8-2017

Linearly Preconditioned Nonlinear Solvers for Phase Field Equations Involving p-Laplacian Terms

Wenqiang Feng
University of Tennessee, Knoxville, wfeng1@vols.utk.edu

To the Graduate Council:

I am submitting herewith a dissertation written by Wenqiang Feng entitled "Linearly Preconditioned Nonlinear Solvers for Phase Field Equations Involving p-Laplacian Terms." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Mathematics.

Steven M. Wise, Major Professor

We have read this dissertation and recommend its acceptance:

Ohannes Karakashian, Tuoc Van Phan, Abner J. Salgado, Wenjun Zhou

Accepted for the Council:
Dixie L. Thompson

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

Linearly Preconditioned Nonlinear Solvers for Phase Field Equations Involving p-Laplacian Terms

A Dissertation Presented for the

Doctor of Philosophy

Degree

The University of Tennessee, Knoxville

Wenqiang Feng

August 2017

To my parents for their endless love, support and encouragement.

Acknowledgements

In retrospect over the last four years, I have been involved in many projects that relate to the Poisson equation [33], Cahn-Hilliard Equation [18], Epitaxial Thin Film Equation with Slope Selection [34, 35, 36], Cahn-Hilliard-Hele-Shaw Equations [14], Phase Field Crystal equation [34, 27] and Functionalized Cahn-Hilliard Equation [31, 32]. I have gained a broad view on the fast solvers and Phase-Field equations topics. But I have to admit that I did not finish all of those work by myself and most of the work are done in cooperation with many researchers. Many main ideas are from Dr. Steven M. Wise and Dr. Cheng Wang. So I have decided to use "we" in stead of "I" in the following chapters of this thesis.

Most of the results in the following chapters were already published in journals or have been prepared for submission. I would like to include them at here to indicate that the research was (and is) always collaborative work by a group of people:

Chapter. 3 :

[34] W. Feng, A.J. Salgado, C. Wang, and S.M. Wise. *Preconditioned steepest descent methods for some nonlinear elliptic equations involving p-Laplacian terms.* J. Comput. Phys., 334:45–67, 2017.

Chapter. 4 :

[36] W. Feng, C. Wang, S. M. Wise, and Z. Zhang. *A second-order energy stable backward differentiation formula method for the epitaxial thin film equation with slope selection.* arXiv preprint arXiv:1706.01943, 2017.

[31] W. Feng, Z. Guan, J. Lowengrub, C. Wang, S.M. Wise and Y. Chen. *An energy stable finite-difference scheme for Functionalized Cahn-Hilliard equation and its convergence analysis.* arXiv preprint arXiv:1610.02473, 2016.

Chapter. 5 :

[35] W. Feng, C. Wang, S. M. Wise, and Z. Zhang. *Linearly preconditioned nonlinear conjugate gradient solvers for the epitaxial thin film equation with slope selection.* In preparation, 2017.

At the completion of my Phd work, I am full of gratitude and would like to thank many people:

First and foremost, I would like to express my deepest appreciation to my thesis advisor, Dr.Steven M.Wise, for his intelligence, diligence, enthusiasm, philosophy and kindness. This dissertation could not have been possible without his broad knowledge, wonderful ideas, and diligent guidance. Dr.Wise's guidance and attitude had a big impact on me, not only in my research but also in my daily life.

I would like to thank my advisory committee members Dr.Ohannes Karakashian, Dr.Tuoc Van Phan, Dr.Abner J. Salgado and Dr.Wenjun Zhou for the valuable discussion and suggestions.

I am also honored to have Dr.Cheng Wang as one of my collaborators. His diverse expertise and perspectives have greatly enriched my educational and research experience. Dr.Wang provided me with very helpful insights on pseudospectral method and implementation of fast Poisson solver. And he also provided the rigorous proofs for the discrete Sobolev inequalities in the following chapters.

I would like to thank my friends too, especially Amanda.E. Diegel, Jin Dong, Bo Gao, Tricia Phillips, Yukun Li, James Ren, Stefan Schnake, Le Yin and Yi Zhang. The great friendship I developed here have made my graduate life very delightful. They gave me great support and encouragement for my research and life here. They also offered their generosity in my job hunting.

Finally, I would like to add a special thanks to my dear family members, especially my parents, my younger sister and my girlfriend, for their endless love, support and encouragement.

The longer I live, the more I realize the impact of attitude on life. Attitude is more important than the past, than education, money, circumstances, failures and successes and much more than what other people think, say or do.

—*Charles R. Swindoll*

Abstract

Phase field models are usually constructed to model certain interfacial dynamics. Numerical simulations of phase-field models require long time accuracy, stability and therefore it is necessary to develop efficient and highly accurate numerical methods. In particular, the unconditionally energy stable , unconditionally solvable, and accurate schemes and fast solvers are desirable.

In this thesis, We describe and analyze preconditioned steepest descent (PSD) solvers for fourth and sixth-order nonlinear elliptic equations that include p-Laplacian terms on periodic domains in 2 and 3 dimensions. Such nonlinear elliptic equations often arise from time discretization of parabolic equations that model various biological and physical phenomena, in particular, liquid crystals, thin film epitaxial growth and phase transformations. The analyses of the schemes involve the characterization of the strictly convex energies associated with the equations. We first give a general framework for PSD in Hilbert spaces. Based on certain reasonable assumptions of the linear pre-conditioner, a geometric convergence rate is shown for the nonlinear PSD iteration. We then apply the general theory to the fourth and sixth-order problems of interest, making use of Sobolev embedding and regularity results to confirm the appropriateness of our pre-conditioners for the regularized p-Lapacian problems. The results include a sharper theoretical convergence result for p-Laplacian systems compared to what may be found in existing works. We demonstrate rigorously how to apply the theory in the finite dimensional setting using finite difference discretization methods.

Based on the PSD framework, we also proposed two efficient and practical Preconditioned Nonlinear Conjugate Gradient (PNCG) solvers. The main idea of the preconditioned solvers is to use a linearized version of the nonlinear operator as a metric for choosing the initial search direction. And the hybrid conjugate directions as the following search direction. In order to make the proposed solvers and scheme much more practical, we also investigate an adaptive time stepping strategy for time dependent problems.

Numerical simulations for some important physical application problems – including thin film epitaxy with slope selection, the square phase field crystal model and functionalized Cahn-Hilliard equation – are carried out to verify the efficiency of the schemes and solvers.

Table of Contents

List of Tables

List of Figures

List of Algorithms

Chapter 1

Introduction

Let $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, be a rectangular domain. In this work we are interested in efficient solution techniques for fourth and sixth-order nonlinear elliptic equations that have p-Laplacian terms. The fourth-order problem reads as follows: given f Ω-periodthesis ic, find u Ω-periodic such that

$$u - s\nabla \cdot (|\nabla u|^{p-2}\nabla u) + s\varepsilon^2\Delta^2 u = f, \tag{1.1}$$

where $0 < \varepsilon \leq 1$ and s is a positive parameter. The sixth-order problem is as follows: given f, g Ω-periodic, find u, w Ω-periodic such that

$$u - \Delta w = g, \tag{1.2a}$$

$$s\lambda u - s\nabla \cdot \left(|\nabla u|^{p-2}\nabla u\right) + s\varepsilon^2\Delta^2 u - w = f, \tag{1.2b}$$

where $0 < \varepsilon \leq 1$, $s > 0$, and $\lambda \geq 0$ are parameters. The highest order positive diffusion term, parameterized by ε, is often referred to as the surface diffusion, following the thin film applications described below.

We will refer to problems (1.1) and (1.2a) – (1.2b) as *regularized p-Laplacian problems*. However, this is primarily for ease of reference. The highest order surface diffusion term, though parameterized by the "small" coefficient $\varepsilon > 0$, must be present

for the related physical models to make sense and is not an artificial regularization. In other words, we will not consider and are not concerned with the singular limit $\varepsilon \searrow 0$.

These model equations arise most commonly from the time discretization for certain time-dependent physical models. For example, consider the thin epitaxial film model with slope selection

$$\partial_t u = \nabla \cdot \left(|\nabla u|^2 \nabla u \right) - \Delta u - \varepsilon^2 \Delta^2 u,$$

in [59, 71, 77, 83], where u is the spatially periodic height of the film. The 4-Laplacian term in combination with the negative Laplacian term gives energetic preference to facets with unit slope, a continuum-level model of the Ehrlich-Schwoebel kinetic barrier. The highest order term models a small amount of surface diffusion, which smooths the corners where the facets merge. In the square Swift-Hohenberg (SSH) equation

$$\partial_t u = (\beta - 1)u + \eta u^3 - u^5 + \alpha \left(|\nabla u|^2 \nabla u \right) - 2\Delta u + \Delta^2 u, \quad \alpha > 0, \quad \beta, \eta \in \mathbb{R},$$

studied in [22, 49, 47, 63], and the square phase field crystal (SPFC) equation

$$\partial_t u = \Delta \left(\gamma_0 u + \gamma_1 \Delta u + \varepsilon^2 \Delta^2 u - \nabla \cdot \left(|\nabla u|^2 \nabla u \right) \right), \quad \gamma_0 \in \mathbb{R}, \quad \gamma_1 > 0,$$

studied in [29, 43, 47, 63], the 4-Laplacian term gives preference to square-symmetry arrays of “dots" in the density field u. In general, such localized structures play important roles in biological, chemical, and physical processes [50].

For these time-dependent problems, convex decomposition schemes have been proposed and analyzed in [71, 77] to obtain unconditional unique solvability and unconditional energy stability. The convex decomposition scheme for the thin film

model is [77]

$$u^m - s\nabla \cdot (|\nabla u^m|^2 \nabla u^m) + s\varepsilon^2 \Delta^2 u^m = u^{m-1} - s\Delta u^{m-1},$$

where $s > 0$ is the time step size, and the superscripts indicate the time discretizations. The convex decomposition scheme for the SPFC model – which can be inferred from the general principles in [77, 82] – is precisely

$$\begin{aligned} u^m - \Delta w^m &= u^{m-1}, \\ s\gamma_0 u^m - s\nabla \cdot \big(|\nabla u^m|^2 \nabla u^m\big) + s\varepsilon^2 \Delta^2 u^m - w^m &= \ - s\gamma_1 \Delta u^{m-1}, \end{aligned}$$

assuming $\gamma_0, \gamma_1 \geq 0$. These schemes are nonlinear and require one to deal with the p-Laplacian term at the implicit time level. We remark that there are also second-order-in-time convex decomposition schemes for such nonlinear parabolic equations, as described in [71], which have similar nonlinear structures. In any case, solving nonlinear elliptic equations with the p-Laplacian term is challenging, because of its highly nonlinear nature. In [71, 77], the authors used a nonlinear conjugate gradient algorithm to solve the nonlinear system at each implicit time step. Such naive gradient methods are guaranteed to converge due to the global convexity of the equations, but are not necessarily efficient.

Several works develop and analyze numerical schemes for nonlinear elliptic equations involving the p-Laplacian operator. The works [6, 8, 53, 62, 74, 87, 88] are based on finite element approximations in space. Recently, the vanishing moment method for the p-Laplacian was proposed in [39]. In that method, the highest order term is purely artificial, whereas, for the models above, the surface diffusion term is small, but non-vanishing. A hybridizable discontinuous Galerkin method for the p-Laplacian was proposed in [21]. Of these works, [53, 87, 88] are primarily focused on efficient solvers for the elliptic equations with p-Laplacian terms, rather than, say, error estimates.

The main goal of this thesis is to design a general framework of preconditioned steepest descent (PSD) methods for certain nonlinear elliptic equations with p-Laplacian terms. The main idea is to use a linearized version of the nonlinear operator as a pre-conditioner, or in other words, as a metric for choosing the search direction. We propose and analyze the preconditioned steepest descent methods for both the fourth- and sixth-order p-Laplacian problems mentioned above. Herein we present numerical simulations for the 6-Laplacian thin film epitaxy and the H^{-1} gradient flow SPFC model by using the proposed method. While we restrict our focus to the p-Laplacian problems herein, the search direction framework is general and can be applied to other nonlinear equations, such as the Cahn-Hilliard (CH) equation [11, 18, 64, 73], functionalized Cahn-Hilliard (FCH) equation [19, 26, 31], for example.

The convergence analyses of the nonlinear iteration algorithms we propose for the p-Laplacian equations are quite challenging, due to the highly nonlinear nature of the problems. However, we are able to recast the equations as equivalent minimization problems involving strictly convex functionals in Hilbert spaces. Once this is done, we are able to characterize the properties of general pre-conditioners that will result in geometric convergence rates. This general approach is applicable to both the 4th and 6th order equations at the space-continuous level, as well as the approximation of these problems in finite dimensions using finite differences. Though we do not explore it here, we remark that the theory is extensible to the pseudo-spectral, spectral-Galerkin, and mixed finite element settings as well, using the appropriate discrete Gagliardo-Nirenberg inequalities. To our knowledge, the only related theoretical results available in the existing literature are to be found in [53], in which finite element PSD solvers were designed and analyzed. Specifically, it was proved in [53] that their method converges with the rate $O(k^{-\beta})$, where k is the iteration index and $\beta = \frac{p}{p-2} > 0$. In this thesis , we provide a theoretical analysis with a geometric convergence rate $O(\alpha^k)$, with $0 < \alpha < 1$, for the finite difference PSD solver applied to the regularized p-Laplacian problems.

For such nonlinear analyses, the essential difficulty has always been associated with the subtle fact that the numerical solution has to be bounded uniformly in certain functional norms, so that a bound for the iteration error could be established. For the p-Laplacian problems, typically a uniform $W^{1,p}$ bound of the numerical solution is available at each iteration stage, and such a bound may be used to derive an $O(k^{-\beta})$ convergence rate for the PSD iteration. However, for the regularized p-Laplacian problems, one observes that a linear operator with higher-order diffusion may be utilized so that a uniform H^2 bound of the numerical solution may be obtained. Specifically, the existence of the surface diffusion term $\varepsilon^2\Delta^2 u$ enables us to derive a geometric convergence rate $O(\alpha^k)$ for the PSD iteration, which gives a sharper theoretical result than the existing one in [53]. Our strategy comes at a cost that we point out at the offset: a linear, positive, constant-coefficient operator of order 4 or 6 must be inverted to obtain the search direction. But, since we are interested in applications involving coarsening processes over periodic domains, the FFT can be utilized to make this process efficient.

By using the similar preconditioned idea, we also proposed two efficient and practical Preconditioned Nonlinear Conjugate Gradient (PNCG) solvers. The main idea of the preconditioned solvers is to use a linearized version of the nonlinear operator as a metric for choosing the search direction for the initial step. And the hybrid conjugate directions as the following search direction. In order to make the proposed solvers and scheme much more practical, we also investigate the adaptive time stepping strategy.

The remainder of the thesis is organized as follows. In Chapter 2, we present some preliminary notations and definitions. In Chapter 3, we present a general preconditioned steepest descent (PSD) method for nonlinear equations in generic Hilbert spaces, and provide the convergence rate estimates for the PSD method. The application of the general theory to the fourth-order and sixth-order regularized p-Laplacian problem are presented in Chapter 4. Based on the framework of PSD solver, we proposed Linearly Preconditioned Nonlinear Conjugate Gradient Solvers

in Chapter 5. The concluding remarks are offered in section 6. In the Appendix A, we give the proof of a few discrete Sobolev inequalities.

Chapter 2

Preliminaries

2.1 Sobolev Spaces over Periodic Domains

2.1.1 Notation

For the remainder of thesis $\Omega \subset \mathbb{R}^d$ with $d = 2, 3$ is a rectangular domain. In what follows, if $d = 2$ we assume $p \in [2, \infty)$; whereas if $d = 3$ we suppose $p \in [2, 6]$. Most of the physically relevant cases correspond to p being an even integer, however, all of our arguments hold for any value of p in the indicated ranges. The Sobolev spaces of periodic functions are defined as follows: for $q \in [1, \infty]$, we set

$$W^{k,q}_{\text{per}}(\Omega) := \left\{ u \in W^{k,q}_{\text{loc}}(\mathbb{R}^d) \;\middle|\; u \text{ is } \Omega - \text{periodic} \right\},$$

where $k \in \mathbb{N}$ is the differentiability index. Observe that $W^{0,q}_{\text{per}}(\Omega) =: L^q_{\text{per}}(\Omega) = L^q(\Omega)$. We denote the norm of $W^{k,q}_{\text{per}}(\Omega)$ by $\|\cdot\|_{W^{k,q}}$, or just $\|\cdot\|_{L^q}$ when $k = 0$. In the case $q = 2$ and $k = 0$, we denote by $(\cdot, \cdot)$ and $\|\cdot\|$ the inner product and norm, respectively. We set $H^k_{\text{per}}(\Omega) = W^{k,2}_{\text{per}}(\Omega)$ and immediately remark that, given the range of p, we have $H^2_{\text{per}}(\Omega) \hookrightarrow W^{1,p}_{\text{per}}(\Omega)$. For $k \in \mathbb{N} + 1$, the continuous dual of $H^k_{\text{per}}(\Omega)$ is denoted by $H^{-k}_{\text{per}}(\Omega)$.

If $L^2_0(\Omega)$ denotes the set of functions in $L^2(\Omega)$ with zero mean, we define

$$\mathring{H}^1_{\text{per}}(\Omega) := H^1_{\text{per}}(\Omega) \cap L^2_0(\Omega), \quad \mathring{H}^{-1}_{\text{per}}(\Omega) := \left\{ v \in H^{-1}_{\text{per}}(\Omega) \;\middle|\; \langle v, 1\rangle = 0 \right\}.$$

We define a linear operator $\mathsf{T} : \mathring{H}^{-1}_{\text{per}}(\Omega) \to \mathring{H}^1_{\text{per}}(\Omega)$ via the following variational problem: given $\zeta \in \mathring{H}^{-1}_{\text{per}}(\Omega)$, $\mathsf{T}[\zeta] \in \mathring{H}^1_{\text{per}}(\Omega)$ solves

$$(\nabla \mathsf{T}[\zeta], \nabla \chi) = \langle \zeta, \chi \rangle, \qquad \forall\, \chi \in \mathring{H}^1_{\text{per}}(\Omega).$$

From the Riesz representation theorem it immediately follows that T is well-defined. We define the inner product

$$(\zeta, \xi)_{\mathring{H}^{-1}_{\text{per}}} := (\nabla \mathsf{T}[\zeta], \nabla \mathsf{T}[\xi]) = \langle \zeta, \mathsf{T}[\xi] \rangle = \langle \xi, \mathsf{T}[\zeta] \rangle, \quad \forall\, \zeta, \xi \in \mathring{H}^{-1}_{\text{per}}(\Omega).$$

The induced norm is denoted $\|\cdot\|_{\mathring{H}^{-1}_{\text{per}}}$. The following facts can be easily established [24]:

Lemma 2.1.1. *On $\mathring{H}^{-1}_{\text{per}}(\Omega)$ the norm $\|\cdot\|_{\mathring{H}^{-1}_{\text{per}}}$ equals the operator norm: for all $\zeta \in \mathring{H}^{-1}_{\text{per}}(\Omega)$,*

$$\|\zeta\|_{\mathring{H}^{-1}_{\text{per}}} = \sup_{0 \neq \chi \in \mathring{H}^1_{\text{per}}(\Omega)} \frac{\langle \zeta, \chi \rangle}{\|\nabla \chi\|}.$$

Consequently, we have $|\langle \zeta, \chi \rangle| \leq \|\zeta\|_{\mathring{H}^{-1}_{\text{per}}} \|\nabla \chi\|$, for all $\chi \in H^1_{\text{per}}(\Omega)$ and $\zeta \in \mathring{H}^{-1}_{\text{per}}(\Omega)$. Furthermore, for all $\zeta \in L^2_0(\Omega)$, we have the Poincaré type inequality: $\|\zeta\|_{\mathring{H}^{-1}_{\text{per}}} \leq C\,\|\zeta\|$, for some $C > 0$.

2.1.2 Interpolation Inequalities

Lemma 2.1.2. *Suppose that $p \in [2, \infty)$ when $d = 2$, and $p \in [2, 6]$, if $d = 3$. For any $\xi \in H^2_{\text{per}}(\Omega)$, we have*

$$\|\nabla \xi\|_{L^p} \leq C_9 \begin{cases} \|\xi\|^{\frac{1}{p}} \cdot \|\Delta \xi\|^{\frac{p-1}{p}}, & \text{if} \quad d = 2, \quad p \in [2, \infty), \\ \|\xi\|^{\frac{3}{2p} - \frac{1}{4}} \cdot \|\Delta \xi\|^{\frac{5}{4} - \frac{3}{2p}}, & \text{if} \quad d = 3, \quad p \in [2, 6], \end{cases} \tag{2.1}$$

for some $C_9 = C_9(d,p) > 0$.

Proof. This follows from the Gagliardo-Nirenberg interpolation inequality and elliptic regularity. □

Lemma 2.1.3. *For every* $\xi \in \mathring{H}^2_{\text{per}}(\Omega)$ *we have*

$$\|\xi\| \leq \|\xi\|_{\mathring{H}^{-1}_{\text{per}}}^{\frac{2}{3}} \|\Delta\xi\|^{\frac{1}{3}}, \tag{2.2}$$

and

$$\|\nabla\xi\| \leq \|\xi\|_{\mathring{H}^{-1}_{\text{per}}}^{\frac{1}{3}} \|\Delta\xi\|^{\frac{2}{3}}. \tag{2.3}$$

Proof. Using integration by parts we get

$$\|\nabla\xi\|^2 = -(\xi, \Delta\xi) \leq \|\xi\| \cdot \|\Delta\xi\|. \tag{2.4}$$

The definition of the $\mathring{H}^{-1}_{\text{per}}(\Omega)$ norm implies that

$$\|\xi\|^2 = (\xi, \xi) \leq \|\xi\|_{\mathring{H}^{-1}_{\text{per}}} \|\nabla\xi\|. \tag{2.5}$$

Therefore, a combination of (2.4) and (2.5) leads to

$$\|\nabla\xi\| \leq \|\xi\|^{\frac{1}{2}} \cdot \|\Delta\xi\|^{\frac{1}{2}} \leq \|\xi\|_{\mathring{H}^{-1}_{\text{per}}}^{\frac{1}{4}} \|\nabla\xi\|^{\frac{1}{4}} \cdot \|\Delta\xi\|^{\frac{1}{2}},$$

so that

$$\|\nabla\xi\|^{\frac{3}{4}} \leq \|\xi\|_{\mathring{H}^{-1}_{\text{per}}}^{\frac{1}{4}} \|\Delta\xi\|^{\frac{1}{2}}.$$

which yields the second inequality. The first may be proved in a similar way. □

Similar to before, the Gagliardo-Nirenberg inequality, together with elliptic regularity, yield the following interpolation result.

Lemma 2.1.4. *Suppose that $p \in [2, \infty)$ when $d = 2$, and $p \in [2, 6]$, if $d = 3$. For any $\xi \in \mathring{H}^2_{\text{per}}(\Omega)$, we have*

$$\|\nabla \xi\|_{L^p} \leq C_9 \begin{cases} \|\xi\|_{\mathring{H}^{-1}_{\text{per}}}^{\frac{2}{3p}} \|\Delta \xi\|^{1-\frac{2}{3p}}, & \text{if} \quad d = 2, \quad p \in [2, \infty), \\ \|\xi\|_{\mathring{H}^{-1}_{\text{per}}}^{\frac{1}{p}-\frac{1}{6}} \|\Delta \xi\|^{\frac{7}{6}-\frac{1}{p}}, & \text{if} \quad d = 3, \quad p \in [2, 6], \end{cases} \tag{2.6}$$

for some $C_9 = C_9(d, p) > 0$.

2.2 Finite Difference Spatial Discretization in 2D

2.2.1 Notation

In this subsection we define the discrete spatial difference operators, function spaces, inner products and norms, following the notation used in [52, 71, 77, 78, 82]. Let $\Omega = (0, L_x) \times (0, L_y)$, where, for simplicity, we assume $L_x = L_y =: L > 0$. We write $L = m \cdot h$, where m is a positive integer. The parameter $h = \frac{L}{m}$ is called the mesh or grid spacing. We define the following two uniform, infinite grids with grid spacing $h > 0$:

$$E := \{x_{i+\frac{1}{2}} \mid i \in \mathbb{Z}\}, \quad C := \{x_i \mid i \in \mathbb{Z}\},$$

where $x_i = x(i) := (i - \frac{1}{2}) \cdot h$. Consider the following 2D discrete periodic function spaces:

$$\begin{aligned}
\mathcal{V}_{\text{per}} &:= \left\{ \nu : E \times E \to \mathbb{R} \;\middle|\; \nu_{i+\frac{1}{2}, j+\frac{1}{2}} = \nu_{i+\frac{1}{2}+\alpha m, j+\frac{1}{2}+\beta m}, \ \forall\, i, j, \alpha, \beta \in \mathbb{Z} \right\}, \\
\mathcal{C}_{\text{per}} &:= \left\{ \nu : C \times C \to \mathbb{R} \mid \nu_{i,j} = \nu_{i+\alpha m, j+\beta m}, \ \forall\, i, j, \alpha, \beta \in \mathbb{Z} \right\}, \\
\mathcal{E}^{\text{ew}}_{\text{per}} &:= \left\{ \nu : E \times C \to \mathbb{R} \;\middle|\; \nu_{i+\frac{1}{2}, j} = \nu_{i+\frac{1}{2}+\alpha m, j+\beta m}, \ \forall\, i, j, \alpha, \beta \in \mathbb{Z} \right\}, \\
\mathcal{E}^{\text{ns}}_{\text{per}} &:= \left\{ \nu : C \times E \to \mathbb{R} \;\middle|\; \nu_{i, j+\frac{1}{2}} = \nu_{i+\alpha m, j+\frac{1}{2}+\beta m}, \ \forall\, i, j, \alpha, \beta \in \mathbb{Z} \right\}.
\end{aligned}$$

The functions of $\mathcal{V}_{\text{per}}$ are called *vertex centered functions*; those of $\mathcal{C}_{\text{per}}$ are called *cell centered functions*. The functions of $\mathcal{E}^{\text{ew}}_{\text{per}}$ are called *east-west edge-centered functions*,

and the functions of $\mathcal{E}^{\rm ns}_{\rm per}$ are called *north-south edge-centered functions*. We also define the mean zero space

$$\mathring{\mathcal{C}}_{\rm per} := \left\{ \nu \in \mathcal{C}_{\rm per} \;\middle|\; \frac{h^2}{|\Omega|} \sum_{i,j=1}^{m} \nu_{i,j} =: \overline{\nu} = 0 \right\}.$$

We now define the important difference and average operators on the spaces:

$$\begin{aligned} A_x \nu_{i+\frac{1}{2},\square} &:= \frac{1}{2}\left(\nu_{i+1,\square} + \nu_{i,\square}\right), \quad D_x \nu_{i+\frac{1}{2},\square} := \frac{1}{h}\left(\nu_{i+1,\square} - \nu_{i,\square}\right), \\ A_y \nu_{\square,i+\frac{1}{2}} &:= \frac{1}{2}\left(\nu_{\square,i+1} + \nu_{\square,i}\right), \quad D_y \nu_{\square,i+\frac{1}{2}} := \frac{1}{h}\left(\nu_{\square,i+1} - \nu_{\square,i}\right), \end{aligned}$$

with A_x, $D_x : \mathcal{C}_{\rm per} \to \mathcal{E}^{\rm ew}_{\rm per}$ if $\square$ is an integer, and A_x, $D_x : \mathcal{E}^{\rm ns}_{\rm per} \to \mathcal{V}_{\rm per}$ if $\square$ is a half-integer, with A_y, $D_y : \mathcal{C}_{\rm per} \to \mathcal{E}^{\rm ns}_{\rm per}$ if $\square$ is an integer, and A_y, $D_y : \mathcal{E}^{\rm ew}_{\rm per} \to \mathcal{V}_{\rm per}$ if $\square$ is a half-integer. Likewise,

$$\begin{aligned} a_x \nu_{i,\square} &:= \frac{1}{2}\left(\nu_{i+\frac{1}{2},\square} + \nu_{i-\frac{1}{2},\square}\right), \quad d_x \nu_{i,\square} := \frac{1}{h}\left(\nu_{i+\frac{1}{2},\square} - \nu_{i-\frac{1}{2},\square}\right), \\ a_y \nu_{\square,j} &:= \frac{1}{2}\left(\nu_{\square,j+\frac{1}{2}} + \nu_{\square,j-\frac{1}{2}}\right), \quad d_y \nu_{\square,j} := \frac{1}{h}\left(\nu_{\square,j+\frac{1}{2}} - \nu_{\square,j-\frac{1}{2}}\right), \end{aligned}$$

with a_x, $d_x : \mathcal{E}^{\rm ew}_{\rm per} \to \mathcal{C}_{\rm per}$ if $\square$ is an integer, and a_x, $d_x : \mathcal{V}_{\rm per} \to \mathcal{E}^{\rm ns}_{\rm per}$ if $\square$ is a half-integer; and with a_y, $d_y : \mathcal{E}^{\rm ns}_{\rm per} \to \mathcal{C}_{\rm per}$ if $\square$ is an integer, and a_y, $d_y : \mathcal{V}_{\rm per} \to \mathcal{E}^{\rm ew}_{\rm per}$ if $\square$ is a half-integer.

Define the 2D center-to-vertex derivatives $\mathfrak{D}_x$, $\mathfrak{D}_y : \mathcal{C}_{\rm per} \to \mathcal{V}_{\rm per}$ component-wise as

$$\begin{aligned} \mathfrak{D}_x \nu_{i+\frac{1}{2},j+\frac{1}{2}} &:= A_y (D_x \nu)_{i+\frac{1}{2},j+\frac{1}{2}} = D_x (A_y \nu)_{i+\frac{1}{2},j+\frac{1}{2}} \\ &= \frac{1}{2h}\left(\nu_{i+1,j+1} - \nu_{i,j+1} + \nu_{i+1,j} - \nu_{i,j}\right), \\ \mathfrak{D}_y \nu_{i+\frac{1}{2},j+\frac{1}{2}} &:= A_x (D_y \nu)_{i+\frac{1}{2},j+\frac{1}{2}} = D_y (A_x \nu)_{i+\frac{1}{2},j+\frac{1}{2}} \\ &= \frac{1}{2h}\left(\nu_{i+1,j+1} - \nu_{i+1,j} + \nu_{i,j+1} - \nu_{i,j}\right). \end{aligned}$$

The utility of these definitions is that the differences $\mathfrak{D}_x$ and $\mathfrak{D}_y$ are collocated on the grid, unlike the case for D_x, D_y. Define the 2D vertex-to-center derivatives $\mathfrak{d}_x$, $\mathfrak{d}_y : \mathcal{V}_{\text{per}} \to \mathcal{C}_{\text{per}}$ component-wise as

$$\begin{aligned}
\mathfrak{d}_x \nu_{i,j} &:= a_y (d_x \nu)_{i,j} = d_x (a_y \nu)_{i,j} \\
&= \frac{1}{2h} \left(\nu_{i+\frac{1}{2},j+\frac{1}{2}} - \nu_{i-\frac{1}{2},j+\frac{1}{2}} + \nu_{i+\frac{1}{2},j-\frac{1}{2}} - \nu_{i-\frac{1}{2},j-\frac{1}{2}} \right), \\
\mathfrak{d}_y \nu_{i,j} &:= a_x (d_y \nu)_{i,j} = d_y (a_x \nu)_{i,j} \\
&= \frac{1}{2h} \left(\nu_{i+\frac{1}{2},j+\frac{1}{2}} - \nu_{i+\frac{1}{2},j-\frac{1}{2}} + \nu_{i-\frac{1}{2},j+\frac{1}{2}} - \nu_{i-\frac{1}{2},j-\frac{1}{2}} \right).
\end{aligned}$$

Now the discrete gradient operator, ∇_h^{v}: $\mathcal{C}_{\text{per}} \to \mathcal{V}_{\text{per}} \times \mathcal{V}_{\text{per}}$, is defined as

$$\nabla_h^{\mathsf{v}} \nu_{i+\frac{1}{2},j+\frac{1}{2}} := (\mathfrak{D}_x \nu_{i+\frac{1}{2},j+\frac{1}{2}}, \mathfrak{D}_y \nu_{i+\frac{1}{2},j+\frac{1}{2}}).$$

The standard 2D discrete Laplacian, $\Delta_h : \mathcal{C}_{\text{per}} \to \mathcal{C}_{\text{per}}$, is given by

$$\Delta_h \nu_{i,j} := d_x (D_x \nu)_{i,j} + d_y (D_y \nu)_{i,j} = \frac{1}{h^2} \left(\nu_{i+1,j} + \nu_{i-1,j} + \nu_{i,j+1} + \nu_{i,j-1} - 4\nu_{i,j} \right).$$

The 2D vertex-to-vertex average, $\mathcal{A} : \mathcal{V}_{\text{per}} \to \mathcal{C}_{\text{per}}$, is defined to be

$$\mathcal{A} \nu_{i,j} := \frac{1}{4} \left(\nu_{i+1,j} + \nu_{i-1,j} + \nu_{i,j+1} + \nu_{i,j-1} \right).$$

The 2D *skew* Laplacian, $\Delta_h^{\mathsf{v}} : \mathcal{C}_{\text{per}} \to \mathcal{C}_{\text{per}}$, is defined as

$$\begin{aligned}
\Delta_h^{\mathsf{v}} \nu_{i,j} &= \mathfrak{d}_x (\mathfrak{D}_x \nu)_{i,j} + \mathfrak{d}_y (\mathfrak{D}_y \nu)_{i,j} \\
&= \frac{1}{2h^2} \left(\nu_{i+1,j+1} + \nu_{i-1,j+1} + \nu_{i+1,j-1} + \nu_{i-1,j-1} - 4\nu_{i,j} \right).
\end{aligned}$$

The 2D discrete p-Laplacian operator is defined as

$$\nabla_h^{\mathsf{v}} \cdot \left(|\nabla_h^{\mathsf{v}} \nu|^{p-2} \nabla_h^{\mathsf{v}} \nu \right)_{ij} := \mathfrak{d}_x (r \, \mathfrak{D}_x \nu)_{i,j} + \mathfrak{d}_y (r \, \mathfrak{D}_y \nu)_{i,j},$$

with

$$r_{i+\frac{1}{2},j+\frac{1}{2}} := \left[(\mathfrak{D}_x u)^2_{i+\frac{1}{2},j+\frac{1}{2}} + (\mathfrak{D}_y u)^2_{i+\frac{1}{2},j+\frac{1}{2}} \right]^{\frac{p-2}{2}}.$$

Clearly, for $p = 2$, $\Delta_h^{\mathsf{v}} \nu = \nabla_h^{\mathsf{v}} \cdot \left(|\nabla_h^{\mathsf{v}} \nu|^{p-2} \nabla_h^{\mathsf{v}} \nu \right)$.

Now we are ready to define the following grid inner products:

$$\begin{aligned}
(\nu, \xi)_2 &:= h^2 \sum_{i=1}^{m} \sum_{j=1}^{n} \nu_{i,j} \psi_{i,j}, \quad \nu,\, \xi \in \mathcal{C}_{\text{per}}, \\
\langle \nu, \xi \rangle &:= (\mathcal{A}(\nu\xi), 1)_2\,, \quad \nu,\, \xi \in \mathcal{V}_{\text{per}}, \\
[\nu, \xi]_{\text{ew}} &:= (A_x(\nu\xi), 1)_2\,, \quad \nu,\, \xi \in \mathcal{E}^{\text{ew}}_{\text{per}}, \\
[\nu, \xi]_{\text{ns}} &:= (A_y(\nu\xi), 1)_2\,, \quad \nu,\, \xi \in \mathcal{E}^{\text{ns}}_{\text{per}}.
\end{aligned}$$

Suppose that $\zeta \in \mathring{\mathcal{C}}_{\text{per}}$, then there is a unique solution $\mathsf{T}_h[\zeta] \in \mathring{\mathcal{C}}_{\text{per}}$ such that $-\Delta_h \mathsf{T}_h[\zeta] = \zeta$. We often write, in this case, $\mathsf{T}_h[\zeta] = -\Delta_h^{-1}\zeta$. The discrete analog of the $\mathring{H}^{-1}_{\text{per}}$ inner product is defined as

$$(\zeta, \xi)_{-1} := (\zeta, \mathsf{T}_h[\xi])_2 = (\mathsf{T}_h[\zeta], \xi)_2\,, \quad \zeta,\, \xi \in \mathring{\mathcal{C}}_{\text{per}}.$$

where summation-by-parts [71, 82] guarantees the symmetry and the second equality.

We now define the following norms for cell-centered functions. If $\nu \in \mathring{\mathcal{C}}_{\text{per}}$, then $\|\nu\|^2_{-1} = (\nu, \nu)_{-1}$. If $\nu \in \mathcal{C}_{\text{per}}$, then $\|\nu\|^2_2 := (\nu, \nu)$; $\|\nu\|^p_p := (|\nu|^p, 1)$ $(1 \le p < \infty)$, and $\|\nu\|_\infty := \max_{\substack{1\le i\le m \\ 1\le j\le n}} |\nu_{i,j}|$. Similarly, we define the gradient norms: for $\nu \in \mathcal{C}_{\text{per}}$,

$$\|\nabla_h^{\mathsf{v}} \nu\|^p_p := \langle |\nabla_h^{\mathsf{v}} \nu|^p, 1 \rangle, \quad |\nabla_h^{\mathsf{v}} \nu|^p := [(\mathfrak{D}_x \nu)^2 + (\mathfrak{D}_y \nu)^2]^{\frac{p}{2}} = [\nabla_h^{\mathsf{v}} \nu \cdot \nabla_h^{\mathsf{v}} \nu]^{\frac{p}{2}} \in \mathcal{V}_{\text{per}},$$

where $2 \le p < \infty$ and

$$\|\nabla_h \nu\|^2_2 := [D_x \nu, D_x \nu]_{\text{ew}} + [D_y \nu, D_y \nu]_{\text{ns}}\,.$$

Consequently, the discrete $\|\cdot\|_{H_h^1}$ and $\|\cdot\|_{H_h^2}$ norms on periodic boundary domain defined as

$$\|\phi\|_{H_h^1}^2 := \|\phi\|_2^2 + \|\nabla_h \phi\|_2^2 , \tag{2.7}$$

$$\|\phi\|_{H_h^2}^2 := \|\phi\|_{H_h^1}^2 + \|\Delta_h \phi\|_2^2 . \tag{2.8}$$

Lemma 2.2.1. *For any $\phi \in \mathcal{C}_{\text{per}}$ with $\overline{\phi} = 0$, we have*

$$\|\nabla_h \phi\|_2^2 \geq \|\nabla_h^{\mathsf{v}} \phi\|_2^2. \tag{2.9}$$

Proof. By the definition of $\mathfrak{D}_x \phi$, we get

$$\mathfrak{D}_x \phi_{i+\frac{1}{2},j+\frac{1}{2}} = \frac{1}{2}\left((D_x\phi)_{i+\frac{1}{2},j} + (D_x\phi)_{i+\frac{1}{2},j+1} \right), \tag{2.10}$$

which in turn implies that

$$\|\mathfrak{D}_x \phi\|_2^2 = h^2 \sum_{i,j=0}^{m-1} (\mathfrak{D}_x \phi_{i+\frac{1}{2},j+\frac{1}{2}})^2 \leq h^2 \sum_{i,j=0}^{m-1} (D_x \phi_{i+\frac{1}{2},j})^2 = \|D_x \phi\|_2^2, \quad \text{i.e., } \|\mathfrak{D}_x \phi\|_2 \leq \|D_x \phi\|_2 \tag{2.11}$$

Using a similar argument, we also obtain $\|\mathfrak{D}_y \phi\|_2 \leq \|D_y \phi\|_2$. These two inequalities lead to the desired estimate; the proof of Lemma. 2.2.1 is complete. □

2.2.2 Discrete Sobolev Inequalities

Lemma 2.2.2. *Suppose that $p \in [2, \infty)$, $d = 2$, we have*

$$\|\nabla_h^v \xi\|_p \leq C_9 \begin{cases} \|\xi\|_2^{\frac{1}{p}} \cdot \|\Delta_h \xi\|_2^{\frac{p-1}{p}}, & \forall \quad \xi \in \mathcal{C}_{\text{per}}, \\ \|\xi\|_{-1}^{\frac{2}{3p}} \cdot \|\Delta_h \xi\|_2^{1-\frac{2}{3p}}, & \forall \quad \xi \in \mathring{\mathcal{C}}_{\text{per}}, \end{cases}$$

for some $C_9 = C_9(p) > 0$.

The proof for $p = 4, d = 2$ can be found in the appendix. Following the similar arguments, the other cases can be proved.

Remark 2.2.3. *Though we have focused on the case $d = 2$ in this section, we can also define our operators and norms, in particular $\nabla_h^v \xi$ and $\|\nabla_h^v \xi\|_p$, in three space dimensions. Then for $p \in [2, 6]$, we expect*

$$\|\nabla_h^v \xi\|_p \le C_9 \begin{cases} \|\xi\|_2^{\frac{3}{2p}-\frac{1}{4}} \|\Delta_h \xi\|_2^{\frac{5}{4}-\frac{3}{2p}}, & \forall \quad \xi \in \mathcal{C}_{\text{per}}, \\ \|\xi\|_{-1}^{\frac{1}{p}-\frac{1}{6}} \|\Delta_h \xi\|_2^{\frac{7}{6}-\frac{1}{p}}, & \forall \quad \xi \in \mathring{\mathcal{C}}_{\text{per}}, \end{cases}$$

for some $C_9 = C_9(d = 3, p) > 0$.

The following preliminary estimates are cited from earlier works. For more details we refer the reader to [52, 82].

Lemma 2.2.4. *For any $f, g \in \mathcal{C}_{\text{per}}$, the following summation by parts formulas are valid:*

$$\begin{aligned} (f, \Delta_h g) &= -(\nabla_h f, \nabla_h g), \\ (f, \Delta_h^2 g) &= (\Delta_h f, \Delta_h g), \\ (f, \Delta_h^3 g) &= -(\nabla_h \Delta_h f, \nabla_h \Delta_h g). \end{aligned} \tag{2.12}$$

Lemma 2.2.5. *Suppose $\phi \in \mathcal{C}_{\text{per}}$. Then*

$$\|\Delta_h \phi\|_2^2 \le \frac{1}{3\alpha^2} \|\phi\|_2^2 + \frac{2\alpha}{3} \|\nabla_h (\Delta_h \phi)\|_2^2, \tag{2.13}$$

is valid for arbitrary $\alpha > 0$.

Lemma 2.2.6. *For $\phi \in \mathcal{C}_{\text{per}}$, we have the estimate*

$$F_h(\phi) \ge C \|\phi\|_{2,2}^2 - \frac{L^3}{4}, \tag{2.14}$$

with C only dependent on Ω, and $F_h(\phi)$ given by (2.18).

The following preliminary estimates are needed in the convergence analysis presented in later sections; the detailed proof is left to Appendix A.4

Proposition 2.2.7. *For any $\phi \in \mathcal{C}_{\text{per}}$ with $\overline{\phi} = 0$, we have*

$$\|\Delta_h \phi\|_2^2 \geq C_1 \|\phi\|_{H_h^2}^2, \tag{2.15}$$

$$\|\phi\|_\infty \leq C \|\phi\|_{H_h^2}, \tag{2.16}$$

$$\|\phi\|_{W_h^{1,6}} := \|\phi\|_6 + \|\nabla_h^{\mathsf{v}} \phi\|_6 \leq C \|\phi\|_{H_h^2}, \tag{2.17}$$

with C and C_1 only dependent on Ω.

2.3 Finite Difference Spatial Discretization in 3D

Similarly, the notation and discrete functions can be easily generated to 3D. For simplicity of presentation, we denote $(\cdot, \cdot)$ as the standard L^2 inner product, and $\|\cdot\|$ as the standard L^2 norm, and $\|\cdot\|_{H^m}$ as the standard H^m norm. We use the notation and results for some discrete functions and operators from [27, 45, 80, 82].

2.3.1 Notation

Let $\Omega = (0, L_x) \times (0, L_y) \times (0, L_z)$, where for simplicity, we assume $L_x = L_y = L_z =: L > 0$. It is also assumed that $h_x = h_y = h_y = h$ and we denote $L = m \cdot h$, where m is a positive integer. The parameter $h = \frac{L}{m}$ is called the mesh or grid spacing. We define the following two uniform, infinite grids with grid spacing $h > 0$:

$$E := \{x_{i+\frac{1}{2}} \mid i \in \mathbb{Z}\}, \quad C := \{x_i \mid i \in \mathbb{Z}\},$$

where $x_i = x(i) := (i - \frac{1}{2}) \cdot h$. Consider the following 3D discrete periodic function spaces:

$$\mathcal{C}_{\text{per}} := \{\nu : C \times C \times C \to \mathbb{R} \mid \nu_{i,j,k} = \nu_{i+\alpha m, j+\beta m, k+\gamma m},\ \forall\, i, j, k, \alpha, \beta, \gamma \in \mathbb{Z}\},$$
$$\mathcal{E}_{\text{per}}^{\text{x}} := \left\{\nu : E \times C \times C \to \mathbb{R} \;\middle|\; \nu_{i+\frac{1}{2}, j, k} = \nu_{i+\frac{1}{2}+\alpha m, j+\beta m, k+\gamma m},\ \forall\, i, j, k, \alpha, \beta, \gamma \in \mathbb{Z}\right\}.$$

The spaces $\mathcal{E}^{\mathrm{y}}_{\mathrm{per}}$ and $\mathcal{E}^{\mathrm{z}}_{\mathrm{per}}$ are analogously defined. The functions of $\mathcal{C}_{\mathrm{per}}$ are called *cell centered functions*. The functions of $\mathcal{E}^{\mathrm{x}}_{\mathrm{per}}$, $\mathcal{E}^{\mathrm{y}}_{\mathrm{per}}$, and $\mathcal{E}^{\mathrm{z}}_{\mathrm{per}}$, are called *east-west face-centered functions*, *north-south face-centered functions*, and *up-down face-centered functions*, respectively. We also define the mean zero space

$$\mathring{\mathcal{C}}_{\mathrm{per}} := \left\{ \nu \in \mathcal{C}_{\mathrm{per}} \,\middle|\, \overline{\nu} := \frac{h^3}{|\Omega|} \sum_{i,j,k=1}^{m} \nu_{i,j,k} = 0 \right\}.$$

We now introduce the important difference and average operators on the spaces:

$$\begin{aligned}
A_x \nu_{i+\frac{1}{2},j,k} &:= \frac{1}{2}\left(\nu_{i+1,j,k} + \nu_{i,j,k}\right), \quad D_x \nu_{i+\frac{1}{2},j,k} := \frac{1}{h}\left(\nu_{i+1,j,k} - \nu_{i,j,k}\right), \\
A_y \nu_{i,j+\frac{1}{2},k} &:= \frac{1}{2}\left(\nu_{i,j+1,k} + \nu_{i,j,k}\right), \quad D_y \nu_{i,j+\frac{1}{2},k} := \frac{1}{h}\left(\nu_{i,j+1,k} - \nu_{i,j,k}\right), \\
A_z \nu_{i,j,k+\frac{1}{2}} &:= \frac{1}{2}\left(\nu_{i,j,k+1} + \nu_{i,j,k}\right), \quad D_z \nu_{i,j,k+\frac{1}{2}} := \frac{1}{h}\left(\nu_{i,j,k+1} - \nu_{i,j,k}\right),
\end{aligned}$$

with A_x, $D_x : \mathcal{C}_{\mathrm{per}} \to \mathcal{E}^{\mathrm{x}}_{\mathrm{per}}$, A_y, $D_y : \mathcal{C}_{\mathrm{per}} \to \mathcal{E}^{\mathrm{y}}_{\mathrm{per}}$, A_z, $D_z : \mathcal{C}_{\mathrm{per}} \to \mathcal{E}^{\mathrm{z}}_{\mathrm{per}}$. Likewise,

$$\begin{aligned}
a_x \nu_{i,j,k} &:= \frac{1}{2}\left(\nu_{i+\frac{1}{2},j,k} + \nu_{i-\frac{1}{2},j,k}\right), \quad d_x \nu_{i,j,k} := \frac{1}{h}\left(\nu_{i+\frac{1}{2},j,k} - \nu_{i-\frac{1}{2},j,k}\right), \\
a_y \nu_{i,j,k} &:= \frac{1}{2}\left(\nu_{i,j+\frac{1}{2},k} + \nu_{i,j-\frac{1}{2},k}\right), \quad d_y \nu_{i,j,k} := \frac{1}{h}\left(\nu_{i,j+\frac{1}{2},k} - \nu_{i,j-\frac{1}{2},k}\right), \\
a_z \nu_{i,j,k} &:= \frac{1}{2}\left(\nu_{i,j,k+\frac{1}{2}} + \nu_{i,j,k-\frac{1}{2}}\right), \quad d_z \nu_{i,j,k} := \frac{1}{h}\left(\nu_{i,j,k+\frac{1}{2}} - \nu_{i,j,k-\frac{1}{2}}\right),
\end{aligned}$$

with a_x, $d_x : \mathcal{E}^{\mathrm{x}}_{\mathrm{per}} \to \mathcal{C}_{\mathrm{per}}$, a_y, $d_y : \mathcal{E}^{\mathrm{y}}_{\mathrm{per}} \to \mathcal{C}_{\mathrm{per}}$, and a_z, $d_z : \mathcal{E}^{\mathrm{z}}_{\mathrm{per}} \to \mathcal{C}_{\mathrm{per}}$. The standard 3D discrete Laplacian, $\Delta_h : \mathcal{C}_{\mathrm{per}} \to \mathcal{C}_{\mathrm{per}}$, is given by

$$\begin{aligned}
\Delta_h \nu_{i,j,k} &:= d_x(D_x\nu)_{i,j,k} + d_y(D_y\nu)_{i,j,k} + d_z(D_z\nu)_{i,j,k} \\
&= \frac{1}{h^2}\left(\nu_{i+1,j,k} + \nu_{i-1,j,k} + \nu_{i,j+1,k} + \nu_{i,j-1,k} + \nu_{i,j,k+1} + \nu_{i,j,k-1} - 6\nu_{i,j,k}\right).
\end{aligned}$$

Now we are ready to define the following grid inner products:

$$(\nu,\xi)_2 := h^3 \sum_{i,j,k=1}^{m} \nu_{i,j,k}\xi_{i,j,k}, \quad \nu,\, \xi \in \mathcal{C}_{\text{per}}, \qquad [\nu,\xi]_{\text{x}} := (a_x(\nu\xi),1)_2\,, \quad \nu,\, \xi \in \mathcal{E}^{\text{x}}_{\text{per}},$$

$$[\nu,\xi]_{\text{y}} := (a_y(\nu\xi),1)_2\,, \quad \nu,\, \xi \in \mathcal{E}^{\text{y}}_{\text{per}}, \qquad [\nu,\xi]_{\text{z}} := (a_z(\nu\xi),1)_2\,, \quad \nu,\, \xi \in \mathcal{E}^{\text{z}}_{\text{per}}.$$

We now define the following norms for cell-centered functions. If $\nu \in \mathcal{C}_{\text{per}}$, then $\|\nu\|_2^2 := (\nu,\nu)_2$; $\|\nu\|_p^p := (|\nu|^p,1)_2$ $(1 \le p < \infty)$, and $\|\nu\|_\infty := \max_{1\le i,j,k\le m} |\nu_{i,j,k}|$. Similarly, we define the gradient norms: for $\nu \in \mathcal{C}_{\text{per}}$,

$$\|\nabla_h \nu\|_2^2 := [D_x\nu, D_x\nu]_{\text{x}} + [D_y\nu, D_y\nu]_{\text{y}} + [D_z\nu, D_z\nu]_{\text{z}}\,.$$

Consequently,

$$\|\nu\|_{2,2}^2 := \|\nu\|_2^2 + \|\nabla_h\nu\|_2^2 + \|\Delta_h\nu\|_2^2\,.$$

In addition, the discrete energy $F_h(\phi) : \mathcal{C}_{\text{per}} \to \mathbb{R}$ is defined as

$$F_h(\phi) = \frac{1}{4}\|\phi\|_4^4 + \frac{1-\epsilon}{2}\|\phi\|_2^2 - \|\nabla_h\phi\|_2^2 + \frac{1}{2}\|\Delta_h\phi\|_2^2. \tag{2.18}$$

Chapter 3

Linearly Preconditioned Steepest Descent Methods

The content in this chapter has been published in [34], for more details please refer to [34].

