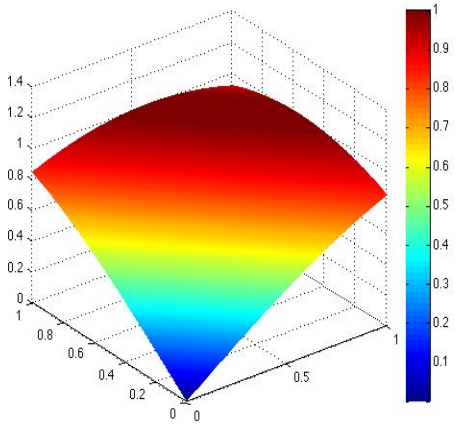

Figure 6.1: Test 1: Surface plot of the Computed pressure p at the terminal time T.

conditions:

$$
\begin{aligned}
p &= 0 && \text{on } \Gamma_j \times (0,T),\ j = 1,3,4,\\
p &= p_2 && \text{on } \Gamma_j \times (0,T),\ j = 2,\\
u_1 &= 0 && \text{on } \Gamma_j \times (0,T),\ j = 1,3,\\
u_2 &= 0 && \text{on } \Gamma_j \times (0,T),\ j = 2,4,\\
\sigma\mathbf{n} - \alpha\mathbf{p}\mathbf{n} &= \mathbf{f}_1 := (0, \alpha p)^T && \text{on } \partial\mathcal{D}_T,
\end{aligned}
$$

where

$$
p_2(x_1,t) = \begin{cases} \sin t & \text{when } x \in [0.2, 0.8) \times (0,T),\\ 0 & \text{others.} \end{cases}
$$

The boundary segments $\Gamma_j, j = 1,2,3,4$, which are defined in **Test 1**, and the above boundary conditions are depicted in Figure 6.3. Also, the initial conditions for Barry-Mercer's problem are $\mathbf{u}(x,0) \equiv \mathbf{0}$ and $p(x,0) \equiv 0$. We remark that Barry-Mercer's problem has a unique solution which is given by an infinite series (cf. [77]).

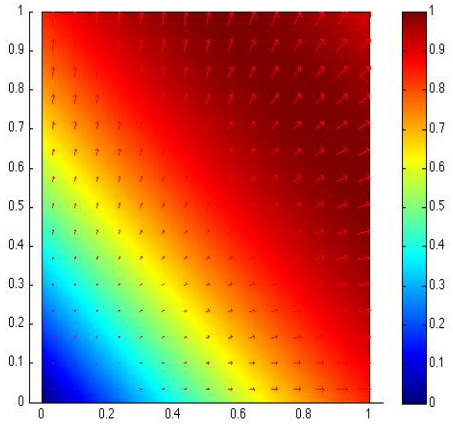

Figure 6.2: Test 1: Computed pressure p (color plot) and displacement $\mathbf{u}$ (arrow plot) at T.

Figures 6.4 and 6.5 display respectively the computed pressure p (surface plot) and the computed displacement $\mathbf{u}$ (arrow plot). We note that the arrows near the boundary match very well with those on the boundary. Our numerical solution approximates the exact solution of Barry-Mercer's problem very well and does not produce any oscillation in computed pressure.

Test 3. This test problem is taken from [77]. Again, we consider problem (6.23)–(6.26) with $\mathcal{D} = [0,1] \times [0,1]$. Let Γ_j be same as in **Test 1** and $c_0 = 0, E = 10^5, \nu = 0.4, \mu = 35714$ and $T = 0.001$. There is no source, that is, $\mathbf{f} \equiv 0$ and $\phi \equiv 0$. The boundary conditions are taken as

$$
\begin{aligned}
-\frac{\kappa}{\mu_f}(\nabla p - \rho_f \mathbf{g}) \cdot \mathbf{n} = 0 \qquad & \text{on } \partial\mathcal{D}_T, \\
\mathbf{u} = \mathbf{0} \qquad & \text{on } \Gamma_3 \times (0,T), \\
\sigma\mathbf{n} - \alpha\mathbf{p}\mathbf{n} = \mathbf{f}_1 \qquad & \text{on } \Gamma_j \times (0,T),\ j = 1,2,4,
\end{aligned}
$$

