A Linear Iteration Algorithm for a Second-Order Energy Stable Scheme for a Thin Film Model without Slope Selection

Wenbin Chen

School of Mathematical Sciences, Fudan University

Shanghai, China 200433

Cheng Wang*

Department of Mathematics, University of Massachusetts Dartmouth

North Dartmouth, MA 02747-2300

Xiaoming Wang

Department of Mathematics, Florida State University

Tallahassee, FL 32306-4510

Steven M. Wise

Department of Mathematics, University of Tennessee

Knoxville, TN 37996-1300

March 3, 2013

Abstract

We present a linear iteration algorithm to implement a second-order energy stable numerical scheme for a model of epitaxial thin film growth without slope selection. The PDE, which is a nonlinear, fourth-order parabolic equation, is the L^2 gradient flow of the energy $\int_\Omega \left(-\frac{1}{2} \ln \left(1 + |\nabla \phi|^2 \right) + \frac{\epsilon^2}{2} |\Delta \phi(\mathbf{x})|^2 \right) d\mathbf{x}$. The energy stability is preserved by a careful choice of the second-order temporal approximation for the nonlinear term, as reported in recent work [18]. The resulting scheme is highly nonlinear, and its implementation is non-trivial. In this paper, we propose a linear iteration algorithm to solve the resulting nonlinear system. To accomplish this we introduce an $O(s^2)$ (with s the time step size) artificial diffusion term, a Douglas-Dupont-type regularization, that leads to a contraction mapping property. As a result, the highly nonlinear system can be decomposed as an iteration of purely linear solvers, which can be very efficiently implemented with the help of FFT in a collocation Fourier spectral setting. We present a careful analysis showing convergence for the numerical scheme in a discrete $L^\infty(0, T; H^1) \cap L^2(0, T; H^3)$ norm. Some numerical simulation results are presented to demonstrate the efficiency of the linear iteration solver and the convergence of the scheme as a whole.

keywords. epitaxial thin film growth, slope selection, energy stability, linear iteration, contraction mapping, Fourier collocation spectral

1 Introduction

In this article we consider an efficient numerical implementation of a second order accurate and energy stable scheme for an epitaxial thin film growth model. (See the review [4] for the recent

---

*Corresponding author, cwang1@umassd.edu

history of such models of thin film growth.) The equation is the gradient flow associated with the following *energy* functional

$$E(\phi) := \int_{\Omega} \left(-\frac{1}{2} \ln \left(1 + |\nabla\phi|^2\right) + \frac{\epsilon^2}{2}|\Delta\phi|^2 \right) d\mathbf{x}, \tag{1}$$

where $\Omega = [0, L_x] \times [0, L_y]$, $\phi : \Omega \to \mathbb{R}$ is a periodic height function, and ϵ is a constant. We note that the first term, which is clearly non-quadratic, represents the Ehrlich-Schwoebel effect, according to which migrating adatoms must overcome a higher energy barrier to stick to a step from an upper rather than from a lower terrace [3, 10, 11, 17]. This results in an uphill atom current in the dynamics and the steepening of mounds in the film. The second term, which is quadratic, but of higher-order, represents the isotropic surface diffusion effect [11, 13]. For the Ehrlich-Schwoebel term we will use the notation

$$E_{c,1}(\phi) = \int_{\Omega} F_{c,1}(\nabla\phi) \, d\mathbf{x}, \quad F_{c,1}(\mathbf{y}) = -\frac{1}{2} \ln \left(1 + |\mathbf{y}|^2\right), \tag{2}$$

where $\mathbf{y} \in \mathbb{R}^2$ and $|\mathbf{y}| = \sqrt{y_1^2 + y_2^2}$. Hence, $E(\phi) = E_{c,1}(\phi) + \frac{\epsilon^2}{2} \|\Delta\phi\|^2$, where $\|\cdot\|$ denotes the L^2 norm. See, for example, [7]. Note that $F_{c,1} : \mathbb{R}^2 \to \mathbb{R}$ is bounded above by 0 and $F_{c,1} \to -\infty$ as $|\mathbf{y}| \to \infty$. Since $F_{c,1}$ has no relative minima, there are no energetically favored values for $|\nabla\phi|$. This implies that there will be no mechanism in any energy-gradient dynamics model that could select a preferred slope of the mounds. See the relevant discussions in [8, 9, 11, 12, 21].