3.1 Linearly Preconditioned Steepest Descent Methods

3.1.1 The Classical Setting: Linear Symmetric Positive Definite Systems in Finite Dimensions

Before we get to the general case, let us quickly review the convergence theory for preconditioned steepest decent methods for solving the linear system $\mathsf{A}\mathbf{u} = \mathbf{f}$, where $\mathsf{A} \in \mathbb{R}^{m\times m}_{\text{sym}}$ is positive definite. This is closely related to the preconditioned conjugate gradient (PCG) method, though may be less familiar to the reader. Solving $\mathsf{A}\mathbf{u} = \mathbf{f}$ is, of course, equivalent to minimizing the quadratic energy $E[\mathbf{v}] := \frac{1}{2}\mathbf{v}^T\mathsf{A}\mathbf{v} - \mathbf{v}^T\mathbf{f}$. Suppose that $\mathsf{L} \in \mathbb{R}^{m\times m}_{\text{sym}}$ is also positive definite. Here A is the *stiffness matrix* and L is the *pre-conditioner*. The idea is that $\mathsf{L} \approx \mathsf{A}$, but the former is "easier to invert." The

preconditioned steepest decent algorithm for approximating the solution to $\mathsf{A}\mathbf{u} = \mathbf{f}$ is given in Algorithm 1 [5, 56].

Algorithm 1 Preconditioned Steepest Descent

```
Input: u_0, f ∈ R^m
Compute residual: r_0 := f − A u_0
Set d_0 := L^{-1} r_0
Set d^0 ← −ḡ_0, k ← 0
for k = 0, ··· , k_max − 1 do
    Compute α_k := (d_k^T r_k)/(d_k^T A d_k)
    u_{k+1} := u_k + α_k d_k
    r_{k+1} := f − A u_{k+1}
    if ‖r_{k+1}‖ < tol or k = k_max − 1 then
        u_⋆ := u_{k+1}
        exit for loop
    else
        d_{k+1} := L^{-1} r_{k+1}
    end if
end for
```

Here $\mathbf{d}_k \in \mathbb{R}^m$ is called the *search direction* and $\mathbf{r}_k \in \mathbb{R}^m$ is called the *residual*. We observe that

$$\alpha_k = \operatorname*{argmin}_{\alpha\in\mathbb{R}} E[\mathbf{u}_k + \alpha\mathbf{d}_k] = \operatorname*{argzero}_{\alpha\in\mathbb{R}} \mathbf{d}_k^T \nabla E[\mathbf{u}_k + \alpha\mathbf{d}_k] = \frac{\mathbf{d}_k^T\mathbf{r}_k}{\mathbf{d}_k^T\mathsf{A}\mathbf{d}_k}.$$

We have the classical convergence result: for some $C > 0$,

$$\|\mathbf{u} - \mathbf{u}_k\|_{\mathsf{A}} \leq C \left(\frac{\kappa - 1}{\kappa + 1}\right)^k \|\mathbf{u} - \mathbf{u}_0\|_{\mathsf{A}},$$

where $\kappa := \frac{\lambda_m}{\lambda_1}$, and λ_m is the largest eigenvalue of $\mathsf{L}^{-1}\mathsf{A}$, and λ_1 is the smallest [5, 56, 70].

3.1.2 Non-Quadratic Energy Functionals in Generic Hilbert Spaces

Here we review the general theory for preconditioned steepest descent in a generic Hilbert space [3, 20, 28, 56]. Suppose that H is a (real) Hilbert space with the inner product $(\,\cdot\,,\,\cdot\,)_H$ and induced norm $\|\cdot\|_H$. We consider an energy functional $E[\,\cdot\,] : H \to \mathbb{R}$ with the following properties:

(E1) E is twice Fréchet differentiable for all points $\nu \in H$. For each fixed $\nu \in H$, $\delta E[\nu](\,\cdot\,) : H \to \mathbb{R}$ is the continuous (bounded) linear functional equal to the first Fréchet derivative at ν, and, for each fixed $\nu \in H$, $\delta^2 E[\nu](\,\cdot\,,\,\cdot\,) : H \times H \to \mathbb{R}$ is the continuous bilinear operator equal to the second Fréchet derivative at ν.

(E2) For every $\nu \in H$,

$$0 \leq \delta^2 E[\nu](\xi, \xi), \quad \forall\, \xi \in H, \tag{3.1}$$

and

$$0 < \delta^2 E[\nu](\xi, \xi), \quad \forall\, \xi \in H \setminus \{0\}\,. \tag{3.2}$$

This implies a strict convexity of E.

(E3) E is coercive with respect to the norm on H, *i.e.*, there exist constants $C_1 > 0$, $C_2 \geq 0$ such that

$$C_1 \|\nu\|_H^2 \leq E[\nu] + C_2, \quad \forall\, \nu \in H.$$

If E satisfies (E1) – (E3), it follows [20] that there is is a unique element $u \in H$ with the property that

$$E[u] \leq E[\nu], \quad \forall\, \nu \in H, \qquad \text{with} \qquad E[u] < E[\nu], \quad \text{for } \nu \neq u,$$

and this *minimizer* further satisfies

$$\delta E[u](\xi) = 0, \quad \forall\, \xi \in H.$$

We wish to construct, via preconditioned steepest descent (PSD), a sequence that converges to the unique minimizer. By H' we denote the continuous dual of H. When it is convenient, we use the symbol $\langle \cdot\,, \cdot \rangle_H : H' \times H \to \mathbb{R}$ to denote the dual pairing between H' and H. Consider a linear operator $\mathcal{L} : H \to H'$. This operator $\mathcal{L}$, which we call the *pre-conditioner* induces a bilinear form on H:

$$(\nu, \xi)_{\mathcal{L}} := \langle \mathcal{L}[\nu], \xi \rangle = \mathcal{L}[\nu](\xi), \quad \forall\, \nu, \xi \in H.$$

We assume that $\mathcal{L}$ satisfies the following properties:

(L1) $(\cdot\,, \cdot)_{\mathcal{L}} : H \times H \to \mathbb{R}$ is symmetric, *i.e.*,

$$(\nu, \xi)_{\mathcal{L}} = (\xi, \nu)_{\mathcal{L}}, \quad \forall\, \nu, \xi \in H;$$

(L2) $(\cdot\,, \cdot)_{\mathcal{L}}$ is continuous with respect to the standard topology of H, *i.e.*, there is some $C_3 > 0$ such that

$$|(\nu, \xi)_{\mathcal{L}}| \leq C_3 \|\nu\|_H \|\xi\|_H, \quad \forall\, \nu, \xi \in H;$$

(L3) $(\cdot\,, \cdot)_{\mathcal{L}}$ is coercive with respect to H, *i.e.*, there is some $C_4 > 0$ such that

$$C_4 \|\nu\|_H^2 \leq (\nu, \nu)_{\mathcal{L}}, \quad \forall\, \nu \in H.$$

It follows that $(\cdot\,, \cdot)_{\mathcal{L}} : H \times H \to \mathbb{R}$ is an inner product on H, equivalent to the primary inner product $(\cdot\,, \cdot)_H$. The induced norm, $\|\nu\|_{\mathcal{L}} := \sqrt{(\nu, \nu)_{\mathcal{L}}}$, is equivalent to the primary norm. By the Riesz Representation Theorem, if $f \in H'$, then there exists a unique $u_f \in H$ such that

$$(u_f, \xi)_{\mathcal{L}} = f(\xi) = \langle f, \xi \rangle, \quad \forall \xi \in H,$$

with

$$\|u_f\|_{\mathcal{L}} = \|f\|_{\mathcal{L}^{-1}} := \sup_{0 \neq \xi \in H} \frac{f(\xi)}{\|\xi\|_{\mathcal{L}}},$$

where the second norm is the $\mathcal{L}$-induced operator norm.

Suppose that $u^k \in H$ is given. We define the following *search direction* problem: find $d^k \in H$ such that

$$\left(d^k, \xi\right)_{\mathcal{L}} = -\delta E\left[u^k\right](\xi), \quad \forall \xi \in H. \tag{3.3}$$

We call d^k the kth *search direction*. In operator form, we write $\mathcal{L}[d^k] = -\delta E[u^k]$ in H'. The functional $-\delta E\left[u^k\right]$ is called the *residual* of u^k. By the Riesz Representation Theorem, we discover that

$$-\delta E\left[u^k\right](d^k) = \left\|d^k\right\|_{\mathcal{L}}^2 = \left\|\delta E\left[u^k\right]\right\|_{\mathcal{L}^{-1}}^2. \tag{3.4}$$

We then define the next iterate u^{k+1} as

$$u^{k+1} := u^k + \alpha_k d^k, \tag{3.5}$$

where $\alpha_k \in \mathbb{R}$ is the unique solution to

$$\alpha_k := \operatorname*{argmin}_{\alpha \in \mathbb{R}} E[u^k + \alpha d^k] = \operatorname*{argzero}_{\alpha \in \mathbb{R}} \delta E[u^k + \alpha d^k](d^k). \tag{3.6}$$

Therefore, we have the *fundamental orthogonality relation*

$$\delta E[u^k + \alpha_k d^k](d^k) = \delta E[u^{k+1}](d^k) = 0. \tag{3.7}$$

It follows that the sequence $\left\{u^k\right\}_{k=0}^{\infty} \subset H$ generated by the preconditioned steepest descent algorithm converges to the unique minimizer $u \in H$. We now wish to estimate the convergence rate.

3.1.3 Estimates of the Convergence Rate for the PSD Method

We summarize some standard results.

Proposition 3.1.1. *Suppose that E satisfies (E1) – (E3). It follows that, for any $\nu, \xi \in H$,*

$$\delta E[\nu](\xi - \nu) \leq E[\xi] - E[\nu] \leq \delta E[\xi](\xi - \nu), \tag{3.8}$$

and, consequently,

$$0 \leq (\delta E[\xi] - \delta E[\nu])\,(\xi - \nu).$$

Proposition 3.1.2. *Suppose that E satisfies (E1) – (E3). Let $\{u^k\}_{k=0}^{\infty} \subset H$ be computed via (3.5). Then, for every $k \geq 0$ we have $E[u^{k+1}] \leq E[u^k]$. Furthermore, $\alpha_k > 0$, as long as $u^k \neq u$.*

Proof. Using the orthogonality relation (3.7) and the convexity inequality (3.8), we find

$$E[u^{k+1}] - E[u^k] \leq \delta E[u^{k+1}](u^{k+1} - u^k) = \alpha_k \delta E[u^{k+1}](d^k) = 0.$$

Now, suppose $d^k \neq 0$. Then, by Taylor's theorem, (3.4), and (3.2),

$$E[u^{k+1}] = E[u^k] - \alpha_k \left\|d^k\right\|_{\mathcal{L}} + \frac{\alpha_k^2}{2}\delta^2 E[\theta^k](d^k, d^k) > E[u^k] - \alpha_k \left\|d^k\right\|_{\mathcal{L}}.$$

Equivalently, we get

$$\alpha_k \left\|d^k\right\|_{\mathcal{L}} > E[u^k] - E[u^{k+1}] \geq 0,$$

which implies that $\alpha_k > 0$. □

Proposition 3.1.3. *Suppose that E satisfies (E1) – (E3) and $u \in H$ is the unique minimizer of E. Then, for any $\xi \in H$,*

$$0 \leq E[\xi] - E[u] \leq (\delta E[\xi] - \delta E[u])\,(\xi - u) = \delta E[\xi](\xi - u),$$

and, consequently,

$$0 \leq E[u^k] - E[u] \leq \left(\delta E[u^k] - \delta E[u]\right)(u^k - u) = \delta E[u^k](u^k - u). \tag{3.9}$$

Proof. This follows immediately from (3.8), because $\delta E[u](\xi) = 0$, for all $\xi \in H$. □

Now, we make the following further assumptions about the pre-conditioner $\mathcal{L}$ with respect to the derivatives of the energy E:

(L4) There is a constant $C_5 > 0$ such that

$$C_5 \|\xi - \nu\|_{\mathcal{L}}^2 \leq (\delta E[\xi] - \delta E[\nu])(\xi - \nu), \tag{3.10}$$

for all $\nu, \xi \in H$.

(L5) Suppose $B := \{\nu \in H \mid E[\nu] \leq E_0\}$ is non-empty. (This is the the case if, for example, one chooses $E_0 = E[0]$.) There is a constant $C_6 = C_6(E_0) > 0$ such that, for all $\nu \in B$, and any $\xi \in H$,

$$\left|\delta^2 E[\nu](\xi, \xi)\right| \leq C_6 \|\xi\|_{\mathcal{L}}^2. \tag{3.11}$$

Remark 3.1.4. *We note that, practically speaking, (L5) is harder of the last two conditions to enforce. In some sense, if the norm induced by $\mathcal{L}$ is not "strong" enough, then there does not exist $C_6 > 0$ so that (L5) is satisfied.*

Theorem 3.1.5. *Suppose that assumptions (E1) – (E3) and (L1) – (L5) are valid. Let $\{u^k\}_{k=0}^{\infty} \subset H$ be the sequence generated by (3.5). Then*

$$0 \leq E[u^k] - E[u] \leq (C_7)^k (E[u^0] - E[u]), \tag{3.12}$$

where

$$0 < C_7 := 1 - \frac{C_5}{2C_6} < 1. \tag{3.13}$$

Proof. Consider the function $g(\alpha) := E[u^k + \alpha d^k] - E[u^k]$, $\alpha \in \mathbb{R}$. Then $g(0) = 0$, and g has a global minimum at $\alpha_k > 0$. By coercivity and continuity of E, there is a β_k, $\alpha_k < \beta_k < \infty$, such that $g(\beta_k) = 0$, and, for all $\alpha \in [0, \beta_k]$,

$$E[u^k + \alpha d^k] \leq E[u^k] \leq E[u^0] =: E_0.$$

By Taylor's theorem, there is a $\gamma = \gamma(u^k, d^k, \alpha) \in (0, 1)$, such that

$$E[u^k + \alpha d^k] - E[u^k] = \alpha \delta E[u^k](d^k) + \frac{\alpha^2}{2} \delta^2 E[\theta^k](d^k, d^k),$$

where $\theta^k := u^k + (1 - \gamma)\alpha d^k$. By convexity of E,

$$E[\theta^k] \leq \gamma E[u^k] + (1 - \gamma) E[u^k + \alpha d^k] \leq E[u^k] \leq E[u^0] = E_0.$$

Using estimate (3.11) – with the set B defined with respect to $E_0 = E[u^0]$ – and norm equality (3.4), we get, for all $\alpha \in [0, \beta_k]$,

$$\begin{aligned} g(\alpha) = E[u^k + \alpha d^k] - E[u^k] &\leq \alpha \delta E[u^k](d^k) + \frac{\alpha^2}{2} C_6 \left\| d^k \right\|_{\mathcal{L}}^2 \\ &= \big(- \alpha + \frac{\alpha^2}{2} C_6 \big) \left\| \delta E[u^k] \right\|_{\mathcal{L}^{-1}}^2 =: f(\alpha). \end{aligned} \tag{3.14}$$

Now, the function $f(\alpha)$ is quadratic, $f(0) = 0$, $f(\beta_k) \geq g(\beta_k) = 0$, and $f'(0) < 0$. See Figure 3.1. Thus f has a minimum in $(0, \beta_k)$. In fact, the minimum is achieved at $0 < \sigma_k := \frac{1}{C_6} < \beta_k$. Then we have

$$E[u^k + \alpha_k d^k] - E[u^k] \leq g(\sigma_k) = E[u^k + \sigma_k d^k] - E[u^k] \leq -\frac{1}{2C_6} \left\| \delta E[u^k] \right\|_{\mathcal{L}^{-1}}^2 = f(\sigma_k),$$

or, equivalently,

$$E[u^k] - E[u^{k+1}] \geq \frac{1}{2C_6} \left\| \delta E[u^k] \right\|_{\mathcal{L}^{-1}}^2 .$$

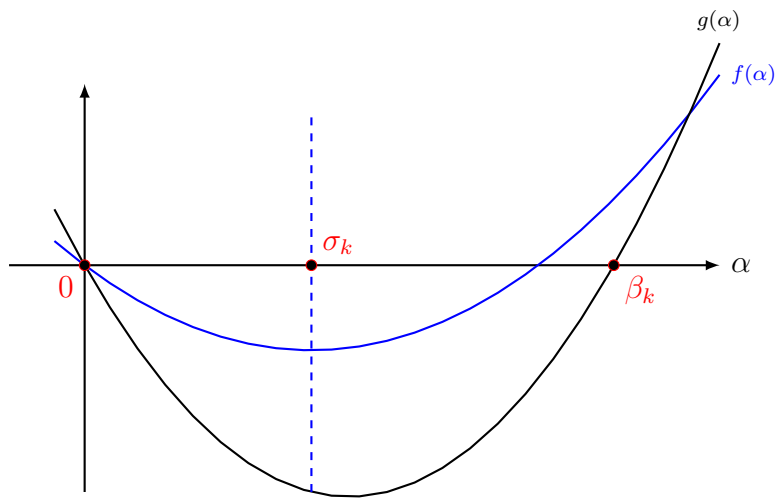


Figure 3.1: The functions $g(\alpha) = E[u^k + \alpha d^k] - E[u^k]$ and $f(\alpha) = \left(- \alpha + \frac{\alpha^2}{2} C_6 \right) \left\| \delta E[u^k] \right\|^2_{\mathcal{L}^{-1}}$ from (3.14). The function g, which is strictly convex, is dominated by the function f, which is quadratic, on the interval $[0, \beta_k]$.

Now, using estimates (3.9) and (3.10) we obtain

$$0 \leq E[u^k] - E[u] \leq \frac{1}{C_5} \left\| \delta E[u^k] \right\|^2_{\mathcal{L}^{-1}} .$$

Combining the last two estimates, we get the result

$$0 \leq E[u^k] - E[u] \leq \frac{2C_6}{C_5} \left(E[u^k] - E[u^{k+1}] \right) ,$$

or, equivalently,

$$0 \leq E[u^{k+1}] - E[u] \leq \left(\frac{2C_6}{C_5} - 1 \right) \left(E[u^k] - E[u^{k+1}] \right) .$$

Since $E[u^{k+1}] > E[u]$, as long as $u^{k+1} \neq u$, and $E[u^k] \geq E[u^{k+1}]$, this last inequality implies that

$$0 < \frac{C_5}{2C_6} < 1.$$

A little more manipulation reveals the equivalent inequality

$$0 \leq E[u^{k+1}] - E[u] \leq \left(1 - \frac{C_5}{2C_6}\right)\left(E[u^k] - E[u]\right),$$

and the result follows. □

If the following property holds, we get a simple corollary of the last theorem.

(L6) There is a constant $C_8 > 0$, such that, for every $v, w \in H$,

$$C_8 \|w\|_{\mathcal{L}}^2 \leq |\delta^2 E[\nu](w, w)|. \tag{3.15}$$

This implies the strict convexity of E and is, therefore, stronger that (E2).

Corollary 3.1.6. *Suppose that assumptions (E1) – (E3) and (L1) – (L6) are valid. Let* $\left\{u^k\right\}_{k=0}^{\infty} \subset H$ *be the sequence generated by* (3.5)*, and define* $e^k := u - u^k$*. Then*

$$\left\|e^k\right\|_{\mathcal{L}}^2 \leq (C_7)^k \frac{E[u^0] - E[u]}{C_8}. \tag{3.16}$$

Proof. By Taylor's theorem and estimate (3.15), we have

$$\begin{aligned} E[u^k] - E[u] &= \delta E[u](e^k) + \frac{1}{2}\delta^2 E[\theta^k](e^k, e^k) \\ &= \frac{1}{2}\delta^2 E[\theta^k](e^k, e^k) \geq C_5 \left\|e^k\right\|_{\mathcal{L}}, \end{aligned} \tag{3.17}$$

where θ^k is in the line segment from u^k to u. The result follows from (3.12). □

3.2 Nonlinear Elliptic Equations on Periodic Domains

3.2.1 A Fourth-Order Regularized p-Laplacian Problem

We consider the following weak formulation of (1.1): given $f \in L^2_{\text{per}}(\Omega)$, find $u \in H^2_{\text{per}}(\Omega)$ such that

$$(u, \xi) + s\left(|\nabla u|^{p-2}\nabla u, \nabla \xi\right) + s\varepsilon^2 \left(\Delta u, \Delta \xi\right) = \left(f, \xi\right), \quad \forall\, \xi \in H^2_{\text{per}}(\Omega), \tag{3.18}$$

where $0 < \varepsilon \leq 1$ and $s > 0$ are parameters. Equation (3.18) is mass conservative in the following sense: $(u - f, 1) = 0$. One can show that the solution of the weak formulation is a minimizer of the following energy: for any $\nu \in H^2_{\text{per}}(\Omega)$,

$$E[\nu] := \frac{1}{2}\|\nu - f\|^2 + \frac{s}{p}\|\nabla \nu\|^p_{L^p} + \frac{s\varepsilon^2}{2}\|\Delta \nu\|^2. \tag{3.19}$$

It is not difficult to show that E satisfies (E1) – (E3). The first derivative of E at a point ν may be calculated as follows: for any $\xi \in H^2_{\text{per}}(\Omega)$,

$$d_\tau E[\nu + \tau\xi]|_{\tau=0} = \delta E[\nu](\xi) = (\nu - f, \xi) + s\left(|\nabla \nu|^{p-2}\nabla \nu, \nabla \xi\right) + s\varepsilon^2 \left(\Delta \nu, \Delta \xi\right).$$

Thus, our original problem is equivalent to the following: find $u \in H^2_{\text{per}}(\Omega)$, such that, for all $\xi \in H^2_{\text{per}}(\Omega)$, $\delta E[u](\xi) = 0$, which is equivalent to (3.18). This problem has a unique solution, which is, in turn, the unique minimizer of the energy (3.19):

$$u := \underset{\nu \in H^2_{\text{per}}(\Omega)}{\text{argmin}}\ E[\nu].$$

The following estimate is holds: for all $\nu, \xi \in H^2_{\text{per}}(\Omega)$,

$$|\delta E[\nu](\xi)| \leq \|\nu - f\| \cdot \|\xi\| + s \|\nabla \nu\|^{p-1}_{L^p} \|\nabla \xi\|_{L^p} + s\varepsilon^2 \|\Delta \nu\| \cdot \|\Delta \xi\| .$$

The second variation is a continuous bilinear operator. Given a fixed $\nu \in H^2_{\text{per}}(\Omega)$, the action of the second variation on the arbitrary pair $(\xi, \eta) \in H^2_{\text{per}}(\Omega) \times H^2_{\text{per}}(\Omega)$ is given by

$$\begin{aligned} \delta^2 E[\nu](\xi, \eta) = \ & (\xi, \eta) + s \left(|\nabla \nu|^{p-2} \nabla \xi, \nabla \eta \right) \\ & + (p-2) s \left(|\nabla \nu|^{p-4} \nabla \nu \cdot \nabla \xi, \nabla \nu \cdot \nabla \eta \right) + s\varepsilon^2 \left(\Delta \xi, \Delta \eta \right), \end{aligned}$$

and we have the bound

$$\begin{aligned} \left| \delta^2 E[\nu](\xi, \eta) \right| \leq \ & \|\xi\| \cdot \|\eta\| + s \|\nabla \nu\|^{p-2}_{L^p} \|\nabla \xi\|_{L^p} \|\nabla \eta\|_{L^p} \\ & + (p-2) s \|\nabla \nu\|^{p-2}_{L^p} \|\nabla \xi\|_{L^p} \|\nabla \eta\|_{L^p} + s\varepsilon^2 \|\Delta \xi\| \cdot \|\Delta \eta\| . \end{aligned} \tag{3.20}$$

For this problem we define the pre-conditioner $\mathcal{L} : H^2_{\text{per}}(\Omega) \to H^{-2}_{\text{per}}(\Omega)$ via

$$\langle \mathcal{L}[\nu], \xi \rangle := (\nu, \xi) + s \left(\nabla \nu, \nabla \xi \right) + s\varepsilon^2 \left(\Delta \nu, \Delta \xi \right), \quad \forall \ \xi \in H^2_{\text{per}}(\Omega).$$

Clearly, this is a positive, symmetric operator, and it satisfies assumptions (L1) – (L3), and one can see the similarities with the nonlinear operator in (3.18). We now wish to find the positive constants C_5, C_6, C_8 such that assumptions (L4) – (L6) are satisfied in addition.

Remark 3.2.1. *We could also consider the possibility of changing the metric in the descent direction calculation by, for example, defining the linear operator* $\mathcal{L}_k : H^2_{\text{per}}(\Omega) \to H^{-2}_{\text{per}}(\Omega)$ *via*

$$\langle \mathcal{L}_k[\nu], \xi \rangle := (\nu, \xi) + s \left(\left| \nabla u^k \right|^{p-2} \nabla \nu, \nabla \xi \right) + s\varepsilon^2 \left(\Delta \nu, \Delta \xi \right), \quad \forall \ \xi \in H^2_{\text{per}}(\Omega).$$

This is similar to the idea in [56]. The search direction *is then found as follows: find* $d^k \in H^2_{\text{per}}(\Omega)$ *such that*

$$\langle \mathcal{L}_k[d^k], \xi \rangle = -\delta E\left[u^k\right](\xi), \quad \forall\, \xi \in H^2_{\text{per}}(\Omega).$$

Our theory does not cover this case, and we will not consider it further here. But we plan to examine this in a future work.

Lemma 3.2.2. *For any* $\nu, \xi \in H^2_{\text{per}}(\Omega)$,

$$C_5 \left\|\xi - \nu\right\|^2_{\mathcal{L}} \leq \left(\delta E[\xi] - \delta E[\nu]\right)(\xi - \nu), \tag{3.21}$$

where $C_5 = \min\left(\frac{1}{2}, \varepsilon s^{-\frac{1}{2}}\right)$. *Let* E_0 *be given, such that* $B := \left\{\nu \in H^2_{\text{per}}(\Omega) \;\middle|\; E[\nu] \leq E_0\right\}$ *is non-empty. For any* $\nu \in B$ *and any* $\xi \in H^2_{\text{per}}(\Omega)$,

$$\left|\delta^2 E[\nu](\xi, \xi)\right| \leq C_6 \left\|\xi\right\|^2_{\mathcal{L}}, \tag{3.22}$$

where

$$C_6 = \begin{cases} 1 + \frac{1}{p}(p-1)^{\frac{2p-1}{p}} \varepsilon^{\frac{-2(p-1)}{p}} s^{\frac{1}{p}} C_9^2 C_{10}^{p-2} & \textit{for} \quad p \in [2, \infty), \quad d = 2, \\ 1 + (p-1)\left(\frac{4p}{6-p}\right)^{\frac{p-6}{4p}} \left(\frac{4p}{5p-6}\right)^{\frac{6-5p}{4p}} s^{\frac{6-p}{4p}} \varepsilon^{\frac{6-5p}{2p}} C_9^2 C_{10}^{p-2} & \textit{for} \quad p \in [2, 6), \quad d = 3, \\ 1 + (p-1)\,\varepsilon^{-2} C_9^2 C_{10}^{p-2} & \textit{for} \quad p = 6, \quad d = 3, \end{cases} \tag{3.23}$$

and $C_{10} = (pE_0)^{\frac{1}{p}}$. *We can take* $C_8 = C_5$ *to satisfy estimate* (3.15) *of assumption (L6).*

Proof. Clearly

$$\begin{aligned} \left(\delta E[\xi] - \delta E[\nu]\right)(\xi - \nu) = \; & \left\|\xi - \nu\right\|^2 + s\varepsilon^2 \left\|\Delta(\xi - \nu)\right\|^2 \\ & + s\left(|\nabla \xi|^{p-2}\nabla\xi - |\nabla\nu|^{p-2}\nabla\nu, \nabla(\xi - \nu)\right). \end{aligned}$$

In addition, the following estimate is available:

$$\left(|\nabla\xi|^{p-2}\nabla\xi - |\nabla\nu|^{p-2}\nabla\nu, \nabla(\xi-\nu)\right) \geq \frac{1}{2^{p-2}} \|\nabla(\xi-\nu)\|_{L^p}^p \geq 0, \quad \text{for } p \geq 2. \quad (3.24)$$

The simple interpolation inequality

$$\|\nabla\xi\|^2 \leq \|\xi\| \cdot \|\Delta\xi\|, \quad \forall \xi \in H^2_{\text{per}}(\Omega),$$

in conjunction with Young's inequality yields

$$\frac{1}{2}\|\xi-\nu\|^2 + \frac{s\varepsilon^2}{2}\|\Delta(\xi-\nu)\|^2 \geq s^{\frac{1}{2}}\varepsilon\|\xi-\nu\| \cdot \|\Delta(\xi-\nu)\| \geq s^{\frac{1}{2}}\varepsilon\|\nabla(\xi-\nu)\|^2.$$

As a consequence, we get

$$\begin{aligned}(\delta E[\xi] - \delta E[\nu])(\xi-\nu) \geq\ & \|\xi-\nu\|^2 + s\varepsilon^2\|\Delta(\xi-\nu)\|^2 \\ \geq\ & \frac{1}{2}\|\xi-\nu\|^2 + \frac{1}{2}s\varepsilon^2\|\Delta(\xi-\nu)\|^2 + s^{\frac{1}{2}}\varepsilon\|\nabla(\xi-\nu)\|^2,\end{aligned}$$

and we conclude that estimate (3.21) is valid by choosing $C_5 = \min(\frac{1}{2}, \varepsilon s^{-\frac{1}{2}})$.

Next we derive (3.22). Suppose $\nu \in B$. From (3.20) we have

$$\left|\delta^2 E[\nu](\xi,\xi)\right| \leq \|\xi\|^2 + (p-1)s\|\nabla\nu\|_{L^p}^{p-2}\|\nabla\xi\|_{L^p}^2 + s\varepsilon^2\|\Delta\xi\|^2. \quad (3.25)$$

Now, since $\nu \in B$,

$$\|\nabla\nu\|_{L^p} \leq (pE_0)^{\frac{1}{p}} =: C_{10}.$$

Suppose that $d = 2$. An application of the Sobolev inequality (2.1) in Lemma 2.1.2 indicates that

$$\begin{aligned}p^{\frac{1}{p}}\left(\frac{p}{p-1}\right)^{\frac{p-1}{p}}\varepsilon^{\frac{2(p-1)}{p}}s^{\frac{(p-1)}{p}}C_9^{-2}\|\nabla\xi\|_{L^p}^2 \leq\ & p^{\frac{1}{p}}\|\xi\|^{\frac{2}{p}} \cdot \left(\frac{p}{p-1}\right)^{\frac{p-1}{p}}\varepsilon^{\frac{2(p-1)}{p}}s^{\frac{p-1}{p}}\|\Delta\xi\|^{\frac{2(p-1)}{p}} \\ \leq\ & \|\xi\|^2 + s\varepsilon^2\|\Delta\xi\|^2,\end{aligned}$$

where Young's inequality is applied in the second step. It follows that,

$$(p-1)s\left\|\nabla\nu\right\|_{L^p}^{p-2}\left\|\nabla\xi\right\|_{L^p}^2 \leq \frac{1}{p}\left(p-1\right)^{\frac{2p-1}{p}}\varepsilon^{\frac{-2(p-1)}{p}}s^{\frac{1}{p}}C_9^2C_{10}^{p-2}\left(\left\|\xi\right\|^2+s\varepsilon^2\left\|\Delta\xi\right\|^2\right). \tag{3.26}$$

Substituting (3.26) in (3.25) yields

$$\left|\delta^2E[\nu](\xi,\xi)\right| \leq \left(1+\frac{1}{p}\left(p-1\right)^{\frac{2p-1}{p}}\varepsilon^{\frac{-2(p-1)}{p}}s^{\frac{1}{p}}C_9^2C_{10}^{p-2}\right)(\left\|\xi\right\|^2+\varepsilon^2\left\|\Delta\xi\right\|^2).$$

We conclude that estimate (3.22) is valid by choosing

$$C_6 = 1+\frac{1}{p}\left(p-1\right)^{\frac{2p-1}{p}}\varepsilon^{\frac{-2(p-1)}{p}}s^{\frac{1}{p}}C_9^2C_{10}^{p-2}.$$

Note that both C_9 and C_{10} are ε and s independent. Following the similar arguments, for $p\in[2,6),\ d=3$, we get

$$C_6 = 1+\left(p-1\right)\left(\frac{4p}{6-p}\right)^{\frac{p-6}{4p}}\left(\frac{4p}{5p-6}\right)^{\frac{6-5p}{4p}}s^{\frac{6-p}{4p}}\varepsilon^{\frac{6-5p}{2p}}C_9^2C_{10}^{p-2}.$$

For the case $p=6,\ d=3$, the Sobolev inequality (2.1) degenerates to $\left\|\nabla\xi\right\|_{L^p}\leq C_9\left\|\Delta\xi\right\|$, for any $\xi\in H^2_{\text{per}}(\Omega)$. Hence, we have

$$\left\|\xi\right\|^2+s\varepsilon^2\left\|\Delta\xi\right\|^2 \geq s\varepsilon^2\left\|\Delta\xi\right\|^2 \geq s\varepsilon^2C_9^{-2}\left\|\nabla\xi\right\|_{L^p}^2,$$

and

$$\left|\delta^2E[\nu](\xi,\xi)\right| \leq \left(1+\left(p-1\right)\varepsilon^{-2}C_9^2C_{10}^{p-2}\right)(\left\|\xi\right\|^2+\varepsilon^2\left\|\Delta\xi\right\|^2).$$

Therefore, estimate (3.22) is valid by choosing

$$C_6 = 1+\left(p-1\right)\varepsilon^{-2}C_9^2C_{10}^{p-2}.$$

That we can take $C_8=C_5$ is the result of a simple calculation that we omit for the sake of brevity. The proof is complete. □

3.2.2 A Sixth-Order Regularized p-Laplacian Problem

We now study problem (1.2a) – (1.2b). A weak formulation is given as follows: for $f, g \in L^2_{\text{per}}(\Omega)$, find $u \in H^2_{\text{per}}(\Omega)$ and $w \in H^1_{\text{per}}(\Omega)$ such that

$$(u, \chi) + (\nabla w, \nabla \chi) = (g, \chi), \quad \forall\, \chi \in H^1_{\text{per}}(\Omega), \tag{3.27a}$$

$$s\lambda\,(u, \xi) + s\left(|\nabla u|^{p-2}\nabla u, \nabla \xi\right) + s\varepsilon^2\,(\Delta u, \Delta \xi) - (w, \xi) = (f, \xi), \quad \forall\, \xi \in H^2_{\text{per}}(\Omega), \tag{3.27b}$$

where $\lambda \geq 0$, and $\varepsilon \in (0, 1]$. This problem is mass-conservative, in the sense that $(u - g, 1) = 0$, and $(w - s\lambda g + f, 1) = 0$, and it can be recast as a minimization problem with an energy that involves the $\mathring{H}^{-1}_{\text{per}}$ norm. In particular, for any $\nu \in \mathring{H}^2_{\text{per}}(\Omega)$ we define

$$\begin{aligned} E[\nu] &= \frac{1}{2}\left(\nu - g + \bar{g}, \mathsf{T}[\nu - g + \bar{g}]\right) + \frac{\lambda s}{2}\left\|\nu + \bar{g}\right\|^2 - (\nu, f) + \frac{s}{p}\left\|\nabla \nu\right\|^p_{L^p} + \frac{s\varepsilon^2}{2}\left\|\Delta \nu\right\|^2 \\ &= \frac{1}{2}\left\|\nu - g + \bar{g}\right\|^2_{\mathring{H}^{-1}_{\text{per}}} + \frac{\lambda s}{2}\left\|\nu + \bar{g}\right\|^2 - (\nu, f) + \frac{s}{p}\left\|\nabla \nu\right\|^p_{L^p} + \frac{s\varepsilon^2}{2}\left\|\Delta \nu\right\|^2. \end{aligned} \tag{3.28}$$

Observe that $\nu - g + \bar{g} \in \mathring{H}^{-1}_{\text{per}}$, which is required for this energy to make sense. It is straightforward to show that E satisfies (E1) – (E3), with respect to the Hilbert space $H = \mathring{H}^2_{\text{per}}(\Omega)$. The first variation of E is given as follows: for any $\xi \in \mathring{H}^2_{\text{per}}(\Omega)$,

$$\begin{aligned} d_\tau E[\nu + \tau\xi]|_{\tau=0} = \delta E[\nu](\xi) = {} & (\mathsf{T}[\nu - g + \bar{g}], \xi) + \lambda s\,(\nu + \bar{g}, \chi) - (f, \xi) \\ & + s\left(|\nabla \nu|^{p-2}\nabla \nu, \nabla \xi\right) + s\varepsilon^2\,(\Delta \nu, \Delta \xi). \end{aligned}$$

The unique minimizer of E – let us call it $u_\star \in \mathring{H}^2_{\text{per}}(\Omega)$ for the moment – satisfies $\delta E[u_\star](\xi) = 0$, for all $\xi \in \mathring{H}^2_{\text{per}}(\Omega)$. By the definition of the T operator, there is a unique element $w_\star \in \mathring{H}^1_{\text{per}}(\Omega)$ such that

$$w_\star := -\mathsf{T}[u_\star - g + \bar{g}].$$

Therefore, we have, for all $\xi \in \mathring{H}^2_{\text{per}}(\Omega)$,

$$s\lambda\left(u_\star + \bar{g}, \xi\right) + s\left(|\nabla u_\star|^{p-2}\nabla u_\star, \nabla\xi\right) + s\varepsilon^2\left(\Delta u_\star, \Delta\xi\right) - \left(w_\star, \xi\right) = \left(f, \xi\right).$$

Setting $u := u_\star + \bar{g}$ and $w := w_\star + s\lambda\bar{g} - \bar{f}$ and using the fact that ξ is of zero mean, we have

$$s\lambda\left(u, \xi\right) + s\left(|\nabla u|^{p-2}\nabla u, \nabla\xi\right) + s\varepsilon^2\left(\Delta u, \Delta\xi\right) - \left(w, \xi\right) = \left(f, \xi\right), \quad \forall\, \xi \in \mathring{H}^2_{\text{per}}.$$

Using the definition of the T operator again, we conclude that $w_\star \in \mathring{H}^1_{\text{per}}(\Omega)$ satisfies

$$\left(\nabla w_\star, \nabla\chi\right) = -\left(u_\star - g + \bar{g}, \chi\right),$$

for all $\chi \in \mathring{H}^1_{\text{per}}(\Omega)$, which implies that

$$\left(\nabla w, \nabla\chi\right) = -\left(u - g, \chi\right).$$

It follows that solving (3.27a) – (3.27b) is equivalent to minimizing the coercive, strictly convex energy (3.28), after the appropriate affine change of variables.

The second variation of E is a continuous bilinear operator. Given a fixed $\nu \in \mathring{H}^2_{\text{per}}(\Omega)$, the action of the second variation on the arbitrary pair $(\xi, \eta) \in \mathring{H}^2_{\text{per}}(\Omega) \times \mathring{H}^2_{\text{per}}(\Omega)$ becomes

$$\begin{aligned}\delta^2 E[\nu](\xi, \eta) = &\ \left(\xi, \mathsf{T}[\eta]\right) + \lambda s\left(\xi, \eta\right) + s\left(|\nabla\nu|^{p-2}\nabla\xi, \nabla\eta\right) \\ &+ (p-2)s\left(|\nabla\nu|^{p-4}\nabla\nu\cdot\nabla\xi, \nabla\nu\cdot\nabla\eta\right) + s\varepsilon^2\left(\Delta\xi, \Delta\eta\right).\end{aligned}$$

Similar to the estimate in the fourth-order case (3.20), we have the bound

$$\begin{aligned}\left|\delta^2 E[\nu](\xi, \eta)\right| \leq &\ \|\xi\|_{\mathring{H}^{-1}_{\text{per}}}\|\eta\|_{\mathring{H}^{-1}_{\text{per}}} + \lambda s\|\xi\|\cdot\|\eta\| + s\|\nabla\nu\|^{p-2}_{L^p}\|\nabla\xi\|_{L^p}\|\nabla\eta\|_{L^p} \\ &+ (p-2)s\|\nabla\nu\|^{p-2}_{L^p}\|\nabla\xi\|_{L^p}\|\nabla\eta\|_{L^p} + s\varepsilon^2\|\Delta\xi\|\cdot\|\Delta\eta\|,\end{aligned}$$

which implies that

$$\left|\delta^2 E[\nu](\xi,\xi)\right| \leq \|\xi\|^2_{\mathring{H}^{-1}_{\text{per}}} + s\lambda \|\xi\|^2 + (p-1)s \|\nabla\nu\|^{p-2}_{L^p} \|\nabla\xi\|^2_{L^p} + s\varepsilon^2 \|\Delta\xi\|^2, \quad (3.29)$$

for all $\nu, \xi \in \mathring{H}^2_{\text{per}}(\Omega)$.

For the sixth order problem, we define the pre-conditioner $\mathcal{L} : \mathring{H}^2_{\text{per}}(\Omega) \to \mathring{H}^{-2}_{\text{per}}(\Omega)$ via

$$\langle \mathcal{L}[\nu], \xi\rangle := s\lambda\,(\nu,\xi) + (\nu,\xi)_{\mathring{H}^{-1}_{\text{per}}} + s\,(\nabla\nu,\nabla\xi) + s\varepsilon^2\,(\Delta\nu,\Delta\xi), \quad \forall\, \xi \in \mathring{H}^2_{\text{per}}(\Omega). \quad (3.30)$$

This operator satisfies (L1) – (L3). To show that it satisfies (L3) – (L6), we need some technical results.

We can now find the coefficients C_5, C_6, and C_8, which establish properties (L4) – (L6) and therefore guarantee the geometric convergence of the PSD method for the sixth-order case.

Lemma 3.2.3. *For any $\nu, \xi \in \mathring{H}^2_{\text{per}}(\Omega)$, we have*

$$C_5 \|\xi - \nu\|^2_{\mathcal{L}} \leq (\delta E[\xi] - \delta E[\nu])\,(\xi - \nu), \quad (3.31)$$

where $C_5 = \min\left(\frac{1}{3}, \varepsilon^{\frac{4}{3}} s^{-\frac{1}{3}}\right)$. Let E_0 be given such that $B := \left\{\xi \in \mathring{H}^2_{\text{per}}(\Omega) \;\middle|\; E[\xi] \leq E_0\right\}$ is non-empty. For any $\nu \in B$ and any $\xi \in \mathring{H}^2_{\text{per}}(\Omega)$, the following estimate is valid:

$$\left|\delta^2 E[\nu](\xi,\xi)\right| \leq C_6 \|\xi\|^2_{\mathcal{L}}, \quad (3.32)$$

where

$$C_6 = \begin{cases} 1 + (p-1)\left(\frac{3p}{2}\right)^{-\frac{2}{3p}} \left(\frac{3p}{3p-2}\right)^{\frac{2-3p}{3p}} \varepsilon^{\frac{4-6p}{3p}} s^{\frac{2}{3p}} C_9^2 C_{10}^{p-2}, & \text{for } \; p \in [2,\infty),\; d = 2, \\ 1 + (p-1)\left(\frac{6p}{6-p}\right)^{\frac{p-6}{6p}} \left(\frac{6p}{7p-6}\right)^{\frac{6-7p}{6p}} \varepsilon^{\frac{6-7p}{3p}} s^{\frac{6-p}{6p}} C_9^2 C_{10}^{p-2}, & \text{for } \; p \in [2,6),\; d = 3, \\ 1 + (p-1)\varepsilon^{-2} C_9^2 C_{10}^{p-2}, & \text{for } \; p = 6,\; d = 3, \end{cases} \quad (3.33)$$

and $C_{10} = (pE_0)^{\frac{1}{p}}$. *We can take* $C_8 = C_5$ *to satisfy estimate* (3.15) *of assumption (L6).*

Proof. The proof is similar to that of Lemma 3.2.2. Using (A.51) again, we have

$$\begin{aligned}
(\delta E[\xi] - \delta E[\nu])(\xi - \nu) &= s\lambda \|\xi - \nu\|^2 + \|\xi - \nu\|^2_{\mathring{H}^{-1}_{\text{per}}} + s\varepsilon^2 \|\Delta(\xi - \nu)\|^2 \\
&\quad + s\left(|\nabla\xi|^{p-2}\nabla\xi - |\nabla\nu|^{p-2}\nabla\nu, \nabla(\xi - \nu)\right). \\
&\geq s\lambda \|\xi - \nu\|^2 + \|\xi - \nu\|^2_{\mathring{H}^{-1}_{\text{per}}} + s\varepsilon^2 \|\Delta(\xi - \nu)\|^2 \\
&\geq s\lambda \|\xi - \nu\|^2 + \frac{2}{3}\|\xi - \nu\|^2_{\mathring{H}^{-1}_{\text{per}}} + \frac{1}{3}s\varepsilon^2 \|\Delta(\xi - \nu)\|^2 \\
&\quad + s^{\frac{2}{3}}\varepsilon^{\frac{4}{3}} \|\nabla(\xi - \nu)\|^2,
\end{aligned}$$

where the last step is a consequence of the interpolation inequality (2.3):

$$\frac{1}{3}\|\xi - \nu\|^2_{\mathring{H}^{-1}_{\text{per}}} + \frac{2}{3}s\varepsilon^2 \|\Delta(\xi - \nu)\|^2 \geq s^{\frac{2}{3}}\varepsilon^{\frac{4}{3}} \|\xi - \nu\|^{\frac{2}{3}}_{\mathring{H}^{-1}_{\text{per}}} \|\Delta(\xi - \nu)\|^{\frac{4}{3}} \geq s^{\frac{2}{3}}\varepsilon^{\frac{4}{3}} \|\nabla(\xi - \nu)\|^2.$$

We conclude that estimate (3.31) holds by choosing $C_5 = \min(\frac{1}{3}, \varepsilon^{\frac{4}{3}} s^{-\frac{1}{3}})$.