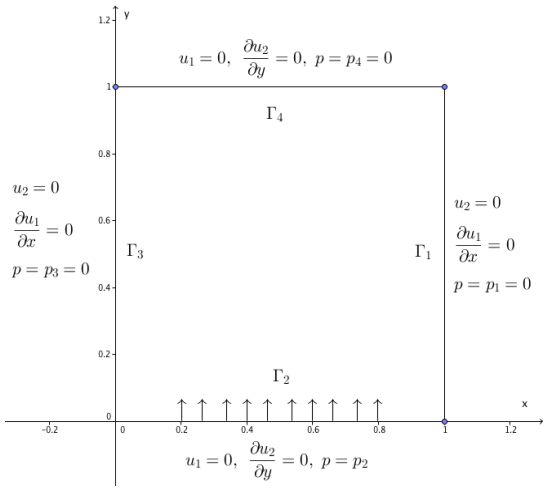

Figure 6.3: Test 2: boundary conditions.

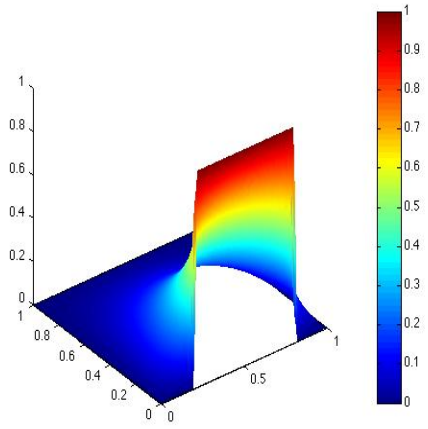

Figure 6.4: Test 2: Surface plot of the computed pressure p at the terminal time T.

where $\mathbf{f}_1 = (f_1^1, f_1^2)$ and

$$f_1^1 \equiv 0 \quad \text{on } \partial\mathcal{D}_T, \qquad f_1^2 = \begin{cases} 0 & \text{on } \Gamma_j \times (0,T),\ j = 1,2,3, \\ -1 & \text{on } \Gamma_4 \times (0,T). \end{cases}$$

The computational domain $\mathcal{D}$ and the above boundary conditions are depicted in Figure 6.6. Also, the zero initial conditions are assigned for both $\mathbf{u}$ and p in this test.

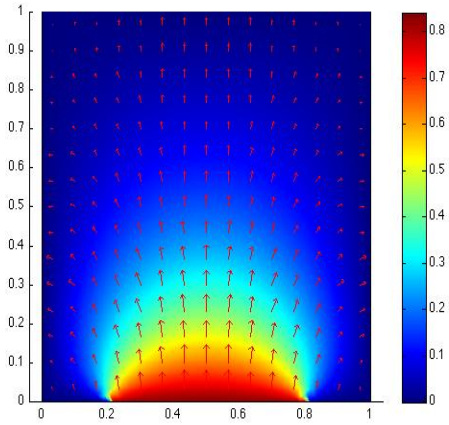

Figure 6.5: Test 2: Computed pressure p (color plot) and displacement (arrow plot) at T.

Figures 6.7–6.8 display respectively the surface and color plot of the computed pressure, the arrow plot of the displacement vector, and the deformation of the whole $\mathcal{D}$. There is no oscillation in the computed pressure and the arrows near the boundary match very well with arrows on the boundary.

We remark that the "locking phenomenon" was observed in the simulation of [77] at $T = 0.001$ for this problem, namely, the computed pressure exhibits some oscillation at $T = 0.001$. The reason for the locking phenomenon was explained as follows: when time step Δt is small, the displacement vector $\mathbf{u}$ is almost divergence free in the short time while the numerical solution does not observe this nearly divergence free property, which results in the locking. However, at later times the displacement vector is no longer divergence free, so no locking exists at later times.

It is clear that our numerical solution does not exhibits the locking phenomenon at $T = 0.001$. This is because our multiphysics reformulation weakly imposes the condition div $\mathbf{u} = q$, hence, $\mathbf{u}$ automatically becomes nearly divergence free when $q \approx 0$ for $0 < t << 1$. Moreover, the pressure p is not a primary variable anymore in our reformulation, instead, p becomes a derivative variable and it is computed using the new primary variables ξ and η. Therefore, our numerical methods are insensitive to the regularity of p.