The chemical potential is defined to be the variational derivative of the energy (1), *i.e.*,

$$\mu := \delta_\phi E = \nabla \cdot \left(\frac{\nabla\phi}{1 + |\nabla\phi|^2} \right) + \epsilon^2 \Delta^2 \phi, \tag{3}$$

assuming such boundary conditions as make the boundary integral vanish. Herein we consider the L^2 gradient flow:

$$\partial_t \phi = -\mu = -\nabla \cdot \left(\frac{\nabla\phi}{1 + |\nabla\phi|^2} \right) - \epsilon^2 \Delta^2 \phi, \tag{4}$$

where the boundary conditions for the height function ϕ are taken to be Ω-periodic. We refer to (4) as the no-slope-selection equation, following most other references. Equation (4) may be rewritten in the form

$$\partial_t \phi = \nabla \cdot \left(\frac{|\nabla\phi|^2}{1 + |\nabla\phi|^2} \nabla\phi \right) - \Delta\phi - \epsilon^2 \Delta^2 \phi. \tag{5}$$

In the small-slope regime, where $|\nabla\phi|^2 \ll 1$, (5) may be replaced by

$$\partial_t \phi = \nabla \cdot \left(|\nabla\phi|^2 \nabla\phi \right) - \Delta\phi - \epsilon^2 \Delta^2 \phi, \tag{6}$$

which we refer to as the slope-selection equation [8, 9, 11, 13]. Solutions to Eq. (6), unlike those of Eq. (4), exhibit pyramidal structures, where the faces of the pyramids have slopes $|\nabla\phi| \approx 1$. Solutions to the no-slope-selection equation (4), on the other hand, exhibit mound-like structures, the slopes of which (on an infinite domain) may grow unbounded [11, 21]. Solutions of (6) and (4) have up-down symmetry in the sense that there is no way to distinguish a hill from a valley. This can be altered by adding adsorption/desorption or other dynamics. Both the slope selection (6) and the no-slope selection (4) models are shown to be well posed in [11].

Numerical simulations of the model (4) with high order accuracy and energy stability have attracted a great deal of attention in recent years. In the paper of Li and Liu [11], a classical second order accurate semi-implicit numerical scheme (implicit for the linear term, explicit for the nonlinear term) is applied. The numerical results showed reasonable stability, though the

theoretical justification for such a scheme may not be easily available. Besides our own work, there have been other efforts to devise and analyze schemes for the slope selection (6) and no-slope selection (4) equations. Numerical schemes and analyses for the slope selection equation (6) can be found in [2, 15, 16, 22]. Recent numerical methods and analyses for the model without slope selection (4) can be found in [2, 14].

In [21], the authors studied an unconditionally energy stable scheme. The derivation of the scheme is based on the idea of a convex splitting of the energy (1) into a purely convex part and a purely concave part, motivated by Eyre's earlier work [5]. But there are two shortcomings of the scheme in [21]: it is only first order accurate (in time), and it is highly nonlinear, due to the implicit treatment of the nonlinear term. In a more recent work [1], we introduced an efficient linear, unconditionally stable, unconditionally solvable scheme for approximating solutions to the no-slope-selection equation (4). The key idea of this linear scheme is an alternate, and more advantageous, way of decomposing the energy into convex and concave terms, so that the nonlinear part of the chemical potential is placed in the concave part instead of the convex part. As a result, the implicit part of the chemical potential is completely linear. And numerical efficiency is greatly improved (over the scheme in [21]) due to the fact that the linear operator involved in the scheme, which is positive elliptic with constant coefficients, can be efficiently inverted by FFT.

It should be noted that the linear scheme in [1] is only first-order accurate in time. Second-order (in time) accurate and energy stable schemes are discussed and analyzed in detailed in another recent paper [18], for both the slope selection (6) and no-slope-selection models (4). These second order schemes come in two varieties, those that inherit the variational structure of the original continuous-in-time gradient flow or those that do not. But in either case, the unconditional energy stability and unique nonlinear solvability are established in detail.