Next we derive (3.32). Inequality (3.29) yields

$$\left|\delta^2 E[\nu](\xi, \xi)\right| \leq s\lambda \|\xi\|^2 + \|\xi\|^2_{\mathring{H}^{-1}_{\text{per}}} + (p-1)s \|\nabla\nu\|^{p-2}_{L^p} \|\nabla\xi\|^2_{L^p} + s\varepsilon^2 \|\Delta\xi\|^2.$$

Since $\nu \in B$, $\|\nabla\nu\|_{L^p} \leq (E(u^0))^{\frac{1}{p}} =: C_{10}$. Suppose that $d = 2$. An application of the Sobolev inequality (2.6) from Lemma 2.1.4 indicates that, for every $\xi \in \mathring{H}^2_{\text{per}}(\Omega)$,

$$\begin{aligned}
&\left(\frac{3p}{2}\right)^{\frac{2}{3p}} \left(\frac{3p}{3p-2}\right)^{\frac{3p-2}{3p}} \varepsilon^{\frac{6p-4}{3p}} s^{\frac{3p-2}{3p}} C_9^{-2} \|\nabla\xi\|^2_{L^p} \\
&\leq \left(\frac{3p}{2}\right)^{\frac{2}{3p}} \|\xi\|^{\frac{4}{3p}}_{\mathring{H}^{-1}_{\text{per}}} \left(\frac{3p}{3p-2}\right)^{\frac{3p-2}{3p}} \varepsilon^{\frac{6p-4}{3p}} s^{\frac{3p-2}{3p}} \|\Delta\xi\|^{\frac{6p-4}{3p}} \\
&\leq \|\xi\|^2_{\mathring{H}^{-1}_{\text{per}}} + s\varepsilon^2 \|\Delta\xi\|^2,
\end{aligned}$$

where, in the last step, we applied Young's inequality. It follows that,

$$
\begin{aligned}
&(p-1)s\,\|\nabla\nu\|_{L^p}^{p-2}\,\|\nabla\xi\|_{L^p}^2 \\
&\quad\le (p-1)\left(\frac{3p}{2}\right)^{-\frac{2}{3p}}\left(\frac{3p}{3p-2}\right)^{\frac{2-3p}{3p}}\varepsilon^{\frac{4-6p}{3p}}s^{\frac{2}{3p}}C_9^2C_{10}^{p-2}\left(\|\xi\|_{\mathring{H}^{-1}_{\mathrm{per}}}^2+s\varepsilon^2\,\|\Delta\xi\|^2\right).
\end{aligned}
$$

As a result, estimate (3.32) is valid by choosing

$$
C_6 = 1+(p-1)\left(\frac{3p}{2}\right)^{-\frac{2}{3p}}\left(\frac{3p}{3p-2}\right)^{\frac{2-3p}{3p}}\varepsilon^{\frac{4-6p}{3p}}s^{\frac{2}{3p}}C_9^2C_{10}^{p-2}.
$$

Similarly, For $p\in[2,6),\ d=3$, we have

$$
\begin{aligned}
&\left(\frac{6p}{6-p}\right)^{\frac{6-p}{6p}}\left(\frac{6p}{7p-6}\right)^{\frac{7p-6}{6p}}\varepsilon^{\frac{7p-6}{3p}}s^{\frac{7p-6}{6p}}C_9^{-2}\,\|\nabla\xi\|_{L^p}^2 \\
&\qquad\le \left(\frac{6p}{6-p}\right)^{\frac{6-p}{6p}}\|\xi\|_{\mathring{H}^{-1}_{\mathrm{per}}}^{\frac{6-p}{3p}}\left(\frac{6p}{7p-6}\right)^{\frac{7p-6}{6p}}\varepsilon^{\frac{7p-6}{3p}}s^{\frac{7p-6}{6p}}\,\|\Delta\xi\|^{\frac{7p-6}{3p}} \\
&\qquad\le \|\xi\|_{\mathring{H}^{-1}_{\mathrm{per}}}^2+s\varepsilon^2\,\|\Delta\xi\|^2.
\end{aligned}
$$

As a result, estimate (3.32) is valid by choosing

$$
C_6 = 1+(p-1)\left(\frac{6p}{6-p}\right)^{\frac{p-6}{6p}}\left(\frac{6p}{7p-6}\right)^{\frac{6-7p}{6p}}\varepsilon^{\frac{6-7p}{3p}}s^{\frac{6-p}{6p}}C_9^2C_{10}^{p-2}.
$$

For the case $p=6,\ d=3$, the Sobolev inequality (2.6) degenerates, as before. But it is straightforward to show that estimate (3.32) is valid upon choosing

$$
C_6 = 1+(p-1)\varepsilon^{-2}C_9^2C_{10}^{p-2}.
$$

As before, we omit the simple argument that one may take $C_8=C_5$ to satisfy (L6). The proof is complete. □

Remark 3.2.4. *We note that a mixed formulation of the sixth-order regularized p-Laplacian problem — expressed in strong form in* (1.2a) – (1.2b) *and in weak form*

in (3.27a) – (3.27b) *— in order to preserve the proper variational structure of the problem. Specifically, observe that the p-Laplacian term appearing in* (1.2b) *is the gradient of a convex energy functional. However, if one applies* $-\Delta$ *to* (1.2b)*, so that the variable* w *is dropped, a composition of the p-Laplacian and regular Laplacian operators yields a nonlinear term that could not be represented as the gradient of a convex energy. In short, the variational/convexity structure would be lost and the theoretical convergence could not be justified.*

3.2.3 Convergence for the Discretized Fourth-Order Problem

The discrete version of (1.1) can be expressed as follows: given $f \in \mathcal{C}_{\text{per}}$, find $u \in \mathcal{C}_{\text{per}}$ such that

$$u - s\nabla_h^{\mathsf{v}} \cdot \left(|\nabla_h^{\mathsf{v}} u|^{p-2} \nabla_h^{\mathsf{v}} u\right) + s\varepsilon^2 \Delta_h u = f. \tag{3.34}$$

This represents a second-order (in space) approximation of the solution of (1.1). As in the space continuous case, we formulate an equivalent minimization problem. Using the definitions from subsection 2.2.1, we have the following discrete energy: given $f \in \mathcal{C}_{\text{per}}$, for any $\nu \in \mathcal{C}_{\text{per}}$, define

$$E_h(\nu) := \frac{1}{2}\|\nu - f\|_2^2 + \frac{s}{p}\|\nabla_h^{\mathsf{v}} \nu\|_p^p + \frac{s\varepsilon^2}{2}\|\Delta_h \nu\|_2^2. \tag{3.35}$$

This (discrete) energy satisfies (E1) – (E3). The discrete variational derivative at $\nu \in \mathcal{C}_{\text{per}}$ is

$$\begin{aligned}
\delta E_h[\nu](\xi) &:= d_\tau E_h(\nu + \tau\xi)|_{\tau=0} \\
&= (\nu - f, \xi)_2 + s\langle |\nabla_h^{\mathsf{v}} \nu|^{p-2} \mathfrak{D}_x \nu, \mathfrak{D}_x \xi\rangle + s\langle |\nabla_h^{\mathsf{v}} \nu|^{p-2} \mathfrak{D}_y \nu, \mathfrak{D}_y \xi\rangle + s\varepsilon^2 (\Delta_h \nu, \Delta_h \xi)_2 \\
&= (\nu - f, w)_2 + s\langle |\nabla_h^{\mathsf{v}} \nu|^{p-2} \nabla_h^{\mathsf{v}} \nu, \nabla_h^{\mathsf{v}} \xi\rangle + s\varepsilon^2 (\Delta_h \nu, \Delta_h \xi)_2 \\
&= \left(\nu - f - s\nabla_h^{\mathsf{v}} \cdot \left(|\nabla_h^{\mathsf{v}} \nu|^{p-2} \nabla_h^{\mathsf{v}} \nu\right) + s\varepsilon^2 \Delta_h^2 \nu, \xi\right)_2,
\end{aligned}$$

for all $\xi \in \mathcal{C}_{\text{per}}$, where we have used summation-by-parts [71, 82] to obtain the last equality. Given a fixed $\nu \in \mathcal{C}_{\text{per}}$, the action of the second variation on the arbitrary pair $(\xi, \eta) \in \mathcal{C}_{\text{per}} \times \mathcal{C}_{\text{per}}$ is given by

$$\begin{aligned}\delta^2 E_h[\nu](\xi,\eta) = \; & (\xi,\eta)_2 + s\langle |\nabla_h^{\mathsf{v}}\nu|^{p-2}\nabla_h^{\mathsf{v}}\xi, \nabla_h^{\mathsf{v}}\eta\rangle \\ & + (p-2)s\langle |\nabla_h^{\mathsf{v}}\nu|^{p-4}\nabla_h^{\mathsf{v}}\nu\cdot\nabla_h^{\mathsf{v}}\xi, \nabla_h^{\mathsf{v}}\nu\cdot\nabla_h^{\mathsf{v}}\eta\rangle + s\varepsilon^2\left(\Delta_h\xi, \Delta_h\eta\right)_2.\end{aligned}$$

We have the bound:

$$\begin{aligned}\left|\delta^2 E_h[\nu](\xi,\eta)\right| \leq \; & \|\xi\|_2 \|\eta\|_2 + s\|\nabla_h^{\mathsf{v}}\nu\|_p^{p-2}\|\nabla_h^{\mathsf{v}}\xi\|_p\|\nabla_h^{\mathsf{v}}\eta\|_p \\ & + (p-2)s\|\nabla_h^{\mathsf{v}}\nu\|_p^{p-2}\|\nabla_h^{\mathsf{v}}\xi\|_p\|\nabla_h^{\mathsf{v}}\eta\|_p + s\varepsilon^2\|\Delta_h\xi\|_2\|\Delta_h\eta\|_2. \quad (3.36)\end{aligned}$$

For this problem, we define the pre-conditioner via

$$\begin{aligned}(\nu,\xi)_{\mathcal{L}_h} = \mathcal{L}_h[\nu](\xi) := & (\nu,\xi)_2 + s\left[D_x\nu, D_x\xi\right]_{\text{ew}} + s\left[D_y\nu, D_y\xi\right]_{\text{ns}} + s\varepsilon^2(\Delta_h\nu, \Delta_h\xi)_2 \\ = & (\nu - s\Delta_h\nu + s\varepsilon^2\Delta_h^2\nu, \xi)_2,\end{aligned}$$

for all $\nu, \xi \in \mathcal{C}_{\text{per}}$, where we have used summation-by-parts to establish the second equality. In other words,

$$\mathcal{L}_h[\nu] = \nu - s\Delta_h\nu + s\varepsilon^2\Delta_h^2\nu.$$

One will notice the similarity of the pre-conditioner with the nonlinear operator in (3.34). The induced norm is

$$\|\nu\|_{\mathcal{L}_h}^2 := (\nu,\nu)_{\mathcal{L}_h} = \|\nu\|_2^2 + s\|\nabla_h\nu\|_2^2 + s\varepsilon^2\|\Delta_h\nu\|_2,$$

defined for every $\nu \in \mathcal{C}_{\text{per}}$.

Mimicking the proofs in the continuous case, using summation-by-parts in place of integration-by-parts, and Lemma 2.2.2, we get the following result, whose proof is omitted:

Lemma 3.2.5. *For any* $\nu, \xi \in \mathcal{C}_{\text{per}}$,

$$C_5 \left\|\xi - \nu\right\|_{\mathcal{L}_h}^2 \leq \left(\delta E_h[\xi] - \delta E_h[\nu]\right)(\xi - \nu), \tag{3.37}$$

where $C_5 = \min\left(\frac{1}{2}, \varepsilon s^{-\frac{1}{2}}\right)$. *Let* E_0 *be given, such that* $B := \{\nu \in \mathcal{C}_{\text{per}} \mid E_h[\nu] \leq E_0\}$ *is non-empty. For any* $\nu \in B$ *and any* $\xi \in \mathcal{C}_{\text{per}}$, *we have*

$$\left|\delta^2 E_h[\nu](\xi, \xi)\right| \leq C_6 \left\|\xi\right\|_{\mathcal{L}_h}^2, \tag{3.38}$$

where

$$C_6 = 1 + \frac{1}{p}(p-1)^{\frac{2p-1}{p}} \varepsilon^{\frac{-2(p-1)}{p}} s^{\frac{1}{p}} C_9^2 C_{10}^{p-2}, \tag{3.39}$$

and $C_{10} = (pE_0)^{\frac{1}{p}}$. *We can take* $C_8 = C_5$ *to satisfy estimate* (3.15) *of assumption (L6).*

3.2.4 Convergence for the Discretized Sixth-Order Problem

The (second-order accurate) discrete version of (1.2a) – (1.2b) can be expressed as follows: given $f, g \in \mathcal{C}_{\text{per}}$, find $u, w \in \mathcal{C}_{\text{per}}$ such that

$$\begin{aligned} u - \Delta_h w &= g, \\ s\lambda u - s\nabla_h^{\mathsf{v}} \cdot \left(\left|\nabla_h^{\mathsf{v}} u\right|^{p-2} \nabla_h^{\mathsf{v}} u\right) + s\varepsilon^2 \Delta_h u - w &= f. \end{aligned}$$

As before, it is convenient to switch to the mean-zero version: find $u_\star, w_\star \in \mathring{\mathcal{C}}_{\text{per}}$ such that

$$u_\star - \Delta_h w_\star = g - \overline{g},$$

$$s\lambda u_\star - s\nabla_h^{\mathsf{v}} \cdot \left(|\nabla_h^{\mathsf{v}} u_\star|^{p-2} \nabla_h^{\mathsf{v}} u_\star\right) + s\varepsilon^2 \Delta_h u_\star - w_\star = f - \overline{f}.$$

Similar to fourth-order regularized p-Laplacian problem, we define the following discrete energy: for every $\nu \in \mathring{\mathcal{C}}_{\text{per}}$

$$E_h(\nu) := \frac{1}{2} \|\nu - g + \bar{g}\|_{-1}^2 + \frac{\lambda s}{2} \|\nu + \bar{g}\|_2^2 - (\nu, f) + \frac{s}{p} \|\nabla_h^{\mathsf{v}} \nu\|_p^p + \frac{s\varepsilon^2}{2} \|\Delta_h \nu\|^2 .$$

For the discrete sixth order problem, we define a linear operator $\mathcal{L}_h : \mathring{\mathcal{C}}_{\text{per}} \to \mathring{\mathcal{C}}_{\text{per}}$ via

$$\begin{aligned}(\nu, \xi)_{\mathcal{L}_h} = \mathcal{L}_h[\nu](\xi) &:= s\lambda\, (\nu, \xi)_2 + (\nu, \xi)_{-1} \\ &\quad + s\, [D_x \nu, D_x \xi]_{\text{ew}} + s\, [D_y \nu, D_y \xi]_{\text{ns}} + s\varepsilon^2 (\Delta_h \nu, \Delta_h \xi)_2 \\ &= \left(s\lambda\nu - s\Delta_h \nu + s\varepsilon^2 \Delta_h^2 \nu - \mathsf{T}_h\, [-\nu]\, , \xi\right)_2 ,\end{aligned}$$

where the second equality may be seen using summation-by-parts [71, 82]. This operator satisfies (L1) – (L3), and the next result, which we give without proof for the sake of brevity, shows that (L4) – (L6) are satisfied as well.

Lemma 3.2.6. *For any $u, v \in \mathring{\mathcal{C}}_{\text{per}}$, the following inequality is valid*

$$C_5 \|u - v\|_{\mathcal{L}_h}^2 \leq (\delta E_h[u] - \delta E_h[\nu])\, (u - v),$$

where $C_5 = \min\left(\frac{1}{3}, \varepsilon^{\frac{4}{3}} s^{-\frac{1}{3}}\right)$. Let E_0 be given such that $B := \left\{\xi \in \mathring{\mathcal{C}}_{\text{per}} \;\middle|\; E_h[\xi] \leq E_0\right\}$ is non-empty. For any $\nu \in B$, we have

$$\left|\delta^2 E_h[\nu](\xi, \xi)\right| \leq C_6 \|\xi\|_{\mathcal{L}_h}^2 ,$$

for all $\xi \in \mathring{\mathcal{C}}_{\text{per}}$, where

$$C_6 = 1 + (p-1) \left(\frac{3p}{2}\right)^{-\frac{2}{3p}} \left(\frac{3p}{3p-2}\right)^{\frac{2-3p}{3p}} \varepsilon^{\frac{4-6p}{3p}} s^{\frac{2}{3p}} C_9^2 C_{10}^{p-2},$$

and $C_{10} = (pE_{h,0})^{\frac{1}{p}}$. *We can take* $C_8 = C_5$ *to satisfy estimate* (3.15) *of assumption (L6).*

Chapter 4

Applications of the Linearly Preconditioned Steepest Descent Methods

In this chapter we perform some numerical experiments to support the theoretical results. The finite difference search direction equations and Poisson equations are solved efficiently using the Fast Fourier Transform (FFT). We would like to point out that the Fourier pseudo-spectral method can be used to discretize space, and, once again, one can utilize the FFT for the inversion of the linear systems. For descriptions of the pseudo-spectral methods, see, for example, [9, 17, 48].

4.1 Application to Epitaxial Thin Film Growth Model with First-Order-In-Time Scheme

The content in this chapter has been published in [34], for more details please refer to [34].

4.1.1 Introduction

In this section we recall the convex splitting numerical scheme in [77] for the thin film epitaxy model with slope selection. Suppose that $\Omega \subset \mathbb{R}^2$ is a rectangular domain. The energy of an epitaxial thin film is given by

$$\mathcal{E}[u] = \int_\Omega \left\{ \frac{1}{p} |\nabla u|^p - \frac{1}{2} |\nabla u|^2 + \frac{\varepsilon^2}{2} |\Delta u|^2 \right\} d\mathbf{x}, \quad \forall\ u \in H^2_{\text{per}}(\Omega),$$

where, $p \geq 4$ is even, $u : \Omega \to \mathbb{R}$ is the height of the film, and ε is a constant. The L^2 gradient flow is

$$\partial_t u = -w, \quad w := \delta\mathcal{E} = -\nabla \cdot \left(|\nabla u|^{p-2} \nabla u \right) + \Delta u + \varepsilon^2 \Delta^2 u, \tag{4.1}$$

and w is called the chemical potential. The model predicts the emergence of a faceted thin film, whose facets have slopes of magnitude approximately one, that coarsens over time. The fully-implicit convex splitting scheme in 2D [77] can be written in operator format as $\mathcal{N}_h[u^{n+1}] = f$, where

$$\mathcal{N}_h[\nu] := \nu - s\nabla_h^{\mathsf{v}} \cdot \left(|\nabla_h^{\mathsf{v}} \nu|^{p-2} \nabla_h^{\mathsf{v}} \nu \right) + \varepsilon^2 s \Delta_h^2 \nu, \quad f = u^n - s\Delta_h^{\mathsf{v}} u^n. \tag{4.2}$$

Hence, the scheme can be reformulated as the fourth-order problem (3.34) with $f = u^n - s\Delta_h^{\mathsf{v}} u^n$ and $p \geq 4$ and even.

In way of summary, to solve $\mathcal{N}_h[u] = f$, suppose that iterate $u^k \in \mathcal{C}_{\text{per}}$ is given. (Note that k is the PSD solver iteration index, not the time step index, the latter of which we usually denote by n.) We first compute the search direction $d^k \in \mathcal{C}_{\text{per}}$ via (3.3):

$$\begin{aligned}
\mathcal{L}[d^k] = d^k - s\Delta_h d^k + s\varepsilon^2 \Delta_h^2 d^k &= -\delta E_h[u^k] \\
&= -\left(u^k - f - s\nabla_h^{\mathsf{v}} \cdot \left(\left|\nabla_h^{\mathsf{v}} u^k\right|^{p-2} \nabla_h^{\mathsf{v}} u^k \right) + s\varepsilon^2 \Delta_h^2 u^k \right) \\
&= f - \mathcal{N}_h[u^k],
\end{aligned}$$

where E_h is as defined in (3.35). This equation is efficiently solved using FFT. Once d^k is found, we perform a line-search according to (4.45): find $\alpha_k \in \mathbb{R}$ such that $q(\alpha_k) = 0$, where

$$\begin{aligned} q(\alpha) &:= \delta E_h[u^k + \alpha d^k](d^k) \\ &= \left(u^k + \alpha d^k - f - s\nabla_h^{\mathsf{v}} \cdot \left(\left|\nabla_h^{\mathsf{v}}(u^k + \alpha d^k)\right|^{p-2} \nabla_h^{\mathsf{v}}(u^k + \alpha d^k)\right) + s\varepsilon^2 \Delta_h^2(u^k + \alpha d^k), d^k\right)_2 \\ &= \left(\mathcal{N}_h[u^k + \alpha d^k] - f, d^k\right)_2 . \end{aligned}$$

The approximation sequence is then updated via $u^{k+1} = u^k + \alpha_k d^k$. When $p = 4$ ($p = 6$), a short calculation shows that q is a cubic (quintic) polynomial whose coefficients can be easily obtained. Moreover, the theory predicts that there is a unique global root for q.

4.1.2 Convergence and complexity of the PSD solver

In this subsection we demonstrate the accuracy and efficiency of the PSD solver by using the epitaxial thin film model with slope selection. We present the results of some convergence tests and perform some sample computations to demonstrate the convergence and near optimal complexity with respect to the grid size h.

Table 4.1: Errors, convergence rates, average iteration numbers and average CPU time for each time step. Parameters are given in the text, and the initial data are defined in (4.3). The refinement path is $s = 0.1h^2$.

		$p = 4$				$p = 6$			
h_c	h_f	$\|\delta_u\|_2$	Rate	$\#_{iter}$	$T_{cpu}(h_f)$	$\|\delta_u\|_2$	Rate	$\#_{iter}$	$T_{cpu}(h_f)$
$\frac{3.2}{16}$	$\frac{3.2}{32}$	6.2192×10^{-3}	-	4	0.0007	9.3074×10^{-3}	-	5	0.0009
$\frac{3.2}{32}$	$\frac{3.2}{64}$	1.2685×10^{-3}	2.29	2	0.0024	1.6392×10^{-3}	2.51	3	0.0032
$\frac{3.2}{64}$	$\frac{3.2}{128}$	2.6046×10^{-4}	2.28	2	0.0114	2.9046×10^{-4}	2.50	2	0.0141
$\frac{3.2}{128}$	$\frac{3.2}{256}$	5.9639×10^{-5}	2.13	2	0.0475	6.5325×10^{-5}	2.15	2	0.0616
$\frac{3.2}{256}$	$\frac{3.2}{512}$	1.4526×10^{-5}	2.04	2	0.3560	1.5886×10^{-5}	2.04	2	0.4636

To simultaneously demonstrate the spatial accuracy and the efficiency of the solver, we perform a typical time-space convergence test for the fully discrete scheme (4.2) for the slope selection model. As in [71, 77], we perform the Cauchy-type

convergence test using the following periodic initial data [71]:

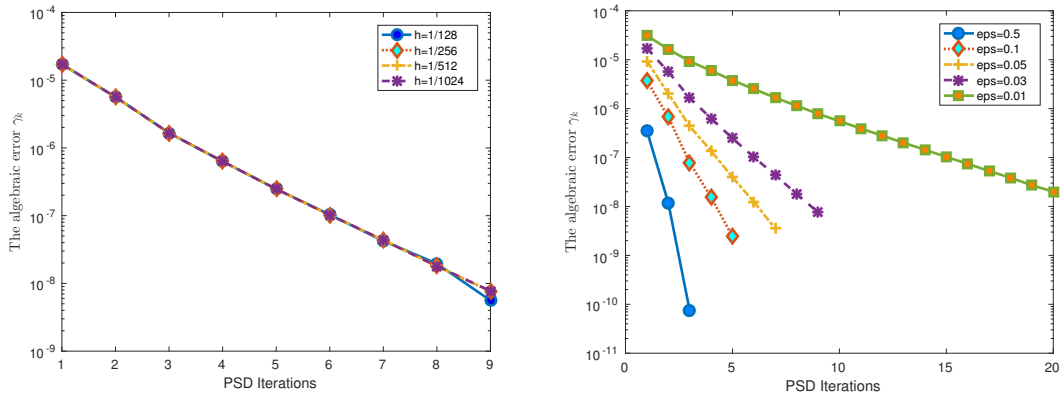
$$\begin{aligned} u(x,y,0) &= 0.1\sin^2\left(\frac{2\pi x}{L}\right)\cdot\sin\left(\frac{4\pi(y-1.4)}{L}\right) \\ &\quad -0.1\cos\left(\frac{2\pi(x-2.0)}{L}\right)\cdot\sin\left(\frac{2\pi y}{L}\right), \end{aligned} \tag{4.3}$$

where $\Omega = (0, 3.2)^2$. In this test, we compute the Cauchy difference, $\delta_u := u_{h_f}(T) - \mathcal{I}_c^f(u_{h_c}(T))$, where $h_c = 2h_f$, and $\mathcal{I}_c^f$ is a bilinear interpolation operator that maps the coarse grid approximation u_{h_c} onto the fine grid. We take a quadratic refinement path, *i.e.*, $s = Ch^2$, to equalize the spatial and temporal error contributions. At the final time, $T = 0.32$, we expect the global error to be $\mathcal{O}(s) + \mathcal{O}(h^2) = \mathcal{O}(h^2)$ in the ℓ^2 and ℓ^∞ norms, as $h, s \to 0$. The other parameters are given by $\varepsilon = 0.1$ and $s = 0.1h^2$. The norms of Cauchy difference, the convergence rates, average iteration number and average CPU time can be found in Table 4.1. Second-order convergence is observed. At the same time, the average iteration count for the solver remains at around 2. Since we are using a quadratic refinement path, increasing the grid size by a factor of two (decreasing the grid spacing by 2) means increasing the number of time-space degrees of freedom by a factor of 16. But the CPU time increases at a much slower rate. The complexity can be offset, of course, by the fact the starting guesses for the solver at each independent time level are better for smaller time step sizes.

To more directly investigate the complexity of the PSD solver we perform another series of tests to determine the dependences of the convergence rates on ε, h, s, and p, in particular. Consider the following spatially periodic function parameterized by s:

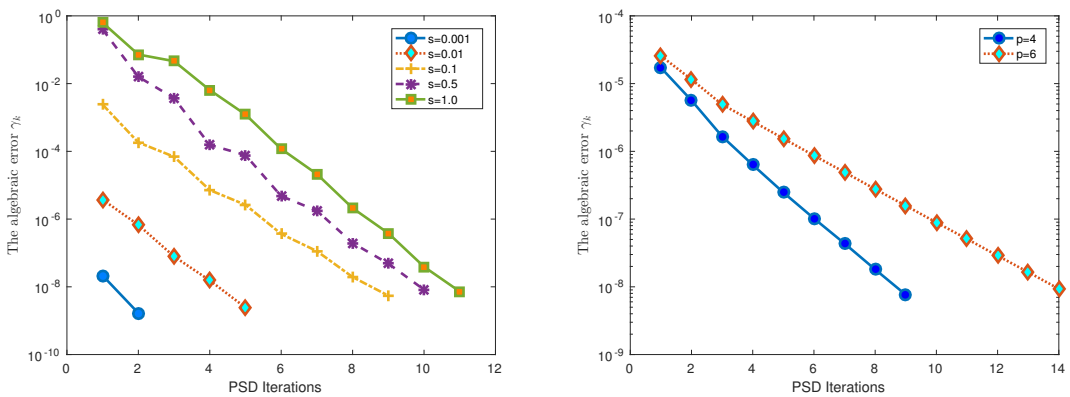
$$\tilde{u}(x,y,s) = \frac{1}{2\pi}\sin\big(2\pi x\big)\cos\big(2\pi y\big)\cos(s). \tag{4.4}$$

First we calculate $f := \mathcal{N}_h\left[\mathcal{I}_h\left(\tilde{u}(\,\cdot\,,\,\cdot\,,s)\right)\right] \in \mathcal{C}_{\text{per}}$, where $\mathcal{I}_h : C^0_{\text{per}}(\Omega) \to \mathcal{C}_{\text{per}}$ is the canonical grid projection operator. Then we compute the sequence $\left\{u^k\right\}_{k=0}^{\infty}$ via the



(a) h-independence: $p = 4$, $s = 0.01$ and $\varepsilon = 0.03$.

(b) ε-dependence: $p = 4$, $s = 0.01$ and $h = 1/512$.



(c) s-dependence: $p = 4$, $h = \frac{1}{512}$ and $\varepsilon = 0.03$.

(d) p-dependence: $h = \frac{1}{512}$, $s = 0.01$ and $\varepsilon = 0.03$.

Figure 4.1: Complexity tests showing the solver performance for changing values of h, ε, s and p. Parameters are given in the text.

PSD algorithm, with the itialization

$$u^0_{i,j} = \tilde{u}(p_i, p_j, 0) + s^2 \sin\big(4\pi p_i\big) \sin\big(6\pi p_j\big).$$

The right-hand-side f is manufactured so that $u = \mathcal{I}_h\left(\tilde{u}(\,\cdot\,,\,\cdot\,,s)\right)$ is the exact algebraic solution to $\mathcal{N}_h[u] = f$, hence $u^k \to \mathcal{I}_h\left(\tilde{u}(\,\cdot\,,\,\cdot\,,s)\right)$, as $k \to \infty$. Define $\gamma_k := \|u^k - \mathcal{I}_h\left(\tilde{u}(\,\cdot\,,\,\cdot\,,s)\right)\|_\infty$. We stop the PSD algorithm when $\gamma_k \le \tau := 1 \times 10^{-8}$.

In Figure 4.12 we plot γ_k versus k, on a semi-log scale, for various choices of h, ε, s and p. In Figure 4.12(a) $p = 4$, $s = 0.01$ and $\varepsilon = 0.03$; in Figure 4.12(b) $p = 4$, $s = 0.01$ and $h = 1/512$; in Figure 4.12(c) $p = 4$, $h = 1/512$ and $\varepsilon = 0.03$; in Figure 4.12(d): $h = 1/512$, $s = 0.01$ and $\varepsilon = 0.03$. As can be seen in Figure 4.12(a), the convergence rate (as gleaned from the error reduction) is nearly uniform and nearly independent of h. Figures 4.12 (b) and (c) indicate that more PSD iterations are required for smaller values of ε and larger values of s, respectively. Figure 4.12(d) shows that the number of PSD iterations increases with the value of p. These general trends are expected form the theory.

4.1.3 Long-time coarsening behavior for the thin film model with $p = 4, 6$

Coarsening processes in thin film systems can take place on very long time scales [57]. In this subsection, we perform (now standard) long time behavior tests for $p = 4, 6$. Such test, which have been performed in many places, will confirm the expected coarsening rates and serve as benchmarks for our solver. See, for example, [71, 77]. The initial data for the simulations are taken as essentially random:

$$u^0_{i,j} = 0.05 \cdot (2r_{i,j} - 1), \tag{4.5}$$

where the $r_{i,j}$ are uniformly distributed random numbers in [0, 1]. Time snapshots of the evolution for the epitaxial thin film growth model with $p = 4$ can be found in Figure 4.2. The coarsening rates for the $p = 4$ case are given in Figure 4.3. These simulation results are consistent with earlier work on this topic in [71, 77, 83], showing the surface roughness, W, grows like $t^{\frac{1}{3}}$ and the energy, E, decays like $t^{-\frac{1}{3}}$. We also present the numerical simulations for the epitaxial thin film growth model with $p = 6$ in Figure 4.4. Notice in Figure 4.4 that the evolution process is significantly different from the process depicted in Figure 4.2.

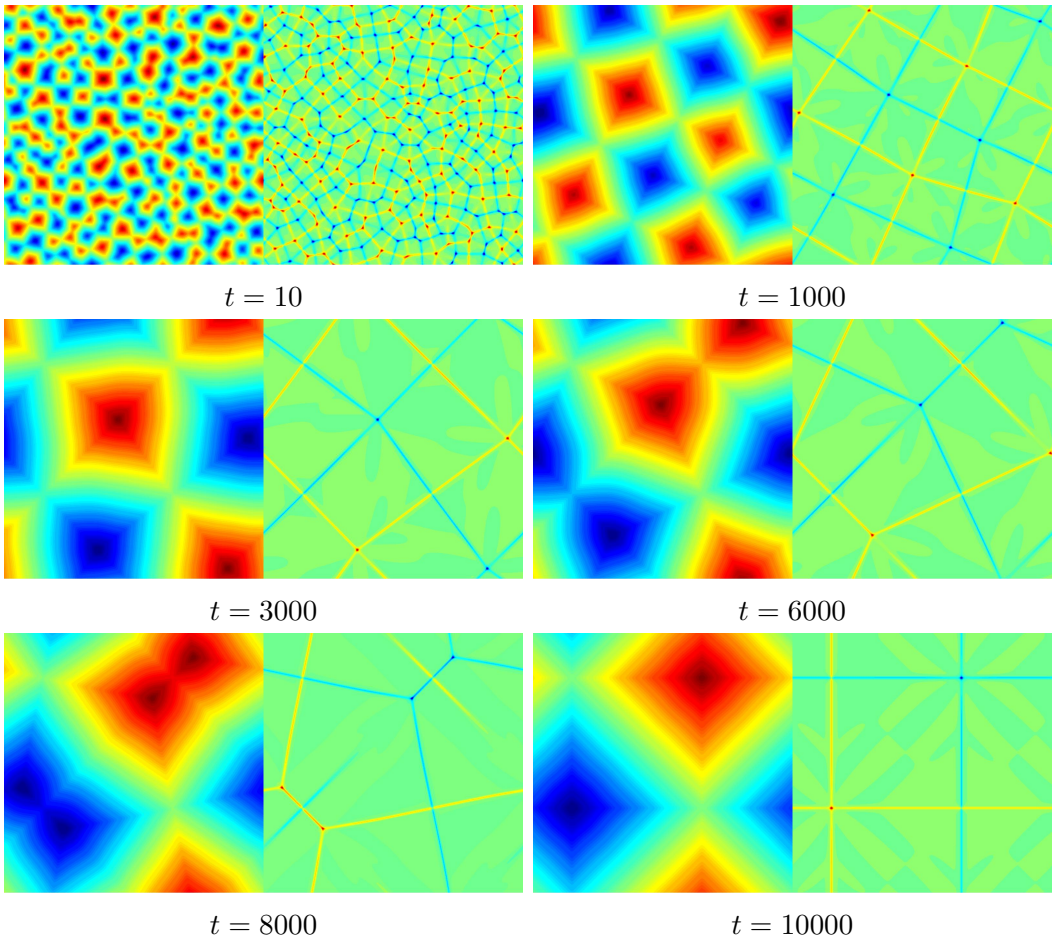
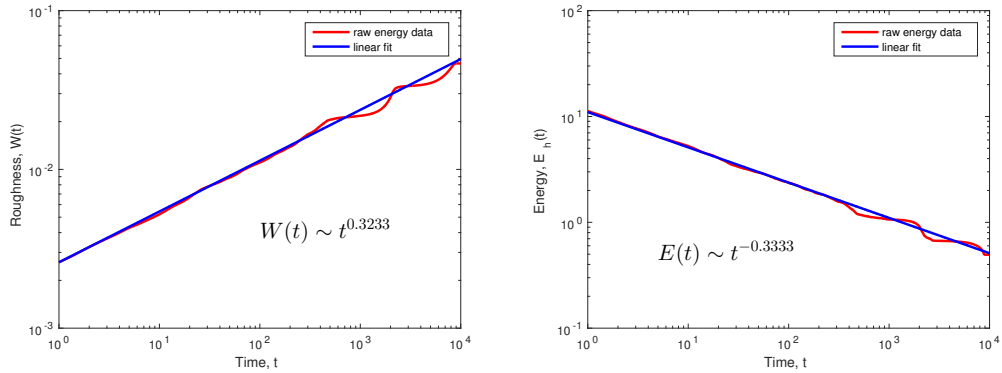


Figure 4.2: Time snapshots of the evolution with PSD solver for the epitaxial thin film growth model with $p = 4$ at $t = 10, 1000, 3000, 6000, 8000$ and 10000. Left: contour plot of u, Right: contour plot of Δu. The parameters are $\varepsilon = 0.03, \Omega = [12.8]^2, s = 0.01$. These simulation results are consistent with earlier work on this topic in [71, 77, 83].



(a) Roughness evolution (b) Energy evolution

Figure 4.3: Log-log plot of Roughness and energy evolution for the simulation depicted in Figure 4.2.

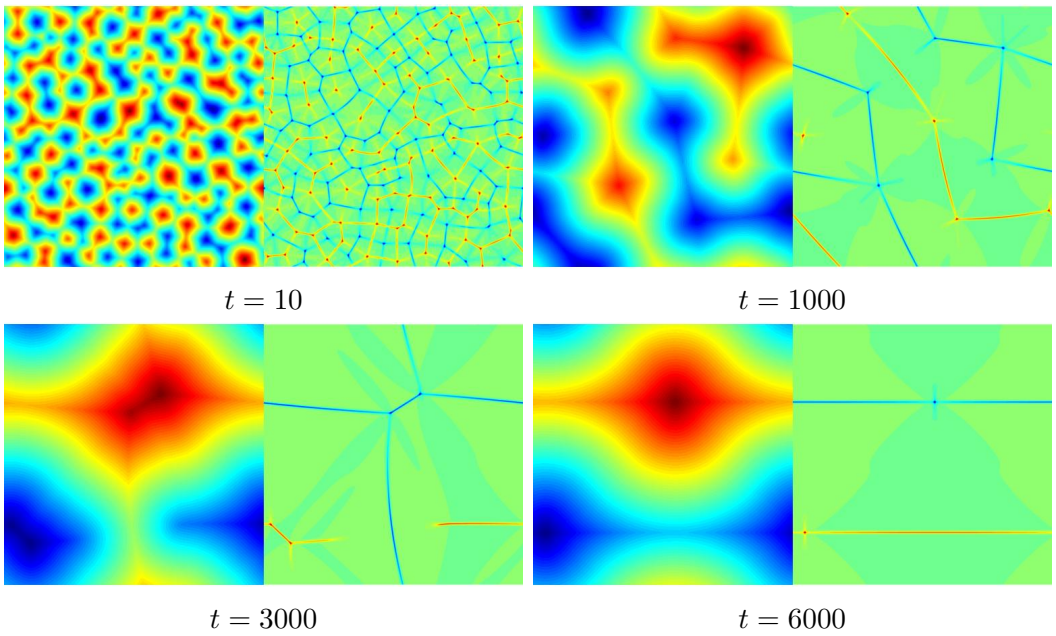


Figure 4.4: Time snapshots of the evolution with PSD solver for the epitaxial thin film growth model with $p = 6$ at $t = 10, 1000, 3000$ and 6000. Left: contour plot of u, Right: contour plot of Δu. The parameters are $\epsilon = 3.0 \times 10^{-2}, \Omega = [12.8]^2, s = 0.01$.

4.2 Application to Epitaxial Thin Film Growth Model with Second-Order-In-Time Backward Differentiation Formula Scheme

The content in this chapter has been published in [36], for more details please refer to [36].

4.2.1 The fully discrete scheme

Let $M \in \mathbb{Z}^+$, and set $s := T/M$, where T is the final time. We define the canonical grid projection operator $\mathsf{P}_h : C^0(\Omega) \to \mathcal{C}_{\text{per}}$ via $[\mathsf{P}_h v]_{i,j} = v(\xi_i, \xi_j)$. Set $u_{h,s} := \mathsf{P}_h u(\cdot, s)$. Then $F_h(u_{h,s}) + \frac{1}{2}\|\nabla_h(u_{h,s} - u_{h,0})\|_2^2 \to F(u(\cdot, 0))$ as $h \to 0$ and $s \to 0$ for sufficiently regular u. We denote ϕ_e as the exact solution to the SS equation (5.9) and take $\Phi_{i,j}^{\ell} = \mathsf{P}_h \phi_e(\cdot, t_\ell)$. In the rest of paper, we shall drop the subscription i, j if no confusion is caused.

With the machinery in last subsection, our second-order-in-time BDF type scheme can be formulated as follows: for $k \geq 1$, given $\phi^{k-1}, \phi^k \in \mathcal{C}_{\text{per}}$, find $\phi^{k+1} \in \mathcal{C}_{\text{per}}$ such that

$$\begin{aligned} \frac{3\phi^{k+1} - 4\phi^k + \phi^{k-1}}{2s} &= \nabla_h^{\mathsf{v}} \cdot (|\nabla_h^{\mathsf{v}} \phi^{k+1}|^2 \nabla_h^{\mathsf{v}} \phi^{k+1}) - \Delta_h^{\mathsf{v}}(2\phi^k - \phi^{k-1}) \\ &\quad - As\Delta_h^2(\phi^{k+1} - \phi^k) - \varepsilon^2 \Delta_h^2 \phi^{k+1}, \end{aligned} \tag{4.6}$$

where $\phi^0 := \Phi^0$, $\phi^1 := \Phi^1$ and A is the constant stability coefficient.

For the SS equation (5.9), we see that the PDE is equivalent if a fixed constant is added or subtracted from the solution. Similar argument could also be applied to the numerical scheme (4.6), since this scheme is mass conservative at a discrete level. For simplicity of presentation, we assume that $\overline{\phi^0} = \overline{\phi^1} = 0$, so that $\overline{\phi^k} = 0$, for any $k \geq 2$.

We now introduce a discrete energy that is consistent with the continuous space energy (5.7) as $h \to 0$. In particular, the discrete energy $F_h : \mathcal{C}_{\text{per}} \to \mathbb{R}$ is defined as:

$$F_h(\phi) = \frac{1}{4} \|\nabla_h^{\mathsf{v}} \phi\|_4^4 - \frac{1}{2} \|\nabla_h^{\mathsf{v}} \phi\|_2^2 + \frac{1}{2} \varepsilon^2 \|\Delta_h \phi\|_2^2 . \tag{4.7}$$

We also denote a modified numerical energy $\tilde{F}_h : \mathcal{C}_{\text{per}} \to \mathbb{R}$ via

$$\tilde{F}_h(\phi, \psi) := F_h(\phi) + \frac{1}{4s} \|\phi - \psi\|_2^2 + \frac{1}{2} \|\nabla_h(\phi - \psi)\|_2^2 . \tag{4.8}$$

Although we can not guarantee that the energy F_h is non-increasing in time, we are able to prove the dissipation of auxiliary energy $\tilde{F}_h$. The unique solvability and the unconditional energy stability of scheme (4.6) is assured by the following theorem.

Theorem 4.2.1. *Suppose that the exact solution ϕ_e is periodic and sufficiently regular, and $\phi^0, \phi^1 \in \mathcal{C}_{\text{per}}$ is obtained via grid projection, as defined above. Given any $(\phi^{k-1}, \phi^k) \in \mathcal{C}_{\text{per}}$, there is a unique solution $\phi^{k+1} \in \mathcal{C}_{\text{per}}$ to the scheme (4.6). And also, the scheme (4.6), with starting values ϕ^0 and ϕ^1, is unconditionally energy stable,* i.e., *for any $\tau > 0$ and $h > 0$, and any positive integer $2 \le k \le M - 1$, The numerical scheme (4.6) has the following energy-decay property:*

$$\tilde{F}_h(\phi^{k+1}, \phi^k) \le \tilde{F}_h(\phi^k, \phi^{k-1}) \le \tilde{F}_h(\phi^1, \phi^0) \le C_0, \tag{4.9}$$

for all $A \ge \frac{1}{16}$, where $C_0 >$ is a constant independent of s, h and T.

Proof. The unique solvability follows from the convexity argument. Taking an inner product with (4.6) by $\phi^{k+1} - \phi^k$ yields

$$\begin{aligned} 0 \;=\; & \left(\frac{3\phi^{k+1} - 4\phi^k + \phi^{k-1}}{2s}, \phi^{k+1} - \phi^k \right) \\ & - \left(\nabla_h^{\mathsf{v}} \cdot (|\nabla_h^{\mathsf{v}} \phi^{k+1}|^2 \nabla_h^{\mathsf{v}} \phi^{k+1}), \phi^{k+1} - \phi^k \right) + \left(\Delta_h^{\mathsf{v}} (2\phi^k - \phi^{k-1}), \phi^{k+1} - \phi^k \right) \\ & + As \left(\Delta_h^2 (\phi^{k+1} - \phi^k), \phi^{k+1} - \phi^k \right) + \varepsilon^2 \left(\Delta_h^2 \phi^{k+1}, \phi^{k+1} - \phi^k \right) \end{aligned}$$

$$:= \quad I_1 + I_2 + I_3 + I_4 + I_5. \tag{4.10}$$

We now establish the estimates for $I_1, \cdots, I_5$. The temporal difference term could be evaluated as follows

$$\left(\frac{3\phi^{k+1} - 4\phi^k + \phi^{k-1}}{2s}, \phi^{k+1} - \phi^k\right) \geq \frac{1}{s}\left(\frac{5}{4}\left\|\phi^{k+1} - \phi^k\right\|_2^2 - \frac{1}{4}\left\|\phi^k - \phi^{k-1}\right\|_2^2\right). \tag{4.11}$$

For the 4-Laplacian term, we have

$$\begin{aligned}\left(-\nabla_h^{\mathsf{v}} \cdot (|\nabla_h^{\mathsf{v}}\phi^{k+1}|^2 \nabla_h^{\mathsf{v}}\phi^{k+1}), \phi^{k+1} - \phi^k\right) &= \left(|\nabla_h^{\mathsf{v}}\phi^{k+1}|^2 \nabla_h^{\mathsf{v}}\phi^{k+1}, \nabla_h^{\mathsf{v}}(\phi^{k+1} - \phi^k)\right) \\ &\geq \frac{1}{4}\left(\|\nabla_h^{\mathsf{v}}\phi^{k+1}\|_4^4 - \|\nabla_h^{\mathsf{v}}\phi^k\|_4^4\right). \end{aligned} \tag{4.12}$$

For the concave diffusive term, the following estimate is valid

$$\begin{aligned}&\left(\Delta_h^{\mathsf{v}}(2\phi^k - \phi^{k-1}), \phi^{k+1} - \phi^k\right) = -\left(\nabla_h^{\mathsf{v}}(2\phi^k - \phi^{k-1}), \nabla_h^{\mathsf{v}}(\phi^{k+1} - \phi^k)\right) \\ =\ & -\left(\nabla_h^{\mathsf{v}}\phi^k, \nabla_h^{\mathsf{v}}(\phi^{k+1} - \phi^k)\right) - \left(\nabla_h^{\mathsf{v}}(\phi^k - \phi^{k-1}), \nabla_h^{\mathsf{v}}(\phi^{k+1} - \phi^k)\right) \\ =\ & -\frac{1}{2}\|\nabla_h^{\mathsf{v}}\phi^{k+1}\|_2^2 + \frac{1}{2}\|\nabla_h^{\mathsf{v}}\phi^k\|_2^2 + \frac{1}{2}\|\nabla_h^{\mathsf{v}}(\phi^{k+1} - \phi^k)\|_2^2 - \left(\nabla_h^{\mathsf{v}}(\phi^k - \phi^{k-1}), \nabla_h^{\mathsf{v}}(\phi^{k+1} - \phi^k)\right) \\ \geq\ & -\frac{1}{2}\left(\|\nabla_h^{\mathsf{v}}\phi^{k+1}\|_2^2 - \|\nabla_h^{\mathsf{v}}\phi^k\|_2^2\right) - \frac{1}{2}\|\nabla_h^{\mathsf{v}}(\phi^k - \phi^{k-1})\|_2^2 \\ \geq\ & -\frac{1}{2}\left(\|\nabla_h^{\mathsf{v}}\phi^{k+1}\|_2^2 - \|\nabla_h^{\mathsf{v}}\phi^k\|_2^2\right) - \frac{1}{2}\|\nabla_h(\phi^k - \phi^{k-1})\|_2^2, \end{aligned} \tag{4.13}$$

where the last step applied the Lemma. 2.2.1.

For the surface diffusion term, we have

$$\left(\Delta_h^2\phi^{k+1}, \phi^{k+1} - \phi^k\right) = \left(\Delta_h\phi^{k+1}, \Delta_h(\phi^{k+1} - \phi^k)\right) \geq \frac{1}{2}\left(\|\Delta_h\phi^{k+1}\|_2^2 - \|\Delta_h\phi^k\|_2^2\right). \tag{4.14}$$

Similarly, the following identity is valid for the stabilizing term:

$$s\left(\Delta_h^2(\phi^{k+1} - \phi^k), \phi^{k+1} - \phi^k\right) = s\|\Delta_h(\phi^{k+1} - \phi^k)\|_2^2. \tag{4.15}$$

Meanwhile, an application of Cauchy inequality indicates the following estimate:

$$\frac{1}{s}\|\phi^{k+1}-\phi^k\|_2^2 + As\|\Delta_h(\phi^{k+1}-\phi^k)\|_2^2 \geq 2A^{1/2}\|\nabla_h(\phi^{k+1}-\phi^k)\|_2^2. \tag{4.16}$$

Therefore, a combination of (4.11)-(4.13) and (4.16) yields

$$\begin{aligned}
&F_h(\phi^{k+1}) - F_h(\phi^k) + \frac{1}{4s}\left(\|\phi^{k+1}-\phi^k\|_2^2 - \|\phi^k-\phi^{k-1}\|_2^2\right)\\
&+\frac{1}{2}\left(\|\nabla_h(\phi^{k+1}-\phi^k)\|_2^2 - \|\nabla_h(\phi^k-\phi^{k-1})\|_2^2\right)\\
&\leq (-2A^{1/2}+\frac{1}{2})\|\nabla_h(\phi^{k+1}-\phi^k)\|_2^2 \leq 0,
\end{aligned} \tag{4.17}$$

provided that $A \geq \frac{1}{16}$. Then the proof follows from the definition of the $\tilde{F}_h$ in (4.8). □

4.2.2 $L_h^\infty(0,T;H_h^2)$ Stability of the Numerical Scheme

The $L_h^\infty(0,T;H_h^2)$ bound of the numerical solution could be derived based on the modified energy stability (4.9).

Theorem 4.2.2. *Let $\phi \in \mathcal{C}_\Omega$, then the $L_h^\infty(0,T;H_h^2)$ bound of the numerical solution is as follows:*

$$\|\phi\|_{H_h^2} \leq \sqrt{2\frac{C_0+|\Omega|}{C_1\varepsilon^2}} := C_2, \tag{4.18}$$

where C_2 is independent of s, h and T.

Proof. Since

$$\frac{1}{8}\psi^4 - \frac{1}{2}\psi^2 \geq -\frac{1}{2}, \tag{4.19}$$

then we have

$$\frac{1}{8}\|\nabla_h^{\mathsf{v}}\phi\|_4^4 - \frac{1}{2}\|\nabla_h^{\mathsf{v}}\phi\|_2^2 \geq -\frac{1}{2}|\Omega|, \tag{4.20}$$

with the discrete H_h^1 norm introduced in (2.7). Then we arrive at the following bound, for any $\phi \in \mathcal{C}_\Omega$:

$$\begin{aligned} F_h(\phi) &\geq \frac{1}{8}\|\nabla_h^{\mathsf{v}}\phi\|_4^4 + \frac{\varepsilon^2}{2}\|\Delta_h\phi\|_2^2 - \frac{1}{2}|\Omega| \\ &\geq \frac{1}{2}\|\nabla_h^{\mathsf{v}}\phi\|_2^2 + \frac{\varepsilon^2}{2}\|\Delta_h\phi\|_2^2 - |\Omega| \\ &\geq \frac{\varepsilon^2}{2}\|\Delta_h\phi\|_2^2 - |\Omega| \\ &\geq \frac{1}{2}C_1\varepsilon^2\|\phi\|_{H_h^2}^2 - |\Omega|, \end{aligned} \tag{4.21}$$

in which C_1 is a constant associated with the discrete elliptic regularity: $\|\Delta_h\phi\|_2^2 \geq C_1\|\phi\|_{H_h^2}^2$, as stated in (2.15) of Proposition 2.2.7. Consequently, its combination with (4.8) finishes the proof. □

Remark 4.2.3. *Note that the constant C_2 is independent of s, h and T, but does depends on ε. In particular, $C_2 = O(\varepsilon^{-1})$.*

4.2.3 Convergence Analysis and Error Estimate

Error equations and consistency analysis

A detailed Taylor expansion implies the following truncation error:

$$\begin{aligned} \frac{3\Phi^{k+1} - 4\Phi^k + \Phi^{k-1}}{2s} &= \nabla_h^{\mathsf{v}} \cdot (|\nabla_h^{\mathsf{v}}\Phi^{k+1}|^2\nabla_h^{\mathsf{v}}\Phi^{k+1}) - \Delta_h^{\mathsf{v}}(2\Phi^k - \Phi^{k-1}) \\ &\quad - As\Delta_h^2(\Phi^{k+1} - \Phi^k) - \varepsilon^2\Delta_h^2\Phi^{k+1} + \tau^k, \end{aligned} \tag{4.22}$$

with $\left\|\tau^k\right\|_2 \leq C(h^2 + s^2)$. Consequently, with an introduction of the error function

$$e^k = \Phi^k - \phi^k, \quad \forall\, k \geq 0, \tag{4.23}$$

we get the following evolutionary equation, by subtracting (4.6) from (4.22):

$$\frac{3e^{k+1} - 4e^k + e^{k-1}}{2s} = \nabla_h^{\mathsf{v}} \cdot (|\nabla_h^{\mathsf{v}}\Phi^{k+1}|^2\nabla_h^{\mathsf{v}}\Phi^{k+1} - |\nabla_h^{\mathsf{v}}\phi^{k+1}|^2\nabla_h^{\mathsf{v}}\phi^{k+1})$$

$$
\begin{aligned}
&-\Delta_h^{\mathsf{v}}(2e^k - e^{k-1}) - As\Delta_h^2(e^{k+1} - e^k) \\
&-\varepsilon^2\Delta_h^2 e^{k+1} + \tau^k,
\end{aligned} \tag{4.24}
$$

In addition, from the PDE analysis for the SS equation in [60, 61] and the global in time H_h^2 stability (4.18) for the numerical solution, we also get the L_h^∞, $W^{1,6}$ and H_h^2 bounds for both the exact solution and numerical solution, uniform in time:

$$
\|\Phi^k\|_\infty,\ \|\Phi^k\|_{W^{1,6}},\ \|\Phi^k\|_{H_h^2} \le C_3, \quad \|\phi^k\|_\infty,\ \|\phi^k\|_{W^{1,6}},\ \|\phi^k\|_{H_h^2} \le C_3, \quad \forall\, k \ge 0, \tag{4.25}
$$

where the 3-D embeddings of H_h^2 into L_h^∞ and into $W^{1,6}$ have been applied, as well as the discrete Sobolev embedding inequalities (2.16), (2.17) in Proposition 2.2.7

Stability and convergence analysis

The convergence result is stated in the following theorem.