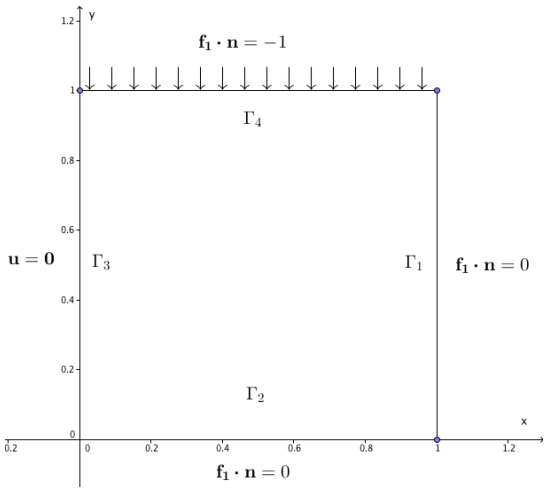

Figure 6.6: Test 3: boundary conditions.

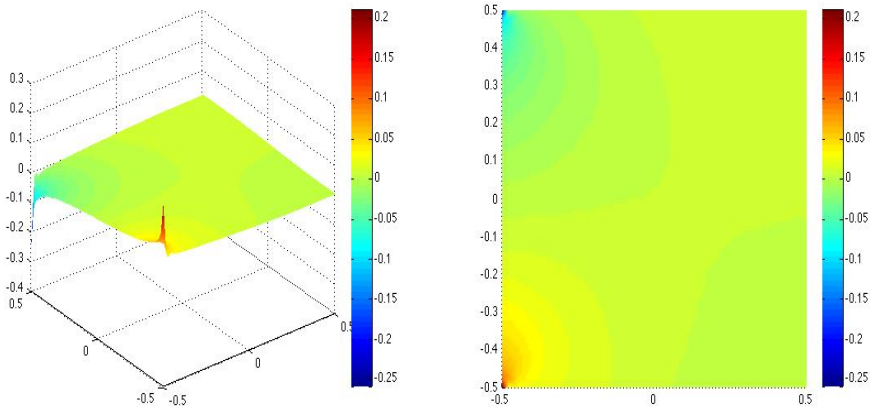

Figure 6.7: Test 3: Computed pressure p: surface plot (left) and color plot (right).

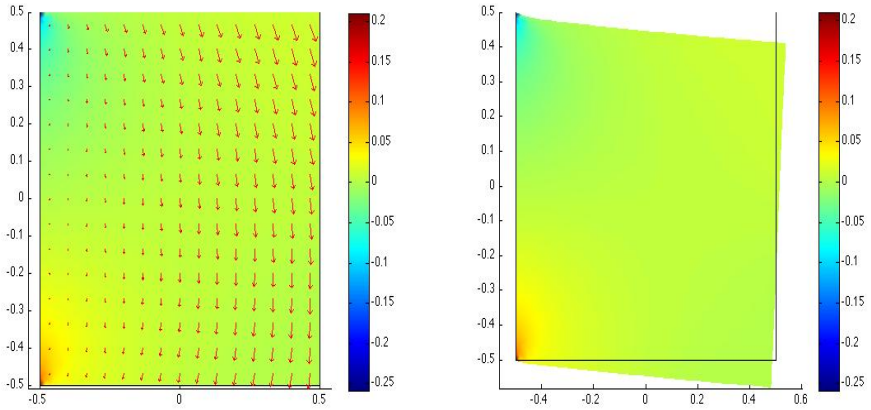

Figure 6.8: Test 3: Arrow plot of the computed displacement (left) and deformation of $\mathcal{D}$ (right).

Chapter 7

Future Works

The work done in this dissertation can be expanded and extended in a number of directions. The following are a few projects which will be pursued in the near future.

- *Mixed finite element methods for the stochastic Cahn-Hilliard equation with gradient-type multiplicative noises.* I plan to carry out the PDE energy law, Hölder continuity in time of the exact solution, the finite element analysis and corresponding error estimates, which are parallel to the analysis in Chapter 5.