Meanwhile, these second order schemes are highly nonlinear, and their numerical implementations are highly non-trivial. In [18], we performed a second order accurate numerical simulation of the slope-selection model (6), using the approach without preserving the variational structure. For the slope-selection model, the nonlinear term is still in the polynomial format so that a nonlinear conjugate gradient solver can be efficiently applied. However, for the no-slope-selection model (4), the numerical difficulty associated with the high degree of nonlinearity is much more prominent, due to the complicated terms appearing in the fractional quotients, either with or without the preservation of the variational structures.

As a result, the numerical implementation of the second order accurate and unconditionally energy stable scheme for the no-slope-selection model (4) has become a very challenging problem. In this paper, we present an efficient linear iteration solver to implement it, with an introduction of second order accurate $O(s^2)$ artificial diffusion term in the form of Douglas-Dupont regularization. In turn, although the numerical scheme itself is highly nonlinear, we treat the nonlinear term explicitly at each iteration stage. Moreover, by a careful nonlinear analysis and using a subtle estimate of the functional bound for the nonlinear quotient term in the no-slope-selection model, a contraction mapping property is theoretically justified if a parameter associated with the artificial diffusion coefficient is greater than a given constant $\frac{5}{4}$. In other words, the highly nonlinear numerical scheme can be very efficiently solved by such a linear iteration algorithm, and a geometric convergence rate is assured for this linear iteration under the given constraint. Similar to the linear splitting scheme reported in [1], the linear operator involved in the scheme, herein denoted $\mathcal{L} : H_{per}^2 \to \left(H_{per}^2\right)^*$, is positive elliptic with constant coefficients, and it can be efficiently inverted at the discrete level by FFT or other existing fast linear solvers.

In addition to the unconditional energy stability and unique solvability of this second order scheme, its convergence analysis is also non-trivial, due to the complicated nonlinear terms in a quotient form. In this paper, we also provide a detailed convergence analysis and numerical error estimate in H^1 norm for this fully discrete scheme (with Fourier spectral differentiation in space),

with a fixed final time. The convergence analysis follows the standard procedure of consistency and stability estimates. In the error estimate, the key part is the control for the nonlinear error terms, in which the bound can be obtained with the help of the specific quotient form in the no-slope-selection model. Subsequently, a careful application of summation by parts in Fourier spectral space leads to the convergence of the numerical scheme in a discrete $L^\infty(0, T; H^1) \cap L^2(0, T; H^3)$ norm.

The rest of the manuscript is organized as follows. In Section 2 we present the numerical scheme. First we recall a second order convex splitting scheme for the no-slope-selection model (4) with unconditional energy stability and unique solvability, as reported in [18]. Then we propose an $O(s^2)$ artificial diffusion term in the form of a Douglas-Dupont-type regularization, and a linear iteration algorithm to implement it. We show that the unconditional energy stability is preserved for such an addition of artificial diffusion, and the corresponding linear iteration algorithm is assured to be a contraction mapping under a condition for the artificial diffusion constant. In Section 3 we present the fully discrete scheme, where Fourier spectral differentiation is utilized in space. The convergence analysis is provided in Section 4. In Section 5 we present some numerical simulation results. We offer our concluding remarks in Section 6.

2 The Numerical Scheme

2.1 A Second-Order Convex Splitting Scheme

Second order accurate convex splitting schemes for the epitaxial thin film growth models, with and without slope selection(see (6) and (4)), were studied in detail in a recent article [18]. Two variants were presented: those schemes that preserve the variational structure of the gradient flow, and those that do not. For simplicity, we recall the one without variational structure. The extension of our method to the scheme with the variational structure can be carried out in a similar manner.

First, consider the energy decomposition given by $E(\phi) = E_{c,1} + E_{c,2} - E_e$, where

$$
E_{c,1} \;=\; -\frac{1}{2}\int_\Omega \ln\left(1 + |\nabla\phi|^2\right)d\mathbf{x}, \quad E_{c,2} = \frac{1}{2}\int_\Omega \left\{ A\,|\nabla\phi|^2 + \left(\epsilon^2 + B\right)(\Delta\phi)^2 \right\}d\mathbf{x}, \tag{7}
$$

$$
E_e \;=\; \frac{1}{2}\int_\Omega \left\{ A|\nabla\phi|^2 + B(\Delta\phi)^2 \right\}d\mathbf{x}, \tag{8}
$$

where A and B are artificial splitting parameters. It is straight forward to show that the *contractive part*, $E_c := E_{c,1} + E_{c,2}$ is convex, provided $A \geq 1$ and $B \geq 0$. The *expansive part*, E_e, is convex for any $A, B \geq 0$. Our second-order scheme will respect this convex splitting, and in doing so we can guarantee unconditional stability and unconditional uni-solvency of the scheme, as we show below.