Theorem 4.2.4. *Let $\Phi \in \mathcal{R}$ be the projection of the exact periodic solution of the SS equation (5.9) with the initial data $\phi^0 := \Phi^0 \in H^2_{\text{per}}(\Omega)$, $\phi^1 := \Phi^1 \in H^2_{\text{per}}(\Omega)$, and the regularity class*

$$
\mathcal{R} = H^3(0,T;C^0(\Omega)) \cap H^2(0,T;C^2(\Omega)) \cap H^1(0,T;C^4(\Omega)) \cap L^\infty(0,T;C^6(\Omega)). \tag{4.26}
$$

Suppose ϕ is the fully-discrete solution of (4.6). Then the following convergence result holds as s, h goes to zero:

$$
\|e^k\|_2 + \left(\frac{3}{16}\varepsilon^2 s \sum_{\ell=0}^{k} \|\Delta_h e^\ell\|^2\right)^{1/2} \le C(s^2 + h^2), \tag{4.27}
$$

where the constant $C > 0$ is independent of s and h.

Proof. Taking an inner product with the numerical error equation (4.24) by e^{k+1} gives

$$
0 \;=\; \left(\frac{3e^{k+1} - 4e^k + e^{k-1}}{2s}, e^{k+1}\right) + \left(|\nabla_h^{\mathsf{v}}\Phi^{k+1}|^2\nabla_h^{\mathsf{v}}\Phi^{k+1} - |\nabla_h^{\mathsf{v}}\phi^{k+1}|^2\nabla_h^{\mathsf{v}}\phi^{k+1}, \nabla_h^{\mathsf{v}} e^{k+1}\right)
$$

$$
\begin{aligned}
& -\left(\nabla_h^{\mathsf{v}}(2e^k - e^{k-1}), \nabla_h^{\mathsf{v}} e^{k+1}\right) + As\left(\Delta_h(e^{k+1} - e^k), \Delta_h e^{k+1}\right) \\
& +\varepsilon^2\left(\Delta_h e^{k+1}, \Delta_h e^{k+1}\right) - \left(\tau^k, e^{k+1}\right) \\
=: \quad & J_1 + J_2 + J_3 + J_4 + J_5 + J_6. \qquad (4.28)
\end{aligned}
$$

For the time difference error term J_1,

$$
\begin{aligned}
\left(\frac{3e^{k+1} - 4e^k + e^{k-1}}{2s}, e^{k+1}\right) &= \frac{3}{4s}\|e^{k+1}\|_2^2 - \frac{1}{s}\|e^k\|_2^2 + \frac{1}{4s}\|e^{k-1}\|_2^2 \\
&\quad + \frac{1}{s}\|e^{k+1} - e^k\|_2^2 - \frac{1}{4s}\|e^{k+1} - e^{k-1}\|_2^2. \qquad (4.29)
\end{aligned}
$$

For the backwards diffusive error term J_3, we have

$$
\begin{aligned}
& -\left(\nabla_h^{\mathsf{v}}(2e^k - e^{k-1}), \nabla_h^{\mathsf{v}} e^{k+1}\right) \\
= \quad & -\frac{1}{2}\|\nabla_h^{\mathsf{v}} e^{k+1}\|_2^2 - \|\nabla_h^{\mathsf{v}} e^k\|_2^2 + \frac{1}{2}\|\nabla_h^{\mathsf{v}} e^{k-1}\|_2^2 \\
& +\|\nabla_h^{\mathsf{v}}(e^{k+1} - e^k)\|_2^2 - \frac{1}{2}\|\nabla_h^{\mathsf{v}}(e^{k+1} - e^{k-1})\|_2^2. \qquad (4.30)
\end{aligned}
$$

And for the stabilizing term J_4,

$$
\begin{aligned}
& As\left(\Delta_h(e^{k+1} - e^k), \Delta_h e^{k+1}\right) \\
= \quad & \frac{As}{2}\left(\|\Delta_h e^{k+1}\|_2^2 - \|\Delta_h e^k\|_2^2 + \|\Delta_h(e^{k+1} - e^k)\|_2^2\right). \qquad (4.31)
\end{aligned}
$$

For the surface diffusion error term J_5 and the local truncation error term J_6, we have

$$
\varepsilon^2\left(\Delta_h e^{k+1}, \Delta_h e^{k+1}\right) = \varepsilon^2\|\Delta_h e^{k+1}\|_2^2, \qquad (4.32)
$$

and

$$
-\left(\tau^k, e^{k+1}\right) \le \|\tau^k\|_2 \cdot \|e^{k+1}\|_2 \le \frac{1}{2}\|\tau^k\|_2^2 + \frac{1}{2}\|e^{k+1}\|_2^2. \qquad (4.33)
$$

For the nonlinear error term J_2, we adopt the same trick in [31], and get

$$\begin{aligned} J_2 &= \left(|\nabla_h^{\mathsf{v}}\Phi^{k+1}|^2\nabla_h^{\mathsf{v}}\Phi^{k+1} - |\nabla_h^{\mathsf{v}}\phi^{k+1}|^2\nabla_h^{\mathsf{v}}\phi^{k+1}, \nabla_h^{\mathsf{v}}e^{k+1}\right) \\ &= \left(\nabla_h^{\mathsf{v}}(\Phi^{k+1}+\phi^{k+1})\cdot\nabla_h^{\mathsf{v}}e^{k+1}\nabla_h^{\mathsf{v}}\Phi^{k+1}, \nabla_h^{\mathsf{v}}e^{k+1}\right) + \left(|\nabla_h^{\mathsf{v}}\phi^{k+1}|^2\nabla_h^{\mathsf{v}}e^{k+1}, \nabla_h^{\mathsf{v}}e^{k+1}\right) \\ &=: J_{2,1} + J_{2,2}. \end{aligned} \tag{4.34}$$

For the first part $J_{2,1}$ of (4.34), we have

$$\begin{aligned} -J_{2,1} &\le C_4\left(\|\nabla_h^{\mathsf{v}}\Phi^{k+1}\|_6 + \|\nabla_h^{\mathsf{v}}\phi^{k+1}\|_6\right)\cdot\|\nabla_h^{\mathsf{v}}\Phi^{k+1}\|_6\cdot\|\nabla_h^{\mathsf{v}}e^{k+1}\|_6\cdot\|\nabla_h^{\mathsf{v}}e^{k+1}\|_2 \\ &\le C_5C_3^2\|\nabla_h^{\mathsf{v}}e^{k+1}\|_6\cdot\|\nabla_h^{\mathsf{v}}e^{k+1}\|_2 \\ &\le C_5C_3^2\|\nabla_h e^{k+1}\|_6\cdot\|\nabla_h e^{k+1}\|_2 \\ &\le C_6\|\Delta_h e^{k+1}\|_2\cdot\|e^{k+1}\|_2^{\frac{1}{2}}\|\Delta_h e^{k+1}\|_2^{\frac{1}{2}} \\ &\le C_7\|e^{k+1}\|_2^{\frac{1}{2}}\cdot\|\Delta_h e^{k+1}\|_2^{\frac{3}{2}} \\ &\le C_8\|e^{k+1}\|_2^2 + \frac{3}{4}\varepsilon^2\|\Delta_h e^{k+1}\|_2^2, \end{aligned} \tag{4.35}$$

in which the $W^{1,6}$ bound (4.25) for the exact and numerical solutions was recalled in the second step, the Sobolev embedding from H_h^2 into $W^{1,6}$ and the estimate (4.25) were used in the last step. The estimate for the second part $J_{2,2}$ of (4.34) is trivial:

$$J_{2,2} \ge 0. \tag{4.36}$$

Then we arrive at

$$-J_2 \le C_9\|e^{k+1}\|_2^2 + \frac{3}{4}\varepsilon^2\|\Delta_h e^{k+1}\|_2^2. \tag{4.37}$$

Finally, a combination of (4.29), (4.30), (4.31), (4.32), (4.33) and (4.37) yields that

$$\begin{aligned} &\frac{3}{4s}\left(\|e^{k+1}\|_2^2 - \|e^k\|_2^2\right) - \frac{1}{4s}\left(\|e^k\|_2^2 - \|e^{k-1}\|_2^2\right) + \frac{1}{2s}\|e^{k+1}-e^k\|_2^2 \\ &-\frac{1}{2s}\|e^k - e^{k-1}\|_2^2 + \frac{As}{2}\left(\|\Delta_h e^{k+1}\|_2^2 - \|\Delta_h e^k\|_2^2\right) + \varepsilon^2\|\Delta_h e^{k+1}\|_2^2 \end{aligned}$$

$$
\begin{aligned}
\leq\ & \frac{1}{2}\|\tau^k\|_2^2 + \frac{1}{2}\|e^{k+1}\|_2^2 + C_9\|e^{k+1}\|_2^2 + \frac{3}{4}\varepsilon^2\|\Delta_h e^{k+1}\|_2^2 \\
& -\|\nabla_h^{\mathsf{v}} e^{k+1}\|_2^2 - 2\|\nabla_h^{\mathsf{v}} e^k\|_2^2 - \|\nabla_h^{\mathsf{v}}(e^{k+1} - e^{k-1})\|_2^2 \\
& +4\varepsilon^{-2}\|e^{k+1}\|_2^2 + 288\varepsilon^{-2}\|e^k\|_2^2 + 72\varepsilon^{-2}\|e^{k-1}\|_2^2 \\
& +\frac{1}{16}\varepsilon^2\left(\|\Delta_h e^{k+1}\|_2^2 + \|\Delta_h e^k\|_2^2 + \|\Delta_h e^{k-1}\|_2^2\right).
\end{aligned} \tag{4.38}
$$

A summation in time implies that

$$
\begin{aligned}
& \frac{3}{4s}\left(\|e^{k+1}\|_2^2 - \|e^1\|_2^2\right) - \frac{1}{4s}\left(\|e^k\|_2^2 - \|e^0\|_2^2\right) + \frac{1}{2s}\|e^{k+1} - e^k\|_2^2 \\
& -\frac{1}{2s}\|e^1 - e^0\|_2^2 + \frac{As}{2}\left(\|\Delta_h e^{k+1}\|_2^2 - \|\Delta_h e^0\|_2^2\right) + \frac{3}{16}\varepsilon^2\sum_{\ell=1}^{k}\|\Delta_h e^{\ell+1}\|_2^2 \\
\leq\ & \frac{1}{2}\sum_{\ell=1}^{n}\|\tau^\ell\|_2^2 + \sum_{\ell=1}^{k}\left(\frac{1}{2} + C_9 + 4\varepsilon^{-2}\right)\|e^{\ell+1}\|_2^2 \\
& +72\varepsilon^{-2}\sum_{\ell=1}^{k}\left(4\|e^\ell\|_2^2 + \|e^{\ell-1}\|_2^2\right) + \frac{1}{16}\varepsilon^2\sum_{\ell=1}^{k}\left(\|\Delta_h e^\ell\|_2^2 + \|\Delta_h e^{\ell-1}\|_2^2\right).
\end{aligned} \tag{4.39}
$$

In turn, an application of discrete Gronwall inequality yields the desired convergence result (4.2.4). This completes the proof of Theorem 4.27. □

4.2.4 Precondition Steepest Descent (PSD) Solver

In this section we describe a preconditioned steepest descent (PSD) algorithm following the practical and theoretical framework in [34]. The fully discrete scheme (4.6) can be recast as a minimization problem: For any $\phi \in \mathcal{C}_{\text{per}}$, the following energy functional is introduced:

$$
E_h[\phi] = \frac{3}{s}\|\phi\|_2^2 + \frac{1}{4}\|\nabla_h^{\mathsf{v}}\phi\|_4^4 + \frac{1}{2}(As + \varepsilon^2)\|\Delta_h\phi\|_2^2. \tag{4.40}
$$

One observes that the fully discrete scheme (4.6) is the discrete variation of the strictly convex energy (4.40) set equal to zero. The nonlinear scheme at a fixed time level

may be expressed as

$$\mathcal{N}_h[\phi] = f, \tag{4.41}$$

with

$$\mathcal{N}_h[\phi] = \frac{3}{2}\phi^{k+1} - s\nabla_h^{\mathsf{v}} \cdot (|\nabla_h^{\mathsf{v}}\phi^{k+1}|^2 \nabla_h^{\mathsf{v}}\phi^{k+1}) + (As^2 + s\varepsilon^2)\Delta_h^2\phi^{k+1}, \tag{4.42}$$

and

$$f = \frac{1}{2}(4\phi^k - \phi^{k-1}) - s\Delta_h^{\mathsf{v}}(2\phi^k - \phi^{k-1}) + As^2\Delta_h^2\phi^k. \tag{4.43}$$

The main idea of the PSD solver is to use a linearized version of the nonlinear operator as a pre-conditioner, or in other words, as a metric for choosing the search direction. A linearized version of the nonlinear operator $\mathcal{N}$, denoted as $\mathcal{L}_h : \mathring{\mathcal{C}}_{\text{per}} \to \mathring{\mathcal{C}}_{\text{per}}$, is defined as follows:

$$\mathcal{L}_h[\psi] := \frac{3}{2}\psi - s\Delta_h\psi + (As^2 + s\varepsilon^2)\Delta_h^2\psi.$$

Clearly, this is a positive, symmetric operator, and we use this as a pre-conditioner for the method. Specifically, this "metric" is used to find an appropriate search direction for the steepest descent solver [34]. Given the current iterate $\phi^n \in \mathcal{C}_{\text{per}}$, we define the following *search direction* problem: find $d^n \in \mathring{\mathcal{C}}_{\text{per}}$ such that

$$\mathcal{L}_h[d^n] = f - \mathcal{N}_h[\phi^n] := r^n,$$

where r^n is the nonlinear residual of the nth iterate ϕ^n. This last equation can be solved efficiently using the Fast Fourier Transform (FFT).

We then obtain the next iterate as

$$\phi^{n+1} = \phi^n + \overline{\alpha} d^n, \tag{4.44}$$

where $\overline{\alpha} \in \mathbb{R}$ is the unique solution to the steepest descent line minimization problem

$$\overline{\alpha} := \underset{\alpha \in \mathbb{R}}{\operatorname{argmax}}\, E_h[\phi^n + \alpha d^n] = \underset{\alpha \in \mathbb{R}}{\operatorname{argzero}}\, \delta E_h[\phi^n + \alpha d^n](d^n). \tag{4.45}$$

The theoretical analysis in [34] suggests that the iteration sequence ϕ^n converges geometrically to ϕ^{k+1}, with ϕ^{k+1} the exact numerical solution of scheme (4.6) at time level $k+1$, *i.e.*, $\mathcal{N}_h[\phi^{k+1}] = f$. And also, this analysis implies a convergence rate independent of h.

Remark 4.2.5. *The Crank-Nicolson version of the second order energy stable scheme for the SS equation (5.9), proposed and analyzed in [71], takes the following (spatially-continuous) form:*

$$\begin{aligned} \frac{\phi^{k+1} - \phi^k}{s} =& \chi(\nabla\phi^{k+1}, \nabla\phi^k) - \Delta\left(\frac{3}{2}\phi^k - \frac{1}{2}\phi^{k-1}\right) - \frac{\varepsilon^2}{2}\Delta^2\left(\phi^{k+1} + \phi^k\right), \\ \chi(\nabla\phi^{k+1}, \nabla\phi^k) :=& \frac{1}{4}\nabla\cdot\left((|\nabla\phi^{k+1}|^2 + |\nabla\phi^k|^2)\nabla(\phi^{k+1} + \phi^k)\right). \end{aligned} \tag{4.46}$$

In this numerical approach, every terms in the chemical potential are evaluated at time instant $t^{k+1/2}$.

Both the CN version (4.46) and the BDF one (4.6) require a nonlinear solver, while the nonlinear term in (4.46) takes a more complicated form than (4.6), which comes from different time instant approximations. As a result, a stronger convexity of the nonlinear term in the BDF one (4.6) is expected to greatly improve the numerical efficiency in the nonlinear iteration.

Such a numerical comparison has been undertaken for the Cahn-Hilliard (CH) model in recent works: the CN and BDF versions of second order accurate, energy stable numerical schemes for the CH equation, proposed in [45], [84], respectively, were tested using the same numerical set-up. The numerical experiments have indicated that, since the nonlinear term in the BDF approach has a stronger convexity than the one in the CN one, a 20 to 25 percent improvement of the computational efficiency is generally available for the CH model.

For the numerical comparison between the BDF and CN approaches for the SS equation (5.9)*, namely* (4.6)*,* (4.46)*, respectively. Such an efficiency improvement is expected to be much greater. This expectation comes from a subtle fact that, the modified CN approximation to the 4-Laplacian term,* $\chi(\nabla\phi^{k+1}, \nabla\phi^k)$*, does not correspond to a convex energy functional, because of the vector gradient form (other than a scalar form) in the 4-Laplacian expansion. As a consequence, the PSD algorithm proposed in this section could hardly be efficiently applied to solve for* (4.46)*, while the PSD application to the BDF approach* (4.6) *has led to a great success. In fact, an application of the Polak-Ribiére variant of NCG method [66] to solve for* (4.46)*, as reported in [71], has shown a fairly poor numerical performance.*

4.2.5 Numerical Experiments

Convergence test and the complexity of the PSD solver

In this subsection we demonstrate the accuracy and complexity of the PSD solver. We present the results of the convergence test and perform some sample computations to investigate the effect of the time step s and stabilized parameter A for the energy $F_h(\phi)$.

To simultaneously demonstrate the spatial accuracy and the efficiency of the solver, we perform a typical time-space convergence test for the fully discrete scheme (4.6) for the slope selection model. As in [71, 77], we perform the Cauchy-type convergence test using the following periodic initial data [71]:

$$\begin{aligned} u(x,y,0) &= 0.1\sin^2\left(\frac{2\pi x}{L}\right)\cdot\sin\left(\frac{4\pi(y-1.4)}{L}\right) \\ &\quad -0.1\cos\left(\frac{2\pi(x-2.0)}{L}\right)\cdot\sin\left(\frac{2\pi y}{L}\right), \end{aligned} \tag{4.47}$$

with $\Omega = [0, 3.2]^2$, $\varepsilon = 0.1$, $s = 0.01h$, $A = 1/16$ and $T = 0.16$. We use a linear refinement path, *i.e.*, $s = Ch$. At the final time $T = 0.16$, we expect the global error to be $\mathcal{O}(s^2) + \mathcal{O}(h^2) = \mathcal{O}(h^2)$, in either the ℓ^2 or ℓ^∞ norm, as $h, s \to 0$. The

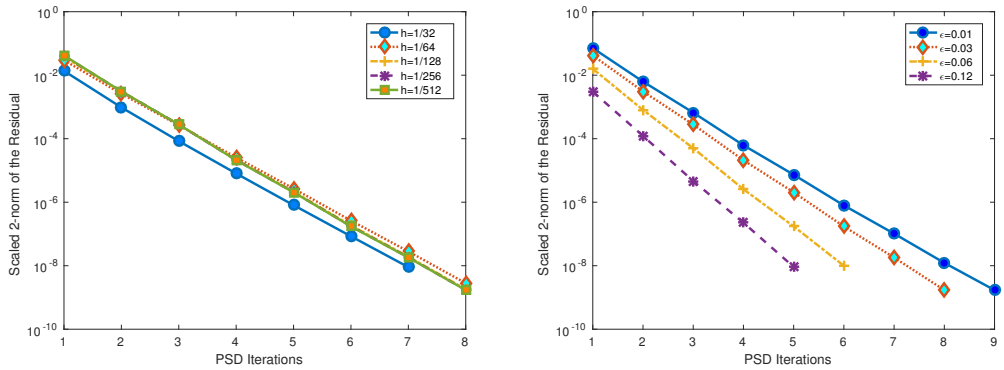
Cauchy difference is defined as $\delta_\phi := \phi_{h_f} - \mathcal{I}_c^f(\phi_{h_c})$, where $\mathcal{I}_c^f$ is a bilinear interpolation operator (with the Nearest Neighbor Interpolation applied in Matlab, which is similar to the 2D case in [31, 34] and the 3D case in [27]). This requires a relatively coarse solution, parametrized by h_c, and a relatively fine solution, parametrized by h_f, in particular $h_c = 2h_f$, at the same final time. The ℓ^2 norms of Cauchy difference and the convergence rates can be found in Table 4.2. The results confirm our expectation for the second-order convergence in both space and time.

Table 4.2: Errors, convergence rates, average iteration numbers and average CPU time (in seconds) for each time step. Parameters are given in the text, and the initial data is defined in (4.47). The refinement path is $s = 0.01h$.

h_c	h_f	$\|\delta_\phi\|_2$	Rate	$\#_{iter}$	$T_{cpu}(h_f)$
$\frac{3.2}{16}$	$\frac{3.2}{32}$	1.2392×10^{-2}	-	8	0.0011
$\frac{3.2}{32}$	$\frac{3.2}{64}$	1.6355×10^{-3}	2.92	6	0.0052
$\frac{3.2}{64}$	$\frac{3.2}{128}$	3.8124×10^{-4}	2.10	5	0.0220
$\frac{3.2}{128}$	$\frac{3.2}{256}$	9.3854×10^{-5}	2.02	4	0.0816
$\frac{3.2}{256}$	$\frac{3.2}{512}$	2.3372×10^{-5}	2.01	4	0.5217

In the second part of this test, we demonstrate the complexity of the PSD solver with initial data (4.3). In Figure 4.5, we plot the semi-log scale of the relative residuals versus PSD iteration numbers for various values of h and ε at $T = 0.02$, with time step $s = 10^{-3}$. The other common parameters are set as $A = 1/16$, $\Omega = [0, 3.2]^2$. Figure 4.5(a) indicates that the convergence rate (as gleaned from the error reduction) is nearly uniform and nearly independent of h for a fixed ε. Figure 4.5(b) shows that the number of PSD iterations increases with a decreasing value of ε, which confirms the theoretical results that the PSD solver is dependent on parameter ε in [34]. Figure 4.5 confirms the expected geometric convergence rate of the PSD solver predicted by the theory in [34].

In the third part of this test, we investigate the effect of the parameters s and A for the energy $F_h(\phi)$ with initial data (4.3). The evolutions of the energy with various time steps s and stabilized parameter A are given in Figure 4.6. As can be seen in Figure 4.6(a), the larger time steps produce inaccurate or nonphysical solutions. In

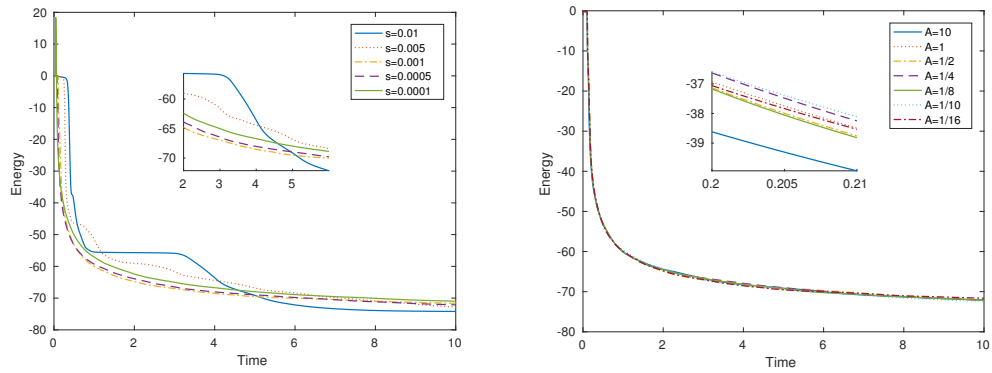


(a) h-independence: $\varepsilon = 3 \times 10^{-2}$.

(b) ε-dependence: $h = 1/512$.

Figure 4.5: Complexity tests showing the solver performance for changing values of h and ε. Parameters are given in the text.

turn, Figure 4.6(a) indicates the proper time steps and provides the motivation of using adaptive time stepping strategy. Figure 4.6(b) shows that the proposed scheme and PSD solver is not that sensitive to the stabilized parameter A when $A \leq 1$.



(a) evolutions of energy w.r.t various s

(b) evolutions of energy w.r.t various A

Figure 4.6: The effect of time steps s and stabilized parameter A for the energy $F_h(\phi)$. Left: the effect of time step s. The other parameters are $\Omega = [0, 3.2]^2$, $\epsilon = 3.0 \times 10^{-2}$ and $A = 1/16$; Right: the effect of stabilized parameter A. The other parameters are $\Omega = [0, 3.2]^2$, $\epsilon = 1.0^{-2}$ and $s = 0.001$.

Long-time coarsening process, energy dissipation and mass conservation

Coarsening processes in thin film system can take place on very long time scales [57]. In this subsection, we perform long time simulation for the SS equation. Such a test, which has been performed in many existing literature, will confirm the expected coarsening rates and serve as a benchmarks for the proposed solver; see, for example, [34, 71, 77]. The initial data for the simulations are taken as essentially random:

$$u_{i,j}^0 = 0.05 \cdot (2r_{i,j} - 1), \tag{4.48}$$

where the $r_{i,j}$ are uniformly distributed random numbers in [0, 1]. Time snapshots of the evolution for the epitaxial thin film growth model can be found in Figure 4.7. The coarsening rates are given in Figure 4.8. The interface width or roughness is defined as

$$W(t_n) = \sqrt{\frac{h^2}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} (\phi_{i,j}^n - \bar{\phi})^2}, \tag{4.49}$$

where m and n are the number of the grid points in x and y direction and $\bar{\phi}$ is the average value of ϕ on the uniform grid. The log-log plots of roughness and energy evolution and the corresponding linear regression are presented in Figure. 4.8. The linear regression in Figure. 4.8 indicates that the surface roughness grows like $t^{1/3}$, while the energy decays like $t^{-1/3}$, which verifies the one-third power law predicted in [58]. More precisely, the linear fits have the form $a_e t^{b_e}$ with $a_e = 3.09870, b_e = -0.33554$ for energy evolution and $a_m t^{b_m}$ with $a_m = -5.35913, b_m = 0.32555$ for roughness evolution. The linear regression is only taken up to $t = 3000$, since the saturation time would be of the order of ε^{-2} under the scaling that we have adopted [71]. These simulation results are consistent with earlier works on this topic in [34, 71, 77, 83]. Moreover, the PSD iteration at each time step demonstrates the efficiency

of the PSD solver and the mass difference indicates that the mass is conservative, up to a tolerance of 10^{-10}, for the simulation depicted in Figure 4.7.

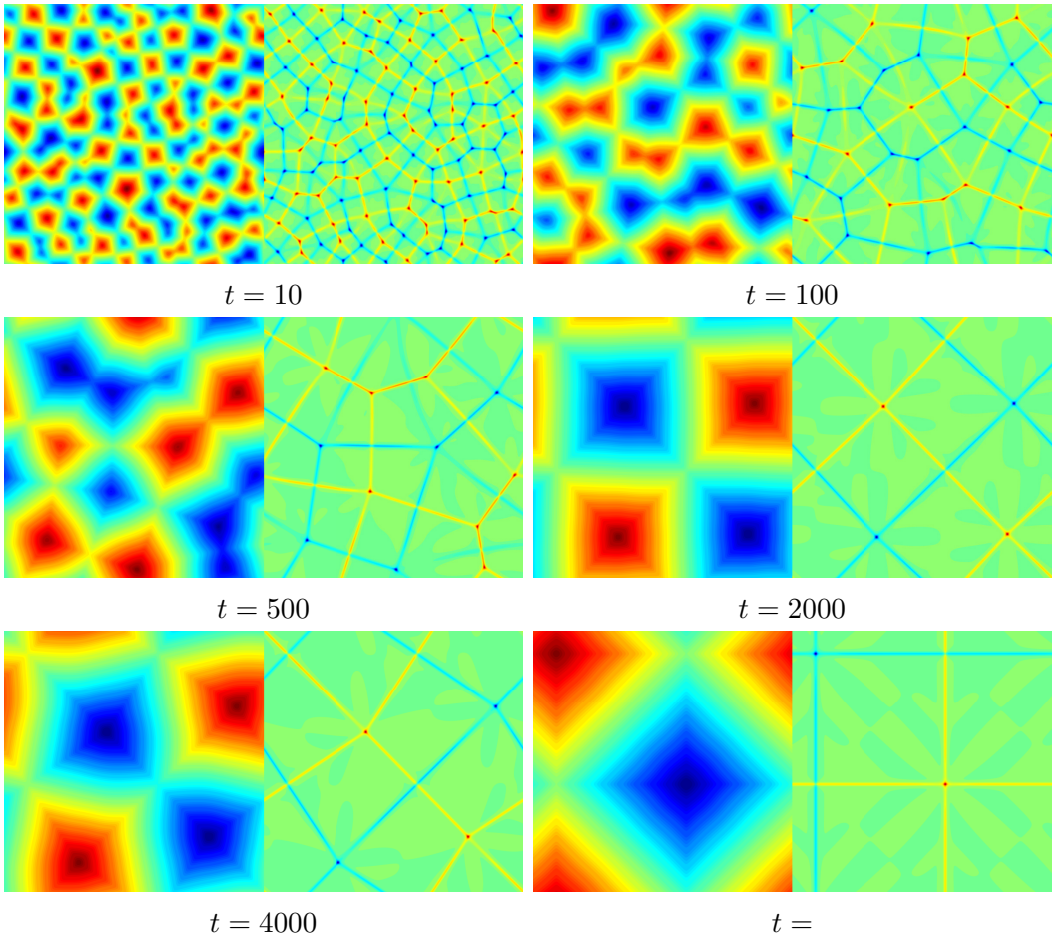
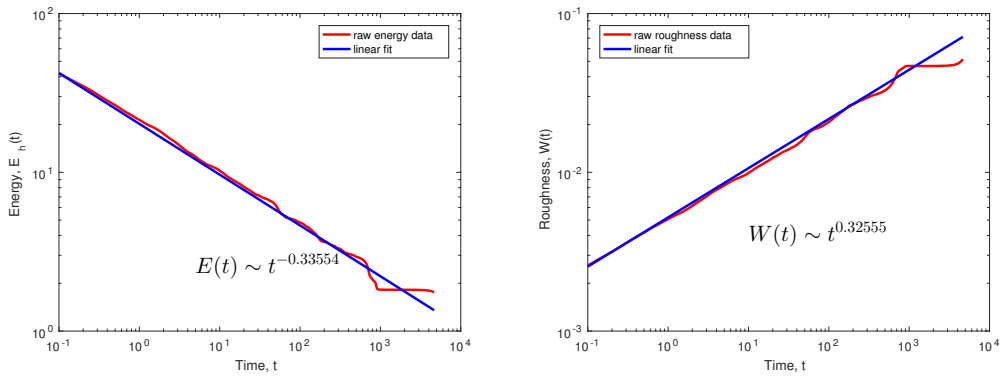
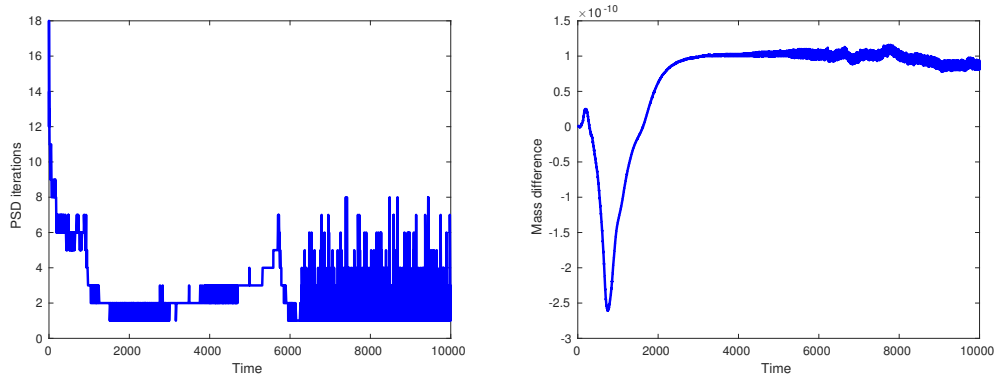


Figure 4.7: Time snapshots of the evolution with PSD solver for the epitaxial thin film growth model at $t = 10, 100, 500, 2000, 4000$ and 10000. Left: contour plot of u, Right: contour plot of Δu. The parameters are $\varepsilon = 0.03, \Omega = [12.8]^2, s = 0.001$, $h = {}^{12.8}/_{512}$ and $A = {}^{1}/_{16}$. These simulation results are consistent with earlier work on this topic in [34, 71, 77, 83].



(a) Energy evolution (b) Roughness evolution

Figure 4.8: The log-log plots of energy and roughness evolution and the corresponding linear regression for the simulation depicted in Figure 4.7.



(a) PSD iterations (b) Mass difference

Figure 4.9: PSD iterations and mass difference at each time steps for the simulation depicted in Figure 4.7.

4.3 Application to Square Phase Field Crystal Model with First-Order-In-Time Scheme

The content in this chapter has been published in [34], for more details please refer to [34].

4.3.1 Introduction

Suppose that $\Omega \subset \mathbb{R}^d$, $d = 2, 3$ is a rectangular domain. The energy of square phase field crystal (SPFC) model is given by [29, 43, 47, 63]:

$$\mathcal{E}[u] = \int_\Omega \left\{ \frac{\gamma_0}{2} u^2 - \frac{\gamma_1}{2} |\nabla u|^2 + \frac{\varepsilon^2}{2} |\Delta u|^2 + \frac{1}{4} |\nabla u|^4 \right\} d\mathbf{x},$$

where $u : \Omega \to \mathbb{R}$ corresponds to the number density field of the atoms, and $\varepsilon > 0$, $\gamma_0, \gamma_1 \geq 0$ are parameters. The SPFC model is the H^{-1} gradient flow of this energy and is given by

$$\partial_t u = \Delta w, \quad w := \delta \mathcal{E} = \gamma_0 u + \gamma_1 \Delta u + \varepsilon^2 \Delta^2 u - \nabla \cdot \left(|\nabla u|^2 \nabla u \right).$$

We propose the following fully-implicit, nonlinear convex-splitting scheme

$$u^{n+1} - \Delta_h w^{n+1} = g, \quad s\gamma_0 u^{n+1} - s\nabla_h^v \cdot \left(\left| \nabla_h^v u^{n+1} \right|^2 \nabla_h^v u^{n+1} \right) + s\varepsilon^2 \Delta_h^2 u^{n+1} - w^{n+1} = f, \tag{4.50}$$

where $g = u^n$ and $f = -\gamma_1 \Delta_h u^n$. Using the techniques of [77, 82], we can prove that this scheme is unconditionally energy stable. The fully discrete scheme can also be rewritten in operator format as $\mathcal{N}_h[u^{n+1}] = g$, where

$$\mathcal{N}_h[\nu] := s\gamma_0 \nu + s\varepsilon^2 \Delta_h^2 \nu - s\nabla_h^v \cdot \left(|\nabla_h^v \nu|^2 \nabla_h^v \nu \right) - T_h[-\nu + f].$$

We can shift the scheme from the affine space of solutions – whose elements ν satisfy $(\nu - \overline{g}, 1)_2 = 0$ – to the mean zero space, but this is not necessary for practical implementation. Otherwise, this scheme is in the scope of our theory, and, according to the prescription in Section 3.2.4, the pre-conditioner should be

$$\mathcal{L}_h[\nu] := s\gamma_0\nu - s\Delta_h\nu + s\varepsilon^2\Delta_h^2\nu - T_h[-\nu].$$

Given $u^k \in \mathcal{C}_{\text{per}}$, with $\left(u^k - \overline{g}, 1\right)_2 = 0$, we compute the search direction $d^k \in \mathring{\mathcal{C}}_{\text{per}}$ by solving the sixth order linear problem $\mathcal{L}[d^k] = f - \mathcal{N}_h[u^k]$ using FFT. Once d^k is found, we perform the line-search: find $\alpha_k \in \mathbb{R}$ such that $q(\alpha_k) = 0$, where

$$q(\alpha) = \left(\mathcal{N}_h[u^k + \alpha d^k] - f, d^k\right)_2.$$

After this, we update the approximation via $u^{k+1} = u^k + \alpha_k d^k$. As before, q is a cubic polynomial (since $p = 4$) whose coefficients can be precomputed. But this time, two of the coefficients involve the $\mathsf{T}_h = -\Delta_h^{-1}$ operator. Specifically, for $q(\alpha)$ we need to compute

$$\begin{aligned}\left(\mathsf{T}_h\left[u^k - f + \alpha d^k\right], d^k\right)_2 &= \left(\mathsf{T}_h\left[u^k - f\right], d^k\right)_2 + \alpha\left(\mathsf{T}_h\left[d^k\right], d^k\right)_2 \\ &= \left(u^k - f, \mathsf{T}_h\left[d^k\right]\right)_2 + \alpha\left(d^k, \mathsf{T}_h\left[d^k\right]\right)_2,\end{aligned}$$

where we have use the linearity and symmetry properties of the T_h operator. These terms have only to be calculated once per line search, and can be efficiently computed using FFT. In fact, observe that we only need to compute $\mathsf{T}_h\left[d^k\right]$, at the cost of a single FFT, per line search!

4.3.2 Numerical Experiments

The 4-Laplacian term in (4.50) gives preference to rotationally invariant patterns with square symmetry. We perform a simple test showing the emergence of these

patterns in this subsection. The initial data for those simulations are similar to (4.5), but we add nucleation sites at specific locations in the domain. The rest of the parameters are given by $\varepsilon = 1.0$; $\lambda = \gamma_0 = 0.5$; $\gamma_1 = 2.0$; $\Omega = (0, 100)^2$; and $s = 0.01$. The time snapshots of the evolution by using the given parameters are presented in Figures 4.10 (one nucleation site) and 4.11 (four nucleation sites). These tests confirm the emergence of the rotationally invariant square-symmetry patterns in the density field u.

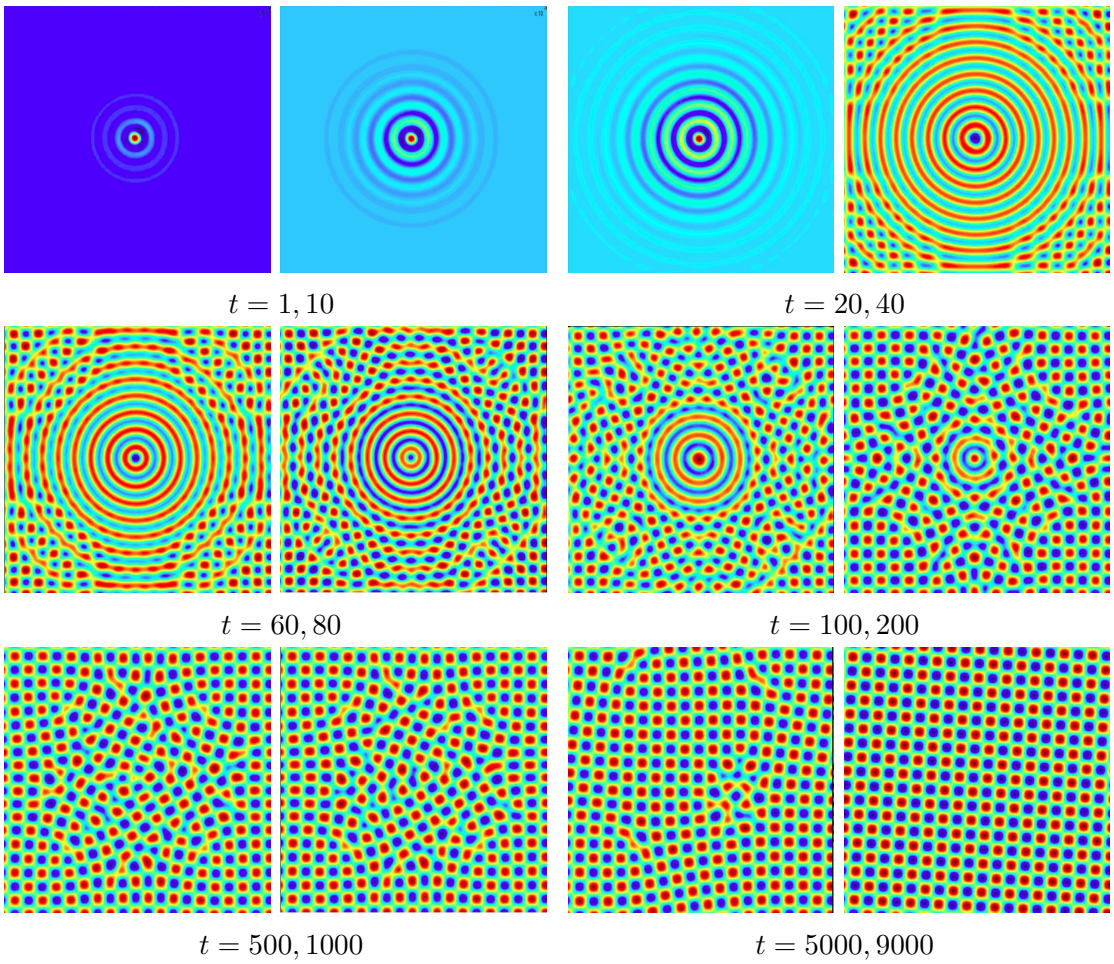


Figure 4.10: Time snapshots of the evolution with PSD solver for squared phase field crystal model at $t = 1, 10, 20, 40, 60, 80, 100, 200, 500, 1000, 5000$ and 9000. The parameters are $\epsilon = 1.0, \lambda = 0.5, \gamma_1 = 2.0, \Omega = [100]^2$ and $s = 0.01$.

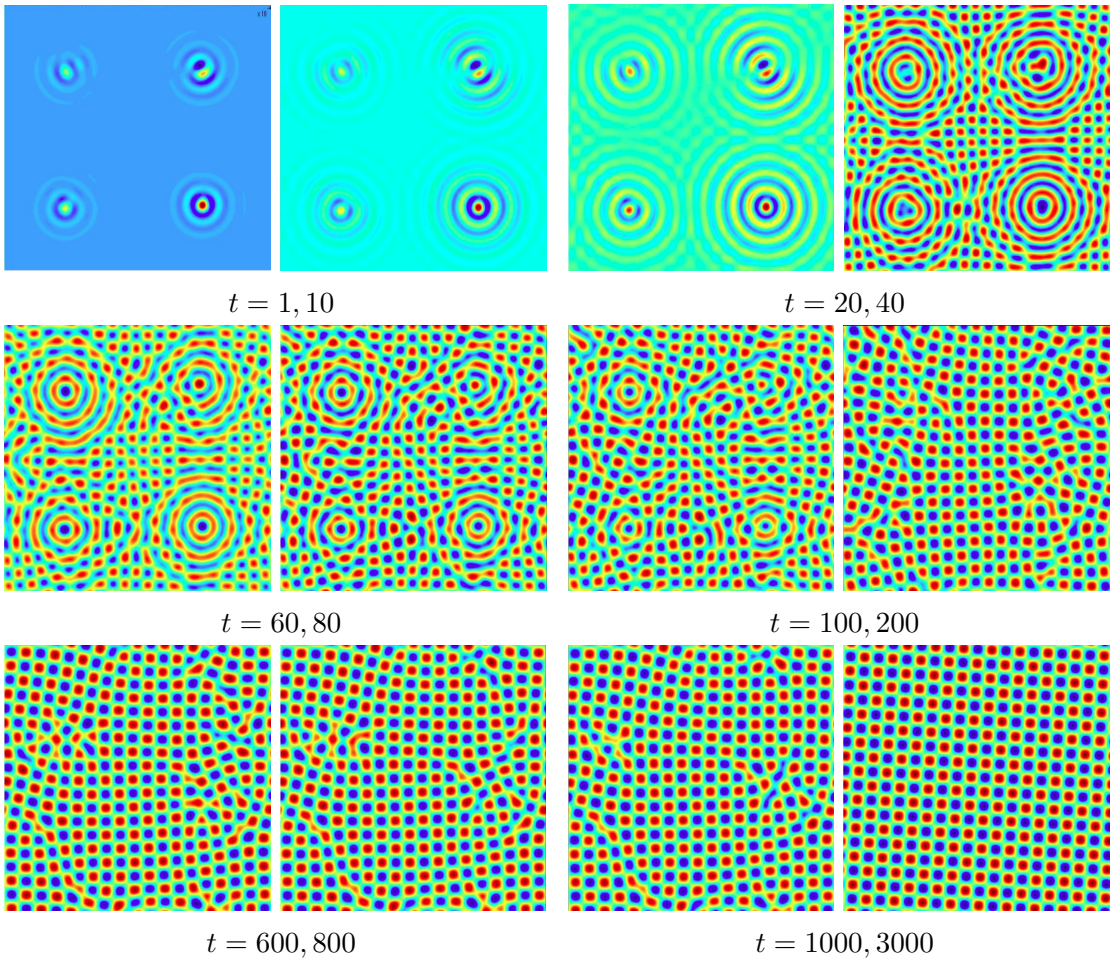


Figure 4.11: Time snapshots of the evolution with PSD solver for squared phase field crystal model at $t = 1, 10, 20, 40, 60, 80, 100, 200, 600, 800, 1000$ and 3000. The parameters are $\epsilon = 1.0, \lambda = 0.5, \gamma_1 = 2.0, \Omega = [100]^2$ and $s = 0.01$.

4.4 Application to Functionalized Cahn-Hilliard Model with First-Order-In-Time Scheme

The content in this chapter has been published in [31], for more details please refer to [31].

4.4.1 Introduction

The Functionalized Cahn-Hilliard (FCH) model was first derived to describe small-angle X-ray scattering data of an amphiphilic mixture in [44]. Recently, the FCH model has been proposed to model the interfacial energy in amphiphilic phase-separated mixtures in [26, 42, 67] where the FCH equations were extended to describe the network morphology of solvated functionalized polymer membranes, such as bilayer in [23, 26], pearling bifurcation in [68, 26], pore-like and micelle network structures in [41, 42, 68]. The FCH energy, which includes a negative multiple of the Cahn-Hilliard energy balanced against the square of its own variational derivative, is highly related to the standard Allen-Cahn (AC) and Cahn-Hilliard (CH) energy [1, 10, 11], given by

$$\mathcal{F}_{CH}(\phi) = \int_\Omega \left\{ \frac{1}{4}\phi^4 - \frac{1}{2}\phi^2 + \frac{\varepsilon^2}{2}\left|\nabla\phi\right|^2 \right\} d\mathbf{x}, \tag{4.51}$$

with $\Omega \subset \mathbf{R}^D$, $D = 2$ or 3. The phase variable $\phi : \Omega \to \mathbf{R}$ is the concentration field, and ε is the width of interface. We assume that $\Omega = (0, L_x) \times (0, L_y) \times (0, L_z)$, ϕ and $\Delta\phi$ are periodic on Ω. In turn, the chemical potential becomes

$$\mu_{CH} := \delta_\phi \mathcal{F}_{CH} = \phi^3 - \phi - \varepsilon^2 \Delta\phi, \tag{4.52}$$

where $\delta_\phi \mathcal{F}_{CH}$ denotes the variational derivative with respect to ϕ. Herein we consider a dimensionless energy of a binary mixture:

$$\mathcal{F}(\phi) \quad = \quad \frac{\varepsilon^{-2}}{2} \int_\Omega \mu_{CH}^2 d\mathbf{x} - \eta \mathcal{F}_{CH}(\phi), \tag{4.53}$$

where η is the switch parameter. When $\eta > 0$ and $\eta < 0$, (4.53) represents the FCH energy [26, 51, 67] and the Cahn-Hilliard-Willmore (CHW) energy [75, 76, 81], respectively. Furthermore, (4.53) represents the strong FCH energy when $\eta = \varepsilon^{-1}$ and weak FCH energy when $\eta = 1$ [26]. By the definition of CH energy in (4.51) and

chemical potential in (4.52), we have

$$\begin{aligned}\mu := \delta_\phi \mathcal{F} &= 3\varepsilon^{-2}\phi^5 - \left(4\varepsilon^{-2}+\eta\right)\phi^3 + \left(\varepsilon^{-2}+\eta\right)\phi + \varepsilon^2\Delta^2\phi + \left(2+\eta\varepsilon^2\right)\Delta\phi \\ &\quad +6\phi\left|\nabla\phi\right|^2 - 6\nabla\cdot\left(\phi^2\nabla\phi\right).\end{aligned}$$

The conserved H^{-1} gradient flow [26, 55, 67] is given by

$$\partial_t\phi = \nabla\cdot\left(M(\phi)\nabla\mu\right), \tag{4.54}$$

where $M(\phi) > 0$ is a diffusion mobility, and where we assume that μ is periodic on Ω.

The FCH equation (4.54) is a sixth-order, highly nonlinear parabolic equation. Numerical approximation of (4.54) is very challenging because of the high derivative order and highly nonlinear nature. One of the biggest challenges is to overcome the numerical stiffness encountered with time-space discretization. Roughly speaking, since the equation is sixth-order parabolic, an explicit numerical scheme is expected to encounter a severe CFL condition: $s \leq Ch^6$, with s and h the time and space step sizes. On the other hand, a fully implicit scheme, such as the backward Euler method, may still be only conditionally stable, and, very likely, will only be conditionally solvable. Ideally, one would like a scheme that preserves some of the time-invariant quantities of the PDE, such as mass conservation and the energy dissipation rate. The first invariant is easily maintained, while the second one is a major challenge. *Often, one attempts only to design a scheme that will dissipate the free energy at the numerical level, without attempting to directly control the rate of dissipation.* In particular, one wants $\mathcal{F}(\phi^{k+1}) \leq \mathcal{F}(\phi^k)$, where ϕ^k is the approximated phase variable at time step k, given some mild CFL condition, or no CFL condition whatever. The energy dissipativity imparts some stability notion for the PDE and the numerical method, as we will see. If $\mathcal{F}(\phi^{k+1}) \leq \mathcal{F}(\phi^k)$, for all $k \geq 1$, with no condition on the time step size, we say that the scheme is *unconditionally strongly energy stable.*

Finally, for large-scale calculations in practice, novel efficient numerical linear and nonlinear solvers have to be carefully developed. We will address this issue in the thesis as well.