- *Adaptive DG methods for phase field models.* Since the mesh size must be very small for small ϵ in order to resolve the diffuse interface, if uniform meshes are used, it is computationally infeasible to compute the solutions of phase field models. As a result, using adaptive meshes is not only efficient but also imperative to solve phase field models, in particular, in high dimensions. One of advantages of DG methods is its ease for adaptivity because unstructured meshes can be easily handled. Hence, I plan to consider adaptive DG methods for the Allen-Cahn and Cahn-Hilliard equations. It is expected that when the interface undergoes a topological change, it will change very fast, so I plan to investigate spectrum-related error indicators which should be sensitive to topological changes. Moreover, I plan to figure out the approximation orders

for the eigenvalues of the linearized Allen-Cahn and Cahn-Hilliard operators on adaptive meshes.

- *Extending the results of my dissertation to other phase field models.* Two-phase flows and crystal growth are two possible applications to be considered.

- *Discontinuous Galerkin methods and adaptivity for stochastic Allen-Cahn and Cahn-Hilliard equations.* I expect that these extension will be very interesting and doable.

- *Analyzing spectrum estimate for linearized stochastic Allen-Cahn and stochastic Cahn-Hilliard operators.* This project is expected to be considerably hard, but it is the only hope at this time for possibly proving convergence of the numerical interfaces to the corresponding sharp interfaces.

Bibliography

[1] A. C. Aristotelous, O. K. and Wise, S. (2013). A mixed discontinuous Galerkin, convex splitting scheme for a modified Cahn-Hilliard equation. *Disc. Cont. Dynamic. Syst. Series B.*, 18(9):2211–2238. 75, 77

[2] Adams, R. A. (2003). *Sobolev Spaces*. Academic Press, New York. 44, 63, 69, 92

[3] Alikakos, N. D., Bates, P. W., and Chen, X. (1994). Convergence of the Cahn-Hilliard equation to the Hele-Shaw model. *Arch. Rational Mech. Anal.*, 128(2):165–205. 9, 12, 61, 79, 99

[4] Allen, S. M. and Cahn, J. W. (1979). A microscopic theory for antiphase boundary motion and its application to antiphase domain coarsening. *Acta Metall.*, 27. 8

[5] Almgren, F., Taylor, J. E., and Wang, L. (1993). Curvature-driven flows: a variational approach. *SIAM J. Control Optim*, 31(2):387–438. 6

[6] Almgren, F. J. and Wang, L. (1991). Mathematical existence of crystal growth with Gibbs-Thomson curvature effects. *Informal Lecture notes for the workshop on Evolving Phase Boundaries at Carnegie Mellon University.* 7

[7] Bartels, S., Muller, R., and Ortner, C. (2011). Robust a priori and a posteriori error analysis for the approximation of Allen-Cahn and Ginzburg-Landau equations past topological changes. *SIAM J. Numer. Anal.* 16, 17

[8] Bellettini, G. and Novaga, M. (1997). Minimal barriers for geometric evolutions. *J. Differential Equations*, 139(1):76–103. 6

[9] Bellettini, G. and Paolini, M. (1995). Quasi-optimal error estimates for the mean curvature flow with a forcing term. *Diff. Integr. Eqns*, 8(4):735–752. 6, 52

[10] Bercovier, J. and Pironneau, O. (1979). Error estimates for finite element solution of the Stokes problem in the primitive variables. *Numer. Math.*, 33:211–224. 201

[11] Brakke, K. A. (1978). The motion of a surface by its mean curvature. *Mathematical Notes, Princeton University Press, Princeton, N.J.*, 20. 6

[12] Brenner, S. C. (1993). A nonconforming mixed multigrid method for the pure displacement problem in planar linear elasticity. *SIAM J. Numer. Anal.*, 30:116–135. 185

[13] Brenner, S. C. and Scott, L. R. (2008). *The Mathematical Theory of Finite Element Methods, third edition.* Springer. 63, 118, 122, 159, 183, 184, 207, 208

[14] Brezzi, F. and Fortin, M. (1992). *Mixed and Hybrid Finite Element Methods.* Springer, New York. 201, 207