To describe the scheme, we first focus on the treatment of the logarithmic, non-quadratic part of the energy density. For the logarithmic term, set $\mathsf{L}(x, y) = \ln\left(1 + x^2 + y^2\right)$. Suppose that $\boldsymbol{\psi}, \boldsymbol{\phi} \in \mathbb{R}^2$ are given, with $\boldsymbol{\psi} = (\psi_x, \psi_y)^T$ and $\boldsymbol{\phi} = (\phi_x, \phi_y)^T$. Define

$$
\mathsf{L}_x(\boldsymbol{\psi}, \boldsymbol{\phi}) \;=\; \begin{cases} -\dfrac{1}{4}\dfrac{\mathsf{L}(\psi_x, \psi_y) - \mathsf{L}(\phi_x, \psi_y)}{\psi_x - \phi_x} - \dfrac{1}{4}\dfrac{\mathsf{L}(\psi_x, \phi_y) - \mathsf{L}(\phi_x, \phi_y)}{\psi_x - \phi_x}, & \text{if} \quad \psi_x \neq \phi_x, \\[4mm] -\dfrac{1}{2}\dfrac{\psi_x}{1 + (\psi_x)^2 + (\psi_y)^2} - \dfrac{1}{2}\dfrac{\psi_x}{1 + (\psi_x)^2 + (\phi_y)^2}, & \text{if} \quad \psi_x = \phi_x, \end{cases} \tag{9}
$$

$$
\mathsf{L}_y(\boldsymbol{\psi}, \boldsymbol{\phi}) \;=\; \begin{cases} -\dfrac{1}{4}\dfrac{\mathsf{L}(\psi_x, \psi_y) - \mathsf{L}(\psi_x, \phi_y)}{\psi_y - \phi_y} - \dfrac{1}{4}\dfrac{\mathsf{L}(\phi_x, \psi_y) - \mathsf{L}(\phi_x, \phi_y)}{\psi_y - \phi_y}, & \text{if} \quad \psi_y \neq \phi_y, \\[4mm] -\dfrac{1}{2}\dfrac{\psi_y}{1 + (\psi_x)^2 + (\psi_y)^2} - \dfrac{1}{2}\dfrac{\psi_y}{1 + (\phi_x)^2 + (\psi_y)^2}, & \text{if} \quad \psi_y = \phi_y. \end{cases} \tag{10}
$$

Now, define

$$\mu_{c,1}^{n+1/2} := \mu_{c,1}(\phi^{n+1}, \phi^n) := -\partial_x \mathsf{L}_x \left(\nabla \phi^{n+1}, \nabla \phi^n \right) - \partial_y \mathsf{L}_y \left(\nabla \phi^{n+1}, \nabla \phi^n \right). \tag{11}$$

The second order convex splitting scheme for the no slope selection model (4) is given by

$$\begin{aligned} \frac{\phi^{n+1} - \phi^n}{s} &= -\mu_{c,1}^{n+1/2} - \frac{\epsilon^2}{2} \Delta^2 \left(\phi^{n+1} + \phi^n \right) + \frac{A}{2} \Delta \left(\phi^{n+1} + \phi^n \right) - A\Delta \left(\frac{3}{2} \phi^n - \frac{1}{2} \phi^{n-1} \right) \\ &\quad - \frac{B}{2} \Delta^2 \left(\phi^{n+1} + \phi^n \right) + B\Delta^2 \left(\frac{3}{2} \phi^n - \frac{1}{2} \phi^{n-1} \right), \end{aligned} \tag{12}$$

in which s is the time step size. In respecting the convex splitting introduced above, we have treated the contribution to the chemical potential from the contractive part, E_c, using an implicit Crank-Nicholson/secant approximation, and the contribution from the expansive part, E_e, using an explicit Adams-Bashforth approximation.