There have been a few previous works on the numerical approximation of the FCH equation. In [12], Chen *et al.* presented an efficient linear, first-order (in time) spectral-Galerkin method for the FCH equation. Their scheme, which utilized linear stabilization terms, are unconditionally solvable, but not necessary energy stable. Jones studied a semi-implicit numerical scheme for the FCH equation in his Ph.D thesis [55]; the energy stability was proved, while the unique solvability has not been theoretically justified. In a more recent work [19], fully implicit schemes with pseudo-spectral approximation in space for the FCH equation are proposed. While they neither proved energy stability or solvability, they did carry out several tests to show the accuracy and efficiency of their methods. In another work [46], Guo *et al.* presented a Local Discontinuous Galerkin (LDG) method to overcome the difficulty associated with the higher order spatial derivatives. Energy stability was established for the semi-discrete (time-continuous) scheme. Their fully discrete scheme was based on the time discretization in [12]. To our knowledge, there has been no rigorous convergence analysis for the FCH model in the existing literature, because of its highly nonlinear nature. In [79] the authors developed a Runge-Kutta exponential time integration (EKR) method for the diffuse Willmore flow, an equation that is closely related to the FCH and CHW models (4.54). This method works well when $M \equiv 1$, but may need to be significantly modified otherwise. It enables one to generate high-order single-step methods, which have a significant advantage over multistep methods when the time step changes adaptively.

In this thesis we propose and analyze an efficient computational scheme for solving the FCH equation primarily, though the theory will be applicable to the CHW equation as well. The convex splitting method, which treats the convex part implicitly and concave part explicitly, has been a popular approach for gradient flows, since it ensures the unique solvability and unconditional stability; see the related works

[2, 7, 13, 15, 16, 24, 25, 45, 52, 78, 77, 82] for a wide class of phase field models. For the FCH equation (4.54), one key difficulty in the energy stability could be observed in the fact that, one nonlinear energy functional in the expansion turns out to be non-convex, non-concave, so that the convex splitting approach is not directly available. To overcome this difficulty, we introduce two auxiliary terms in the energy functional, so that their combination with the original term become convex. In turn, a convex-concave decomposition for the FCH energy is available, and the first order in time convex-splitting scheme could be appropriately designed. Because of its convex splitting nature, both the unique solvability and unconditional energy stability could be theoretically justified.

As a result of the proposed numerical scheme, a 4-Laplacian term has to be solved in an H^{-1} gradient flow at each time step in the finite difference approximation, which turns out to be very challenging. We apply the Preconditioned Steepest Descent (PSD) solver, recently proposed and analyzed in [34], to solve the nonlinear system. The main idea is to use a linearized version of the nonlinear operator as a pre-conditioner, or in other words, as a metric for choosing the search direction. Furthermore, the convexity of the nonlinear energy functional assures a geometric convergence of such a PSD iteration. In practice, only a Poisson-like equation needs to be solved at each iteration stage, and the geometric convergence of the nonlinear iteration greatly improves the numerical efficiency.

On the theoretical side, we also present a global in time H^2_{per} stability of the numerical scheme. This uniform in time bound enables us to derive the full order convergence analysis, with first order temporal accuracy and second order spatial accuracy. In addition, such a convergence is unconditional, without any requirement between the time step size s and the spatial mesh h. In the authors' knowledge, this is the first such theoretical result for the FCH/CHW model.

4.4.2 The first order convex splitting scheme

The convex-concave energy decomposition with auxiliary terms

For any $\phi \in H^2_{\text{per}}(\Omega)$, the FCH energy in (4.53) could be expanded as

$$\begin{aligned}\mathcal{F}(\phi) &= \frac{\varepsilon^{-2}}{2}\|\phi\|_{L^6}^6 - \left(\varepsilon^{-2}+\frac{\eta}{4}\right)\|\phi\|_{L^4}^4 + \left(\frac{\varepsilon^{-2}}{2}+\frac{\eta}{2}\right)\|\phi\|^2 + \frac{\varepsilon^2}{2}\|\Delta\phi\|^2 \\ &\quad - \left(1+\frac{\eta\varepsilon^2}{2}\right)\|\nabla\phi\|^2 + 3\int_\Omega \phi^2\,|\nabla\phi|^2\,d\mathbf{x}. \end{aligned} \tag{4.55}$$

Unlike the energies for the AC in [37], CH in [2, 25, 30, 38, 45], Phase Field Crystal (PFC) and the modified version in [7, 52, 78, 82], epitaxial thin film growth in [13, 16, 72, 77], the convex splitting idea cannot be directly applied to the FCH energy (4.53). The main difficulty is associated with the last term in (4.55):

$$\mathcal{G}(\phi) := \int_\Omega 3\phi^2\,|\nabla\phi|^2\,d\mathbf{x}, \tag{4.56}$$

which is neither convex nor concave. To overcome this difficulty, we perform a careful analysis for the following energy functional:

$$\mathcal{H}(\phi) := \int_\Omega \left(A(\phi^4+|\nabla\phi|^4) + 3\phi^2\,|\nabla\phi|^2\right)d\mathbf{x}. \tag{4.57}$$

Lemma 4.4.1. *$\mathcal{H} : W^{1,4}_{\text{per}}(\Omega) \to \mathbb{R}$ is convex provided that $A \geq 1$.*

Proof. We denote $g(\phi) := 3\phi^2\,|\nabla\phi|^2$ and $h(\phi) := A(\phi^4+|\nabla\phi|^4) + g(\phi)$, so that $\mathcal{G}(\phi) = \int_\Omega g(\phi)\,d\mathbf{x}$ and $\mathcal{H}(\phi) = \int_\Omega h(\phi)\,d\mathbf{x}$. Based on the following inequalities, which come from the convexity of $q_2(x) = x^2$ and $r_2(\chi) = \chi\cdot\chi$):

$$\left(\frac{\phi_1+\phi_2}{2}\right)^2 \leq \frac{\phi_1^2+\phi_2^2}{2}, \quad \left|\nabla\left(\frac{\phi_1+\phi_2}{2}\right)\right|^2 \leq \frac{|\nabla\phi_1|^2+|\nabla\phi_2|^2}{2}, \quad \forall \phi_1,\,\phi_2,$$

we get

$$g\left(\frac{\phi_1+\phi_2}{2}\right) = 3\left(\frac{\phi_1+\phi_2}{2}\right)^2 \left|\nabla\left(\frac{\phi_1+\phi_2}{2}\right)\right|^2 \le 3\frac{\phi_1^2+\phi_2^2}{2}\cdot\frac{|\nabla\phi_1|^2+|\nabla\phi_2|^2}{2}.$$

A careful comparison with $\frac{g(\phi_1)+g(\phi_2)}{2} = \frac{3\phi_1^2|\nabla\phi_1|^2+3\phi_2^2|\nabla\phi_2|^2}{2}$ shows that

$$\begin{aligned}\frac{g(\phi_1)+g(\phi_2)}{2} - g\left(\frac{\phi_1+\phi_2}{2}\right) &\ge \frac{3(\phi_1^2-\phi_2^2)(|\nabla\phi_1|^2-|\nabla\phi_2|^2)}{4} \\ &\ge -\frac{3}{8}\left((\phi_1^2-\phi_2^2)^2+(|\nabla\phi_1|^2-|\nabla\phi_2|^2)^2\right). \quad (4.58)\end{aligned}$$

Meanwhile, the convexity of $q_4(x) = x^4$ and $r_4(\chi) = |\chi|^4$ indicates the following inequalities:

$$\frac{\phi_1^4+\phi_2^4}{2} - \left(\frac{\phi_1+\phi_2}{2}\right)^4 \ge \frac{3}{8}(\phi_1^4+\phi_2^4-2\phi_1^2\phi_2^2) = \frac{3}{8}(\phi_1^2-\phi_2^2)^2, \quad (4.59)$$

and

$$\begin{aligned}\frac{|\nabla\phi_1|^4+|\nabla\phi_2|^4}{2} - \left|\nabla\left(\frac{\phi_1+\phi_2}{2}\right)\right|^4 &\ge \frac{3}{8}(|\nabla\phi_1|^4+|\nabla\phi_2|^4-2|\nabla\phi_1|^2\cdot|\nabla\phi_2|^2) \\ &= \frac{3}{8}(|\nabla\phi_1|^2-|\nabla\phi_2|^2)^2. \quad (4.60)\end{aligned}$$

A combination of (4.58), (4.59) and (4.60) implies that

$$\frac{h(\phi_1)+h(\phi_2)}{2} - h\left(\frac{\phi_1+\phi_2}{2}\right) \ge 0, \quad \forall\phi_1,\, \phi_2,$$

provided that $A \ge 1$. As a result, an integration over Ω leads to the following fact:

$$\frac{\mathcal{H}(\phi_1)+\mathcal{H}(\phi_2)}{2} - \mathcal{H}\left(\frac{\phi_1+\phi_2}{2}\right) \ge 0, \quad \forall\phi_1,\, \phi_2, \quad \text{if } A \ge 1.$$

The convexity of H is assured under the condition $A \ge 1$. Lemma 1 is proved. □

Corollary 4.4.2. *The energy $\mathcal{F} : H^2_{\text{per}}(\Omega) \to \mathbb{R}$ possesses a convex splitting over $H^2_{\text{per}}(\Omega)$. In particular,*

$$\mathcal{F}(\phi) = \mathcal{F}_c(\phi) - \mathcal{F}_e(\phi), \tag{4.61}$$

with

$$\mathcal{F}_c(\phi) := \int_\Omega \left\{ \frac{\varepsilon^{-2}}{2}\phi^6 + \left(\frac{\varepsilon^{-2}}{2} + \frac{\eta}{2}\right)\phi^2 + \frac{\varepsilon^2}{2}(\Delta\phi)^2 + A(\phi^4 + |\nabla\phi|^4) + 3\phi^2 |\nabla\phi|^2 \right\} d\mathbf{x}, \tag{4.62}$$

and

$$\mathcal{F}_e(\phi) := \int_\Omega \left\{ \left(\epsilon^{-2} + \frac{\eta}{4}\right)\phi^4 + \left(1 + \frac{\eta\varepsilon^2}{2}\right)|\nabla\phi|^2 + A(\phi^4 + |\nabla\phi|^4) \right\} d\mathbf{x}, \tag{4.63}$$

where both $\mathcal{F}_c, \mathcal{F}_e : H^2_{\text{per}}(\Omega) \to \mathbb{R}$ are strictly convex.

We recall the following proposition from [82]:

Proposition 4.4.3. *Suppose that $\phi, \psi \in H^4_{\text{per}}(\Omega)$ and that $\mathcal{F}$ admits a (not necessarily unique) convex splitting into $\mathcal{F} = \mathcal{F}_c - \mathcal{F}_e$ then*

$$\mathcal{F}(\phi) - \mathcal{F}(\psi) \leq \left(\delta_\phi \mathcal{F}_c(\phi) - \delta_\phi \mathcal{F}_e(\psi), \phi - \psi\right). \tag{4.64}$$

If $\phi, \psi \in H^2_{\text{per}}(\Omega)$ only, then (4.64) can be understood in the weak sense.

The first order convex splitting scheme

Based on the convex-concave decomposition in (4.62) and (4.63) for the physical energy $\mathcal{F}(\phi)$, we consider the following semi-implicit, first-order-in-time, convex splitting scheme:

$$\phi^{k+1} - \phi^k = s\nabla \cdot \left(M(\phi^k)\nabla\tilde{\mu}\right), \quad \tilde{\mu}\left(\phi^{k+1}, \phi^k\right) := \delta_\phi \mathcal{F}_c(\phi^{k+1}) - \delta_\phi \mathcal{F}_e(\phi^k). \tag{4.65}$$

where, precisely,

$$\begin{aligned}\tilde{\mu}\left(\phi^{k+1},\phi^{k}\right) &= 3\varepsilon^{-2}(\phi^{k+1})^5 + 4A(\phi^{k+1})^3 + (\varepsilon^{-2}+\eta)\phi^{k+1} + \varepsilon^2\Delta^2\phi^{k+1} \\ &\quad + 6\phi^{k+1}\left|\nabla\phi^{k+1}\right|^2 - 6\nabla\cdot\left((\phi^{k+1})^2\nabla\phi^{k+1}\right) - 4A\nabla\cdot\left(|\nabla\phi^{k+1}|^2\nabla\phi^{k+1}\right) \\ &\quad - (4\varepsilon^{-2}+\eta)(\phi^k)^3 + (2+\eta\varepsilon^2)\Delta\phi^k - 4A(\phi^k)^3 + 4A\nabla\cdot\left(|\nabla\phi^k|^2\nabla\phi^k\right).\end{aligned} \tag{4.66}$$

The scheme may be expressed in a weak form as follows: find the pair $(\phi,\mu)\in H^2_{\text{per}}(\Omega)\times H^1_{\text{per}}(\Omega)$ such that

$$(\phi,\nu) + s(M\nabla\mu,\nabla\nu) = (g,\nu), \tag{4.67}$$

$$\left(3\varepsilon^{-2}\phi^5 + 4A\phi^3 + (\varepsilon^{-2}+\eta)\phi,\psi\right) + \varepsilon^2(\Delta\phi,\Delta\psi) + 6(\phi\,|\nabla\phi|^2,\psi) \tag{4.68}$$

$$+6\left(\phi^2\nabla\phi,\nabla\psi\right) + 4A\left(|\nabla\phi|^2\nabla\phi,\nabla\psi\right) - (\mu,\psi) = (f,\psi), \tag{4.69}$$

where $g=\phi^k$, $M=M(\phi^k)$, and

$$f = \delta_\phi\mathcal{F}_e(\phi^k) = (4\varepsilon^{-2}+\eta)(\phi^k)^3 - (2+\eta\varepsilon^2)\Delta\phi^k + 4A(\phi^k)^3 - 4A\nabla\cdot\left(|\nabla\phi^k|^2\nabla\phi^k\right).$$

Observe that, if $\phi^k\in H^2_{\text{per}}(\Omega)$ is given, we have $g\in L^2_{\text{per}}(\Omega)=L^2(\Omega)$.

Theorem 4.4.4. *The convex splitting scheme (4.65) is uniquely solvable and unconditionally energy stable: $\mathcal{F}(\phi^{k+1})\le\mathcal{F}(\phi^k)$. In particular, if $\phi^k\in H^2_{\text{per}}(\Omega)$, then $\phi^{k+1}\in H^2_{\text{per}}(\Omega)$.*

Proof. The existence and unique solvability follows from standard convexity analyses. For the stability, let $\phi=\phi^{k+1}$ and $\psi=\phi^k$ in (4.64) to find

$$\begin{aligned}\mathcal{F}(\phi^{k+1})-\mathcal{F}(\phi^k) &\le \left(\delta_\phi\mathcal{F}_c(\phi^{k+1})-\delta_\phi\mathcal{F}_e(\phi^k),\phi^{k+1}-\phi^k\right) \\ &= s\left(\tilde{\mu},\nabla\cdot\left(M(\phi^k)\nabla\tilde{\mu}\right)\right) = -s\left(\nabla\tilde{\mu},M(\phi^k)\nabla\tilde{\mu}\right)\le 0,\end{aligned}$$

where we have interpreted the right-hand-side of (4.64) in the weak sense. □

Global-in-time H^2_{per} stability of the numerical scheme

For simplicity, we will take the mobility $M \equiv 1$ in the remainder of the thesis.

Lemma 4.4.5. *There are constants $C_0, C_1 > 0$ such that, for all $\phi \in H^2_{\text{per}}(\Omega)$,*

$$\frac{\varepsilon^{-2}}{6} \|\phi\|_{L^6}^6 + C_0 \varepsilon^2 \|\phi\|_{H^2_{\text{per}}}^2 \leq \mathcal{F}(\phi) + C_1. \tag{4.70}$$

Proof. For the concave diffusion term in (4.55), an application of Cauchy's inequality shows that

$$\|\nabla\phi\|^2 = \int_\Omega \phi \cdot \Delta\phi \, d\mathbf{x} \leq \|\phi\| \cdot \|\Delta\phi\| \leq \frac{\varepsilon^2}{4(1+\frac{\eta\varepsilon^2}{2})} \|\Delta\phi\|^2 + \frac{1+\frac{\eta\varepsilon^2}{2}}{\varepsilon^2} \|\phi\|^2, \quad \forall \eta > 0. \tag{4.71}$$

Then we obtain

$$(1+\frac{\eta\varepsilon^2}{2}) \|\nabla\phi\|^2 \leq \frac{\varepsilon^2}{4} \|\Delta\phi\|^2 + C_2 \|\phi\|^2, \tag{4.72}$$

with $C_2 := (1+\frac{\eta\varepsilon^2}{2})^2 \varepsilon^{-2} = O(\varepsilon^{-2})$. Applications of Hölder's inequality imply that

$$\|\phi\|_{L^6} \geq \frac{1}{|\Omega|^{1/12}} \|\phi\|_{L^4}, \quad \|\phi\|_{L^6} \geq \frac{1}{|\Omega|^{1/3}} \|\phi\|.$$

Now, define $C_3 := C_2 - \left(\frac{\varepsilon^{-2}}{2} + \frac{\eta}{2}\right) + 1 > 0$; we note that $C_3 = O(\varepsilon^{-2})$. As a consequence of the last two inequalities, we get

$$\frac{1}{6} \|\phi\|_{L^6}^6 \geq \frac{1}{6|\Omega|^{1/2}} \|\phi\|_{L^4}^6 \geq (1+\frac{\eta\varepsilon^2}{4}) \|\phi\|_{L^4}^4 - C_4, \tag{4.73}$$

$$\frac{1}{6} \|\phi\|_{L^6}^6 \geq \frac{1}{6|\Omega|^2} \|\phi\|^6 \geq \varepsilon^2 C_3 \|\phi\|^2 - C_5, \tag{4.74}$$

for some constants $C_4, C_5 > 0$, which are of order 1, where Young's inequality was repeated applied. Therefore, a combination of (4.55), (4.72), (4.73) and (4.74) yields

$$\begin{aligned} \mathcal{F}(\phi) &\geq \frac{\varepsilon^{-2}}{6} \|\phi\|_{L^6}^6 + \|\phi\|^2 + \frac{\varepsilon^2}{4} \|\Delta\phi\|^2 - C_1, \\ &\geq \frac{\varepsilon^{-2}}{6} \|\phi\|_{L^6}^6 + C_0 \varepsilon^2 \|\phi\|_{H^2_{\text{per}}}^2 - C_1, \end{aligned} \tag{4.75}$$

where $C_1 := \varepsilon^{-2}(C_4 + C_5) = O(\varepsilon^{-2})$ and the elliptic regularity estimate $\|\phi\|_{H^2}^2 \leq C_0(\|\phi\|^2 + \|\Delta\phi\|^2)$ was applied in the second step. □

Corollary 4.4.6. *Suppose that $\phi_0 \in H^2_{\text{per}}(\Omega)$. For any positive integer k, we have*

$$\left\|\phi^k\right\|_{H^2_{\text{per}}} \leq C_6 := \frac{\mathcal{F}(\phi^0) + C_1}{C_0 \varepsilon^2}. \tag{4.76}$$

Proof. The unconditional energy stability in Theorem 4.4.4 implies that, for any positive integer k,

$$\mathcal{F}(\phi^k) \leq \mathcal{F}(\phi^0). \tag{4.77}$$

A combination of (4.70) and (4.77) yields the result. □

Remark 4.4.7. *Note that the constant C_6 is independent of k and s, but does depends on ε. In particular, $C_6 = O(\varepsilon^{-4})$.*

4.4.3 Convergence analysis

Main result

The convergence result is stated in the following theorem.

Theorem 4.4.8. *Let ϕ_e be the exact solution of the FCH equation (4.54) with the periodic boundary condition and let ϕ be the numerical solution of (4.65). Then the following convergence result holds as s goes to zero:*

$$\|\phi_e - \phi\|_{\ell^\infty(0,T;\mathring{H}^{-1}_{\text{per}})} + \|\phi_e - \phi\|_{\ell^2(0,T;H^2_{\text{per}})} \leq Cs, \tag{4.78}$$

where the constant C depends only on the regularity of the exact solution.

Proof of the main result

Consistency analysis

We denote the exact solution $\Phi(x,y,z,t) = \phi_e(x,y,z,t)$. A detailed Taylor expansion implies the following truncation error:

$$\begin{aligned}
\frac{\Phi^{k+1}-\Phi^k}{s} &= \Delta\Big(3\varepsilon^{-2}(\Phi^{k+1})^5 - (4\varepsilon^{-2}+\eta)(\Phi^k)^3 + (\varepsilon^{-2}+\eta)\Phi^{k+1} + \varepsilon^2\Delta^2\Phi^{k+1} \\
&\quad +(2+\eta\varepsilon^2)\Delta\Phi^k + 6\Phi^{k+1}\left|\nabla\Phi^{k+1}\right|^2 - 6\nabla\cdot\left((\Phi^{k+1})^2\nabla\Phi^{k+1}\right) \\
&\quad +4A(\Phi^{k+1})^3 - 4A\nabla\cdot\left(|\nabla\Phi^{k+1}|^2\nabla\Phi^{k+1}\right) \\
&\quad -4A(\Phi^k)^3 + 4A\nabla\cdot\left(|\nabla\Phi^k|^2\nabla\Phi^k\right)\Big) + \tau^k,
\end{aligned} \tag{4.79}$$

with $\left\|\tau^k\right\| \le Cs$. Consequently, with an introduction of the error function

$$e^k = \Phi^k - \phi^k, \quad \forall\, k \ge 0, \tag{4.80}$$

we get the following evolutionary equation, by subtracting (4.66) from (4.79):

$$\begin{aligned}
\frac{e^{k+1}-e^k}{s} &= \Delta\Big(3\varepsilon^{-2}\Big((\Phi^{k+1})^4 + (\Phi^{k+1})^3\phi^{k+1} + (\Phi^{k+1})^2(\phi^{k+1})^2 \\
&\quad +\Phi^{k+1}(\phi^{k+1})^3 + (\phi^{k+1})^4\Big)e^{k+1} \\
&\quad -(4\varepsilon^{-2}+\eta+4A)\left((\Phi^k)^2 + \Phi^k\phi^k + (\phi^k)^2\right)e^k \\
&\quad +(\varepsilon^{-2}+\eta)e^{k+1} + \varepsilon^2\Delta^2 e^{k+1} \\
&\quad +(2+\eta\varepsilon^2)\Delta e^k + 6e^{k+1}\left|\nabla\Phi^{k+1}\right|^2 \\
&\quad +6\phi^{k+1}\left(\nabla(\Phi^{k+1}+\phi^{k+1})\cdot\nabla e^{k+1}\right) \\
&\quad -6\nabla\cdot\left((\Phi^{k+1}+\phi^{k+1})e^{k+1}\nabla\Phi^{k+1} + (\phi^{k+1})^2\nabla e^{k+1}\right) \\
&\quad +4A(\left((\Phi^{k+1})^2 + \Phi^{k+1}\phi^{k+1} + (\phi^{k+1})^2\right)e^{k+1} \\
&\quad -4A\nabla\cdot\left((\nabla(\Phi^{k+1}+\phi^{k+1})\cdot\nabla e^{k+1})\nabla\Phi^{k+1} + |\nabla\phi^{k+1}|^2\nabla e^{k+1}\right) \\
&\quad +4A\nabla\cdot\left((\nabla(\Phi^k+\phi^k)\cdot\nabla e^k)\nabla\Phi^k + |\nabla\phi^k|^2\nabla e^k\right)\Big) + \tau^k.
\end{aligned} \tag{4.81}$$

In addition, from the PDE analysis for the FCH equation and the global in time $H^2_{\rm per}$ stability (4.76) for the numerical solution, we also get the L^∞, $W^{1,6}$ and $H^2_{\rm per}$ bounds for both the exact solution and numerical solution, uniform in time:

$$\|\Phi^k\|_{L^\infty},\ \|\Phi^k\|_{W^{1,6}},\ \|\Phi^k\|_{H^2_{\rm per}} \le C_7, \quad \|\phi^k\|_{L^\infty},\ \|\phi^k\|_{W^{1,6}},\ \|\phi^k\|_{H^2_{\rm per}} \le C_7, \quad \forall\, k \ge 0, \tag{4.82}$$

where the 3-D embeddings of $H^2_{\rm per}$ into L^∞ and into $W^{1,6}$ have been applied. Also note that C_7 and C_8 are time independent constants, that depend on ε as $O(\varepsilon^{-4})$.

Stability and convergence analysis

First, we recall that the exact solution to the FCH equation (4.54) is mass conservative:

$$\int_\Omega \Phi(\cdot,t)\,d\mathbf{x} \equiv \int_\Omega \Phi(\cdot,0)\,d\mathbf{x}, \quad \forall t > 0.$$

On the other hand, the numerical solution (4.65) is also mass conservative. In turn, we conclude that the numerical error function $e^k \in \mathring{H}^2_{\rm per}(\Omega)$:

$$\overline{e^k} := \int_\Omega e^k\,d\mathbf{x} = \int_\Omega e^0 = 0, \quad \text{since } e^0 \equiv 0.$$

Consequently, we define $\psi^k := (-\Delta)^{-1} e^k \in \mathring{H}^{-1}_{\rm per}(\Omega)$ as

$$-\Delta\psi^k = e^k, \quad \text{with} \int_\Omega \psi^k\,d\mathbf{x} = 0.$$

Define $I_i, i = 1, \cdots, 10$ by

$$\begin{aligned}
I_1 :&= -6\varepsilon^{-2} s \int_\Omega \Big((\Phi^{k+1})^4 + (\Phi^{k+1})^3 \phi^{k+1} + (\Phi^{k+1})^2 (\phi^{k+1})^2 \\
&\quad + \Phi^{k+1} (\phi^{k+1})^3 + (\phi^{k+1})^4 \Big) \left| e^{k+1} \right|^2 d\mathbf{x}, \\
I_2 :&= -8As \int_\Omega \left((\Phi^{k+1})^2 + \Phi^{k+1}\phi^{k+1} + (\phi^{k+1})^2 \right) \left| e^{k+1} \right|^2 d\mathbf{x}, \\
I_3 :&= 2(2+\eta\varepsilon^2) s (\nabla e^k, \nabla e^{k+1}),
\end{aligned}$$

$$
\begin{aligned}
I_4: &= 2(4\varepsilon^{-2}+\eta+4A)s\int_\Omega \left((\Phi^k)^2+\Phi^k\phi^k+(\phi^k)^2\right)e^k e^{k+1}d\mathbf{x},\\
I_5: &= -12s\int_\Omega |\nabla\Phi^{k+1}|^2(e^{k+1})^2 d\mathbf{x},\\
I_6: &= -12s\int_\Omega \phi^{k+1}\left(\nabla(\Phi^{k+1}+\phi^{k+1})\cdot\nabla e^{k+1}\right)e^{k+1}d\mathbf{x},\\
I_7: &= -12s\left((\Phi^{k+1}+\phi^{k+1})e^{k+1}\nabla\Phi^{k+1}+(\phi^{k+1})^2\nabla e^{k+1},\nabla e^{k+1}\right),\\
I_8: &= -8As\left((\nabla(\Phi^{k+1}+\phi^{k+1})\cdot\nabla e^{k+1})\nabla\Phi^{k+1}+|\nabla\phi^{k+1}|^2\nabla e^{k+1},\nabla e^{k+1}\right),\\
I_9: &= 8As\left((\nabla(\Phi^{k}+\phi^{k})\cdot\nabla e^{k})\nabla\Phi^{k}+|\nabla\phi^{k}|^2\nabla e^{k},\nabla e^{k+1}\right),\\
I_{10}: &= -2s(\tau^k,e^{k+1}).
\end{aligned}
$$

Therefore, taking an L^2 inner product with the numerical error equation (4.81) by $2\psi^k$ gives

$$
\begin{aligned}
\|e^{k+1}\|^2_{\mathring{H}^{-1}_{\rm per}} &- \|e^k\|^2_{\mathring{H}^{-1}_{\rm per}}+\|e^{k+1}-e^k\|^2_{\mathring{H}^{-1}_{\rm per}}\\
&+ 2(\varepsilon^{-2}+\eta)s\|e^{k+1}\|^2+2\varepsilon^2 s\|\Delta e^{k+1}\|^2=\sum_{i=1}^{10} I_i,
\end{aligned}
\tag{4.83}
$$

where integration-by-parts has been repeatedly applied.

The local truncation error term I_{10} can be bounded by the Cauchy inequality:

$$
-2(\tau^k,e^{k+1})\le 2\|\tau^k\|\cdot\|e^{k+1}\|\le\|\tau^k\|^2+\|e^{k+1}\|^2. \tag{4.84}
$$

Meanwhile, an application of weighted Sobolev inequality shows that

$$
\|e^{k+1}\|\le C_8\|e^{k+1}\|^{2/3}_{\mathring{H}^{-1}_{\rm per}}\cdot\|e^{k+1}\|^{1/3}_{\mathring{H}^{2}_{\rm per}}\le C_9\|e^{k+1}\|^{2/3}_{\mathring{H}^{-1}_{\rm per}}\cdot\|\Delta e^{k+1}\|^{1/3}, \tag{4.85}
$$

where a standard estimate of elliptic regularity was applied at the second step, considering the fact that $\overline{e^{k+1}}=0$. Subsequently, an application of Young's inequality gives

$$
\|e^{k+1}\|^2\le C_{10}\varepsilon^{-1}\|e^{k+1}\|^2_{\mathring{H}^{-1}_{\rm per}}+\frac{\varepsilon^2}{8}\|\Delta e^{k+1}\|^2,
$$

and its combination with (4.84) yields

$$-2(\tau^k, e^{k+1}) \le \|\tau^k\|^2 + C_{10}\varepsilon^{-1}\|e^{k+1}\|^2_{\mathring{H}^{-1}_{\rm per}} + \frac{\varepsilon^2}{8}\|\Delta e^{k+1}\|^2. \tag{4.86}$$

The first integral term I_1 turns out to be non-positive,

$$I_1 \le 0, \tag{4.87}$$

due to the fact that

$$(\Phi^{k+1})^4 + (\Phi^{k+1})^3\phi^{k+1} + (\Phi^{k+1})^2(\phi^{k+1})^2 + \Phi^{k+1}(\phi^{k+1})^3 + (\phi^{k+1})^4 \ge 0.$$

Since $(\Phi^{k+1})^2 + \Phi^{k+1}\phi^{k+1} + (\phi^{k+1})^2 \ge 0$, similar estimates can be derived for I_2 and I_5:

$$I_2 = -8As\int_\Omega \left((\Phi^{k+1})^2 + \Phi^{k+1}\phi^{k+1} + (\phi^{k+1})^2\right)\left|e^{k+1}\right|^2 d\mathbf{x} \le 0, \tag{4.88}$$

$$I_5 = -12s\int_\Omega |\nabla\Phi^{k+1}|^2(e^{k+1})^2 d\mathbf{x} \le 0. \tag{4.89}$$

For the term I_3, we denote $C_{11} = 2 + \eta\varepsilon^2$ and observe that

$$I_3 = 2C_{11}s(\nabla e^k, \nabla e^{k+1}) \le C_{11}s(\|\nabla e^k\|^2 + \|\nabla e^{k+1}\|^2). \tag{4.90}$$

Meanwhile, a similar estimate as (4.85) could be carried out to bound $\|\nabla e^{k+1}\|$:

$$\|\nabla e^{k+1}\| \le C_{12}\|e^{k+1}\|^{1/3}_{\mathring{H}^{-1}_{\rm per}} \cdot \|e^{k+1}\|^{2/3}_{H^2_{\rm per}} \le C_{13}\|e^{k+1}\|^{1/3}_{\mathring{H}^{-1}_{\rm per}} \cdot \|\Delta e^{k+1}\|^{2/3}, \tag{4.91}$$

so that an application of Young's inequality leads to

$$\|\nabla e^{k+1}\|^2 \le C_{14}\varepsilon^{-4}\|e^{k+1}\|^2_{\mathring{H}^{-1}_{\rm per}} + \frac{\varepsilon^2}{8C_{11}}\|\Delta e^{k+1}\|^2. \tag{4.92}$$

The term $\|\nabla e^k\|$ can be bounded in the same fashion:

$$\|\nabla e^k\|^2 \le C_{15}\varepsilon^{-4}\|e^k\|^2_{\mathring{H}^{-1}_{\rm per}} + \frac{\varepsilon^2}{8C_{11}}\|\Delta e^k\|^2. \tag{4.93}$$

Substituting (4.92) and (4.93) into (4.90), we get

$$I_3 \le C_{16}s(\|e^{k+1}\|^2_{\mathring{H}^{-1}_{\rm per}} + \|e^k\|^2_{\mathring{H}^{-1}_{\rm per}}) + \frac{\varepsilon^2}{8}s(\|\Delta e^{k+1}\|^2 + \|\Delta e^k\|^2). \tag{4.94}$$

For the term I_4, we denote $C_{17} = 4\varepsilon^{-2} + \eta + 4A$. By the L^∞ bound in (4.82) for both the exact and numerical solutions, we see that

$$\|(\Phi^k)^2 + \Phi^k\phi^k + (\phi^k)^2\|_{L^\infty} \le 3C_7^2. \tag{4.95}$$

This in turn implies that

$$\begin{aligned} I_4 &\le 2C_{17}s\|(\Phi^k)^2 + \Phi^k\phi^k + (\phi^k)^2\|_{L^\infty}\cdot\|e^k\|\cdot\|e^{k+1}\| \\ &\le 6C_{17}C_7^2 s\|e^k\|\cdot\|e^{k+1}\| \le 3C_{17}C_7^2 s(\|e^k\|^2 + \|e^{k+1}\|^2). \end{aligned} \tag{4.96}$$

Meanwhile, the estimate (4.86) can be performed with alternate coefficients, so that the following inequalities are available:

$$\|e^j\|^2 \le C_{18}\|e^j\|^2_{\mathring{H}^{-1}_{\rm per}} + \frac{\varepsilon^2}{24C_{17}C_7^2}\|\Delta e^j\|^2, \quad \text{for } j = k, k+1. \tag{4.97}$$

Subsequently, its combination with (4.96) yields

$$I_4 \le C_{19}s(\|e^k\|^2_{\mathring{H}^{-1}_{\rm per}} + \|e^{k+1}\|^2_{\mathring{H}^{-1}_{\rm per}}) + \frac{\varepsilon^2}{8}s(\|\Delta e^{k+1}\|^2 + \|\Delta e^k\|^2). \tag{4.98}$$

For the term I_6, we start from an application of Hölder inequality:

$$\begin{aligned} I_6 &= -12s\int_\Omega \phi^{k+1}\left(\nabla(\Phi^{k+1} + \phi^{k+1})\cdot\nabla e^{k+1}\right)e^{k+1}d\mathbf{x} \\ &\le C_{20}s\|\phi^{k+1}\|_{L^\infty}\cdot(\|\nabla\Phi^{k+1}\|_{L^6} + \|\nabla\phi^{k+1}\|_{L^6})\cdot\|\nabla e^{k+1}\|_{L^{3/2}}\cdot\|e^{k+1}\|_{L^6} \end{aligned}$$

$$\leq \quad C_{21}C_7^2 s \cdot \|\nabla e^{k+1}\|_{L^{3/2}} \cdot \|e^{k+1}\|_{L^6}, \tag{4.99}$$

in which the L^∞ and $W^{1,6}$ stability bounds for the exact and numerical solutions were recalled in the second step of (4.82). Moreover, the first term $\|\nabla e^{k+1}\|_{L^{3/2}}$ can be bounded in the following way:

$$\|\nabla e^{k+1}\|_{L^{3/2}} \leq C_{22}\|\nabla e^{k+1}\| \leq C_{23}\|e^{k+1}\|_{\mathring{H}^{-1}_{\rm per}}^{1/3} \cdot \|\Delta e^{k+1}\|^{2/3}, \tag{4.100}$$

with an earlier estimate (4.91) recalled. For the second term $\|e^{k+1}\|_{L^6}$, a 3-D Sobolev embedding could be applied so that

$$\|e^{k+1}\|_{L^6} \leq C_{24}\|\nabla e^{k+1}\| \leq C_{25}\|e^{k+1}\|_{\mathring{H}^{-1}_{\rm per}}^{1/3} \cdot \|\Delta e^{k+1}\|^{2/3}. \tag{4.101}$$

We also note that the zero-mean property for e^{k+1} was used in the first step. Therefore, a combination of (4.99)-(4.101) results in

$$I_6 \leq C_{26}C_7^2 s\|e^{k+1}\|_{\mathring{H}^{-1}_{\rm per}}^{2/3} \cdot \|\Delta e^{k+1}\|^{4/3} \leq C_{27}s\|e^{k+1}\|_{\mathring{H}^{-1}_{\rm per}}^2 + \frac{\varepsilon^2}{8}s\|\Delta e^{k+1}\|^2, \tag{4.102}$$

with the Young's inequality applied in the last step.

For the term I_7, we decompose it into two parts: $I_7 = I_{7,1} + I_{7,2}$, with

$$I_{7,1} \quad = \quad -12s\left((\Phi^{k+1} + \phi^{k+1})e^{k+1}\nabla\Phi^{k+1}, \nabla e^{k+1}\right), \tag{4.103}$$

$$I_{7,2} \quad = \quad -12s\left((\phi^{k+1})^2\nabla e^{k+1}, \nabla e^{k+1}\right). \tag{4.104}$$

It is clear that the second part is always non-positive:

$$I_{7,2} = -12s\int_\Omega (\phi^{k+1})^2|\nabla e^{k+1}|^2 d\mathbf{x} \leq 0. \tag{4.105}$$

For the first part, an application of Hölder inequality shows that

$$I_{7,1} \quad \leq \quad C_{28}s(\|\Phi^{k+1}\|_{L^\infty} + \|\phi^{k+1}\|_{L^\infty}) \cdot \|\nabla\Phi^{k+1}\|_{L^6} \cdot \|\nabla e^{k+1}\|_{L^{3/2}} \cdot \|e^{k+1}\|_{L^6}$$

$$\leq \quad C_{29}C_7^2 s\|\nabla e^{k+1}\|_{L^{3/2}} \cdot \|e^{k+1}\|_{L^6}. \tag{4.106}$$

Again, the L^∞ and $W^{1,6}$ bounds (4.25) for the exact and numerical solutions were recalled in the second step. Furthermore, by repeating the same analyses as (4.100)-(4.101), we are able to arrive at the following estimate, similar to (4.102):

$$I_{7,1} \leq C_{30}C_7^2 s \cdot \|e^{k+1}\|_{\mathring{H}_{\rm per}^{-1}}^{1/3} \cdot \|\Delta e^{k+1}\|^{2/3} \leq C_{31}s\|e^{k+1}\|_{\mathring{H}_{\rm per}^{-1}}^2 + \frac{\varepsilon^2}{8}s\|\Delta e^{k+1}\|^2. \tag{4.107}$$

Consequently, a combination of (4.104)), (4.105) and (4.107) leads to

$$I_7 \leq C_{31}s\|e^{k+1}\|_{\mathring{H}_{\rm per}^{-1}}^2 + \frac{\varepsilon^2}{8}s\|\Delta e^{k+1}\|^2. \tag{4.108}$$

Similarly, the term I_8 is also decomposed into two parts: $I_8 = I_{8,1} + I_{8,2}$, with

$$\begin{aligned} I_{8,1} &= -8As\left((\nabla(\Phi^{k+1}+\phi^{k+1})\cdot\nabla e^{k+1})\nabla\Phi^{k+1}, \nabla e^{k+1}\right), \\ I_{8,2} &= -8As\left(|\nabla\phi^{k+1}|^2\nabla e^{k+1}, \nabla e^{k+1}\right) = -8As\int_\Omega |\nabla e^{k+1}|^4 d\mathbf{x} \leq 0. \end{aligned}$$

For the first part $I_{8,1}$, the following estimate is available, in a similar way as (4.106)-(4.107):

$$\begin{aligned} I_{8,1} &\leq C_{32}s(\|\nabla\Phi^{k+1}\|_{L^6} + \|\nabla\phi^{k+1}\|_{L^6}) \cdot \|\nabla\Phi^{k+1}\|_{L^6} \cdot \|\nabla e^{k+1}\|_{L^6} \cdot \|\nabla e^{k+1}\| \\ &\leq C_{33}C_7^2 s\|\nabla e^{k+1}\|_{L^6} \cdot \|e^{k+1}\| \\ &\leq C_{34}C_7^2 s\|\Delta e^{k+1}\| \cdot \|e^{k+1}\|_{\mathring{H}_{\rm per}^{-1}}^{1/3} \cdot \|\Delta e^{k+1}\|^{2/3} \\ &\leq C_{35}C_7^2 s\|\Delta e^{k+1}\|^{5/3} \cdot \|e^{k+1}\|_{\mathring{H}_{\rm per}^{-1}}^{1/3} \leq C_{36}s\|e^{k+1}\|_{\mathring{H}_{\rm per}^{-1}}^2 + \frac{\varepsilon^2}{8}s\|\Delta e^{k+1}\|^2, \end{aligned}$$

in which the $W^{1,6}$ bound (4.25) for the exact and numerical solutions was recalled in the second step, the 3-D Sobolev embedding from $H_{\rm per}^2$ into $W^{1,6}$ and the estimate (4.91) were used in the third step, and the Young inequality was applied at the last

step. Then we arrive at

$$I_8 = I_{8,1} + I_{8,2} \le I_{8,1} \le C_{36} s \|e^{k+1}\|^2_{\mathring{H}^{-1}_{\rm per}} + \frac{\varepsilon^2}{8} s \|\Delta e^{k+1}\|^2. \tag{4.109}$$

The term I_9 can be handled in the same way as I_8. We begin with a decomposition $I_9 = I_{9,1} + I_{9,2}$, with

$$\begin{aligned} I_{9,1} &= 8As\left((\nabla(\Phi^k + \phi^k)\cdot\nabla e^k)\nabla\Phi^k, \nabla e^{k+1}\right), \\ I_{9,2} &= 8As\left(|\nabla\phi^{k+1}|^2\nabla e^k, \nabla e^{k+1}\right). \end{aligned}$$

The following estimates can be carried out:

$$\begin{aligned} I_{9,1} &\le C_{37} s(\|\nabla\Phi^k\|_{L^6} + \|\nabla\phi^k\|_{L^6})\cdot\|\nabla\Phi^k\|_{L^6}\cdot\|\nabla e^k\|_{L^6}\cdot\|\nabla e^{k+1}\| \\ &\le C_{38} C_7^2 s\|\nabla e^k\|_{L^6}\cdot\|e^{k+1}\| \\ &\le C_{39} C_7^2 s\|\Delta e^k\|\cdot\|e^{k+1}\|^{1/3}_{\mathring{H}^{-1}_{\rm per}}\cdot\|\Delta e^{k+1}\|^{2/3} \\ &\le C_{40} s\|e^{k+1}\|^2_{\mathring{H}^{-1}_{\rm per}} + \frac{\varepsilon^2}{16} s(\|\Delta e^{k+1}\|^2 + \|\Delta e^k\|^2), \\ I_{9,2} &\le C_{41} s\|\nabla\phi^k\|^2_{L^6}\cdot\|\nabla e^k\|_{L^6}\cdot\|\nabla e^{k+1}\| \le C_{42} C_7^2 s\|\nabla e^k\|_{L^6}\cdot\|e^{k+1}\| \\ &\le C_{43} C_7^2 s\|\Delta e^k\|\cdot\|e^{k+1}\|^{1/3}_{\mathring{H}^{-1}_{\rm per}}\cdot\|\Delta e^{k+1}\|^{2/3} \\ &\le C_{44} s\|e^{k+1}\|^2_{\mathring{H}^{-1}_{\rm per}} + \frac{\varepsilon^2}{16} s(\|\Delta e^{k+1}\|^2 + \|\Delta e^k\|^2). \end{aligned}$$

Consequently, we get

$$I_9 = I_{9,1} + I_{9,2} \le C_{45} s\|e^{k+1}\|^2_{\mathring{H}^{-1}_{\rm per}} + \frac{\varepsilon^2}{8} s(\|\Delta e^{k+1}\|^2 + \|\Delta e^k\|^2). \tag{4.110}$$

Finally, a combination of (4.83), (4.86), (4.87), (4.88), (4.89), (4.94), (4.98), (4.102), (4.108), (4.109) and (4.110) yields that

$$\|e^{k+1}\|^2_{\mathring{H}^{-1}_{\rm per}} - \|e^k\|^2_{\mathring{H}^{-1}_{\rm per}} + 2(\varepsilon^{-2} + \eta)s\|e^{k+1}\|^2 + \frac{9}{8}\varepsilon^2 s\|\Delta e^{k+1}\|^2$$

$$\leq \quad C_{46} s(\|e^{k+1}\|^2_{\mathring{H}^{-1}_{\rm per}} + \|e^k\|^2_{\mathring{H}^{-1}_{\rm per}}) + \frac{3}{8}\varepsilon^2 s\|\Delta e^k\|^2 + s\|\tau^k\|^2. \tag{4.111}$$

Subsequently, an application of discrete Gronwall inequality leads to an $\ell^\infty(0,T;\mathring{H}^{-1}_{\rm per})\cap \ell^2(0,T;H^2_{\rm per})$ convergence of the numerical scheme (4.65):

$$\|e^k\|^2_{\mathring{H}^{-1}_{\rm per}} + \frac{3}{4}\varepsilon^2 s\sum_{l=0}^{k}\|\Delta e^l\|^2 \leq Cs^2, \tag{4.112}$$

for any $1 \leq k \leq K$. Note that the constant C depends on the exact solution, the physical parameter ε, and final time T, independent on s. The proof of Theorem 4.4.8 is finished.

Fully discrete finite difference scheme

With the machinery in last subsection, the discrete energy of FCH can be rewritten as:

$$\mathcal{F}_h(\phi) = \mathcal{F}_{c,h}(\phi) - \mathcal{F}_{e,h}(\phi) \tag{4.113}$$

where

$$\mathcal{F}_{c,h}(\phi) \quad = \quad \frac{\varepsilon^{-2}}{2}\|\phi\|_6^6 + \left(\frac{\varepsilon^{-2}}{2} + \frac{\eta}{2}\right)\|\phi\|_2^2 + \frac{\varepsilon^2}{2}\|\Delta_h\phi\|_2^2 + \mathcal{H}_h(\phi), \tag{4.114}$$

$$\mathcal{F}_{e,h}(\phi) \quad = \quad \left(\varepsilon^{-2} + \frac{\eta}{4}\right)\|\phi\|_4^4 + \left(1 + \frac{\eta\varepsilon^2}{2}\right)\|\nabla_h^{\mathsf{v}}\phi\|_2^2 + A\|\phi\|_4^4 + A\|\nabla_h^{\mathsf{v}}\phi\|_4^4, \tag{4.115}$$

and

$$\mathcal{H}_h(\phi) \quad = \quad A\|\phi\|_4^4 + A\|\nabla_h^{\mathsf{v}}\phi\|_4^4 + 3\left(\phi^2, \mathfrak{A}(|\nabla_h^{\mathsf{v}}\phi|^2)\right)_2. \tag{4.116}$$

Proposition 4.4.9. *Suppose $\phi \in \mathcal{C}_{\rm per}$. The first variational derivative of $\mathcal{H}_h(\phi)$ is*

$$\delta\mathcal{H}_h(\phi) \quad = \quad 4A\phi^3 - 4A\left(\mathfrak{d}_x\left([(\mathfrak{D}_x\phi)^2 + (\mathfrak{D}_y\phi)^2]\mathfrak{D}_x\phi\right) + \mathfrak{d}_y\left([(\mathfrak{D}_x\phi)^2 + (\mathfrak{D}_y\phi)^2]\mathfrak{D}_y\phi\right)\right)$$

$$+ 6\phi\mathfrak{A}[(\mathfrak{D}_x\phi)^2 + (\mathfrak{D}_y\phi)^2] - 6\left(\mathfrak{d}_x\left(\mathfrak{a}\left(\phi^2\right)\mathfrak{D}_x\phi\right) + \mathfrak{d}_y\left(\mathfrak{a}\left(\phi^2\right)\mathfrak{D}_y\phi\right)\right).$$

Lemma 4.4.10. *Suppose that $\phi \in \mathcal{C}_{\text{per}}$ and $A \geq 1$ then $\mathcal{H}_h(\phi)$, $\mathcal{F}_{c,h}(\phi)$ and $\mathcal{F}_{e,h}(\phi)$ are strictly convex.*

Proof. The convexity proof of $\mathcal{H}_h(\phi)$ is similar to Lemma 4.4.1. The convexities of $\mathcal{F}_{c,h}(\phi)$ and $\mathcal{F}_{e,h}(\phi)$ follow from the convexity of $\mathcal{H}_h(\phi)$. □

According to Proposition 4.4.9 and some other standard calculations [72], the fully discretized finite difference convex splitting scheme can be rewritten as: given $f, g \in \mathcal{C}_{\text{per}}$, find $\phi^{k+1}, \tilde{\mu}^{k+1} \in \mathcal{C}_{\text{per}}$ such that

$$\phi^{k+1} - s\Delta_h\tilde{\mu}^{k+1} = g, \tag{4.117}$$

where

$$\begin{aligned}
\tilde{\mu}^{k+1} &= \delta_\phi\mathcal{F}_{c,h}(\phi^{k+1}) - \delta_\phi\mathcal{F}_{e,h}(\phi^k) \\
&= 3\varepsilon^{-2}(\phi^{k+1})^5 + 4A(\phi^{k+1})^3 + (\varepsilon^{-2} + \eta)\phi^{k+1} \\
&\quad + 6(\phi^{k+1})^2\mathfrak{A}(|\nabla_h^{\mathsf{v}}\phi^{k+1}|^2) + \varepsilon^2\Delta_h^2\phi^{k+1} \\
&\quad - 6\nabla_h^{\mathsf{v}} \cdot (\mathfrak{a}\big(\big(\phi^{k+1}\big)^2\big)\nabla_h^{\mathsf{v}}\phi^{k+1}) \\
&\quad -4A\nabla_h^{\mathsf{v}} \cdot (|\nabla_h^v\phi^{k+1}|^2\nabla_h^{\mathsf{v}}\phi^{k+1}) - f,
\end{aligned} \tag{4.118}$$

with

$$\begin{aligned}
g := \phi^k,\ f : &= -(4\varepsilon^{-2} + \eta)(\phi^k)^3 + (2 + \eta\varepsilon^2)\Delta_h^{\mathsf{v}}\phi^k - 4A(\phi^k)^3 \\
&\quad +4A\nabla_h^{\mathsf{v}} \cdot (|\nabla_h^v\phi^k|^2\nabla_h^{\mathsf{v}}\phi^k).
\end{aligned} \tag{4.119}$$

This scheme is mass-conservative in the sense that $\phi - g \in \mathring{\mathcal{C}}_{\text{per}}$.