[15] Brzezniak, Z., Carelli, E., and Prohl, A. (2013). Finite element based discretisations of the incompressible Navier-Stokes equations with multiplicative random forcing. *IMA J. Num. Anal.* 111

[16] Cahn, J. W. and Hilliard, J. E. (1958). Free energy of a nonuniform system i, interfacial free energy. *J. Chem. Phys.*, 28:258–267. 9

[17] Carelli, E. and Prohl, A. (2012). Rates of convergence for discretizations of the stochastic incompressible Navier-Stokes equations. *SIAM J. Num. Anal.*, 50:2467–2496. 111

[18] Chen, X. (1993). Hele-Shaw problem and area-preserving, curve shorting motion. *Arch. Rational Mech. Anal.*, 123:117–151. 7

[19] Chen, X. (1994). Spectrum for the Allen-Cahn, Cahn-Hilliard and phase field equations for generic interfaces. *Comm. PDEs*, 19(7-8):1371–1395. 18, 25, 26, 63, 70, 71

[20] Chen, X. (1996). Global asymptotic limit of solutions of the Cahn-Hilliard equation. *J. Diff. Geom.*, 44(2):262–311. 9, 61

[21] Chen, Y. G., Giga, Y., and Goto, S. (1989). Uniqueness and existence of viscosity solutions of generalized mean curvature flow equations. *Proc. Japan Acad. Ser. A Math. Sci.*, 65(7):207–210. 6

[22] Chen, Z. and Chen, H. (2004). Pointwise error estimates of discontinuous Galerkin methods with penalty for second-order elliptic problems. *SIAM J.Numer. Anal.*, 42:1146–1166. 36, 74

[23] Ciarlet, P. (1978). *The Finite Element Method for Elliptic Problems*. North-Holland, Amsterdam. 159, 183, 184, 207

[24] Coussy, O. (2004). *Poromechanics*. Wiley and Sons. 181

[25] Dauge, M. (1988). *Elliptic Boundary Value Problems on Corner Domains, Lecture Notes in Math.*, volume 1341. Springer-Verlag, Berlin. 185

[26] Dautray, R. and Lions, J. L. (1990). *Mathematical Analysis and Numerical Methods for Science and Technology*, volume 1. Springer-Verlag. 184, 185, 197

[27] Deckelnick, K., Dziuk, G., and Elliott, C. M. (2005). Computation of geometric partial differential equations and mean curvature flow. *Acta Numer.*, 14:139–232. 6

[28] Du, Q. and Nicolaides, R. A. (1991). Numerical analysis of a continuum model of phase transition. *SIAM J. Numer. Anal.*, 28:1310–1322. 61

[29] Dupont, T. (1972). Some L^2 error estimates for parabolic Galerkin methods. *In the mathematical foundations of the finite element method with applications to partial differential equations(Proc. Sympos., Univ. Maryland, Baltimore, Md., 1972). Academic Press, New York.*, pages 491–504. 90

[30] Elliott, C. M. (1997). Approximation of curvature dependent interface motion. *Art in Numerical Analysis*, pages 407–440. 16

[31] Elliott, C. M. and French, D. A. (1989). A nonconforming finite-element method for the two-dimensional Cahn-Hilliard equation. *SIAM J. Numer. Anal.*, 26:884–903. 61

[32] Es-Sarhir, A. and von Renesse, M. (2013). Ergodicity of stochastic curve shortening flow in the plane. *SIAM J. Math. Anal.*, 44:224–244. 109, 110, 111, 112, 113, 114

[33] Evans, L. C., Soner, H. M., and Souganidis, P. E. (1992). Phase transitions and generalized motion by mean curvature. *Pure Appl. Math.*, 45(9):1097–1123. 5, 6, 8, 11, 16, 51

[34] Evans, L. C. and Spruck, J. (1991). Motion of level sets by mean curvature. *J. Differential Geom.*, 33:635–681. 3, 6

[35] Eyre, D. J. (1998). *An unconditional stable one-step scheme for gradient systems.* 18

[36] Feng, X. (2006). Fully discrete finite element approximations of the Navier-Stokes-Cahn-Hilliard diffuse interface model for two phase fluid flows. *SIAM J. Numer. Anal.*, 44:1049–1072. 10