Employing the notation

$$\widetilde{\Delta}_s \phi^n := \phi^{n+1} - 2\phi^n + \phi^{n-1}, \tag{13}$$

the scheme writes as

$$\frac{\phi^{n+1} - \phi^n}{s} = -\mu_{c,1}^{n+1/2} - \frac{\epsilon^2}{2} \Delta^2 \left(\phi^{n+1} + \phi^n \right) + \frac{A}{2} \Delta \left(\widetilde{\Delta}_s \phi^n \right) - \frac{B}{2} \Delta^2 \left(\widetilde{\Delta}_s \phi^n \right). \tag{14}$$

Theorem 2.1. *For any $A \geq 0$ and any $B \geq 0$ the scheme (14) is second order, i.e., its local truncation error is $O\left(s^2\right)$, and is unconditionally strongly energy stable with respect to the discrete energy*

$$\mathcal{E}(\phi^n, \phi^{n-1}) := E(\phi^n) + \frac{A}{4} \left\| \nabla \left(\phi^n - \phi^{n-1} \right) \right\|_{L^2}^2 + \frac{B}{4} \left\| \Delta \left(\phi^n - \phi^{n-1} \right) \right\|_{L^2}^2, \tag{15}$$

i.e., $\mathcal{E}(\phi^{n+1}, \phi^n) \leq \mathcal{E}(\phi^n, \phi^{n-1})$, for any $n \geq 1$, and any $s > 0$. Furthermore, if $A \geq 1$ the scheme is unconditionally uniquely solvable.

Proof. The assertion regarding the order of the method can be verified by Taylor expansions, assuming sufficient regularity. We omit the details for brevity. For energy stability, we use the identities

$$-\left(\phi^{n+1} - \phi^n, \mu_{c,1}^{n+1/2} \right) = -E_{c,1}(\phi^{n+1}) + E_{c,1}(\phi^n), \tag{16}$$

which follows from the careful construction of $\mu_{c,1}^{n+1/2}$ [18],

$$\begin{aligned} \frac{A}{2} \left(\phi^{n+1} - \phi^n, \Delta \left(\widetilde{\Delta}_s \phi^n \right) \right) &= -\frac{A}{4} \left\| \nabla \left(\phi^{n+1} - \phi^n \right) \right\|_{L^2}^2 + \frac{A}{4} \left\| \nabla \left(\phi^n - \phi^{n-1} \right) \right\|_{L^2}^2 \\ &\quad - \frac{A}{4} \left\| \nabla \left(\widetilde{\Delta}_s \phi^n \right) \right\|_{L^2}^2, \end{aligned} \tag{17}$$

$$\begin{aligned} -\frac{B}{2} \left(\phi^{n+1} - \phi^n, \Delta^2 \left(\widetilde{\Delta}_s \phi^n \right) \right) &= -\frac{B}{4} \left\| \nabla \left(\phi^{n+1} - \phi^n \right) \right\|_{L^2}^2 + \frac{B}{4} \left\| \Delta \left(\phi^n - \phi^{n-1} \right) \right\|_{L^2}^2 \\ &\quad - \frac{B}{4} \left\| \Delta \left(\widetilde{\Delta}_s \phi^n \right) \right\|_{L^2}^2, \end{aligned} \tag{18}$$

and

$$-\frac{\epsilon^2}{2} \left(\phi^{n+1} - \phi^n, \Delta^2 \left(\phi^{n+1} + \phi^n \right) \right) = -\frac{\epsilon^2}{2} \left\| \Delta \phi^{n+1} \right\|_{L^2}^2 + \frac{\epsilon^2}{2} \left\| \Delta \phi^n \right\|_{L^2}^2. \tag{19}$$