Theorem 4.4.11. *The fully discrete scheme (4.117) – (4.119) is unconditionally discrete energy stable: $\mathcal{F}_h(\phi^{k+1}) \leq \mathcal{F}_h(\phi^k)$.*

Proof. The proof follows from Lemma 4.4.10 and the discrete version of (4.4.3) found in [82]. □

Following similar ideas as in the analyses for the semi-discrete case, we are able to derive the unique solvability, unconditional energy stability and the $\ell^\infty(0,T;H^{-1}) \cap \ell^2(0,T;H^2)$ convergence for the fully discrete scheme (4.117) – (4.119). The detailed proofs are skipped for the sake of brevity and are left to interested readers.

Theorem 4.4.12. *The fully discrete scheme* (4.117) – (4.119) *is uniquely solvable. Let ϕ_e be the exact solution of the FCH equation* (4.54) *with the periodic boundary condition and let ϕ be the numerical solution of* (4.117) – (4.119)*. Then the following convergence result holds as s, h goes to zero:*

$$\|\phi_e(t^k) - \phi^k\|_{-1} + \left(\varepsilon^2 s \sum_{l=0}^{k} \|\Delta_h(\phi_e(t^l) - \phi^l)\|^2\right)^{1/2} \leq C(s + h^2), \tag{4.120}$$

where the constant C depends only on the regularity of the exact solution.

4.4.4 Preconditioned steepest descent (PSD) solver

In this section we describe a preconditioned steepest descent (PSD) algorithm for advancing the convex splitting scheme in time, following the practical and theoretical framework in [34]. The fully discrete scheme (4.117) – (4.119) can be recast as a minimization problem with an energy that involves the $\|\cdot\|_{-1}^2$ norm: For any $\phi \in \mathcal{C}_{\text{per}}$,

$$\begin{aligned} E_h[\phi] \;=\; & \frac{1}{2}\|\phi - g\|_{-1}^2 + \frac{s\varepsilon^{-2}}{2}\|\phi\|_6^6 + \frac{s(\varepsilon^{-2}+\eta)}{2}\|\phi\|_2^2 \\ & + As\|\phi\|_4^4 + As\|\nabla_h^{\mathsf{v}} u\|_4^4 + 3\left(\phi^2, \mathfrak{A}\left(|\nabla_h^{\mathsf{v}}\phi|^2\right)\right)_2 \\ & + \frac{s\varepsilon^2}{2}\|\Delta_h\phi\|_2^2 + s\left(g,\phi\right)_2, \end{aligned} \tag{4.121}$$

which is strictly convex provided that $A \geq 1$. One will observe that the fully discrete scheme (4.117) – (4.119) is the discrete variation of the strictly convex energy (4.121)

set equal to zero. The nonlinear scheme at a fixed time level may be expressed as

$$\mathcal{N}_h[\phi] = f, \tag{4.122}$$

where

$$\begin{aligned}\mathcal{N}_h[\phi] &= -\Delta_h^{-1}(\phi - g) + 3s\varepsilon^{-2}\phi^5 + 4sA\phi^3 + s(\varepsilon^{-2} + \eta)\phi + 6s\phi^2\mathfrak{A}(|\nabla_h^{\mathsf{v}}\phi|^2) \\ &\quad -6s\nabla_h^{\mathsf{v}} \cdot \left(\mathfrak{a}\left(\phi^2\right)\nabla_h^{\mathsf{v}}\phi\right) - 4sA\nabla_h^{\mathsf{v}} \cdot (|\nabla_h^{\mathsf{v}}\phi|^2\nabla_h^{\mathsf{v}}\phi) + s\varepsilon^2\Delta_h^2\phi. \end{aligned} \tag{4.123}$$

The main idea of the PSD solver is to use a linearized version of the nonlinear operator as a pre-conditioner, or in other words, as a metric for choosing the search direction. A linearized version of the nonlinear operator $\mathcal{N}$ is defined as follows: $\mathcal{L}_h : \mathring{\mathcal{C}}_{\text{per}} \to \mathring{\mathcal{C}}_{\text{per}}$,

$$\mathcal{L}_h[\psi] := -\Delta_h^{-1}\psi + s(4\varepsilon^{-2} + \eta + 4A + 6)\psi - s(6 + 4A)\Delta_h\psi + s\varepsilon^2\Delta_h^2\psi.$$

Clearly, this is a positive, symmetric operator, and we use this as a pre-conditioner for the method. Specifically, this "metric" is used to find an appropriate search directtion for our steepest descent solver. Given the current iterate $\phi^n \in \mathcal{C}_{\text{per}}$, we define the following *search direction* problem: find $d^n \in \mathring{\mathcal{C}}_{\text{per}}$ such that

$$\mathcal{L}_h[d^n] = f - \mathcal{N}_h[\phi^n] := r^n,$$

where r^n is the nonlinear residual of the nth iterate ϕ^n. This last equation can be solved efficiently using the Fast Fourier Transform (FFT).

We then define the next iterate as

$$\phi^{n+1} = \phi^n + \overline{\alpha}d^n, \tag{4.124}$$

where $\overline{\alpha} \in \mathbb{R}$ is the unique solution to the steepest descent line minimization problem

$$\overline{\alpha} := \underset{\alpha \in \mathbb{R}}{\operatorname{argmax}}\, E_h[\phi^n + \alpha d^n] = \underset{\alpha \in \mathbb{R}}{\operatorname{argzero}}\, \delta E_h[\phi^n + \alpha d^n](d^n). \tag{4.125}$$

We have the convergence $\phi^n \to \phi^{k+1}$, as $n \to \infty$, where $\mathcal{N}_h[\phi^{k+1}] = f$, *i.e.*, ϕ^{k+1} is the solution of the scheme (4.117) – (4.119) at time level $k+1$.

4.4.5 Numerical results

We perform some numerical experiments with PSD solver to support the theoretical results in this section. The finite difference search direction equations and Poisson equations are solved efficiently using the Fast Fourier Transform (FFT). Though we do not present it here, we also implement the scheme by using pseudo-spectral method; see the related descriptions in [9, 17, 34, 48].

Convergence test

In this numerical experiment, we apply the benchmark problem in [19, 55] to show that our scheme is first order accurate in time. The convergence test is performed with the initial data given by

$$\phi(x, y, 0) = 2e^{\sin(\frac{2\pi x}{L_x}) + \sin(\frac{2\pi y}{L_y}) - 2} + 2.2e^{-\sin(\frac{2\pi x}{L_x}) - \sin(\frac{2\pi y}{L_y}) - 2} - 1. \tag{4.126}$$

We use a quadratic refinement path, *i.e.*, $s = Ch^2$. At the final time $T = 0.32$, we expect the global error to be $\mathcal{O}(s) + \mathcal{O}(h^2) = \mathcal{O}(h^2)$ under either the ℓ^2 or ℓ^∞ norm, as $h, s \to 0$. Since an exact solution is not available, instead of calculating the error at the final time, we compute the Cauchy difference, which is defined as $\delta_\phi := \phi_{h_f} - \mathcal{I}_c^f(\phi_{h_c})$, where $\mathcal{I}_c^f$ is a bilinear interpolation operator. This requires having a relatively coarse solution, parametrized by h_c, and a relatively fine solution, parametrized by h_f, where $h_c = 2h_f$, at the same final time. The Cauchy difference is also expected to be $\mathcal{O}(s) + \mathcal{O}(h^2) = \mathcal{O}(h^2)$, as $h, s \to 0$. The other parameters are

given by $L_x = L_y = 3.2$, $\varepsilon = 0.18$, $A = 1.0$, $\eta = 1.0$, $s = 0.1h^2$. The norms of Cauchy difference, the convergence rates, average iteration number and average CPU time (in seconds) can be found in Table 4.3. The results confirm our expectation for the convergence order and also demonstrate the efficiency of our algorithm. Moreover, the semi-log scale of the residual $\|r^n\|_\infty$ with respect to the PSD iterations can be found in Fig. 4.12, which confirms the expected geometric convergence rate of the PSD solver predicted by the theory in [34].

Table 4.3: Errors, convergence rates, average iteration numbers and average CPU time (in seconds) for each time step. Parameters are given in the text, and the initial data is defined in (4.126). The refinement path is $s = 0.1h^2$.

h_c	h_f	$\|\delta_\phi\|_2$	Rate	$\#_{iter}$	$T_{cpu}(h_f)$
$\frac{3.2}{16}$	$\frac{3.2}{32}$	1.8131×10^{-2}	-	27	0.0136
$\frac{3.2}{32}$	$\frac{3.2}{64}$	4.2725×10^{-3}	2.09	25	0.0493
$\frac{3.2}{64}$	$\frac{3.2}{128}$	7.7211×10^{-4}	2.47	19	0.1534
$\frac{3.2}{128}$	$\frac{3.2}{256}$	1.7075×10^{-4}	2.18	11	0.4809
$\frac{3.2}{256}$	$\frac{3.2}{512}$	4.0134×10^{-5}	2.09	05	2.1579

Long time simulation of benchmark problem

Time snapshots of the benchmark problem in [19, 55] for the long time test can be found in Fig. 4.13. The initial data is defined in (4.126) and the other parameters are given by $L_x = L_y = 6.4$, $\varepsilon = 0.18$, $A = 1.0$, $\eta = 1.0$, $s = 1 \times 10^{-4}$ and $h = 6.4/256$. The numerical results in Fig. 4.13 are consistent with earlier work on this topic in [19, 55].

Spinodal decomposition, energy dissipation and mass conservation

In the second test, we simulate the spinodal decomposition, energy-dissipation and mass-conservation. We start with the following random initial condition:

$$\phi(x, y, 0) = 0.5 + 0.05(2r - 1), \tag{4.127}$$

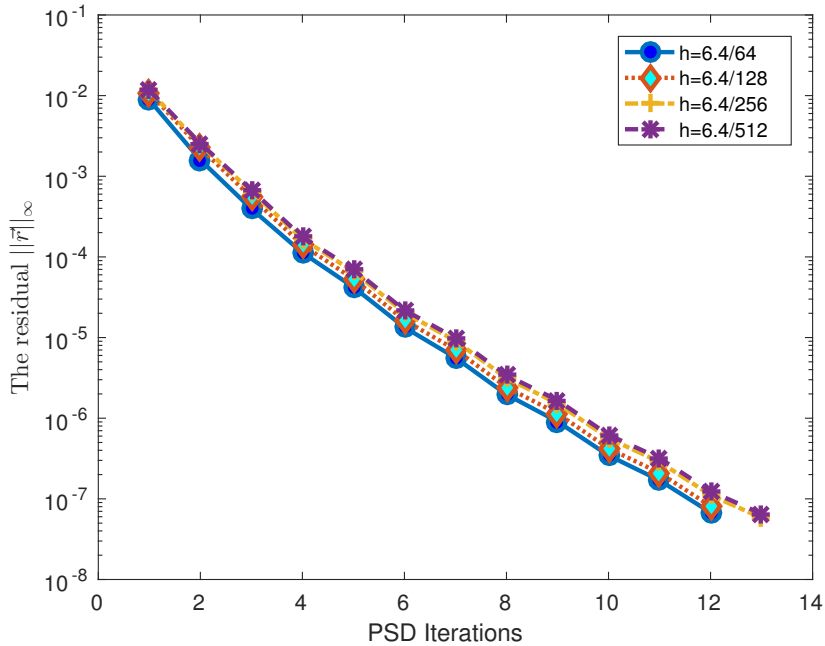


Figure 4.12: Solver convergence (complexity) test for the problem defined in Section 5.2.1. The only difference is that for this test, we use a fixed time step size, $s = 1.0 \times 10^{-5}$ for all runs. We plot on a Semi-log scale of the residual $\|r^n\|_\infty$ with respect to the PSD iteration count n at the 20th time step, *i.e.*, $t = 2.0 \times 10^{-4}$. The initial data is defined in (4.126), $L_x = L_y = 6.4$, $\varepsilon = 0.18$, $A = 1.0$, $\eta = 1.0$, and the grid sizes are as specified in the legend. We observe that the residual is decreasing by a nearly constant factor for each iteration.

where r are the real random numbers in $(0, 1)$. The rest of parameters are given by $L_x = L_y = 12.8$, $\varepsilon = 0.1$, $A = 1.0$, $\eta = 1.0$, $s = 1 \times 10^{-4}$ and $h = 12.8/256$. The snapshots of spinodal decomposition with initial data in (4.127) can be found in Fig. 4.14. This experiment also simulates the amphiphilic di-block copolymer mixtures of polyethylene. The numerical results are consistent with chemical experiments on this topic in [54]. Fig. 4.15 indicates that the simulation has captured all the structural elements with hyperbolic (saddle) surfaces identified in this work, such as short cylinders with one and two beads, cylinder undulation, Y-junction and bilayer-cylinder junction can be found in zoom boxes.

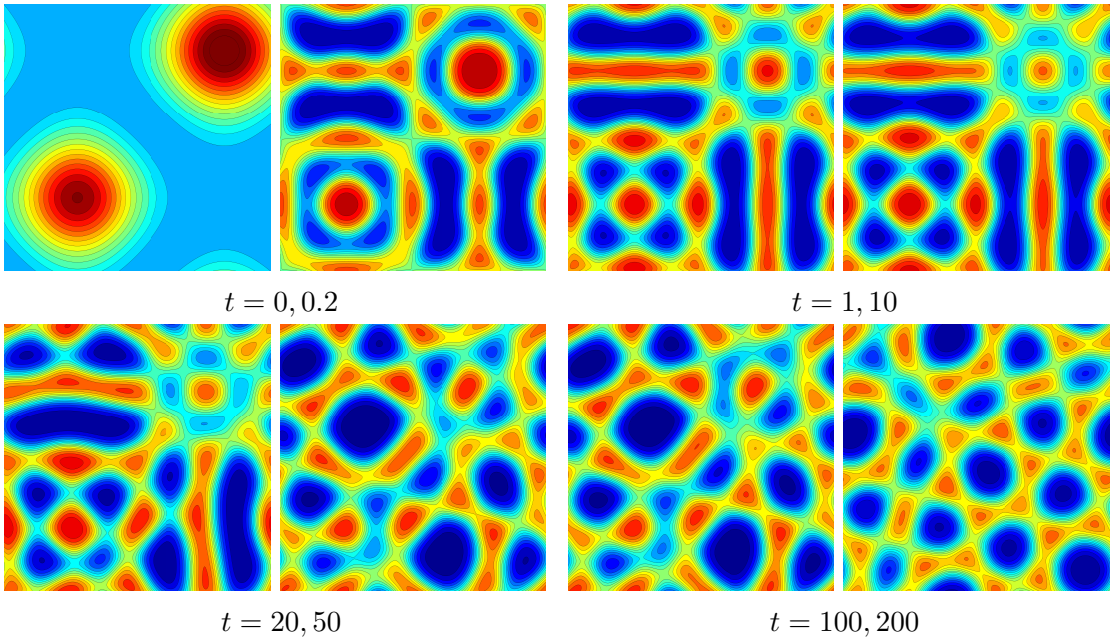


Figure 4.13: Time snapshots of the benchmark problem with initial data in (4.126) at $t = 0, 0.2, 1, 10, 20, 50, 100$ and 200. The parameters are $\varepsilon = 0.18$, $\Omega = (0, 6.4)^2$, $A = 1.0$, $\eta = 1.0$, $s = 1 \times 10^{-4}$ and $h = 6.4/256$. The numerical results are consistent with earlier work on this topic in [19, 55].

The evolutions of discrete energy and mass for the simulation depicted in Fig. 4.14 are presented in Fig. 4.16. The evolution of discrete energy in Fig. 4.16 demonstrates the energy dissipation property, and the evolution of discrete mass clearly indicates the mass conservation property.

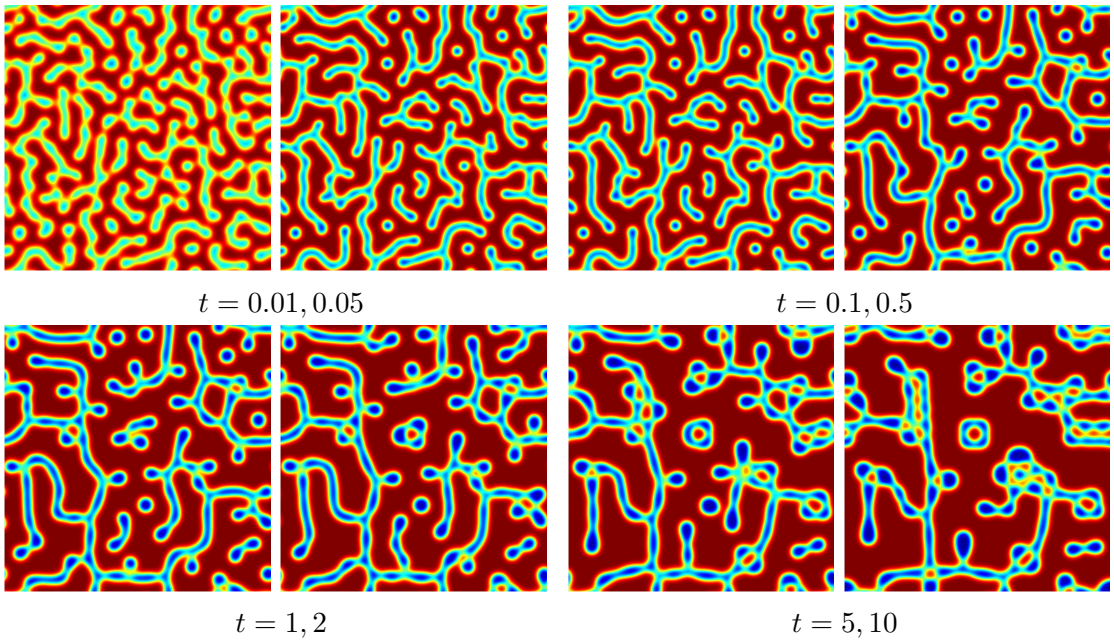


Figure 4.14: Snapshots of spinodal decomposition with initial data in (4.127) at $t = 0.01, 0.05, 0.1, 0.5, 1, 2, 5$ and 10. The parameters are $\varepsilon = 0.1, \Omega = [12.8]^2$, $A = 1.0$, $\eta = 1.0$, $s = 1 \times 10^{-4}$ and $h = 12.8/256$.

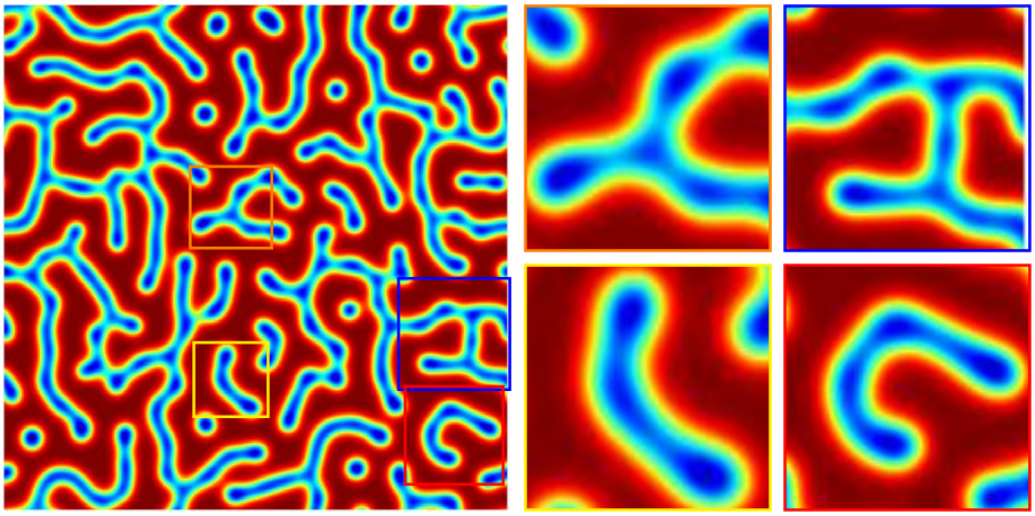


Figure 4.15: Left: Snapshots of spinodal decomposition at $t = 0.05$. Right: Zoom boxes. Yellow box: Short cylinders with an undulation; Red box: Short cylinders with two undulations; Blue box: Bilayer- Cylinder junction; Orange box: Y-junction. Those numerical results are consistent with chemical experiments on this topic in [54].

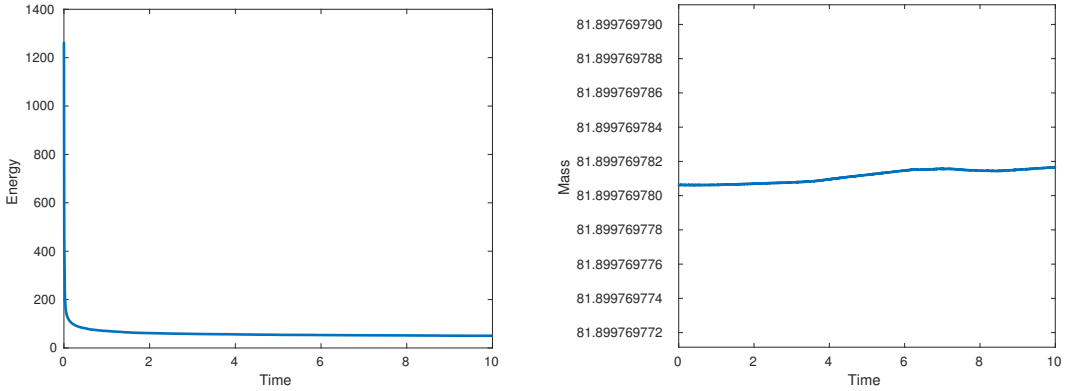


Figure 4.16: The evolutions of discrete energy and mass for the simulation depicted in Fig. 4.14. Left: Energy Dissipation; Right: Mass Conservation.

Chapter 5

Linearly Preconditioned Nonlinear Conjugate Gradient Solvers

The content in this chapter has been published in [35], for more details please refer to [35].

5.1 Linearly Preconditioned Nonlinear Conjugate Gradient Methods

Based on the PSD algorithm, we define $\overline{g}_k = \mathcal{L}_h^{-1}(r^k)$. Then our PNCG algorithms are given by the following equations:

$$\phi^{k+1} = \phi^k + \overline{\alpha}_k d^k \tag{5.1}$$

$$d^{k+1} = -\overline{g}_{k+1} + \overline{\beta}_{k+1} d^k, d^0 = -\overline{g}_0. \tag{5.2}$$

And more details can be found in Algorithm 2.

However, there several different ways to choose the scaling parameter $\overline{\beta}_{k+1}$. And two of the best known formulas for $\overline{\beta}_{k+1}$ are named after their deveiops:

Algorithm 2 Linearly Preconditioned Nonlinear Conjugate Gradient (PNCG) Method

Compute residual: $r^0 := f - \mathcal{N}_h(\phi^0)$
Set $\overline{g}_0 = \mathcal{L}_h^{-1}(r^0)$
Set $d^0 \leftarrow -\overline{g}_0, k \leftarrow 0$
while $\overline{g}_k \neq 0$ **do**
    Compute $\overline{\alpha}_k$ ▷ secant search
    $\phi^{k+1} \leftarrow \phi^k + \overline{\alpha}_k d^k$ ▷ steepest descent algorithm
    $\overline{g}_{k+1} \leftarrow \mathcal{L}_h^{-1}(r^{k+1}) = \mathcal{L}_h^{-1}(f - \mathcal{N}_h(\phi^{k+1}))$
    Compute $\overline{\beta}_{k+1}$
    $d^{k+1} \leftarrow -\overline{g}_{k+1} + \overline{\beta}_{k+1} d^k$
    $k \leftarrow k + 1$
end while

Fletcher-Reeves [40]:

$$\overline{\beta}_{k+1}^{FR} = \frac{\overline{g}_{k+1}^T \overline{g}_{k+1}}{\overline{g}_k^T \overline{g}_k} \tag{5.3}$$

Polak-Ribière [65]:

$$\overline{\beta}_{k+1}^{PR} = \frac{\overline{g}_{k+1}^T (\overline{g}_{k+1} - \overline{g}_k)}{\overline{g}_k^T \overline{g}_k} \tag{5.4}$$

Based on those two best known formulas, we proposed the following two PNCG solvers:

PNCG1:

$$\overline{\beta}_{k+1} = \max\{0, \overline{\beta}_{k+1}^{PR}\} \tag{5.5}$$

PNCG2 :

$$\overline{\beta}_{k+1} = \max\{0, \min\{\overline{\beta}_{k+1}^{FR}, \overline{\beta}_{k+1}^{PR}\}\} \tag{5.6}$$

Remark 5.1.1. *The PNCG2 is also called hybrid conjugate gradient algorithm in [88].*

5.2 Application to Epitaxial Thin Film Growth Model with First-Order-In-Time Scheme

In this section we provide a brief introduction about SS model and recall the first-order-in-time unconditional energy stable numerical schemes in [34, 71].

This epitaxial thin film model was first proposed by P. Aviles and Y. Giga to study the dynamics of smectic liquid crystals in [4]. The energy of the SS model is as follows:

$$F(\phi) := \int_\Omega \left(\frac{1}{4} \left(|\nabla \phi|^2 - 1 \right)^2 + \frac{\varepsilon^2}{2} (\Delta \phi)^2 \right) \mathrm{d}\mathbf{x} \ , \tag{5.7}$$

where $\Omega = (0, L_x) \times (0, L_y)$, $\phi : \Omega \to \mathbb{R}$ is a scaled height function of thin film and ε is a constant which represents the width of the rounded corner. The corresponding chemical potential is defined to be the variational derivative of the energy (5.7), *i.e.*,

$$\mu := \delta_\phi F = -\nabla \cdot (|\nabla \phi|^2 \nabla \phi) + \Delta \phi + \varepsilon^2 \Delta^2 \phi. \tag{5.8}$$

And the SS equation becomes the L^2 gradient flow associated with the energy (5.7):

$$\partial_t \phi = -\mu = \nabla \cdot (|\nabla \phi|^2 \nabla \phi) - \Delta \phi - \varepsilon^2 \Delta^2 \phi. \tag{5.9}$$

Periodic boundary conditions are assumed for ϕ and μ in both spatial directions for simplicity.

With the machinery in Section 2.2, the first-order-in-time unconditional energy stable scheme in [34, 71] can be formulated as follows: for $n \geq 0$, given $\phi^n \in \mathcal{C}_{\text{per}}$, find $\phi^{n+1} \in \mathcal{C}_{\text{per}}$ such that

$$\phi^{n+1} - s \nabla_h \cdot \left(\left| \nabla_h \phi^{n+1} \right|^2 \nabla^{n+1} \phi \right) + s \epsilon^2 \Delta_h^2 \phi^{n+1} \quad = \quad \phi^n - s \Delta_h \phi^n. \tag{5.10}$$

We now define a fully discrete energy that is consistent with the continuous space energy (5.7) as $h \to 0$. In particular, the discrete energy $F_h : \mathcal{C}_{\text{per}} \to \mathbb{R}$ is defined as:

$$F_h(\phi) = \frac{1}{4} \|\nabla_h \phi\|_4^4 - \frac{1}{2} \|\nabla_h \phi\|_2^2 + \frac{1}{2}\varepsilon^2 \|\Delta_h \phi\|_2^2 . \tag{5.11}$$

Theorem 5.2.1. *The numerical scheme (5.10) is unconditionally energy stable, i.e. the discrete energy F_h satisfies the following energy dissipation law:*

$$F_h(\phi^{n+1}) - F_h(\phi^n) \leq -s \left\|\mu^{n+1}\right\|_2^2 . \tag{5.12}$$

Proof. The numerical scheme (5.10) can be rewritten as a nonlinear system:

$$\begin{aligned} \phi^{n+1} - \phi^n &= -s\mu^{n+1} && (5.13) \\ \mu^{n+1} &= -\nabla_h \cdot (|\nabla_h \phi^{n+1}|^2 \nabla_h \phi^{n+1}) + \Delta_h \phi^n + \varepsilon^2 \Delta_h^2 \phi^{n+1}. && (5.14) \end{aligned}$$

By taking the L^2 inner product of (5.13) with μ^{n+1}, we obtain

$$-s \left\|\mu^{n+1}\right\|_2^2 = \left(\mu^{n+1}, \phi^{n+1} - \phi^n\right) . \tag{5.15}$$

By taking the L^2 inner product of (5.14) with $\phi^{n+1} - \phi^n$ yields

$$\begin{aligned} \left(\mu^{n+1}, \phi^{n+1} - \phi^n\right) &= -\left(\nabla_h \cdot (|\nabla_h \phi^{n+1}|^2 \nabla_h \phi^{n+1}), \phi^{n+1} - \phi^n\right) \\ &\quad + \left(\Delta_h \phi^n, \phi^{n+1} - \phi^n\right) + \varepsilon^2 \left(\Delta_h^2 \phi^{n+1}, \phi^{n+1} - \phi^n\right) \\ &=: I_1 + I_2 + I_3. \end{aligned} \tag{5.16}$$

For the term I_1 which involves with 4-Laplacian term, we have

$$\begin{aligned} -\left(\nabla_h \cdot (|\nabla_h \phi^{n+1}|^2 \nabla_h \phi^{n+1}), \phi^{n+1} - \phi^n\right) &= \left(|\nabla_h \phi^{n+1}|^2 \nabla_h \phi^{n+1}, \nabla_h(\phi^{n+1} - \phi^n)\right) \\ &\geq \frac{1}{4}\left(\|\nabla_h \phi^{n+1}\|_4^4 - \|\nabla_h \phi^n\|_4^4\right) . \end{aligned} \tag{5.17}$$

For the explicit linear term I_2, we have

$$\begin{aligned}\left(\Delta_h\phi^n, \phi^{n+1}-\phi^n\right) &= -\left(\nabla_h\phi^n, \nabla_h(\phi^{n+1}-\phi^n)\right)\\ &= -\frac{1}{2}\left\|\nabla_h\phi^{n+1}\right\|_2^2 + \frac{1}{2}\left\|\nabla_h\phi^n\right\|_2^2 + \frac{1}{2}\left\|\nabla_h(\phi^{n+1}-\phi^n)\right\|_2^2. \quad (5.18)\end{aligned}$$

For the highest-order diffusion term I_3, we have

$$\begin{aligned}\varepsilon^2\left(\Delta_h^2\phi^{n+1}, \phi^{n+1}-\phi^n\right) &= \varepsilon^2\left(\Delta_h\phi^{n+1}, \Delta_h(\phi^{n+1}-\phi^n)\right)\\ &= \frac{1}{2}\varepsilon^2\left\|\Delta_h\phi^{n+1}\right\|_2^2 - \frac{1}{2}\varepsilon^2\left\|\Delta_h\phi^n\right\|_2^2\\ &\quad + \frac{1}{2}\varepsilon^2\left\|\Delta_h(\phi^{n+1}-\phi^n)\right\|_2^2. \quad (5.19)\end{aligned}$$

A combination of (5.15), (5.17)-(5.19) yields

$$F_h(\phi^{n+1}) - F_h(\phi^n) + \frac{1}{2}\left\|\nabla_h(\phi^{n+1}-\phi^n)\right\|_2^2 + \frac{1}{2}\varepsilon^2\left\|\Delta_h(\phi^{n+1}-\phi^n)\right\|_2^2 \le -s\left\|\mu^{n+1}\right\|_2^2. \quad (5.20)$$

The desired result (5.12) follows from dropping some positive difference terms from (5.20). □

5.2.1 Numerical Experiments

In this section we demonstrate the accuracy, complexity and efficiency of the PNCG solvers. We present the results of the convergence tests and perform some sample computations to demonstrate the complexity and the efficiency of PNCG solvers. Moreover, We also provide some comparison results between the proposed PNCG solvers and the PSD solver in [34]. The stop tolerances for all of the following simulations are 1.0×10^{-10}.

Convergence tests and the complexity of the PNCG solvers

To simultaneously demonstrate the spatial accuracy and the efficiency of the solver, we perform a typical time-space convergence test for the fully discrete scheme (5.10) for the SS model. As in [34, 71, 77], we perform the Cauchy-type convergence test

using the following periodic initial data [71]:

$$
\begin{aligned}
u(x,y,0) &= 0.1\sin^2\left(\frac{2\pi x}{L}\right)\cdot\sin\left(\frac{4\pi(y-1.4)}{L}\right) \\
&\quad -0.1\cos\left(\frac{2\pi(x-2.0)}{L}\right)\cdot\sin\left(\frac{2\pi y}{L}\right),
\end{aligned}
\tag{5.21}
$$

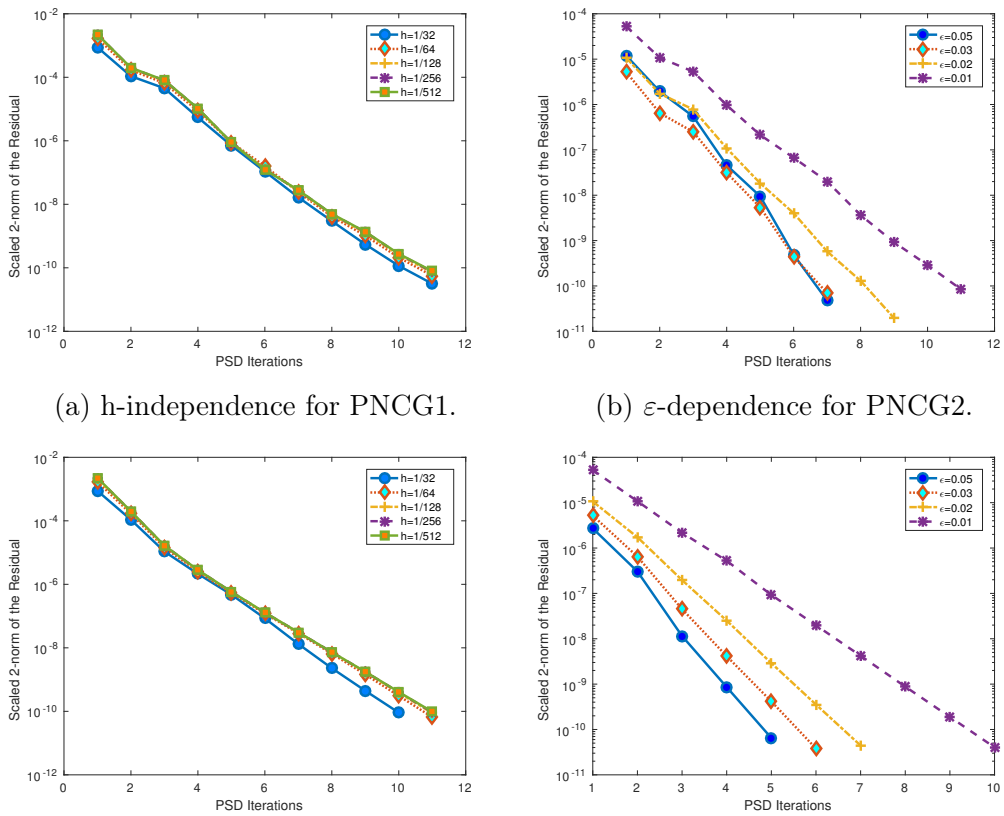
with $\Omega = [0, 3.2]^2$, $\varepsilon = 1 \times 10^{-1}$, $s = 0.1h^2$ and $T = 0.32$. We use a quadratic refinement path, *i.e.*, $s = Ch^2$. At the final time $T = 0.32$, we expect the global error to be $\mathcal{O}(s) + \mathcal{O}(h^2) = \mathcal{O}(h^2)$ under either the ℓ^2 or ℓ^∞ norm, as $h, s \to 0$. The Cauchy difference is defined as $\delta_\phi := \phi_{h_f} - \mathcal{I}_c^f(\phi_{h_c})$, where $\mathcal{I}_c^f$ is a bilinear interpolation operator (We applied Nearest Neighbor Interpolation in Matlab, which is similar to the 2D case in [31, 34] the 3D case in [27]). This requires having a relatively coarse solution, parametrized by h_c, and a relatively fine solution, parametrized by h_f, in particular $h_c = 2h_f$, at the same final time T. The ℓ^2 norms of Cauchy difference and the convergence rates can be found in Table 5.1 which confirms our expectation for the first order in time and second order in space convergence.

Table 5.1: Errors, convergence rates, average iteration numbers and average CPU time (in seconds) for each time step. Parameters are given in the text, and the initial data is defined in (5.21). The refinement path is $s = 0.01h$.

		PNCG1				PNCG2			
h_c	h_f	$\lVert\delta_\phi\rVert_2$	Rate	$\#_{iter}$	$T_{cpu}(h_f)$	$\lVert\delta_\phi\rVert_2$	Rate	$\#_{iter}$	$T_{cpu}(h_f)$
$\frac{3.2}{16}$	$\frac{3.2}{32}$	5.9944×10^{-3}	-	4	0.0007	5.9944×10^{-3}	-	4	0.0007
$\frac{3.2}{32}$	$\frac{3.2}{64}$	1.1500×10^{-3}	2.38	2	0.0026	1.1500×10^{-3}	2.38	2	0.0025
$\frac{3.2}{64}$	$\frac{3.2}{128}$	2.4689×10^{-4}	2.22	2	0.0220	2.4688×10^{-4}	2.22	2	0.0129
$\frac{3.2}{128}$	$\frac{3.2}{256}$	5.8656×10^{-5}	2.07	2	0.0674	5.8656×10^{-5}	2.07	2	0.0653
$\frac{3.2}{256}$	$\frac{3.2}{512}$	1.4463×10^{-5}	2.02	2	0.4542	1.4463×10^{-5}	2.02	2	0.4533

In the second part of this test, we demonstrate the complexity of the PNCG solvers with initial data (5.21). In Figure 5.1 (a) and (c), we plot the semi-log scale of the relative residuals versus PNCG1 and PNCG2 iteration numbers for various values of h, respectively. The other common parameters for the h-independence are $\Omega = [0, 3.2]^2$, $\varepsilon = 1 \times 10^{-1}$, $s = 0.001$, $T = 1 \times 10^{-2}$ with time steps $s = 1 \times 10^{-3}$ and the initial data (5.21). And the semi-log scale of the relative residuals versus PNCG1

and PNCG2 iteration numbers for various values of ε can be found in Figure 5.1 (b) and (d), respectively. The other common parameters for the ε-dependence are $\Omega = [0, 3.2]^2$, $h = {}^{3.2}/_{512}$, $s = 0.001$, $T = 1 \times 10^{-2}$ with time steps $s = 1 \times 10^{-3}$ and the initial data (5.21). Figure 5.1(a) and (c) indicate that the convergence rates of PNCG solvers (as gleaned from the error reduction) are nearly uniform and nearly independent of h for a fixed ε. Figure 5.1(b) and (d) show that the number of PNCG iterations increases with the decreasing of ε. Moreover, Figure 5.1(b) and (d) indicate that the PNCG2 is more efficient and robust than PNCG1. Figure 5.1 provides the similar geometric convergence rate of the PNCG solvers predicted by the theory in [34].



(a) h-independence for PNCG1.

(b) ε-dependence for PNCG2.

(c) h-independence for PNCG1.

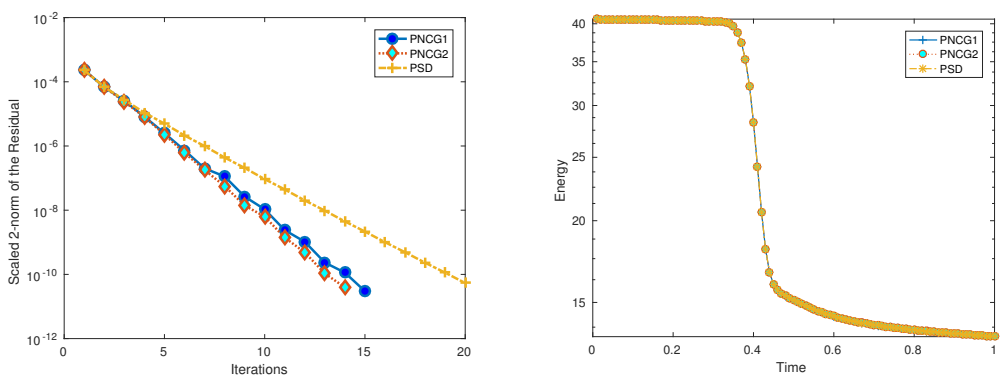
(d) ε-dependence for PNCG2.

Figure 5.1: Complexity tests showing the solver performance for changing values of h and ε. Parameters are given in the text, and the initial data is defined in (5.21).

In the third part of this test, we perform some comparisons between the proposed PNCG solvers and the PSD solver. The parameters for the comparison simulations are $\Omega = [0, 12.8]^2$, $\varepsilon = 3\times10^{-2}$, $h = {}^{12.8}/_{512}$, $s = 0.01$, and $T = 0.32$. The average iteration numbers, total CPU time (in seconds) and speedups for the preconditioned methods can be found in Table 5.2. The Table 5.2 indicates that the PNCG1 solver and PNCG2 solver have provided a 1.34x and 1.39x speedup over PSD solver, respectively. The error reductions at $T = 1$ (100th iteration) in Figure 5.2 (a) indicates the PNCG solvers have less iteration numbers and faster error reduction at each time iteration. And energy evolutions in Figure 5.2 (a) show that the preconditioned solvers have the same energy evolutions.

Table 5.2: The average iteration numbers and total CPU time (in seconds) for the preconditioned methods with fixed time steps $s = 0.01$ and initial data (5.21). Parameters are given in the text.

Methods	PSD	PNCG1	PNCG2
$\#_{iter}$	17	13	12
T_{cpu}(s)	232.2665	173.2407	167.6591
Speedup	-	1.34	1.39



(a) The error reduction at $T = 1$.

(b) Energy evolutions.

Figure 5.2: The error reductions at $T = 1$ and energy evolutions for the preconditioned solvers. Parameters are given in the text, and the initial data is defined in (5.21).

Long-time coarsening process, energy dissipation and mass conservation

Coarsening processes in thin film systems can take place on very long time scales [57]. In this subsection, we perform long time behavior tests for SS model. Such test, which have been performed in many places, will confirm the expected coarsening rates and serve as benchmarks for our solver. See, for example, [71, 77]. The initial data for the simulations are taken as essentially random:

$$u_{i,j}^0 = 0.05 \cdot (2r_{i,j} - 1), \tag{5.22}$$

where the $r_{i,j}$ are uniformly distributed random numbers in [0, 1]. Since all of the solvers give similar results, we only present the results from PNCG2 in the fellowing content of this subsection. Time snapshots of the evolution for the epitaxial thin film growth model can be found in Figure 5.3. The coarsening rates are given in Figure 5.4. The interface width or roughness is defined as

$$W(t_n) = \sqrt{\frac{h^2}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} (\phi_{i,j}^n - \bar{\phi})^2}, \tag{5.23}$$

where m and n are the number of the grid points in x and y direction and $\bar{\phi}$ is the average value of ϕ on the uniform grid. The log-log plots of roughness and energy evolution and the corresponding linear regression are presented in Figure. 5.4. The linear regressions in Figure. 5.4 indicate that the surface roughness grows like $t^{1/3}$ and the energy decays like $t^{1/3}$. In particular, the linear fits have the form $a_e t^{b_e}$ with $a_e = 11.2, b_e = -0.3315$ for energy evolution and $a_r t^{b_r}$ with $a_r = 0.0025, b_r = 0.3255$ for roughness evolution. Those decay properties confirm the one-third power law in [58]. Moreover, these simulation results are consistent with earlier work on this topic in [34, 71, 77, 83].

The PNCG2 iterations and mass difference at each time steps for the simulation depicted in Figure 5.3 are presented in 5.5. The PNCG2 iterations indicate that the

average iteration number of PNCG2 solver is only 3. And the mass difference at each time steps clearly shows the mass Conservative property.

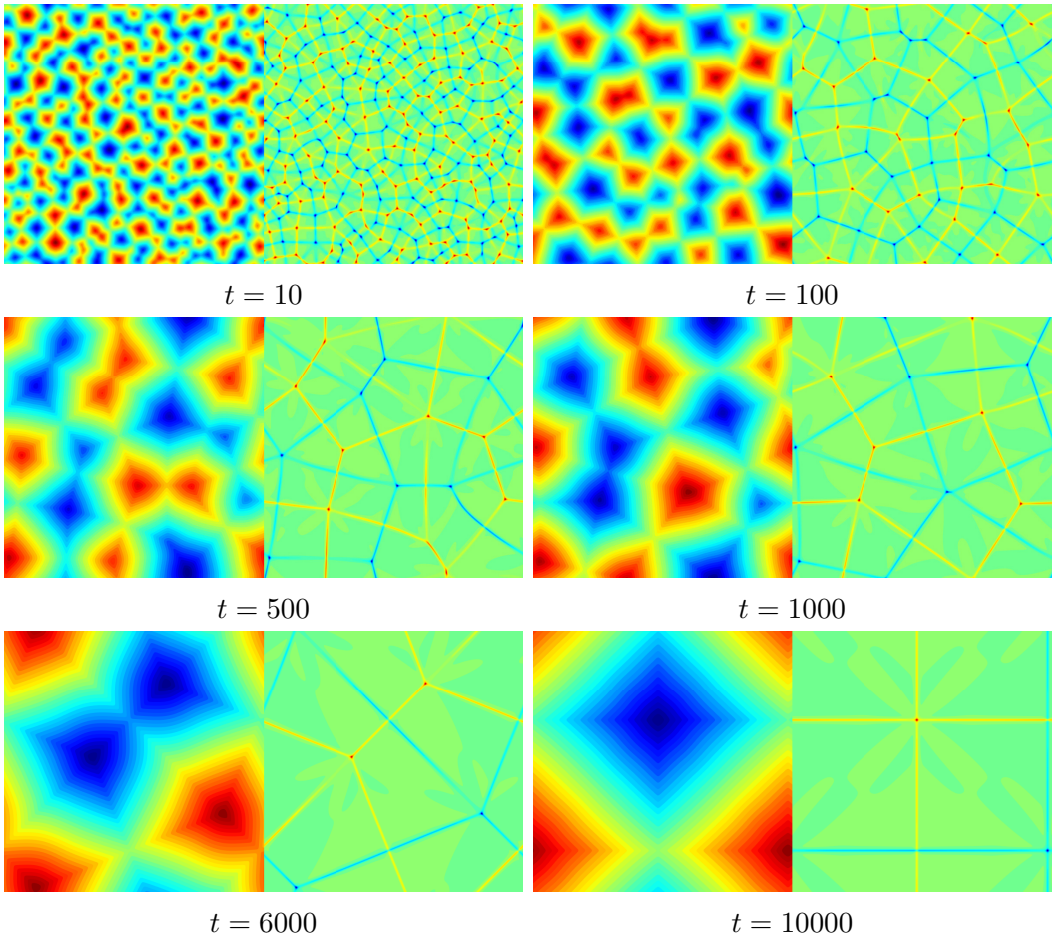
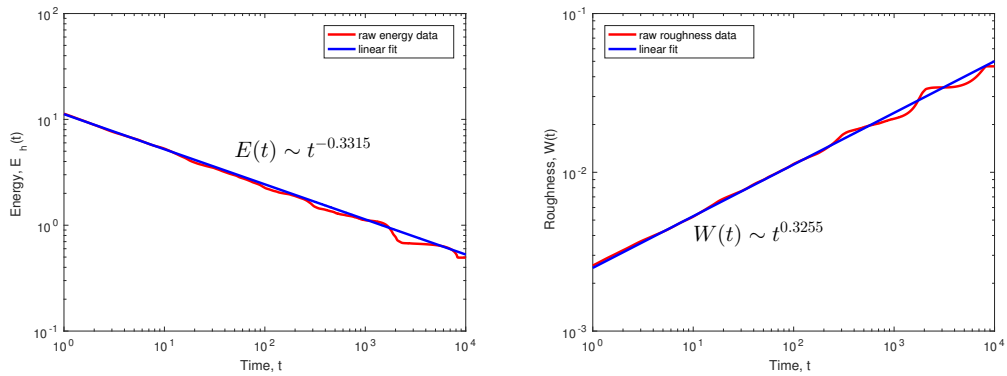
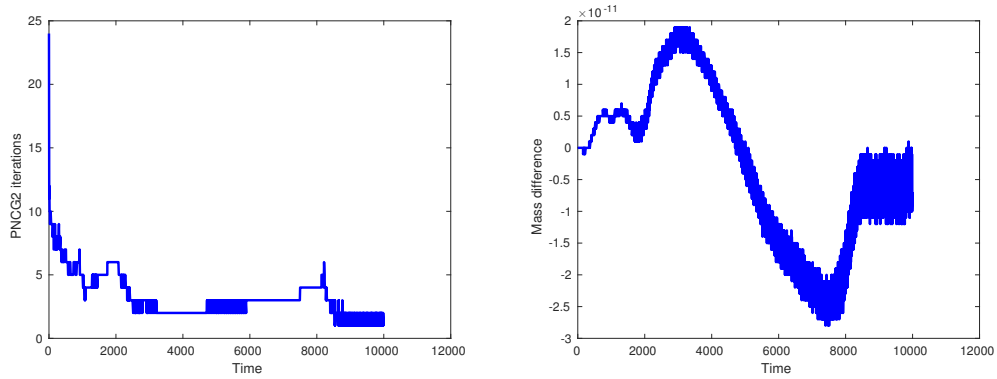


Figure 5.3: Time snapshots of the evolution with PSD solver for the epitaxial thin film growth model at $t = 10, 100, 500, 1000, 6000$ and 10000. Left: contour plot of u, Right: contour plot of Δu. The parameters are $\varepsilon = 0.03, \Omega = [12.8]^2$ and $s = 0.01$. These simulation results are consistent with earlier work on this topic in [34, 36, 71, 77, 83].



(a) Energy evolution (b) Roughness evolution

Figure 5.4: The log-log plots of energy and roughness evolution and the corresponding linear regression for the simulation depicted in Figure 5.3.



(a) PNCG2 iterations (b) Mass difference

Figure 5.5: PNCG2 iterations and mass difference at each time steps for the simulation depicted in Figure 5.3.