[37] Feng, X., Ge, Z., and Li, Y. (2014). Multiphysics finite element methods for a poroelasticity model. *Summitted.* 14, 182

[38] Feng, X. and He, Y. (2010). Fully discrete finite element approximations of a polymer gel model. *SIAM J. Numer. Anal.*, 48(2186-2217). 181, 196, 198, 202

[39] Feng, X., He, Y., and Liu, C. (2007). Analysis of finite element approximations of a phase field model for two phase fluids. *Math. Comp.*, 76:539–571. 31

[40] Feng, X. and Li, Y. (2014). Analysis of symmetric interior penalty discontinuous Galerkin methods for the Allen-Cahn equation and the mean curvature flow. *IMA J. Numer. Anal.*, doi:10.1093/imanum/dru058. 9, 13, 62, 73, 77, 97, 146, 176

[41] Feng, X., Li, Y., and Prohl, A. (2014). Finite element approximations of the stochastic mean curvature flow of planar curves of graphs. *Stoch. PDE: Anal. Comp.*, 2:54–83. 10, 14, 146

[42] Feng, X., Li, Y., and Xing, Y. (2014). Analysis of mixed discontinuous Galerkin methods for the Cahn-Hilliard equation and the Hele-Shaw flow. *Submitted*. 9, 13, 64, 97

[43] Feng, X., Li, Y., and Zhang, Y. (2015). Finite element methods for the stochastic Allen-Cahn equation with gradient-type multiplicative noises. *Submitted*. 10, 14

[44] Feng, X., Li, Y., and Zhang, Y. (2015). Mixed finite element methods for the stochastic Cahn-Hilliard equation with gradient-type multiplicative noises. *In preparation*. 11, 14

[45] Feng, X. and Prohl, A. (2003). Numerical analysis of the Allen-Cahn equation and approximation for mean curvature flows. *Numer. Math.*, 94:33–65. 6, 11, 16, 17, 18, 19, 25, 30, 31, 34, 35, 36, 37, 40, 49, 146, 169

[46] Feng, X. and Prohl, A. (2004). Error analysis of a mixed finite element method for the Cahn-Hilliard equation. *Numer. Math.*, 74:47–84. 61, 63, 64, 71, 73, 74, 78, 81

[47] Feng, X. and Prohl, A. (2005). Numerical analysis of the Cahn-Hilliard equation and approximation for the Hele-Shaw problem. *Inter. and Free Bound.*, 7:1–28. 12, 62, 63, 64, 67, 68, 69, 70, 71, 73, 74, 78

[48] Feng, X. and Wu, H. (2005). A posteriori error estimates and an adaptive finite element algorithm for the Allen-Cahn equation and the mean curvature flow. *J. Sci. Comput.*, 24(2):121–146. 17, 26

[49] Feng, X. and Wu, H. (2008). A posteriori error estimates and an adaptive finite element algorithm for the Cahn-Hiliard equation and the Hele-Shaw flow. *J. Comp. Math.*, 26:767–796. 4, 62

[50] Flandoli, F. (1996). Stochastic flows for nonlinear second-order parabolic SPDE. *Ann. Probab.*, 24:547–558. 146

[51] Funaki, T. (1995). The sacling limit for a stochastic PDE and the separation of phases. *Probab. Theory Relat. Fields*, 102:221–288. 146

[52] Giga, Y. (2006). Surface evolution equations, a level let approach. *Monographs in Mathematics, Birkhäuser Verlag, Basel*, 99. 3, 6

[53] Giorgi, E. D. (1994). Barriers, boundaries, motion of manifolds. *Conference held at Department of Mathematics of Pavia.* 6

[54] Girault, V. and Raviart, P. (1981). *Finite Element Method for Navier-Stokes Equations: theory and algorithms.* Springer-Verlag, Berlin, Heidelberg, New York. 185

[55] Gyöngy, I. and Millet, A. (2009). Rate of convergence of space time approximations for stochastic evolution equations. *Potential Anal.*, 30:29–64. 111

[56] Huisken, G. (1984). Flow by mean curvature of convex surfaces into spheres. *J. Differential Geom.*, 20(1):237–266. 6