Testing the scheme (14) with $\phi^{n+1} - \phi^n$, using the identities above, and rearranging terms, we have

$$\mathcal{E}\left(\phi^{n+1}, \phi^n \right) + \frac{A}{4} \left\| \nabla \left(\widetilde{\Delta}_s \phi^n \right) \right\|_{L^2}^2 + \frac{B}{4} \left\| \Delta \left(\widetilde{\Delta}_s \phi^n \right) \right\|_{L^2}^2 + s \left\| \frac{\phi^{n+1} - \phi^n}{s} \right\|_{L^2}^2 = \mathcal{E}\left(\phi^n, \phi^{n-1} \right), \tag{20}$$

which proves the unconditional energy stability. The assertion regarding the unconditional solvability follows from similar arguments in [18], and we omit the details here. $\square$

2.2 A Linear Iteration Scheme

The unconditional energy stability and unique solvability of the second order scheme (14) having been established, it remains to efficiently "invert" the method at each time step. However, the numerical implementation of the scheme is clearly a challenge, due to its highly nonlinear nature. In this section, we propose a linear iteration method to solve the scheme, and prove that the iteration in question always converges to the unique solution of (14) if the splitting parameter A is chosen judiciously.

From here let us fix the value of one of the splitting parameters, setting $B = \frac{\epsilon^2}{2}$. The scheme (14) then becomes

$$\frac{\phi^{n+1} - \phi^n}{s} = -\mu_{c,1}^{n+1/2} + \frac{A}{2}\Delta\left(\widetilde{\Delta}_s\phi^n\right) - \epsilon^2\Delta^2\left(\frac{3}{4}\phi^{n+1} + \frac{1}{4}\phi^{n-1}\right). \tag{21}$$

The splitting parameter A will be chosen later; for now we only assume that $A \geq 1$, so that the scheme enjoys unconditional solvability. The purpose for taking $B > 0$ is to place heavier weight on the highest order linear term at the time step t^{n+1}. This has the effect to simplify the convergence analysis that will follow later. Our choice of B fixes the discrete energy for the scheme to be

$$\mathcal{E}(\phi^n, \phi^{n-1}) := E(\phi^n) + \frac{A}{4}\left\|\nabla\left(\phi^n - \phi^{n-1}\right)\right\|_{L^2}^2 + \frac{\epsilon^2}{8}\left\|\Delta\left(\phi^n - \phi^{n-1}\right)\right\|_{L^2}^2. \tag{22}$$

Now, the scheme (21) can be rewritten as

$$\mathcal{L}\left(\phi^{n+1}\right) := \left(\frac{1}{s}I - \frac{A}{2}\Delta + \frac{3\epsilon^2}{4}\Delta^2\right)\phi^{n+1} = -\mu_{c,1}(\phi^{n+1}, \phi^n) + F\left(\phi^n, \phi^{n-1}\right), \tag{23}$$

where

$$F\left(\phi^n, \phi^{n-1}\right) := \frac{1}{s}\phi^n + \frac{A}{2}\Delta\left(-2\phi^n + \phi^{n-1}\right) - \frac{\epsilon^2}{4}\Delta^2\phi^{n-1}. \tag{24}$$

Recall, $\mu_{c,1}(\phi^{n+1}, \phi^n) = \mu_{c,1}^{n+1/2}$, as defined in (11). $\mathcal{L}$ is a positive, linear, constant coefficient differential operator.

Now, we propose the following linear iteration method to solve the scheme (23): given ϕ^n, ϕ^{n-1}, and ψ^k smooth enough and periodic, find the unique periodic solution, ψ^{k+1}, that satisfies

$$\mathcal{L}\left(\psi^{k+1}\right) = -\mu_{c,1}(\psi^k, \phi^n) + F\left(\phi^n, \phi^{n-1}\right). \tag{25}$$

Here k stands for the iteration index, not the time step index. The method is initialized via $\psi^0 := \phi^n$. Clearly, $\psi = \phi^{n+1}$ is the unique fixed point solution:

$$\mathcal{L}(\psi) = -\mu_{c,1}(\psi, \phi^n) + F\left(\phi^n, \phi^{n-1}\right). \tag{26}$$

We now prove that the linear fixed point iteration (25) must converge, and therefore to the unique fixed point, provided A is sufficiently large.

Theorem 2.2. *The linear iteration (25) is a contraction mapping provided that* $\alpha := \frac{\sqrt{3}\epsilon}{\sqrt{s}} + \frac{A}{2} > \frac{5}{4}$.

Two preliminary estimates are needed for the proof the theorem and are given in the next two lemmas.