5.3 Adaptive time-stepping method for the First-Order-In-Time Scheme

As we have proven in section 5.2, the proposed scheme is unconditionally energy stable which allows us to adopt large time steps to the simulations. However, for the sake of accuracy, large time steps are not proper for a rapidly phase transformation. In order to make the proposed scheme much more practical, we apply the adaptive time stepping strategy in [69, 85, 86] based on the change ratio of the free energy. The adaptive time steps is defined as follows:

$$s = \max\left(s_{\min}, \frac{s_{\max}}{\sqrt{1+\alpha|\delta_t F_h(\phi)|^2}}\right), \tag{5.24}$$

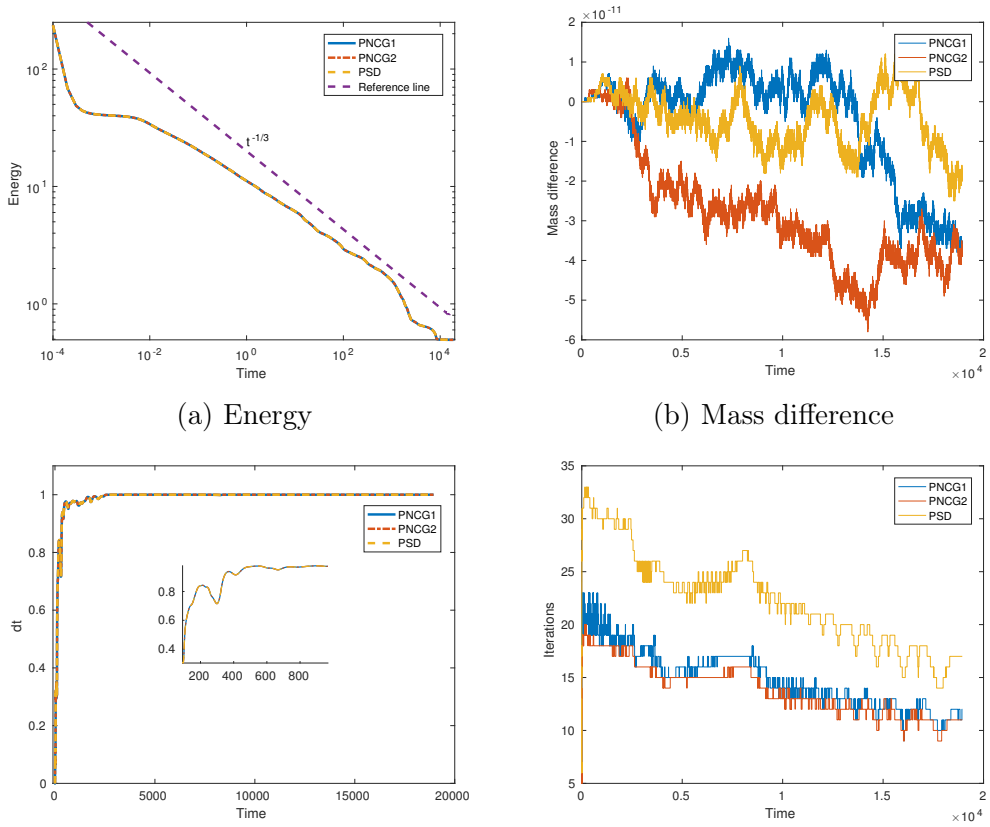
where α is a constant, δ_t is the derivative w.r.t. time and $s_{\min}$ and $s_{\max}$ are pre-set lower and upper bound of the time steps, respectively. By introducing the pre-set time steps, the $s_{\min}$ can force the adaptive time steps bounded from below to avoid too small time steps and the $s_{\max}$ gives the upper bound of the time steps to guarantee the accuracy. Consequently,

$$s_{\min} \leq s \leq s_{\max}.$$

5.3.1 Numerical Experiments

In order to make the proposed scheme much more practical, we investigate the adaptive time stepping strategy in this subsection. The Figure 5.24 (a) and (c) indicate that the energy decays and adaptive time steps are the same, which demonstrate the robustness of the proposed preconditioned methods. Moreover, according to the $t^{-1/3}$ reference line in Figure 5.24 (a), we can conclude that the energy E_h decays like $t^{-1/3}$. From the Figure 5.24 (b), we can observe that the masses are conservative in sense of 10^{-11}. The adaptive time steps and the number of iterations at each time steps are presented in Figure 5.24 (c) and (d), respective.

From Figure 5.24 (d), we can clearly see that the PNCG solvers have less iteration numbers than the PSD solver. Moreover, the total CPU time in Table 5.3 shows that the adaptive time stepping approach can greatly save CPU time without losing accuracy.



(a) Energy

(b) Mass difference

(c) Adaptive time steps

(d) Number of iterations

Figure 5.6: Adaptive time-stepping methods with adaptive time steps (5.24). The rest of the parameters are $\Omega = [0, 12.8]^2$, $\varepsilon = 3.0 \times 10^{-2}$, $h = {}^{12.8}/_{512}$, $smin = 1.0^{-4}$, $smax = 1.0$.

Table 5.3: The total CPU time (in seconds) and speedups for the preconditioned methods with adaptive time steps for the simulation depicted in Figure 5.6. Parameters are given in the text.

Methods	PSD	PNCG1	PNCG2
T_{cpu}(s)	106968.8356	75029.2439	70657.7581
Speedup	-	1.43	1.51

5.4 Application to Epitaxial Thin Film Growth Model with Second-Order-In-Time Backward Differentiation Formula Scheme

The content in this chapter has been published in [36], for more details please refer to [36].

5.4.1 Convergence test and the complexity of the Preconditioned solvers

In this subsection we demonstrate the accuracy and complexity of the preconditioned solvers. We present the results of the convergence test and perform some sample computations to investigate the effect of the time step s and stabilized parameter A for the energy $F_h(\phi)$.

To simultaneously demonstrate the spatial accuracy and the efficiency of the solver, we perform a typical time-space convergence test for the fully discrete scheme (4.6) for the slope selection model. As in [14, 71, 77], we perform the Cauchy-type convergence test using the following periodic initial data [71]:

$$\begin{aligned} u(x,y,0) &= 0.1\sin^2\left(\frac{2\pi x}{L}\right)\cdot\sin\left(\frac{4\pi(y-1.4)}{L}\right) \\ &\quad -0.1\cos\left(\frac{2\pi(x-2.0)}{L}\right)\cdot\sin\left(\frac{2\pi y}{L}\right), \end{aligned} \tag{5.25}$$

with $\Omega = [0, 3.2]^2$, $\varepsilon = 0.1$, $s = 0.01h$, $A = 1/16$ and $T = 0.32$. We use a linear refinement path, *i.e.*, $s = Ch$. At the final time $T = 0.32$, we expect the global error to be $\mathcal{O}(s^2) + \mathcal{O}(h^2) = \mathcal{O}(h^2)$, in either the ℓ^2 or ℓ^∞ norm, as $h, s \to 0$. The Cauchy difference is defined as $\delta_\phi := \phi_{h_f} - \mathcal{I}_c^f(\phi_{h_c})$, where $\mathcal{I}_c^f$ is a bilinear interpolation operator (with the Nearest Neighbor Interpolation applied in Matlab, which is similar to the 2D case in [31, 34] and the 3D case in [27]). This requires a relatively coarse solution, parametrized by h_c, and a relatively fine solution, parametrized by h_f, in particular $h_c = 2h_f$, at the same final time. The ℓ^2 norms of Cauchy difference and the convergence rates can be found in Table 5.4. The results confirm our expectation for the second-order convergence in both space and time.

Table 5.4: Errors, convergence rates, average iteration numbers and average CPU time (in seconds) for each time step. Parameters are given in the text, and the initial data is defined in (5.25). The refinement path is $s = 0.01h$.

				PSD		PNCG1		PNCG2	
h_c	h_f	$\lVert\delta_\phi\rVert_2$	Rate	$\#_{iter}$	$T_{cpu}(h_f)$	$\#_{iter}$	$T_{cpu}(h_f)$	$\#_{iter}$	$T_{cpu}(h_f)$
$\frac{3.2}{16}$	$\frac{3.2}{32}$	1.3938×10^{-2}	-	11	0.0019	9	0.0016	9	0.0015
$\frac{3.2}{32}$	$\frac{3.2}{64}$	1.7192×10^{-3}	3.02	10	0.0103	9	0.0093	8	0.0085
$\frac{3.2}{64}$	$\frac{3.2}{128}$	3.8734×10^{-4}	2.15	08	0.0529	8	0.0486	7	0.0454
$\frac{3.2}{128}$	$\frac{3.2}{256}$	9.4766×10^{-5}	2.03	07	0.2512	7	0.2038	6	0.2046
$\frac{3.2}{256}$	$\frac{3.2}{512}$	2.3564×10^{-5}	2.01	07	1.6650	7	1.6268	6	1.5207

In the second part of this test, we demonstrate the complexity of the preconditioned solvers with initial data (5.25). In Figure 5.7, we plot the semi-log scale of the relative residuals versus preconditioned solvers' iteration numbers for various values of h and ε at $T = 0.02$, with time step $s = 10^{-3}$. The other common parameters are set as $A = 1/16$, $\Omega = [0, 3.2]^2$. The figures in the top row of Figure 5.7 indicate that the convergence rate (as gleaned from the error reduction) is nearly uniform and nearly independent of h for a fixed ε. And the plots in the bottom row of Figure 5.7 show that the number of preconditioned solvers' iterations increases with a decreasing value of ε, which confirms the theoretical results that the PSD solver is dependent on parameter ε in [34]. Figure 5.7 confirms the expected geometric convergence rate of

the PSD solver predicted by the theory in [34]. Moreover, the number of the interation steps in Figure 5.7 also indicate that PNCG2 is the most efficient one and PNCG1 is better than PSD, especially when ε is small.

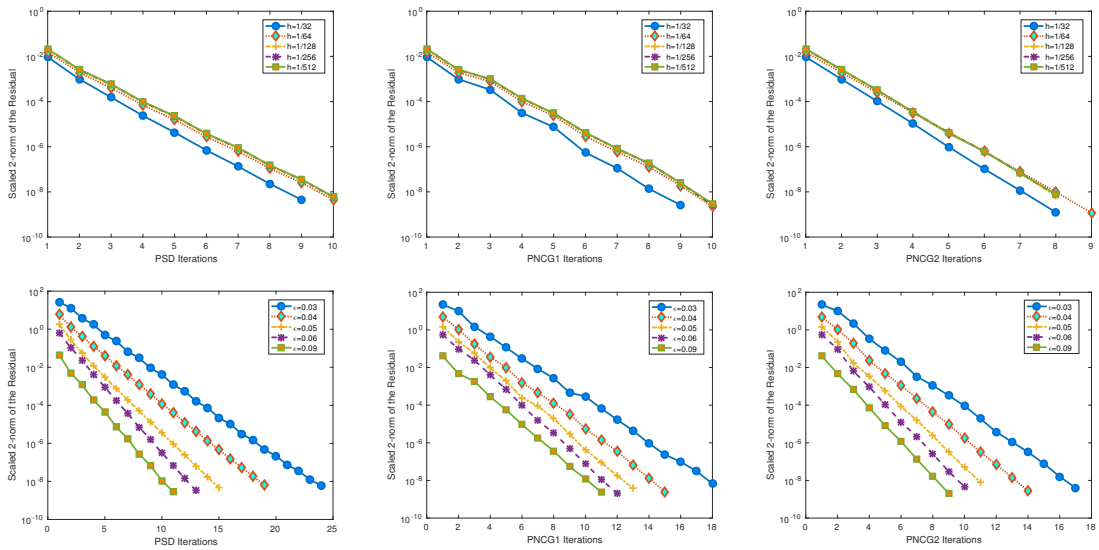


Figure 5.7: Complexity tests showing the solvers' performance for changing values of h and ε. Top row: h-independence with $\varepsilon = 0.1$; Bottom row: ε-dependence with $h = {}^{3.2}/_{512}$. The rest of the parameters are given in the text.

In the third part of this test, we perform CPU time comparison between the proposed preconditioned solvers and the PSD solver with random initial data. The initial data for the simulations are taken as essentially random:

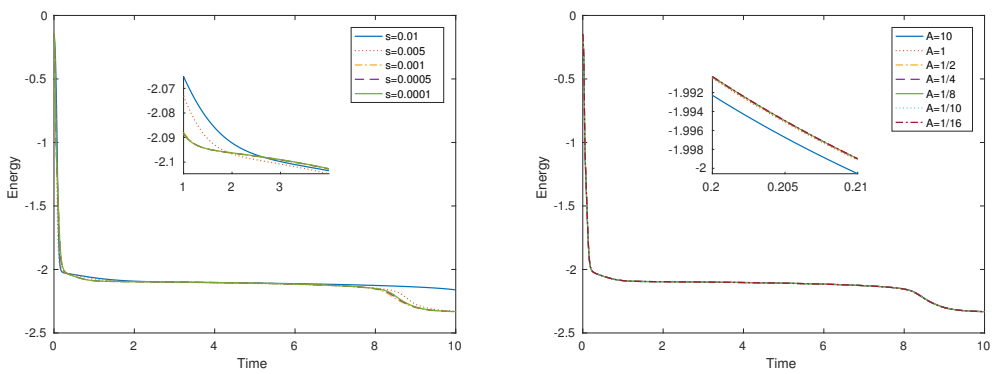
$$u^0_{i,j} = 0.05 \cdot (2r_{i,j} - 1), \tag{5.26}$$

where the $r_{i,j}$ are uniformly distributed random numbers in [0, 1]. The parameters for the comparison simulations are $\Omega = [0, 12.8]^2$, $\varepsilon = 3 \times 10^{-2}$, $h = {}^{12.8}/_{512}$, $s = 0.001$ and $T = 1$. The average iteration numbers, total CPU time (in seconds) and speedups for the preconditioned methods can be found in Table 5.5. The Table 5.5 indicates that the PNCG1 solver and PNCG2 solver have provided a 1.37x and 1.45x speedup over PSD solver, respectively.

Table 5.5: The average iteration numbers and total CPU time (in seconds) for the preconditioned methods with fixed time steps $s = 0.001$. Parameters are given in the text.

Methods	PSD	PNCG1	PNCG2
$\#_{iter}$	20	14	13
T_{cpu}(s)	4406.1764	3212.2898	3035.4369
Speedup	-	1.37	1.45

In the fourth part of this test, we investigate the effect of the parameters s and A for the energy $F_h(\phi)$ with initial data (5.25). Since the proposed solvers give the same results, we only present the results from PSD solver in the rest of the thesis. The evolutions of the energy with various time steps s and stabilized parameter A are given in Figure 5.8. As can be seen in Figure 5.8(a), the larger time steps produce inaccurate or nonphysical solutions. In turn, Figure 5.8(a) indicates the proper time steps and provides the motivation of using adaptive time stepping strategy. Figure 5.8(b) shows that the proposed scheme and PSD solver is not that sensitive to the stabilized parameter A when $A \leq 1$.



(a) evolutions of energy w.r.t various s (b) evolutions of energy w.r.t various A

Figure 5.8: The effect of time steps s and stabilized parameter A for the energy $F_h(\phi)$. Left: the effect of time step s. The other parameters are $\Omega = [0, 3.2]^2$, $\epsilon = 3.0 \times 10^{-2}$ and $A = 1/16$; Right: the effect of stabilized parameter A. The other parameters are $\Omega = [0, 3.2]^2$, $\epsilon = 1.0^{-2}$ and $s = 0.001$.

5.4.2 Long-time coarsening process, energy dissipation and mass conservation

Coarsening processes in thin film system can take place on very long time scales [57]. In this subsection, we perform long time simulation for the SS equation. Such a test, which has been performed in many existing literature, will confirm the expected coarsening rates and serve as a benchmarks for the proposed solver; see, for example, [34, 71, 77].

The initial data for this simulations are taken as (5.26). Time snapshots of the evolution for the epitaxial thin film growth model can be found in Figure 5.9. The coarsening rates are given in Figure 5.10. The interface width or roughness is defined as

$$W(t_n) = \sqrt{\frac{h^2}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}(\phi_{i,j}^n - \bar{\phi})^2}, \tag{5.27}$$

where m and n are the number of the grid points in x and y direction and $\bar{\phi}$ is the average value of ϕ on the uniform grid. The log-log plots of roughness and energy evolution and the corresponding linear regression are presented in Figure. 5.10. The linear regression in Figure. 5.10 indicates that the surface roughness grows like $t^{1/3}$, while the energy decays like $t^{-1/3}$, which verifies the one-third power law predicted in [58]. More precisely, the linear fits have the form $a_e t^{b_e}$ with $a_e = 3.09870, b_e = -0.33554$ for energy evolution and $a_m t^{b_m}$ with $a_m = -5.35913, b_m = 0.32555$ for roughness evolution. The linear regression is only taken up to $t = 3000$, since the saturation time would be of the order of ε^{-2} under the scaling that we have adopted [71]. These simulation results are consistent with earlier works on this topic in [34, 71, 77, 83].

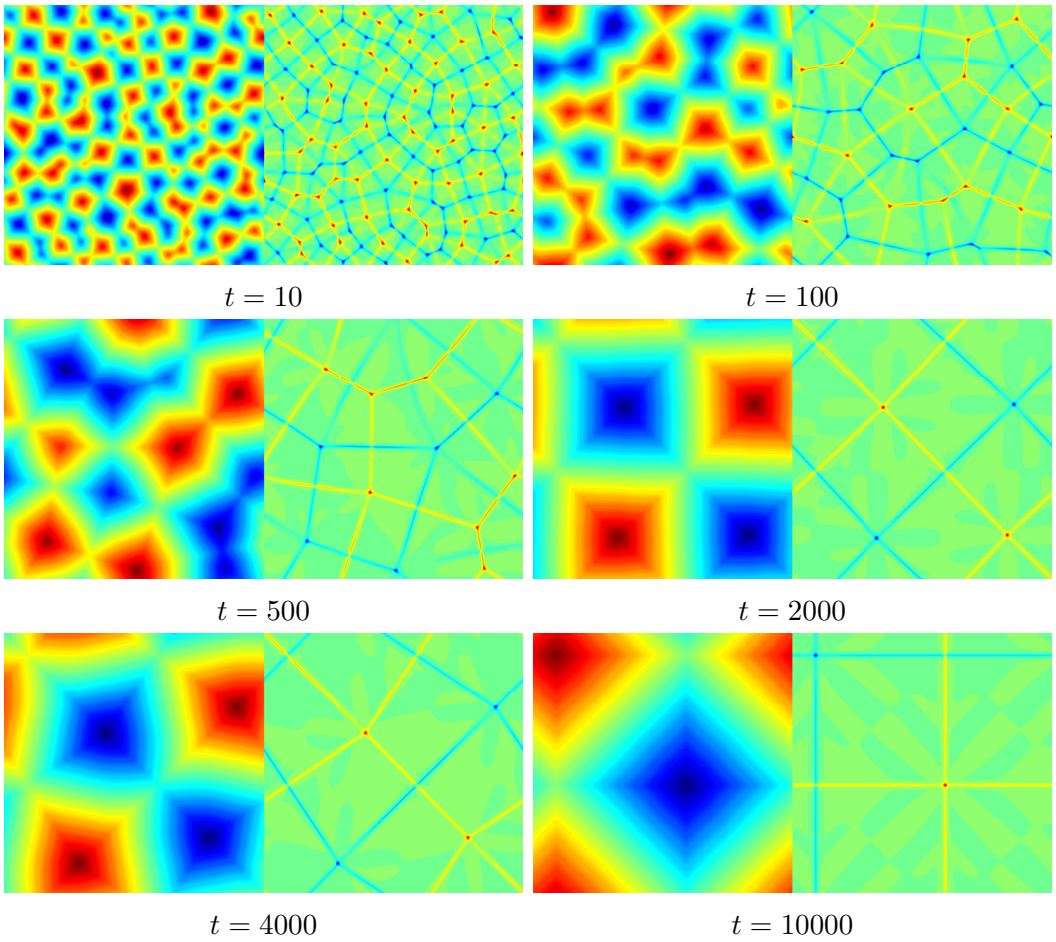
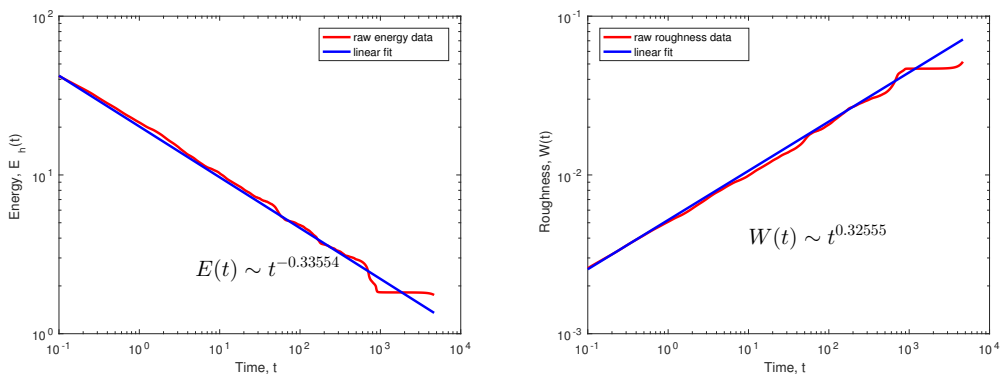


Figure 5.9: Time snapshots of the evolution with preconditioned solvers for the epitaxial thin film growth model at $t = 10, 100, 500, 2000, 4000$ and 10000. Left: contour plot of u, Right: contour plot of Δu. The parameters are $\varepsilon = 0.03, \Omega = [12.8]^2, s = 0.001$, $h = {}^{12.8}/_{512}$ and $A = {}^{1}/_{16}$. These simulation results are consistent with earlier work on this topic in [34, 71, 77, 83].



(a) Energy evolution

(b) Roughness evolution

Figure 5.10: The log-log plots of energy and roughness evolution and the corresponding linear regression for the simulation depicted in Figure 4.7.

Chapter 6

Conclusions

A preconditioned steepest descent (PSD) solver is proposed and analyzed for fourth and sixth-order regularized p-Laplacian equations. Solution of the highly nonlinear finite difference equations are equivalent to the minimizations of associated strictly convex energies. The energy dissipation property of the PSD solver leads to a bound for the numerical solution at each iteration stage. This fact, coupled with an upper-bound for the second derivative of the energy with respect to the metric induced by the pre-conditioner, leads to a geometric convergence rate for our (PSD) solver, which is proved rigorously for both the continuous and discrete space cases. In the present setting the pre-conditioner is a linear, constant-coefficient, positive, and symmetric finite difference operator. The key to the efficiency of our method is that this pre-conditioner can be efficiently inverted using the FFT. Various numerical results are presented in this thesis, including a convergence test and a complexity analysis for the PSD solver, as well as long-time simulation results for the thin film epitaxy model with slope selection (both $p = 4$ and $p = 6$) and the square phase field crystal model.

Since we have shown rigorously that our equations result as the gradients of strictly convex functionals, it also possible to use Newton's method (or a quasi Newton's method) to solve the nonlinear equations. One will still obtain global convergence, and in fact, we expect the convergence rate to be faster than geometric. On the

other hand, in the case of (3.34), say, Newton's method requires one to invert a complicated, non-constant coefficient, fourth-order linear finite difference equation. One could not use FFT for the inversion of this operator but would have to design an efficient solver for this purpose. This is a non-trivial task. So, in summary, although the (quasi) Newton's method would give a faster convergence rate than the PSD solver – in particular, super-linear convergence $\|e^{k+1}\| \leq C\|e^k\|^\beta$, $\beta > 1$, versus a geometric convergence rate – the PSD solver is, at least currently, much more efficient.

Based on the PSD solvers' framework, we also proposed two efficient and practical Preconditioned Nonlinear Conjugate Gradient (PNCG) solvers. In order to make the proposed solvers and scheme much more practical, we also investigate the adaptive time stepping strategy.

Numerical simulations for some important physical application problems – including thin film epitaxy with slope selection, the square phase field crystal model and functionalized Cahn-Hilliard equation – are carried out to verify the efficiency of the schemes and solvers.

Bibliography

[1] S. M. Allen and J. W. Cahn. A microscopic theory for antiphase boundary motion and its application to antiphase domain coursening. *Acta. Metall.*, 27:1085, 1979. 73

[2] A. Aristotelous, O. Karakasian, and S.M. Wise. A mixed discontinuous Galerkin, convex splitting scheme for a modified Cahn-Hilliard equation and an efficient nonlinear multigrid solver. *Discrete Contin. Dyn. Sys. B*, 18:2211–2238, 2013. 76, 77

[3] K. Atkinson and W. Han. *Theoretical numerical analysis*, volume 39. Springer, 2005. 21

[4] P. Aviles, Y. Giga, et al. A mathematical problem related to the physical theory of liquid crystal configurations. In *Proc. Centre Math. Anal. Austral. Nat. Univ*, volume 12, pages 1–16, 1987. 103

[5] O. Axelsson. *Iterative solution methods*. Cambridge university press, 1996. 20

[6] J.W. Barrett and W. Liu. Finite element approximation of the parabolic p-Laplacian. *SIAM J. Numer. Anal.*, 31:413–428, 1994. 3

[7] A. Baskaran, Z. Hu, J.S. Lowengrub, C. Wang, S.M. Wise, and P. Zhou. Energy stable and efficient finite-difference nonlinear multigrid schemes for the modified phase field crystal equation. *J. Comput. Phys.*, 250:270–292, 2013. 76, 77

[8] R. Bermejo and J.A. Infante. A multigrid algorithm for the p-Laplacian. *SIAM J. Sci. Comput.*, 21(5):1774–1789, 2000. 3

[9] J.P. Boyd. *Chebyshev and Fourier spectral methods*. Courier Corporation, 2001. 44, 95

[10] J.W. Cahn. On spinodal decomposition. *Acta Metall.*, 9:795, 1961. 73

[11] J.W. Cahn and J.E. Hilliard. Free energy of a nonuniform system. I. Interfacial free energy. *J. Chem. Phys.*, 28:258, 1958. 4, 73

[12] F. Chen and J. Shen. Efficient spectral-Galerkin methods for systems of coupled second-order equations and their applications. *J. Comput. Phys.*, 231(15):5016–5028, 2012. 75

[13] W. Chen, S. Conde, C. Wang, X. Wang, and S.M. Wise. A linear energy stable scheme for a thin film model without slope selection. *J. Sci. Comput.*, 52:546–562, 2012. 76, 77

[14] W. Chen, W. Feng, Y. Liu, C. Wang, and S.M. Wise. A second order energy stable scheme for the Cahn-Hilliard-Hele-Shaw equations. *arXiv preprint arXiv:1611.02967*, 2016. iv, 114

[15] W. Chen, Y. Liu, C. Wang, and S.M. Wise. An optimal-rate convergence analysis of a fully discrete finite difference scheme for Cahn-Hilliard-Hele-Shaw equation. *Math. Comput.*, 85:2231–2257, 2016. 76

[16] W. Chen, C. Wang, X. Wang, and S.M. Wise. A linear iteration algorithm for energy stable second order scheme for a thin film model without slope selection. *J. Sci. Comput.*, 59:574–601, 2014. 76, 77

[17] K. Cheng, W. Feng, S. Gottlieb, and C. Wang. A Fourier pseudospectral method for the "Good" Boussinesq equation with second-order temporal accuracy. *Numer. Methods Partial Differ. Equ.*, 31(1):202–224, 2015. 44, 95

[18] K. Cheng, W. Feng, C. Wang, and S.M. Wise. A refined truncation error estimate for long stencil fourth order finite difference approximation and its application to the Cahn-Hilliard equation. *in preparation*, 2016. iv, 4

[19] A. Christlieb, J. Jones, K. Promislow, B. Wetton, and M. Willoughby. High accuracy solutions to energy gradient flows from material science models. *J. Comput. Phys.*, 257, Part A:193 – 215, 2014. xvii, 4, 75, 95, 96, 98

[20] P.G. Ciarlet. *Introduction to Numerical Linear Algebra and Optimisation.* Cambridge University Press, New York, NY, USA, 1989. 21

[21] B. Cockburn and J. Shen. A hybridizable discontinuous Galerkin method for the p-Laplacian. *SIAM J. Sci. Comput.*, 38(1):A545–A566, 2016. 3

[22] M.C. Cross and P.C. Hohenberg. Pattern formation outside of equilibrium. *Rev. Mod. Phys.*, 65(3):851, 1993. 2

[23] S. Dai and K. Promislow. Geometric evolution of bilayers under the Functionalized Cahn–Hilliard equation. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 469(2153), 2013. 73

[24] A.E. Diegel, X. Feng, and S.M. Wise. Analysis of a mixed finite element method for a Cahn–Hilliard–Darcy–Stokes system. *SIAM J. Numer. Anal.*, 53(1):127–152, 2015. 8, 76

[25] A.E Diegel, C. Wang, and S.M. Wise. Stability and convergence of a second order mixed finite element method for the Cahn-Hilliard equation. *IMA J. Numer. Anal.*, 36(4):1867–1897, 2016. 76, 77

[26] A. Doelman, G. Hayrapetyan, K. Promislow, and B. Wetton. Meander and pearling of single-curvature bilayer interfaces in the functionalized cahn–hilliard equation. *SIAM J. Math. Anal.*, 46(6):3640–3677, 2014. 4, 73, 74

[27] L. Dong, W. Feng, C. Wang, S. M. Wise, and Z. Zhang. Convergence analysis and numerical implementation of a second order numerical scheme for the three-dimensional phase field crystal equation. *arXiv preprint arXiv:1611.06288*, 2016. iv, 16, 64, 106, 115

[28] I. Ekeland and R. Temam. *Convex analysis and 9 variational problems*. SIAM, 1976. 21

[29] K.R. Elder, M. Katakowski, M. Haataja, and M. Grant. Modeling elastic and plastic deformations in nonequilibrium processing using phase field crystals. *Phys. Rev. E*, 70:051605, 2004. 2, 69

[30] D. Eyre. Unconditionally gradient stable time marching the Cahn-Hilliard equation. In J. W. Bullard, R. Kalia, M. Stoneham, and L.Q. Chen, editors, *Computational and Mathematical Models of Microstructural Evolution*, volume 53, pages 1686–1712, Warrendale, PA, USA, 1998. Materials Research Society. 77

[31] W. Feng, Z. Guan, J. Lowengrub, C. Wang, and S.M. Wise. An energy stable finite-difference scheme for Functionalized Cahn-Hilliard equation and its convergence analysis. *arXiv preprint arXiv:1610.02473*, 2016. iv, v, 4, 59, 64, 72, 106, 115

[32] W. Feng, Z. Guo, J. Lowengrub, and S.M. Wise. Mass-conservative cell-centered finite difference methods and an efficient multigrid solver for the diffusion equation on block-structured, locally cartesian adaptive grids. *In preparation*, 2016. iv

[33] W. Feng, T.L. Lewis, and S.M. Wise. Discontinuous galerkin derivative operators with applications to second-order elliptic problems and stability. *Math Meth. Appl. Sci.*, 38(18):5160–5182, 2015. iv

[34] W. Feng, A.J. Salgado, C. Wang, and S.M. Wise. Preconditioned steepest descent methods for some nonlinear elliptic equations involving p-Laplacian terms. *J. Comput. Phys.*, 334:45–67, 2017. iv, xvi, xvii, xviii, 19, 44, 60, 61, 62, 64, 66, 67, 69, 76, 93, 95, 96, 103, 105, 106, 107, 109, 110, 115, 116, 118, 119

[35] W. Feng, C. Wang, and S. M. Wise. Linearly preconditioned nonlinear conjugate gradient solvers for the epitaxial thin film equation with slope selection. *In preparation*, 2017. iv, v, 101

[36] W. Feng, C. Wang, S. M. Wise, and Z. Zhang. A second-order energy stable backward differentiation formula method for the epitaxial thin film equation with slope selection. *arXiv preprint arXiv:1706.01943*, 2017. iv, xvii, 52, 110, 114

[37] X. Feng and Y. Li. Analysis of symmetric interior penalty discontinuous galerkin methods for the Allen–Cahn equation and the mean curvature flow. *IMA J. Numer. Anal.*, page dru058, 2014. 77

[38] X. Feng, Y. Li, and Y. Xing. Analysis of mixed interior penalty discontinuous galerkin methods for the Cahn–Hilliard equation and the Hele–Shaw flow. *SIAM J. Numer. Anal.*, 54(2):825–847, 2016. 77

[39] X. Feng and M. Neilan. Vanishing moment method and moment solution for second order fully nonlinear partial differential equations. *J. Scient. Comp.*, 38(1):74–98, 2009. 3

[40] R. Fletcher and C.M. Reeves. Function minimization by conjugate gradients. *Comput. J.*, 7(2):149–154, 1964. 102

[41] N. Gavish, G. Hayrapetyan, K. Promislow, and L. Yang. Curvature driven flow of bi-layer interfaces. *Physica D.*, 240(7):675–693, 2011. 73

[42] N. Gavish, J. Jones, Z. Xu, A. Christlieb, and K. Promislow. Variational models of network formation and ion transport: applications to perfluorosulfonate ionomer membranes. *Polymers*, 4(1):630–655, 2012. 73

[43] A.A Golovin and A.A. Nepomnyashchy. Disclinations in square and hexagonal patterns. *Phys. Rev. E*, 67:056202, 2003. 2, 69

[44] G. Gompper and M. Schick. Correlation between structural and interfacial properties of amphiphilic systems. *Phys. Rev. Lett.*, 65:1116–1119, Aug 1990. 73

[45] J. Guo, C. Wang, S.M. Wise, and X. Yue. An H^2 convergence of a second-order convex-splitting, finite difference scheme for the three-dimensional Cahn-Hilliard equation. *Commu. Math. Sci.*, 14:489–515, 2016. 16, 62, 76, 77

[46] R. Guo, Y. Xu, and Z. Xu. Local discontinuous Galerkin methods for the Functionalized Cahn-Hilliard equation. *J. Sci. Comput.*, 63(3):913–937, 2015. 75

[47] J.A.M. Hernández, F.G. Castañeda, and J.A.M. Cadenas. Formation of square patterns using a model alike Swift-Hohenberg. In *2014 11th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE)*, pages 1–6. IEEE, 2014. 2, 69

[48] J.S. Hesthaven, S. Gottlieb, and D. Gottlieb. *Spectral methods for time-dependent problems*, volume 21. Cambridge University Press, 2007. 44, 95

[49] R.B. Hoyle. Steady squares and hexagons on a subcritical ramp. *Phys. Rev. E*, 51(1):310, 1995. 2

[50] R.B. Hoyle. *Pattern formation: an introduction to methods.* Cambridge University Press, 2006. 2

[51] W. Y. Hsu and T. D. Gierke. Ion transport and clustering in nafion perfluorinated membranes. *J. Membr. Sci.*, 13(3):307 – 326, 1983. 73

[52] Z. Hu, S.M. Wise, C. Wang, and J.S. Lowengrub. Stable and efficient finite-difference nonlinear-multigrid schemes for the phase-field crystal equation. *J. Comput. Phys.*, 228:5323–5339, 2009. 10, 15, 76, 77

[53] Y. Huang, R. Li, and W. Liu. Preconditioned descent algorithms for p-Laplacian. *J. Sci. Comput.*, 32(2):343–371, 2007. 3, 4, 5

[54] S. Jain and F. S. Bates. Consequences of nonergodicity in aqueous binary PEO-PB micellar dispersions. *Macromolecules*, 37(4):1511–1523, 2004. xvii, 97, 100

[55] J. Jones. *Development of a fast and accurate time stepping scheme for the Functionalized Cahn-Hilliard equation and application to a graphics processing unit*. PhD thesis, Michigan State University, 2013. xvii, 74, 75, 95, 96, 98

[56] A.V. Knyazev and I. Lashuk. Steepest descent and conjugate gradient methods with variable preconditioning. *SIAM J. Matrix Anal. Appl.*, 29(4):1267–1280, 2007. 20, 21, 31

[57] R.V. Kohn. Energy-driven pattern formation. In *International Congress of Mathematicians*, volume 1, pages 359–383, 2006. 49, 66, 109, 118

[58] R.V. Kohn and X. Yan. Upper bound on the coarsening rate for an epitaxial growth model. *Comm. Pure Appl. Math.*, 56:1549–1564, 2003. 66, 109, 118

[59] B. Li and J. Liu. Epitaxial growth without slope selection: energetics, coarsening, and dynamic scaling. *J. Nonlinear Sci.*, 14(5):429–451, 2004. 2

[60] B. Li and J.G. Liu. Thin film epitaxy with or without slope selection. *Euro. J. Appl. Math.*, 14:713–743, 2003. 57

[61] B. Li and J.G. Liu. Epitaxial growth without slope selection: Energetics, coarsening, and dynamic scaling. *J. Nonlinear Sci.*, 14:429–451, 2004. 57

[62] W. Liu and N. Yan. Quasi-norm local error estimators for p-Laplacian. *SIAM J. Numer. Anal.*, 39(1):100–127, 2001. 3

[63] D.J. Lloyd, B. Sandstede, D. Avitabile, and A.R. Champneys. Localized hexagon patterns of the planar Swift-Hohenberg equation. *SIAM J Appl. Dyn. Syst.*, 7(3):1049–1100, 2008. 2, 69

[64] R.L. Pego. Front migration in the nonlinear Cahn-Hilliard equation. *Proc. R. Soc. Lond. A*, 422:261–278, 1989. 4

[65] E. Polak and G. Ribiere. Note sur la convergence de directions conjuges. *ESAIM Math. Model Num.*, 3(R1):35–43, 1969. 102

[66] E. Polak and G. Ribiére. Note sur la convergence des méthodes de directions conjuguées. *Rev. Fr. Imform. Rech. Oper.*, 16:35–43, 1969. 63

[67] K. Promislow and B. Wetton. Pem fuel cells: A mathematical overview. *SIAM J. Appl. Math.*, 70(2):369–409, 2009. 73, 74

[68] K. Promislow and Q. Wu. Existence of pearled patterns in the planar functionalized Cahn–Hilliard equation. *J. Differ. Equations*, 259(7):3298–3343, 2015. 73

[69] Z. Qiao, Z. Zhang, and T. Tang. An adaptive time-stepping strategy for the molecular beam epitaxy models. *SIAM J. Sci. Comput.*, 33(3):1395–1414, 2011. 112

[70] Y. Saad. *Iterative methods for sparse linear systems*. SIAM, 2003. 20

[71] J. Shen, C. Wang, X. Wang, and S.M. Wise. Second-order convex splitting schemes for gradient flows with Ehrlich-Schwoebel type energy: application to thin film epitaxy. *SIAM J. Numer. Anal.*, 50(1):105–125, 2012. xv, xvi, xvii, xviii, 2, 3, 10, 13, 40, 42, 46, 47, 49, 50, 62, 63, 66, 67, 103, 105, 106, 109, 110, 114, 118, 119

[72] J. Shen, C. Wang, X. Wang, and S.M. Wise. Second-order convex splitting schemes for gradient flows with Ehrlich-Schwoebel type energy: Application to thin film epitaxy. *SIAM J. Numer. Anal.*, 50:105–125, 2012. 77, 92

[73] J. Shen and X. Yang. Numerical approximations of Allen-Cahn and Cahn-Hilliard equations. *Discrete Contin. Dyn. Sys. A*, 28:1669–1691, 2010. 4

[74] X. Tai and J. Xu. Global and uniform convergence of subspace correction methods for some convex optimization problems. *Math. Comp.*, 71(237):105–124, 2002. 3

[75] S. Torabi, J.S. Lowengrub, A. Voigt, and S.M. Wise. A new phase-field model for strongly anisotropic systems. In *Proc. R. Soc. A*, pages rspa–2008. The Royal Society, 2009. 73

[76] S. Torabi, S.M. Wise, J.S. Lowengrub, A. Ratz, and A. Voigt. A new method for simulating strongly anisotropic Cahn-Hilliard equations. *MST 2007 Conference Proceedings*, 3:1432, 2007. 73

[77] C. Wang, X. Wang, and S.M. Wise. Unconditionally stable schemes for equations of thin film epitaxy. *Discrete Contin. Dyn. Sys. A*, 28:405–423, 2010. xv, xvi, xvii, xviii, 2, 3, 10, 45, 46, 49, 50, 63, 66, 67, 69, 76, 77, 105, 109, 110, 114, 118, 119

[78] C. Wang and S.M. Wise. An energy stable and convergent finite-difference scheme for the modified phase field crystal equation. *SIAM J. Numer. Anal.*, 49(3):945–969, 2011. 10, 76, 77

[79] X. Wang, L. Ju, and Q. Du. Efficient and stable exponential time differencing runge-kutta methods for phase field elastic bending energy models. *J. Comput. Phys*, 316:21–38, 2016. 75

[80] S.M. Wise. Unconditionally stable finite difference, nonlinear multigrid simulation of the Cahn-Hilliard-Hele-Shaw system of equations. *J. Sci. Comput.*, 44:38–68, 2010. 16

[81] S.M. Wise, J. Kim, and J. Lowengrub. Solving the regularized, strongly anisotropic Cahn–Hilliard equation by an adaptive nonlinear multigrid method. *J. Comput. Phys.*, 226(1):414–446, 2007. 73

[82] S.M. Wise, C. Wang, and J. Lowengrub. An energy stable and convergent finite-difference scheme for the phase field crystal equation. *SIAM J. Numer. Anal.*, 47:2269–2288, 2009. 3, 10, 13, 15, 16, 40, 42, 69, 76, 77, 79, 93

[83] C. Xu and T. Tang. Stability analysis of large time-stepping methods for epitaxial growth models. *SIAM J. Numer. Anal.*, 44(4):1759–1779, 2006. xv, xvi, xvii, xviii, 2, 49, 50, 66, 67, 109, 110, 118, 119

[84] Y. Yan, W. Chen, C. Wang, and S.M. Wise. A second-order energy stable BDF numerical scheme for the Cahn-Hilliard equation. *Commun. Comput. Phys.*, 2016. Submitted and in review. 62

[85] Z. Zhang, Y. Ma, and Z. Qiao. An adaptive time-stepping strategy for solving the phase field crystal model. *J. Comput. Phys.*, 249:204–215, 2013. 112

[86] Z. Zhang and Z. Qiao. An adaptive time-stepping strategy for the Cahn-Hilliard equation. *Commun. Comput. Phys.*, 11(04):1261–1278, 2012. 112

[87] G. Zhou and C. Feng. The steepest descent algorithm without line search for p-Laplacian. *Appl. Math. Comput.*, 224:36–45, 2013. 3

[88] G. Zhou, Y. Huang, and C. Feng. Preconditioned hybrid conjugate gradient algorithm for p-Laplacian. *Int. J. Num. Anal. Modell*, 2:123–130, 2005. 3, 103

Appendix

Appendix A

Proof of discrete Sobolev inequality

A.1 Proof of Lemma 2.2.2

Herein we only present the proof of (2.2.2) in Lemma 2.2.2 with $d = 2$ and $p = 4$. The other cases can be handled in the same way. Without loss of generality, we assume that $m = N = 2K + 1$ is odd and $L_x = L_y = L$, so that $h = \frac{L}{N} = \frac{L}{2K+1}$. We use N, rather than m, for the mesh size, as it is more standard.

For simplicity of presentation, we are focused on the estimate of $\|D_x u\|_4$, and we aim to establish the following estimate:

$$\|D_x u\|_4 \le C_0^{(1)} \|u\|_2^{\frac{1}{4}} \cdot \|\Delta_h u\|_2^{\frac{3}{4}}, \quad \forall\, u \in \mathring{\mathcal{C}}_{\text{per}} \tag{A.1}$$

where $C_0^{(1)} > 0$ depends upon L, but is independent of h and u. Due to the periodic boundary conditions for u and its cell-centered representation, it has a corresponding discrete Fourier transformation:

$$u_{i,j} \quad = \quad \sum_{\ell,m=-K}^{K} \hat{u}_{\ell,m}^{N} \mathrm{e}^{2\pi i(\ell x_i + m y_j)/L}, \tag{A.2}$$

where $x_i = (i - \frac{1}{2})h$, $y_j = (j - \frac{1}{2})h$, and $\hat{u}^N_{\ell,m}$ are discrete Fourier coefficients. Then we make its extension to a continuous function:

$$u_{\mathbf{F}}(x, y) = \sum_{\ell,m=-K}^{K} \hat{u}^N_{\ell,m} \mathrm{e}^{2\pi i(\ell x + m y)/L}. \tag{A.3}$$

Similarly, we denote the grid function $f := D_x u \in \mathcal{E}^{\mathrm{ew}}_{\mathrm{per}}$. The periodic boundary conditions for f and its (east-west-edge-centered) mesh location indicates the following discrete Fourier transformation:

$$f_{i+1/2,j} \quad = \quad \sum_{\ell,m=-K}^{K} \hat{f}^N_{\ell,m} \mathrm{e}^{2\pi i(\ell x_{i+1/2} + m y_j)/L}, \tag{A.4}$$

with $\hat{f}^N_{\ell,m}$ the discrete Fourier coefficients. Its extension to a continuous function is given by

$$f_{\mathbf{F}}(x, y) = \sum_{\ell,m=-K}^{K} \hat{f}^N_{\ell,m} \mathrm{e}^{2\pi i(\ell x + m y)/L}. \tag{A.5}$$

Meanwhile, we observe that $\hat{u}^N_{0,0} = 0$ and $\hat{f}^N_{0,0} = 0$. The first identity comes from the fact that $\overline{u} = 0$, while the second one is due to the fact that $\overline{f} = \overline{D_x u} = 0$, for any periodic grid function u.

The following preliminary estimates will play a very important role in the later analysis.

Lemma A.1.1. *We have*

$$\|u\|_2 = \|u_{\mathbf{F}}\|, \tag{A.6}$$

$$\frac{4}{\pi^2}\|\Delta u_{\mathbf{F}}\| \le \|\Delta_h u\|_2 \le \|\Delta u_{\mathbf{F}}\|, \tag{A.7}$$

$$\|\partial_x f_{\mathbf{F}}\| \le \left\|\partial_x^2 u_{\mathbf{F}}\right\|, \quad \|\partial_y f_{\mathbf{F}}\| \le \|\partial_x \partial_y u_{\mathbf{F}}\|, \tag{A.8}$$

$$\|f_{\mathbf{F}}\|_{\mathring{H}^{-1}_{\mathrm{per}}} \le \|u_{\mathbf{F}}\|. \tag{A.9}$$

Proof. Parseval's identity (at both the discrete and continuous levels) implies that

$$\sum_{i,j=0}^{N-1} |u_{i,j}|^2 = N^2 \sum_{\ell,m=-K}^{K} |\hat{u}^N_{\ell,m}|^2, \tag{A.10}$$

$$\|u_{\mathbf{F}}\|^2 = L^2 \sum_{\ell,m=-K}^{K} |\hat{u}^N_{\ell,m}|^2. \tag{A.11}$$

Based on the fact that $hN = L$, this in turn results in

$$\|u\|_2^2 = \|u_{\mathbf{F}}\|^2 = L^2 \sum_{\ell,m=-K}^{K} |\hat{u}^N_{\ell,m}|^2, \tag{A.12}$$

so that (A.6) is proven.

For the comparison between $f = D_x u$ and $\partial_x u_{\mathbf{F}}$, we look at the following Fourier expansions:

$$\begin{aligned} f_{i+1/2,j} = (D_x u)_{i+1/2,j} &= \frac{u_{i+1,j} - u_{i,j}}{h} \\ &= \sum_{\ell,m=-K}^{K} w_\ell \hat{u}^N_{\ell,m} \mathrm{e}^{2\pi i(\ell x_{i+1/2} + m y_j)/L}, \end{aligned} \tag{A.13}$$

$$f_{\mathbf{F}}(x,y) = \sum_{\ell,m=-K}^{K} w_\ell \hat{u}^N_{\ell,m} \mathrm{e}^{2\pi i(\ell x + m y)/L}, \tag{A.14}$$

$$\partial_x u_{\mathbf{F}}(x,y) = \sum_{\ell,m=-K}^{K} \nu_\ell \hat{u}^N_{\ell,m} \mathrm{e}^{2\pi i(\ell x + m y)/L}, \tag{A.15}$$

with

$$w_\ell = -\frac{2i \sin \frac{\ell \pi h}{L}}{h}, \quad \nu_\ell = -\frac{2\ell \pi i}{L}. \tag{A.16}$$

A comparison of Fourier eigenvalues between $|w_\ell|$ and $|\nu_\ell|$ shows that

$$\frac{2}{\pi}|\nu_\ell| \le |w_\ell| \le |\nu_\ell|, \quad \text{for} \quad -\mathrm{K} \le \ell \le \mathrm{K}. \tag{A.17}$$

For the estimate (A.7), we look at similar Fourier expansions:

$$(\Delta_h u)_{i,j} = \sum_{\ell,m=-K}^{K} \left(w_\ell^2 + w_m^2\right) \hat{u}_{\ell,m}^N \mathrm{e}^{2\pi i(\ell x_i + m y_j)/L}, \tag{A.18}$$

$$\Delta u_{\mathbf{F}}(x,y) = \sum_{\ell,m=-K}^{K} \left(\nu_\ell^2 + \nu_m^2\right) \hat{u}_{\ell,m}^N \mathrm{e}^{2\pi i(\ell x + m y)/L}. \tag{A.19}$$

In turn, an application of Parseval's identity yields

$$\|\Delta_h u\|_2^2 = L^2 \sum_{\ell,m=-K}^{K} \left|w_\ell^2 + w_m^2\right|^2 |\hat{u}_{\ell,m}^N|^2, \tag{A.20}$$

$$\|\Delta u_{\mathbf{F}}\|^2 = L^2 \sum_{\ell,m=-K}^{K} \left|\nu_\ell^2 + \nu_m^2\right|^2 |\hat{u}_{\ell,m}^N|^2. \tag{A.21}$$

The eigenvalue comparison estimate (A.17) implies the following inequality:

$$\frac{4}{\pi^2}\left|\nu_\ell^2 + \nu_m^2\right| \leq \left|w_\ell^2 + w_m^2\right| \leq \left|\nu_\ell^2 + \nu_m^2\right|, \quad \text{for} \quad -\mathrm{K} \leq \ell, \mathrm{m} \leq \mathrm{K}. \tag{A.22}$$

As a result, inequality (A.7) comes from a combination of A.20, (A.21) and (A.22).