[57] Ilmanen, T. (1993). Convergence of the Allen-Cahn equation to Brakke's motion by mean curvature. *J. Diff. Geom.*, 38(2):417–461. 6, 8, 16

[58] Karakashian, O. and Pascal, F. (2004). Adaptive discontinuous Galerkin approximations of second order elliptic problems. *Proceedings of European Congress on Computational Methods in Applied Sciences and Engineering.* 50, 100

[59] Kawasaki, K. and Ohta, T. (1982). Kinetic drumhead model of interface. I. *Progr. Theoret. Phys.*, 67:147–163. 10, 142

[60] Kessler, D., Nochetto, R. H., and Schmidt, A. (2004). A posteriori error control for the Allen-Cahn problem: circumventing Gronwall's inequality. *Math. Model. Numer. Anal.*, 38:129–142. 17

[61] Kovacs, M., Larsson, S., and Mesforush, A. (2011). Finite element approximation of the Cahn-Hilliard-Cook equation. *SIAM J. Num. Anal.*, 49:2407–2429. 111

[62] Krylov, N. (2010). Itô's formula for the L_p-norm of stochastic W_p^1-valued processes. *Probab. Theory Relat. Fields*, 147:583–605. 114, 115, 117, 120, 123

[63] Krylov, N. and Rozovskii, B. (2007). Stochastic evolution equations. *In Stochastic Differential Equations: Theory and Applications: Interdisciplinary Math and Science, World Science Publication, Hackensack*, 2:1–69. 145

[64] LeVeque, R. J. and Li, Z. (1994). The immersed interface method for elliptic equations with discontinuous coefficients and singular sources. *SIAM J. Numer. Anal.*, 31:1019–1044. 2

[65] Luckhaus, S. (1990). Solutions for the two-phase Stefan problem with Gibbs-Thomson law for the melting temperature. *Euro. J. Appl. Math.*, 1:101–111. 7

[66] M. Katsoulakis, G. K. and Lakkis, O. (2007). Noise regularization and computations for the 1-dimensional stochastic Allen-Cahn problem. *Interfaces Free Boundaries*, 9:1–30. 146

[67] McFadden, G. B. (2002). Phase field models of solidification. *Contemporary Mathematics*, 295:107–145. 10

[68] Murad, M. A. and Loula, A. F. D. (1992). Improved accuracy in finite element analysis of Biot's consolidation problem. *Comput. Methods in Appl. Mech. and Engr*, 95:359–382. 181, 187, 196

[69] Nochetto, R. H. and Verdi, C. (1997). Convergence past singularities for a fully discrete approximation of curvature-driven interfaces. *SIAM J. Numer. Anal.*, 34(2):490–512. 146

[70] O. A. Ladyzenskaja, V. A. S. and Uarlceva, N. N. (1967). *Linear and Quasilinear Equations of Parabolic type, Translations of Mathematical Monographs*, volume 23. American Mathematical Society, Providence, R.I. 190

[71] Osher, S. and Fedkiw, R. (2003). *Level Set Methods and Dynamic Implicit Surfaces*, volume 153. Springer-Verlag, New York. 3, 6

[72] Osher, S. and Sethian, J. A. (1988). Fronts propagating with curvature-dependent speed: algorithms based on hamilton-jacobi formulations. *J. Comput. Phys.*, 79:12–49. 2, 6

[73] Pachpatte, B. G. (2001). *Inequalities for Finite Difference Equations*, volume 247. CRC Press. 18, 26, 97

[74] Pego, R. L. (1989). Front migration in the nonlinear Cahn-Hilliard equation. *Proc. Roy. Soc. London Ser. A, 422(1863)*, pages 107–145. 9, 61

[75] Phillips, P. J. and Wheeler, M. F. (2007). A coupling of mixed and continuous Galerkin finite element methods for poroelasticity I: the continuous in time case. *Comput. Geosci.*, 11:131–144. 13, 181, 185, 187, 196

[76] Phillips, P. J. and Wheeler, M. F. (2007). A coupling of mixed and continuous Galerkin finite element methods for poroelasticity II: the discrete in time case. *Comput. Geosci.*, 11:145–158. 181, 216