For the estimate (A.8), we observe the following Fourier expansions:

$$\partial_x f_{\mathbf{F}}(x,y) = \sum_{\ell,m=-K}^{K} \nu_\ell w_\ell \hat{u}_{\ell,m}^N \mathrm{e}^{2\pi i(\ell x + m y)/L}, \tag{A.23}$$

$$\partial_x^2 u_{\mathbf{F}}(x,y) = \sum_{\ell,m=-K}^{K} \nu_\ell^2 \hat{u}_{\ell,m}^N \mathrm{e}^{2\pi i(\ell x + m y)/L}, \tag{A.24}$$

which in turn leads to (with an application of Parseval's identity)

$$\|\partial_x f_{\mathbf{F}}\|^2 = L^2 \sum_{\ell,m=-K}^{K} |\nu_\ell w_\ell|^2 |\hat{u}_{\ell,m}^N|^2, \tag{A.25}$$

$$\left\|\partial_x^2 u_{\mathbf{F}}\right\|^2 = L^2 \sum_{\ell,m=-K}^{K} |\nu_\ell|^4 |\hat{u}_{\ell,m}^N|^2. \tag{A.26}$$

Similarly, the following inequality could be derived, based on the eigenvalue comparison estimate (A.17):

$$|\nu_\ell w_\ell|^2 \le |\nu_\ell|^4, \quad \text{for} \quad -\mathrm{K} \le \ell, \mathrm{m} \le \mathrm{K}. \tag{A.27}$$

Consequently, a combination of (A.25), (A.26) and (A.27) leads to the first inequality in (A.8). The second inequality, $\|\partial_y f_{\mathbf{F}}\| \le \|\partial_x \partial_y u_{\mathbf{F}}\|$, could be derived in the same manner.

For the last estimate (A.9), we observe that

$$\|f_{\mathbf{F}}\|^2_{\mathring{H}^{-1}_{\mathrm{per}}} = L^2 \sum_{(\ell,m)\neq \mathbf{0}, \ell,m=-K}^{K} \frac{1}{|\nu_\ell^2 + \nu_m^2|} \cdot |w_\ell|^2 |\hat{u}^N_{\ell,m}|^2. \tag{A.28}$$

Meanwhile, the derivation of the following inequality is straight forward:

$$\frac{1}{|\nu_\ell^2 + \nu_m^2|} \cdot |w_\ell|^2 = \frac{|w_\ell|^2}{|\nu_\ell^2 + \nu_m^2|} \le \frac{|\nu_\ell|^2}{|\nu_\ell|^2} \le 1, \quad \forall (\ell, m) \neq \mathbf{0}, \tag{A.29}$$

in which the eigenvalue estimate (A.17) was used again in the second step. In comparison with (A.11), we arrive at (A.9). The proof of Lemma A.1.1 is complete. □

The following lemma gives a bound of the discrete ℓ^4 norm of the grid function f, in terms of the continuous L^4 norm of its continuous version $f_{\mathbf{F}}$.

Lemma A.1.2. *We have*

$$\|f\|_4 \le \sqrt{2} \|f_{\mathbf{F}}\|_{L^4}. \tag{A.30}$$

Proof. We denote the following grid function

$$g_{i+1/2,j} = \left(f_{i+1/2,j}\right)^2. \tag{A.31}$$

A direct calculation shows that

$$\|f\|_4 = (\|g\|_2)^{\frac{1}{2}}. \tag{A.32}$$

Note that both norms are discrete in the above identity. Moreover, we assume the grid function g has a discrete Fourier expansion as

$$g_{i+1/2,j} = \sum_{\ell,m=-K}^{K} (\hat{g}_c^N)_{\ell,m} \mathrm{e}^{2\pi \mathrm{i}(\ell x_{i+1/2} + m y_j)}, \tag{A.33}$$

and denote its continuous version as

$$G(x,y) = \sum_{\ell,m=-K}^{K} (\hat{g}_c^N)_{\ell,m} \mathrm{e}^{2\pi \mathrm{i}(\ell x + m y)} \in \mathcal{P}_K. \tag{A.34}$$

With an application of the Parseval equality at both the discrete and continuous levels, we have

$$\|g\|_2^2 = \|G\|^2 = \sum_{\ell,m=-K}^{K} \left|(\hat{g}_c^N)_{\ell,m}\right|^2. \tag{A.35}$$

On the other hand, we also denote

$$H(x,y) = (f_{\mathbf{F}}(x,y))^2 = \sum_{\ell,m=-2K}^{2K} (\hat{h}^N)_{\ell,m} \mathrm{e}^{2\pi \mathrm{i}(\ell x + m y)} \in \mathcal{P}_{2K}. \tag{A.36}$$

The reason for $H \in \mathcal{P}_{2K}$ is because $f_{\mathbf{F}} \in \mathcal{P}_K$. We note that $H \neq G$, since $H \in \mathcal{P}_{2K}$, while $G \in \mathcal{P}_K$, although H and G have the same interpolation values on at the numerical grid points $(x_i, y_{j+1/2})$. In other words, g is the interpolation of H onto the numerical grid point and G is the continuous version of g in $\mathcal{P}_K$. As a result, collocation coefficients $\hat{g}_c^N$ for G are not equal to $\hat{h}^N$ for H, due to the aliasing error.

In more detail, for $-K \le \ell, m \le K$, we have the following representations:

$$(\hat{g}_c^N)_{\ell,m} = \begin{cases} (\hat{h}^N)_{\ell,m} + (\hat{h}^N)_{\ell+N,m} + (\hat{h}^N)_{\ell,m+N} + (\hat{h}^N)_{\ell+N,m+N}, \ \ell < 0, m < 0, \\ (\hat{h}^N)_{\ell,m} + (\hat{h}^N)_{\ell+N,m}, \ \ell < 0, m = 0, \\ (\hat{h}^N)_{\ell,m} + (\hat{h}^N)_{\ell+N,m} + (\hat{h}^N)_{\ell,m-N} + (\hat{h}^N)_{\ell+N,m-N}, \ \ell < 0, m > 0, \\ (\hat{h}^N)_{\ell,m} + (\hat{h}^N)_{\ell-N,m} + (\hat{h}^N)_{\ell,m-N} + (\hat{h}^N)_{\ell-N,m-N}, \ \ell > 0, m > 0, \\ (\hat{h}^N)_{\ell,m} + (\hat{h}^N)_{\ell-N,m}, \ \ell > 0, m = 0, \\ (\hat{h}^N)_{\ell,m} + (\hat{h}^N)_{\ell-N,m} + (\hat{h}^N)_{\ell,m+N} + (\hat{h}^N)_{\ell-N,m+N}, \ \ell > 0, m < 0, \\ (\hat{h}^N)_{\ell,m} + (\hat{h}^N)_{\ell,m+N}, \ \ell = 0, m < 0, \\ (\hat{h}^N)_{\ell,m}, \ \ell = 0, m = 0, \\ (\hat{h}^N)_{\ell,m} + (\hat{h}^N)_{\ell,m-N}, \ \ell = 0, m > 0. \end{cases} \tag{A.37}$$

With an application of Cauchy inequality, it is clear that

$$\sum_{\ell,m=-K}^{K} \left|(\hat{g}_c^N)_{\ell,m}\right|^2 \le 4 \left| \sum_{\ell,m=-2K}^{2K} (\hat{h}^N)_{\ell,m} \right|^2 . \tag{A.38}$$

Meanwhile, an application of Parseval's identity to the Fourier expansion (A.36) gives

$$\|H\|^2 = \left| \sum_{\ell,m=-2K}^{2K} (\hat{h}^N)_{\ell,m} \right|^2 . \tag{A.39}$$

Its comparison with (A.35) indicates that

$$\|g\|_2^2 = \|G\|^2 \le 4 \|H\|^2 , \quad \text{i.e. } \|g\|_2 \le 2 \|H\| , \tag{A.40}$$

with the estimate (A.38) applied. Meanwhile, since $H(x,y) = (f_{\mathbf{F}}(x,y))^2$, we have

$$\|f_{\mathbf{F}}\|_{L^4} = (\|H\|)^{\frac{1}{2}} . \tag{A.41}$$

Therefore, a combination of (A.32), (A.40) and (A.41) results in

$$\|f\|_4 = (\|g\|_2)^{\frac{1}{2}} \le (2\,\|H\|)^{\frac{1}{2}} \le \sqrt{2}\,\|f_{\mathbf{F}}\|_{L^4}\,. \tag{A.42}$$

This finishes the proof of (A.30). □

Now we proceed into the proof of Proposition 2.2.2.

Proof. We begin with an application of (A.30) in Lemma A.1.2:

$$\|D_x u\|_4 = \|f\|_4 \le \sqrt{2}\|f_{\mathbf{F}}\|_{L^4}. \tag{A.43}$$

Meanwhile, using the fact that $\overline{f_{\mathbf{F}}} = 0$, we apply the 2-D Sobolev inequality and get

$$\|f_{\mathbf{F}}\|_{L^4} \le C\|f_{\mathbf{F}}\|_{H^{\frac{1}{2}}} \le C\|f_{\mathbf{F}}\|_{\mathring{H}^{-1}_{\mathrm{per}}}^{\frac{1}{4}} \cdot \|\nabla f_{\mathbf{F}}\|^{\frac{3}{4}}. \tag{A.44}$$

Moreover, the estimates (A.6) – (A.9) (in Lemma A.1.1) indicate that

$$\|f_{\mathbf{F}}\|_{\mathring{H}^{-1}_{\mathrm{per}}} \le \|u_{\mathbf{F}}\| = \|u\|_2, \tag{A.45}$$

$$\|\partial_x f_{\mathbf{F}}\| \le \left\|\partial_x^2 u_{\mathbf{F}}\right\| \le M_0\,\|\Delta u_{\mathbf{F}}\| \le \frac{\pi^2 M_0}{4}\|\Delta_h u\|_2, \tag{A.46}$$

$$\|\partial_y f_{\mathbf{F}}\| \le \|\partial_x\partial_y u_{\mathbf{F}}\| \le M_0\,\|\Delta u_{\mathbf{F}}\| \le \frac{\pi^2 M_0}{4}\|\Delta_h u\|_2, \tag{A.47}$$

so that

$$\|\nabla f_{\mathbf{F}}\| \le \frac{\sqrt{2}\pi^2 M_0}{4}\|\Delta_h u\|_2, \tag{A.48}$$

where the following elliptic regularity estimate is applied:

$$\left\|\partial_x^2 u_{\mathbf{F}}\right\|, \|\partial_x\partial_y u_{\mathbf{F}}\| \le M_0\,\|\Delta u_{\mathbf{F}}\|\,.$$

Therefore, a substitution of (A.102), (A.105) and (A.101) into (A.100) results in

$$\|D_x u\|_4 \le C_0^{(1)}\|u\|_2^{\frac{1}{4}} \cdot \|\Delta_h u\|_2^{\frac{3}{4}}, \quad \text{with } C_0^{(1)} = 2^{-5/8} M_0^{\frac{3}{4}} \pi^{3/2}.$$

The estimate for $\|D_y u\|_4$ could be derived in the same fashion. The result is stated below; its proof is skipped for the sake of brevity.

$$\|D_y u\|_4 \le C_0^{(1)} \|u\|_2^{\frac{1}{4}} \cdot \|\Delta_h u\|_2^{\frac{3}{4}}.$$

Moreover, by the definition of $\mathfrak{D}_x u$ and $\mathfrak{D}_y u$ we get

$$\|\mathfrak{D}_x u\|_4 = \|A_y(D_x u)\|_4 \le \|D_x u\|_4, \quad \|\mathfrak{D}_y u\|_4 = \|A_x(D_y u)\|_4 \le \|D_y u\|_4.$$

As a consequence, the first case of (2.2.2) (with $d = 2$, $p = 4$) is valid, by setting $C_0 = \sqrt{2} C_0^{(1)}$. The other cases could be analyzed in the same way. This finishes the proof of Proposition 2.2.2. □

A.2 Proof of Lemma 4.4.10

We will need the following discrete inequality:

Lemma A.2.1. *For any vertex-centered grid function $f \in \mathcal{V}_{\text{per}}$, we have*

$$\langle f, f \rangle = \langle 1, f^2 \rangle \ge \left(1, (\mathfrak{A} f)^2\right)_2. \tag{A.49}$$

Proof. Based on the definition of the average operator $\mathfrak{A}$, we have the following expansion and estimate:

$$\begin{aligned} (\mathfrak{A} f)_{i,j}^2 &= \left(\frac{1}{4}(f_{i-1/2,j-1/2} + f_{i+1/2,j-1/2} + f_{i-1/2,j+1/2} + f_{i+1/2,j+1/2})\right)^2 \\ &\le \frac{1}{4}\left(f_{i-1/2,j-1/2}^2 + f_{i+1/2,j-1/2}^2 + f_{i-1/2,j+1/2}^2 + f_{i+1/2,j+1/2}^2\right). \end{aligned} \tag{A.50}$$

Therefore, by summing over the grid index, in combination with the index counting, we arrive at

$$\sum_{i,j} (\mathfrak{A} f)_{i,j}^2 \le \sum_{i,j} f_{i+1/2,j+1/2}^2. \tag{A.51}$$

In turn, estimate (3.21) is a direct consequence of this inequality. This finishes the proof of this lemma. □

Using the last lemma, we are ready to prove Lemma 4.4.10:

Proof. We denote $g(\phi) := 3\phi^2 \mathcal{A} \left|\nabla_h^v \phi\right|^2$, at each cell center grid point (i,j). Consequently, we obtain $\mathcal{G}_h(\phi) = (1, g(\phi))_2$ and $\mathcal{H}_h(\phi) = \mathcal{G}_h(\phi) + A(\|\phi\|_4^4 + \|\nabla_h^v \phi\|_4^4)$. The following inequalities are evaluated at a point-wise level, for any ϕ_1, ϕ_2:

$$\begin{aligned}
&\left(\frac{\phi_1+\phi_2}{2}\right)^2 \le \frac{\phi_1^2+\phi_2^2}{2}, \quad \text{at } (i,j), \\
&\left|\nabla_h^v \left(\frac{\phi_1+\phi_2}{2}\right)\right|^2 \le \frac{|\nabla_h^v \phi_1|^2 + |\nabla_h^v \phi_2|^2}{2}, \quad \text{at } (i+1/2, j+1/2),
\end{aligned}$$

which come from the convexity of $q_2(x) = x^2$ and $r_2(\chi) = \chi \cdot \chi$. Moreover, taking an average operator $\mathfrak{A}$ to the second inequality leads to the following estimate:

$$\mathfrak{A}\Big(\left|\nabla_h^v \left(\frac{\phi_1+\phi_2}{2}\right)\right|^2\Big) \le \frac{\mathfrak{A}(|\nabla_h^v \phi_1|^2) + \mathfrak{A}(|\nabla_h^v \phi_2|^2)}{2}, \quad \text{at } (i,j).$$

These inequalities in turn imply that

$$\begin{aligned}
g\left(\frac{\phi_1+\phi_2}{2}\right) &= 3\left(\frac{\phi_1+\phi_2}{2}\right)^2 \cdot \mathfrak{A}\Big(\left|\nabla_h^v \left(\frac{\phi_1+\phi_2}{2}\right)\right|^2\Big) \\
&\le 3\frac{\phi_1^2+\phi_2^2}{2} \cdot \frac{\mathfrak{A}(|\nabla_h^v \phi_1|^2) + \mathfrak{A}(|\nabla_h^v \phi_2|^2)}{2},
\end{aligned}$$

at a point-wise level. A careful comparison with $\frac{g(\phi_1)+g(\phi_2)}{2} = \frac{3\phi_1^2\left|\nabla_h^v\phi_1\right|^2+3\phi_2^2\left|\nabla_h^v\phi_2\right|^2}{2}$ shows that

$$\begin{aligned}
& \frac{g(\phi_1)+g(\phi_2)}{2} - g\left(\frac{\phi_1+\phi_2}{2}\right) \\
\geq\ & \frac{3(\phi_1^2-\phi_2^2)(\mathfrak{A}(|\nabla_h^v\phi_1|^2)-\mathfrak{A}(|\nabla_h^v\phi_2|^2))}{4} \\
\geq\ & -\frac{3}{8}\left((\phi_1^2-\phi_2^2)^2+(\mathfrak{A}(|\nabla_h^v\phi_1|^2-|\nabla_h^v\phi_2|^2))^2\right). \qquad \text{(A.52)}
\end{aligned}$$

Similarly, the convexity of $q_4(x) = x^4$ and $r_4(\chi) = |\chi|^4$ indicates the following inequalities:

$$\frac{\phi_1^4+\phi_2^4}{2} - \left(\frac{\phi_1+\phi_2}{2}\right)^4 \geq \frac{3}{8}(\phi_1^4+\phi_2^4-2\phi_1^2\phi_2^2) = \frac{3}{8}(\phi_1^2-\phi_2^2)^2, \quad \text{at } (i,j) \qquad \text{(A.53)}$$

and

$$\begin{aligned}
& \frac{|\nabla_h^v\phi_1|^4+|\nabla_h^v\phi_2|^4}{2} - \left|\nabla_h^v\left(\frac{\phi_1+\phi_2}{2}\right)\right|^4 \qquad \text{(A.54)} \\
\geq\ & \frac{3}{8}(|\nabla_h^v\phi_1|^4+|\nabla_h^v\phi_2|^4-2|\nabla_h^v\phi_1|^2\cdot|\nabla_h^v\phi_2|^2) \\
\geq\ & \frac{3}{8}(|\nabla_h^v\phi_1|^2-|\nabla_h^v\phi_2|^2)^2, \quad \text{at } (i+1/2, j+1/2). \qquad \text{(A.55)}
\end{aligned}$$

Meanwhile, the following estimate is available, with an application of inequality (A.49) in Lem. A.2.1, by taking $f = |\nabla_h^v\phi_1|^2 - |\nabla_h^v\phi_2|^2$:

$$\langle 1, (|\nabla_h^v\phi_1|^2-|\nabla_h^v\phi_2|^2)^2\rangle \geq \left(1, \mathfrak{A}(|\nabla_h^v\phi_1|^2-|\nabla_h^v\phi_2|^2))^2\right)_2. \qquad \text{(A.56)}$$

As a result, a combination of (A.52), (A.53), (A.55) and (A.56) yields

$$\frac{\mathcal{H}_h(\phi_1)+\mathcal{H}_h(\phi_2)}{2} - \mathcal{H}_h\left(\frac{\phi_1+\phi_2}{2}\right) \geq 0, \quad \forall \phi_1,\, \phi_2, \quad \text{if } A \geq 1.$$

The convexity of $\mathcal{H}_h(\phi)$ is assured under the condition $A \geq 1$. □

A.3 Proof of Proposition 4.4.9

We will need the following average-shift identity:

Lemma A.3.1. $(\phi, \mathfrak{A}\psi)_2 = \langle \mathfrak{a}\phi, \psi \rangle$ *for any* $\phi \in \mathcal{C}_{\text{per}}, \psi \in \mathcal{V}_{\text{per}}$.

Now, we are ready to prove Proposition. 4.4.9.

Proof. From the definitions of the discrete norms, we have

$$\begin{aligned}
\mathcal{H}_h[\phi] &= A\,\|\phi\|_4^4 + A\,\|\nabla_h^v \phi\|_4^4 + 3\left(\phi^2, \mathfrak{A}(|\nabla_h^v \phi|^2)\right)_2 \\
&= A\left(|\phi|^4, 1\right)_2 + A\langle |\nabla_h^v \phi|^4, 1\rangle + 3\left(\phi^2, \mathfrak{A}(|\nabla_h^v \phi|^2)\right)_2 \\
&= A\left(|\phi|^4, 1\right)_2 + A\left\langle [(\mathfrak{D}_x\phi)^2 + (\mathfrak{D}_y\phi)^2]^2, 1\right\rangle + 3\left(\phi^2, \mathfrak{A}[(\mathfrak{D}_x\phi)^2 + (\mathfrak{D}_y\phi)^2]\right)_2 .
\end{aligned}$$

So the perturbation of the discrete energy is

$$\begin{aligned}
\mathcal{H}_h[\phi + s\psi] &= A\left(|\phi + s\psi|^4, 1\right)_2 + A\left\langle [(\mathfrak{D}_x(\phi + s\psi))^2 + (\mathfrak{D}_y(\phi + s\psi))^2]^2, 1\right\rangle \\
&\quad + 3\left((\phi + s\psi)^2, \mathfrak{A}[(\mathfrak{D}_x(\phi + s\psi))^2 + (\mathfrak{D}_y(\phi + s\psi))^2]\right)_2 .
\end{aligned}$$

and the first variational derivative is

$$\begin{aligned}
\frac{\delta \mathcal{H}_h}{\delta s}[\phi + s\psi] &= 4A\left((\phi + s\psi)^3, \psi\right)_2 \\
&+ 4A\langle [(\mathfrak{D}_x\phi + s\mathfrak{D}_x\psi)^2 + (\mathfrak{D}_y\phi + s\mathfrak{D}_y\psi)^2] \\
&\quad \times [(\mathfrak{D}_x\phi + s\mathfrak{D}_x\psi)\mathfrak{D}_x\psi + (\mathfrak{D}_y\phi + s\mathfrak{D}_y\psi)\mathfrak{D}_y\psi], 1\rangle \\
&+ 6\left((\phi + s\psi)\psi, \mathfrak{A}[(\mathfrak{D}_x\phi + s\mathfrak{D}_x\psi)^2 + (\mathfrak{D}_y\phi + s\mathfrak{D}_y\psi)^2]\right)_2 \\
&+ 6\left((\phi + s\psi)^2, \mathfrak{A}[(\mathfrak{D}_x\phi + s\mathfrak{D}_x\psi)\mathfrak{D}_x\psi + (\mathfrak{D}_y\phi + s\mathfrak{D}_y\psi)\mathfrak{D}_y\psi]\right)_2 .
\end{aligned}$$

setting $s = 0$ yields

$$\begin{aligned}
\frac{\delta \mathcal{H}_h}{\delta s}[\phi + s\psi]|_{s=0} &= 4A\left(\phi^3, \psi\right)_2 \\
&+ 4A\langle [(\mathfrak{D}_x\phi)^2 + (\mathfrak{D}_y\phi)^2][\mathfrak{D}_x\phi\mathfrak{D}_x\psi + \mathfrak{D}_y\phi\mathfrak{D}_y\psi], 1\rangle
\end{aligned}$$

$$
\begin{aligned}
&+ \quad 6\left(\phi\psi, \mathfrak{A}[(\mathfrak{D}_x\phi)^2 + (\mathfrak{D}_y\phi)^2]\right)_2 + 6\left(\phi^2, \mathfrak{A}[\mathfrak{D}_x\phi\mathfrak{D}_x\psi + \mathfrak{D}_y\phi\mathfrak{D}_y\psi]\right)_2 \\
&= \quad 4A\left(\phi^3, \psi\right)_2 \\
&+ \quad 4A\langle[(\mathfrak{D}_x\phi)^2 + (\mathfrak{D}_y\phi)^2][\mathfrak{D}_x\phi, \mathfrak{D}_x\psi\rangle + 4A\langle[(\mathfrak{D}_x\phi)^2 + (\mathfrak{D}_y\phi)^2]\mathfrak{D}_y\phi, \mathfrak{D}_y\psi\rangle \\
&+ \quad 6\left(\phi\mathfrak{A}[(\mathfrak{D}_x\phi)^2 + (\mathfrak{D}_y\phi)^2], \psi\right)_2 \\
&+ \quad 6\left(\phi^2\mathfrak{A}\left(\mathfrak{D}_x\phi\mathfrak{D}_x\psi\right), 1\right)_2 + 6\left(\phi^2\mathfrak{A}\left(\mathfrak{D}_y\phi\mathfrak{D}_y\psi\right), 1\right)_2 .
\end{aligned}
$$

Applying Lem.A.3.1 yields

$$
\begin{aligned}
\frac{\delta\mathcal{H}_h}{\delta s}[\phi] &= 4A\left(\phi^3, \psi\right)_2 \\
&+ 4A\langle[(\mathfrak{D}_x\phi)^2 + (\mathfrak{D}_y\phi)^2]\mathfrak{D}_x\phi, \mathfrak{D}_x\psi\rangle + 4A\langle[(\mathfrak{D}_x\phi)^2 + (\mathfrak{D}_y\phi)^2]\mathfrak{D}_y\phi, \mathfrak{D}_y\psi\rangle \\
&+ 6\left(\phi\mathfrak{A}[(\mathfrak{D}_x\phi)^2 + (\mathfrak{D}_y\phi)^2], \psi\right)_2 + 6\left\langle\mathfrak{a}\left(\phi^2\right)\mathfrak{D}_x\phi, \mathfrak{D}_x\psi\right\rangle + 6\left\langle\mathfrak{a}\left(\phi^2\right)\mathfrak{D}_y\phi, \mathfrak{D}_y\psi\right\rangle .
\end{aligned}
$$

By using the summation-by-parts formula, we have

$$
\begin{aligned}
\frac{\delta\mathcal{H}_h}{\delta s}[\phi] &= 4A\left(\phi^3, \psi\right)_2 \\
&- 4A\left(\mathfrak{d}_x\left([(\mathfrak{D}_x\phi)^2 + (\mathfrak{D}_y\phi)^2]\mathfrak{D}_x\phi\right), \psi\right)_2 - 4A\left(\mathfrak{d}_y\left([(\mathfrak{D}_x\phi)^2 + (\mathfrak{D}_y\phi)^2]\mathfrak{D}_y\phi\right), \psi\right)_2 \\
&+ 6\left(\phi\mathfrak{A}[(\mathfrak{D}_x\phi)^2 + (\mathfrak{D}_y\phi)^2], \psi\right)_2 - 6\left(\mathfrak{d}_x\left(\mathfrak{a}\left(\phi^2\right)\mathfrak{D}_x\phi\right), \psi\right)_2 - 6\left(\mathfrak{d}_y\left(\mathfrak{a}\left(\phi^2\right)\mathfrak{D}_y\phi\right), \psi\right)_2 .
\end{aligned}
$$

The proof of Proposition. 4.4.9 is complete. Hence the discrete variational derivative of $\mathcal{H}_h$ is

$$
\begin{aligned}
\delta\mathcal{H}_h[\phi] &= 4A\phi^3 - 4A\left(\mathfrak{d}_x\left([(\mathfrak{D}_x\phi)^2 + (\mathfrak{D}_y\phi)^2]\mathfrak{D}_x\phi\right) + \mathfrak{d}_y\left([(\mathfrak{D}_x\phi)^2 + (\mathfrak{D}_y\phi)^2]\mathfrak{D}_y\phi\right)\right) \\
&+ 6\phi\mathfrak{A}[(\mathfrak{D}_x\phi)^2 + (\mathfrak{D}_y\phi)^2] - 6\left(\mathfrak{d}_x\left(\mathfrak{a}\left(\phi^2\right)\mathfrak{D}_x\phi\right) + \mathfrak{d}_y\left(\mathfrak{a}\left(\phi^2\right)\mathfrak{D}_y\phi\right)\right) .
\end{aligned}
$$

□

A.4 Proof of Proposition 2.2.7

For simplicity of presentation, in the analysis of $\|\nabla_h \phi\|_6$, we are focused on the estimate of $D_x\phi\|_6$. Due to the periodic boundary conditions for ϕ and its cell-centered representation, it has a corresponding discrete Fourier transformation:

$$\phi_{i,j} = \sum_{\ell,m=-K}^{K} \hat{\phi}_{\ell,m}^N \mathrm{e}^{2\pi i(\ell x_i + m y_j)/L}, \tag{A.57}$$

where $x_i = (i - \frac{1}{2})h$, $y_j = (j - \frac{1}{2})h$, and $\hat{\phi}_{\ell,m}^N$ are discrete Fourier coefficients. Then we make its extension to a continuous function:

$$\phi_{\mathbf{F}}(x, y) = \sum_{\ell,m=-K}^{K} \hat{\phi}_{\ell,m}^N \mathrm{e}^{2\pi i(\ell x + m y)/L}. \tag{A.58}$$

Similarly, we denote a grid function $f_{i+\frac{1}{2},j+\frac{1}{2}} = \mathfrak{D}_x \phi_{i+\frac{1}{2},j+\frac{1}{2}} = A_y(D_x\phi)_{i+\frac{1}{2},j+\frac{1}{2}}$. The periodic boundary conditions for f and its mesh location indicates the following discrete Fourier transformation:

$$f_{i+\frac{1}{2},j+\frac{1}{2}} = \sum_{\ell,m=-K}^{K} \hat{f}_{\ell,m}^N \mathrm{e}^{2\pi i(\ell x_{i+\frac{1}{2}} + m y_{j+\frac{1}{2}})/L}, \tag{A.59}$$

with $\hat{f}_{\ell,m}^N$ the discrete Fourier coefficients. And also, its extension to a continuous function is given by

$$f_{\mathbf{F}}(x, y) = \sum_{\ell,m=-K}^{K} \hat{f}_{\ell,m}^N \mathrm{e}^{2\pi i(\ell x + m y)/L}. \tag{A.60}$$

Meanwhile, we also observe that $\hat{\phi}_{0,0}^N = 0$ and $\hat{f}_{0,0}^N = 0$. The first identity comes from the fact that $\overline{\phi} = 0$, while the second one is due to the fact that $\overline{f} = \overline{\mathfrak{D}_x \phi} = 0$, for any periodic grid function ϕ.

The following preliminary estimates will play a very important role in the later analysis.

Lemma A.4.1. *We have*

$$\|\phi\|_2 = \|\phi_{\mathbf{F}}\|, \tag{A.61}$$

$$\frac{2}{\pi}\|\nabla\phi_{\mathbf{F}}\| \le \|\nabla_h\phi\|_2 \le \|\nabla\phi_{\mathbf{F}}\|, \quad \frac{4}{\pi^2}\|\Delta\phi_{\mathbf{F}}\| \le \|\Delta_h\phi\|_2 \le \|\Delta\phi_{\mathbf{F}}\|, \tag{A.62}$$

$$\|\partial_x f_{\mathbf{F}}\| \le \left\|\partial_x^2\phi_{\mathbf{F}}\right\|, \quad \|\partial_y f_{\mathbf{F}}\| \le \|\partial_x\partial_y\phi_{\mathbf{F}}\|. \tag{A.63}$$

Proof. Parseval's identity (at both the discrete and continuous levels) implies that

$$\sum_{i,j=0}^{N-1}|\phi_{i,j}|^2 = N^2\sum_{\ell,m=-K}^{K}|\hat{\phi}^N_{\ell,m,n}|^2, \quad \|\phi_{\mathbf{F}}\|^2 = L^2\sum_{\ell,m=-K}^{K}|\hat{\phi}^N_{\ell,m}|^2. \tag{A.64}$$

Based on the fact that $hN = L$, this in turn results in

$$\|\phi\|_2^2 = \|\phi_{\mathbf{F}}\|^2 = L^2\sum_{\ell,m=-K}^{K}|\hat{\phi}^N_{\ell,m}|^2, \tag{A.65}$$

so that (A.61) is proven.

For the comparison between $f = \mathfrak{D}_x\phi$ and $\partial_x\phi_{\mathbf{F}}$, we look at the following Fourier expansions:

$$f_{i+\frac{1}{2},j+\frac{1}{2}} = \frac{\phi_{i+1,j} - \phi_{i,j} + \phi_{i+1,j+1} - \phi_{i,j+1}}{2h} = \sum_{\ell,m=-K}^{K}\mu_{\ell,m}\hat{\phi}^N_{\ell,m}\mathrm{e}^{2\pi i(\ell x_{i+\frac{1}{2}}+my_{j+\frac{1}{2}})/L}, \tag{A.66}$$

$$f_{\mathbf{F}}(x,y) = \sum_{\ell,m=-K}^{K}\mu_{\ell,m}\hat{\phi}^N_{\ell,m}\mathrm{e}^{2\pi i(\ell x+my)/L}, \tag{A.67}$$

$$\partial_x\phi_{\mathbf{F}}(x,y) = \sum_{\ell,m=-K}^{K}\nu_{\ell}\hat{\phi}^N_{\ell,m}\mathrm{e}^{2\pi i(\ell x+my)/L}, \tag{A.68}$$

with

$$\mu_{\ell,m} = -\frac{2i\sin\frac{\ell\pi h}{L}}{h}\cos(m\pi h), \quad \nu_\ell = -\frac{2\ell\pi i}{L}. \tag{A.69}$$

A comparison of Fourier eigenvalues between $|\mu_{\ell,m}|$ and $|\nu_\ell|$ shows that

$$\frac{2}{\pi}|\nu_\ell| \leq |\mu_{\ell,m}| \leq |\nu_\ell|, \quad \text{for} \quad -\mathrm{K} \leq \ell, \mathrm{m} \leq \mathrm{K}, \tag{A.70}$$

which in turn leads to

$$\frac{2}{\pi}\|\partial_x \phi_{\mathbf{F}}\| \leq \|\mathfrak{D}_x \phi\|_2 \leq \|\partial_x \phi_{\mathbf{F}}\|. \tag{A.71}$$

A similar estimate could also be derived:

$$\frac{2}{\pi}\|\partial_y \phi_{\mathbf{F}}\| \leq \|\mathfrak{D}_y \phi\|_2 \leq \|\partial_y \phi_{\mathbf{F}}\|. \tag{A.72}$$

A combination of (A.71) and (A.72) yields the first inequality of (A.62).

For the second estimate of (A.62), we look at similar Fourier expansions:

$$(\Delta_h \phi)_{i,j} = \sum_{\ell,m=-K}^{K} \left(\mu_\ell^2 + \mu_m^2\right) \hat{\phi}_{\ell,m}^N \mathrm{e}^{2\pi i(\ell x_i + m y_j)/L}, \tag{A.73}$$

$$\Delta \phi_{\mathbf{F}}(x,y) = \sum_{\ell,m=-K}^{K} \left(\nu_\ell^2 + \nu_m^2\right) \hat{\phi}_{\ell,m}^N \mathrm{e}^{2\pi i(\ell x + m y)/L}, \tag{A.74}$$

with $\mu_\ell = -\frac{2i\sin\frac{\ell\pi h}{L}}{h}$, $\mu_m = -\frac{2i\sin\frac{m\pi h}{L}}{h}$. It is also clear that $\frac{2}{\pi}|\nu_\ell| \leq |\mu_\ell| \leq |\nu_\ell|$, for any $-K \leq \ell \leq K$. In turn, an application of Parseval's identity yields

$$\|\Delta_h \phi\|_2^2 = L^2 \sum_{\ell,m=-K}^{K} \left|\mu_\ell^2 + \mu_m^2\right|^2 |\hat{\phi}_{\ell,m}^N|^2, \tag{A.75}$$

$$\|\Delta \phi_{\mathbf{F}}\|^2 = L^2 \sum_{\ell,m=-K}^{K} \left|\nu_\ell^2 + \nu_m^2\right|^2 |\hat{\phi}_{\ell,m}^N|^2. \tag{A.76}$$

The eigenvalue comparison estimate (A.70) implies the following inequality:

$$\frac{4}{\pi^2}\left|\nu_\ell^2 + \nu_m^2\right| \leq \left|\mu_\ell^2 + \mu_m^2\right| \leq \left|\nu_\ell^2 + \nu_m^2\right|, \quad \text{for} \quad -\mathrm{K} \leq \ell, \mathrm{m} \leq \mathrm{K}. \tag{A.77}$$

As a result, inequality (A.62) comes from a combination of (A.75), (A.76) and (A.77).

For the estimate (A.63), we observe the following Fourier expansions:

$$\partial_x f_{\mathbf{F}}(x,y) = \sum_{\ell,m=-K}^{K} \nu_\ell \mu_{\ell,m} \hat{\phi}^N_{\ell,m} \mathrm{e}^{2\pi i(\ell x+my)/L}, \tag{A.78}$$

$$\partial_x^2 \phi_{\mathbf{F}}(x,y) = \sum_{\ell,m=-K}^{K} \nu_\ell^2 \hat{\phi}^N_{\ell,m} \mathrm{e}^{2\pi i(\ell x+my)/L}, \tag{A.79}$$

which in turn leads to (with an application of Parseval's identity)

$$\|\partial_x f_{\mathbf{F}}\|^2 = L^2 \sum_{\ell,m=-K}^{K} |\nu_\ell \mu_{\ell,m}|^2 \, |\hat{\phi}^N_{\ell,m}|^2, \tag{A.80}$$

$$\left\|\partial_x^2 \phi_{\mathbf{F}}\right\|^2 = L^2 \sum_{\ell,m=-K}^{K} |\nu_\ell|^4 |\hat{\phi}^N_{\ell,m}|^2. \tag{A.81}$$

Similarly, the following inequality could be derived, based on the eigenvalue comparison estimate (A.70):

$$|\nu_\ell \mu_{\ell,m}|^2 \le |\nu_\ell|^4, \quad \text{for} \quad -\mathrm{K} \le \ell, \mathrm{m} \le \mathrm{K}. \tag{A.82}$$

Consequently, a combination of (A.80), (A.81) and (A.82) leads to the first inequality in (A.63). The second inequality, $\|\partial_y f_{\mathbf{F}}\| \le \|\partial_x \partial_y \phi_{\mathbf{F}}\|$, could be derived in the same manner. The proof of Lemma A.4.1 is complete. □

With the estimates in Lemma A.4.1, we are able to make the following derivations:

$$\|\phi\|^2_{H_h^2} = \|\phi\|_2^2 + \|\nabla_h \phi\|_2^2 + \|\Delta_h \phi\|_2^2 \le \|\phi_{\mathbf{F}}\|^2 + \|\nabla \phi_{\mathbf{F}}\|^2 + \|\Delta \phi_{\mathbf{F}}\|^2 \le \|\phi_{\mathbf{F}}\|^2_{H_h^2}, \tag{A.83}$$

$$\|\phi_{\mathbf{F}}\|^2_{H_h^2} \le B_0 \|\Delta \phi_{\mathbf{F}}\|^2, \quad \text{(elliptic regularity, since } \textstyle\int_\Omega \phi_{\mathbf{F}}\, d\mathbf{x} = 0), \tag{A.84}$$

$$\text{so that} \quad \|\Delta_h \phi\|_2^2 \ge \frac{4}{\pi^2} \|\Delta \phi_{\mathbf{F}}\|^2 \ge \frac{4}{\pi^2 B_0} \|\phi_{\mathbf{F}}\|^2_{H_h^2} \ge \frac{4}{\pi^2 B_0} \|\phi\|^2_{H_h^2}, \tag{A.85}$$

so that (2.15) (in Proposition 2.2.7) is proved with $C_1 = \frac{4}{\pi^2 B_0}$.

Inequality (2.16) could be proved in a similar way. The following fact is observed:

$$\|\phi\|_\infty \le \|\phi_{\mathbf{F}}\|_{L^\infty} \le C\|\phi_{\mathbf{F}}\|_{H_h^2} \le C\|\phi\|_{H_h^2}, \tag{A.86}$$

in which the first step is based on the fact that, ϕ is the grid interpolation of the continuous function $\phi_{\mathbf{F}}$, the second step comes from the Sobolev embedding, while the last step comes from the the estimates in Lemma A.4.1.

For the proof of (2.17), the last inequality in Proposition 2.2.7, the following lemma is needed, which gives a bound of the discrete ℓ^p (with $p = 4, 6$) norm of the grid functions ϕ and f, in terms of the continuous L^p norm of its continuous version $f_{\mathbf{F}}$.

Lemma A.4.2. *For $\phi \in \mathcal{C}_{\text{per}}$, $f \in \mathcal{V}_{\text{per}}$, we have*

$$\|\phi\|_p \le \sqrt{\frac{p}{2}}\|\phi_{\mathbf{F}}\|_{L^p}, \quad \|f\|_p \le \sqrt{\frac{p}{2}}\|f_{\mathbf{F}}\|_{L^p}, \quad \textit{with } p = 4, 6. \tag{A.87}$$

Proof. For simplicity of presentation, we only present the analysis for $\|f\|_p \le \sqrt{\frac{p}{2}}\|f_{\mathbf{F}}\|_{L^p}$; the analysis for ϕ could be carried out in the same fashion. And also, we are focused on the case of $p = 4$. The case with $p = 6$ could be handled in a similar, yet more tedious way.

We denote the following grid function

$$g_{i+\frac{1}{2},j+\frac{1}{2}} = \left(f_{i+\frac{1}{2},j+\frac{1}{2}}\right)^2. \tag{A.88}$$

A direct calculation shows that

$$\|f\|_4 = (\|g\|_2)^{\frac{1}{2}}. \tag{A.89}$$

Note that both norms are discrete in the above identity. Moreover, we assume the grid function g has a discrete Fourier expansion as

$$g_{i+\frac{1}{2},j+\frac{1}{2}} = \sum_{\ell,m=-K}^{K} (\hat{g}_c^N)_{\ell,m} \mathrm{e}^{2\pi \mathrm{i}(\ell x_{i+1/2} + m y_{j+\frac{1}{2}})}, \tag{A.90}$$

and denote its continuous version as

$$G(x,y) = \sum_{\ell,m=-K}^{K} (\hat{g}_c^N)_{\ell,m} \mathrm{e}^{2\pi \mathrm{i}(\ell x + m y)} \in \mathcal{P}_K. \tag{A.91}$$

With an application of the Parseval equality at both the discrete and continuous levels, we have

$$\|g\|_2^2 = \|G\|^2 = \sum_{\ell,m=-K}^{K} \left|(\hat{g}_c^N)_{\ell,m}\right|^2. \tag{A.92}$$

On the other hand, we also denote

$$H(x,y) = (f_{\mathbf{F}}(x,y))^2 = \sum_{\ell,m=-2K}^{2K} (\hat{h}^N)_{\ell,m} \mathrm{e}^{2\pi \mathrm{i}(\ell x + m y)} \in \mathcal{P}_{2K}. \tag{A.93}$$

The reason for $H \in \mathcal{P}_{2K}$ is because $f_{\mathbf{F}} \in \mathcal{P}_K$. We note that $H \neq G$, since $H \in \mathcal{P}_{2K}$, while $G \in \mathcal{P}_K$, although H and G have the same interpolation values on at the numerical grid points $(x_{i+\frac{1}{2}}, y_{j+\frac{1}{2}})$. In other words, g is the interpolation of H onto the numerical grid point and G is the continuous version of g in $\mathcal{P}_K$. As a result, collocation coefficients $\hat{g}_c^N$ for G are not equal to $\hat{h}^N$ for H, due to the aliasing error.

In more detail, for $-K \le \ell, m \le K$, we have the following representations:

$$(\hat{g}_c^N)_{\ell,m} = \begin{cases} (\hat{h}^N)_{\ell,m} + (\hat{h}^N)_{\ell+N,m} + (\hat{h}^N)_{\ell,m+N} + (\hat{h}^N)_{\ell+N,m+N}, \ \ell < 0, m < 0, \\ (\hat{h}^N)_{\ell,m} + (\hat{h}^N)_{\ell+N,m}, \ \ell < 0, m = 0, \\ (\hat{h}^N)_{\ell,m} + (\hat{h}^N)_{\ell+N,m} + (\hat{h}^N)_{\ell,m-N} + (\hat{h}^N)_{\ell+N,m-N}, \ \ell < 0, m > 0, \\ (\hat{h}^N)_{\ell,m} + (\hat{h}^N)_{\ell-N,m} + (\hat{h}^N)_{\ell,m-N} + (\hat{h}^N)_{\ell-N,m-N}, \ \ell > 0, m > 0, \\ (\hat{h}^N)_{\ell,m} + (\hat{h}^N)_{\ell-N,m}, \ \ell > 0, m = 0, \\ (\hat{h}^N)_{\ell,m} + (\hat{h}^N)_{\ell-N,m} + (\hat{h}^N)_{\ell,m+N} + (\hat{h}^N)_{\ell-N,m+N}, \ \ell > 0, m < 0, \\ (\hat{h}^N)_{\ell,m} + (\hat{h}^N)_{\ell,m+N}, \ \ell = 0, m < 0, \\ (\hat{h}^N)_{\ell,m}, \ \ell = 0, m = 0, \\ (\hat{h}^N)_{\ell,m} + (\hat{h}^N)_{\ell,m-N}, \ \ell = 0, m > 0. \end{cases} \tag{A.94}$$

With an application of Cauchy inequality, it is clear that

$$\sum_{\ell,m=-K}^{K} \left|(\hat{g}_c^N)_{\ell,m}\right|^2 \le 4 \left| \sum_{\ell,m=-2K}^{2K} (\hat{h}^N)_{\ell,m} \right|^2 . \tag{A.95}$$

Meanwhile, an application of Parseval's identity to the Fourier expansion (A.93) gives

$$\|H\|^2 = \left| \sum_{\ell,m=-2K}^{2K} (\hat{h}^N)_{\ell,m} \right|^2 . \tag{A.96}$$

Its comparison with (A.92) indicates that

$$\|g\|_2^2 = \|G\|^2 \le 4 \|H\|^2 , \quad \text{i.e. } \|g\|_2 \le 2 \|H\| , \tag{A.97}$$

with the estimate (A.95) applied. Meanwhile, since $H(x,y) = (f_{\mathbf{F}}(x,y))^2$, we have

$$\|f_{\mathbf{F}}\|_{L^4} = (\|H\|_{L^2})^{\frac{1}{2}} . \tag{A.98}$$

Therefore, a combination of (A.89), (A.97) and (A.98) results in

$$\|f\|_4 = (\|g\|_2)^{\frac{1}{2}} \leq (2\,\|H\|_{L^2})^{\frac{1}{2}} \leq \sqrt{2}\,\|f_{\mathbf{F}}\|_{L^4}\,. \tag{A.99}$$

This finishes the proof of (A.87) with $p = 4$, the inequality with $p = 6$ could be proved in the same fashion. □

Now we proceed into the proof of (2.17) in Proposition 2.2.7.

Proof. We begin with an application of (A.87) in Lemma A.4.2:

$$\|\mathfrak{D}_x\phi\|_6 = \|f\|_6 \leq \sqrt{3}\|f_{\mathbf{F}}\|_{L^6}. \tag{A.100}$$

Meanwhile, using the fact that $\overline{f_{\mathbf{F}}} = 0$, we apply the 2-D Sobolev inequality and get

$$\|f_{\mathbf{F}}\|_{L^6} \leq B_0^{(1)}\|f_{\mathbf{F}}\|_{H^1} \leq C(\|f_{\mathbf{F}}\| + \|\nabla f_{\mathbf{F}}\|). \tag{A.101}$$

Moreover, the estimates (A.61)-(A.63) (in Lemma A.4.1) indicate that

$$\|f_{\mathbf{F}}\| \leq \|\partial_x\phi_{\mathbf{F}}\| \leq \frac{\pi}{2}\|\nabla_h\phi\|_2, \tag{A.102}$$

$$\|\partial_x f_{\mathbf{F}}\| \leq \left\|\partial_x^2\phi_{\mathbf{F}}\right\| \leq M_0\,\|\Delta\phi_{\mathbf{F}}\| \leq \frac{\pi^2 M_0}{4}\|\Delta_h\phi\|_2, \tag{A.103}$$

$$\|\partial_y f_{\mathbf{F}}\| \leq \|\partial_x\partial_y\phi_{\mathbf{F}}\| \leq M_0\,\|\Delta\phi_{\mathbf{F}}\| \leq \frac{\pi^2 M_0}{4}\|\Delta_h\phi\|_2, \tag{A.104}$$

$$\text{so that}\quad \|f_{\mathbf{F}}\| + \|\nabla f_{\mathbf{F}}\| \leq \frac{\sqrt{2}\pi^2 M_0}{4}(\|\nabla_h\phi\|_2 + \|\Delta_h\phi\|_2), \tag{A.105}$$

in which the following elliptic regularity estimate is applied:

$$\left\|\partial_x^2\phi_{\mathbf{F}}\right\|, \|\partial_x\partial_y\phi_{\mathbf{F}}\| \leq M_0\,\|\Delta\phi_{\mathbf{F}}\|\,. \tag{A.106}$$

Therefore, a substitution of (A.103), (A.105) and (A.101) into (A.100) results in

$$\|\mathfrak{D}_x\phi\|_6 \leq \frac{\sqrt{6}\pi^2 M_0 B_0^{(1)}}{4}\|\phi\|_{H_h^2}. \tag{A.107}$$

The estimate for $\|D_y\phi\|_6$ could be derived in the same fashion:

$$\|\mathfrak{D}_y\phi\|_6 \leq \frac{\sqrt{6}\pi^2 M_0 B_0^{(1)}}{4}\|\phi\|_{H_h^2}. \tag{A.108}$$

As a consequence, (2.17) is valid, by setting $C = \sqrt{2}B_0^{(1)}$. The proof of Proposition 2.2.7 is complete. □

Publication

1. W. Feng, C. Wang, S.M. Wise and Z. Zhang. *Linearly preconditioned nonlinear conjugate gradient solvers for the epitaxial thin film equation with slope selection.* In preparation, 2017.

2. L. Dong, W. Feng, C. Wang, S.M. Wise and Z. Zhang. *Convergence analysis and numerical implementation of a second order numerical schemes for the three-dimensional phase field crystal equation.* arXiv preprint arXiv:1611.06288, 2016.

3. W. Chen, W. Feng, Y. Liu, C. Wang, S.M. Wise.*A Second Order Energy Stable Scheme for the Cahn-Hilliard-Hele-Shaw Equations.* arXiv preprint arXiv:1611.02967, 2016.

4. W. Feng, Z. Guan, J. Lowengrub, C. Wang, S.M. Wise and Y. Chen. *An energy stable finite-difference scheme for Functionalized Cahn-Hilliard equation and its convergence analysis.* arXiv preprint arXiv:1610.02473, 2016.

5. W. Feng, A.J. Salgado, C. Wang and S.M. Wise. *Preconditioned steepest descent methods for some nonlinear elliptic equations involving p-Laplacian terms.* J. Comput. Phys., 334:45–67, 2017.

6. W. Feng, T.L Lewis and S.M. Wise. *Discontinuous galerkin derivative operators with applications to second order elliptic problems and stability.* Math. meth. appl. sci. , 38(18):5160-5182, 2015.

Vita

Wenqiang Feng was born in Hubei province, China. He got the bachelor degree under the supervision of Prof. Kelong Zheng and Yiping Xu with thesis *The Application of Algebraic Multigrid Algorithms in Image Denoising* at Southwest University of Science and Technology in 2009. Then, he was recommended to Department of mathematics, University of Science and Technology of China. He got the master degree under the supervision of Prof. Shumin Li with thesis *The Application of the Non-Local Algorithm in Image Denoising* at Department of mathematics, University of science and technology of China in 2011. In May 2013, he received his Master Degree in Applied Mathematics from the Missouri University of science and Technology , Rolla, Missouri, USA. In the same year, he finished his thesis *Immersed finite element method for interface problems with algebraic multigrid solver* under the supervision of Prof. Xiaoming He. In August 2017, He will receive his Phd degree in Mathematics from the University of Tennessee, Knoxville under the advisement of Dr. Steven M. Wise.

As a student, besides taking the courses and participating in research projects, he has not given up any chance to improve himself. He also served as the president of the Chinese Scholar and Student Association at Missouri University of Science and Technology.