[77] Phillips, P. J. and Wheeler, M. F. (2009). Overcoming the problem of locking in linear elasticity and poroelasticity: an heuristic approach. *Comput. Geosci.*, 13:1–15. 182, 215, 216, 217, 218, 220

[78] Prato, G. D. and Zabczyk, J. (1992). *Stochastic Equations in Infinite Dimensions*. Cambridge University Press. 114, 145

[79] Prohl, A. (2015). Strong rates of convergence for a space-time discretization of the stochastic Allen-Cahn equation with multiplicative noise. *preprint*. 146

[80] Radkevitch, E. (1991). The Gibbs-Thomson correction and conditions for the classical solution of the modified Stefan problem. *Soviet Mathematics Doklady*, 43. 7

[81] Rayleigh, L. (1892). On the theory of surface forces II. *Compressible fluids. Phil. Mag.*, 33:209–220. 2, 3

[82] Rivière, B. (2008). *Discontinuous Galerkin Methods for Solving Elliptic and Parabolic Equations*. SIAM. 30, 32, 36, 40, 82

[83] Röger, M. and Weber, H. (2013). Tightness for a stochastic Allen-Cahn equation. *Stoch. PDE: Anal. Comp.*, 1(1):175–203. 10, 11, 142, 144, 145, 147

[84] S. Brassesco, A. D. M. and Presutti, E. (1995). Brownian fluctuations of the interface in the $d = 1$ Ginzburg-Landau equation with noise. *Ann. Inst. H. Poincare Probab. Stat.*, 31:81–118. 146

[85] Schowalter, R. E. (2000). Diffusion in poro-elastic media. *J. Math. Anal.*, 251:310–340. 181

[86] Schreyer-Bennethum, L. (2007). Theory of flow and deformation of swelling porous materials at the macroscale. *Computers and Geotechnics*, 34:267–278. 181

[87] Sethian, J. A. (1999). *Level Set Methods and Fast Marching Methods. Evolving Interfaces in Computational Geometry, Fluid Mechanics, Computer Vision, and Materials Science, Second edition.* Cambridge Monographs on Applied and Computational Mathematics, 3. Cambridge University Press, Cambridge. 6

[88] Stoth, B. (1996). Convergence of the Cahn-Hilliard equation to the Mullins-Sekerka problem in spherical symmetry. *J. Diff. Eqs.*, 125(1):154–183. 61

[89] Temam, R. (1977). *Navier-Stokes Equations*, volume 2. Studies in Mathematics and its Applications, North-Holland. 183, 184, 196

[90] Terzaghi, K. (1943). *Theoretical Soil Mechanics*. John Wiley and Sons, New York. 181

[91] Unverdi, S. o. and Tryggvason, G. (1992). A front-tracking method for viscous, incompressible multi-fluid flows. *Journal of Computational Physics*, 100:25–37. 2

[92] van der Waals (1893). The thermodynamic theory of capillarity under the hypothesis of a continuous density variation. *J. Stat. Phys.*, 20:197–244. 2, 3

[93] Yamaue, T. and Doi, M. (2004). Swelling dynamics of constrained thin-plate under an external force. *Phys. Rev. E*, 70(011401). 181, 196, 198

[94] Yip, N. K. (1998). Stochastic motion by mean curvature. *Arch. Rational Mech. Anal.*, 144(4):313–355. 109, 145

[95] Zabusky, N. and Overman, E. (1983). Regularization and contour dynamical algorithms i. tangential regularization. *Journal of Computational Physics*, 52:351–373. 2

[96] Zhu, X. P. (2002). *Lectures on Mean Curvature Flows*, volume 32. AMS/IP Studies in Advanced Mathematics, American Mathematical Society, Providence, RI; International Press, Somerville, MA. 6

Vita

Yukun Li was born in Chuzhou, Anhui, China on December 06, 1986. He attended Southwest Jiaotong University in 2004, and obtained his bachelor of science degree in mathematics in 2008. Upon his graduation, he was admitted to Zhejiang University, and obtained his master of science degree in mathematics in 2010. In August 2010, he attended the University of Tennessee, Knoxville in the United States as a graduate student in Mathematics Department with concentration on Computational and Applied Mathematics.