Lemma 2.3. *Let $a, b \in \mathbb{R}$ be arbitrary but fixed. Define $h_1 : \mathbb{R} \to \mathbb{R}$ via*

$$h_1(x) = \begin{cases} \dfrac{\mathsf{L}(x, a) - \mathsf{L}(b, a)}{x - b}, & \text{if } x \neq b, \\[2ex] \dfrac{2b}{1 + b^2 + a^2}, & \text{if } x = b. \end{cases} \tag{27}$$

Then $h_1 \in C^1(\mathbb{R})$ and $|h_1'(x)| \leq 2$, for all $x \in \mathbb{R}$.

Proof. Define $f(x) := \mathsf{L}(x, a) = \ln(1 + x^2 + a^2)$ for any $x \in \mathbb{R}$. It is clear that $h_1(x) = \frac{f(x) - f(b)}{x - b}$ for $x \neq b$ and h_1 is at least twice continuously differentiable for all such x. A direct calculation yields

$$h_1'(x) = \frac{f'(x) \cdot (x - b) - (f(x) - f(b))}{(x - b)^2} = \frac{f'(x) - \frac{f(x) - f(b)}{x - b}}{x - b}. \tag{28}$$

Without loss of generality, suppose $x < b$. An application of the mean value theorem gives

$$\frac{f(x) - f(b)}{x - b} = f'(\xi), \tag{29}$$

for some $\xi \in (x, b)$. We obtain

$$h_1'(x) = \frac{f'(x) - f'(\xi)}{x - b} = \frac{f''(\eta)(x - \xi)}{x - b}, \tag{30}$$

for some $\eta \in (x, \xi)$, where we have invoked the mean value theorem a second time in the last step. We clearly have $|x - \xi| \leq |x - b|$, and, as a result, we arrive at

$$|h_1'(x)| \leq |f''(\eta)| \leq \max_{\eta \in (x, b)} |f''(\eta)|. \tag{31}$$

Meanwhile, we have

$$f'(x) = \frac{2x}{1 + x^2 + a^2}, \quad f''(x) = \frac{2(1 - x^2 + a^2)}{1 + x^2 + a^2}, \quad \text{and} \quad \max_{x \in \mathbb{R}} |f''(x)| \leq 2. \tag{32}$$

The result then follows from (31).

Now, a detailed Taylor expansion for f shows that

$$f(x) = \ln(1 + x^2 + a^2) = f(b) + \frac{2b}{1 + b^2 + a^2} \cdot (x - b) + \frac{1 - \zeta^2 + a^2}{(1 + \zeta^2 + a^2)^2} \cdot (x - b)^2, \tag{33}$$

for some $\zeta = \zeta(x)$ between x and b. This in turn yields

$$h_1(x) = \frac{f(x) - f(b)}{x - b} = \frac{2b}{1 + b^2 + a^2} + \frac{1 - \zeta^2 + a^2}{(1 + \zeta^2 + a^2)^2} \cdot (x - b), \tag{34}$$

so that

$$\frac{h_1(x) - h_1(b)}{x - b} = \frac{1 - \zeta^2 + a^2}{(1 + \zeta^2 + a^2)^2}. \tag{35}$$

Since $\zeta \to b$ as $x \to b$, we have

$$h_1'(b) = \frac{1 - b^2 + a^2}{(1 + b^2 + a^2)^2}. \tag{36}$$

From here it is easy to show that $|h_1'(b)| \leq 1$, and the result is proven. $\qquad \square$

Lemma 2.4. *Let $a, b \in \mathbb{R}$ be arbitrary but fixed. Define $h_2 : \mathbb{R} \to \mathbb{R}$ via*

$$h_2(x) = \begin{cases} \dfrac{\mathsf{L}(a, x) - \mathsf{L}(b, x)}{a - b}, & \text{if } a \neq b, \\[2ex] \dfrac{2a}{1 + a^2 + x^2}, & \text{if } a = b. \end{cases} \tag{37}$$

Then $h_2 \in C^1(\mathbb{R})$ and $|h_2'(x)| \leq 1$, for all $x \in \mathbb{R}$.

Proof. If $a \neq b$, a direct calculation shows that

$$h_2'(x) = \frac{-2x(a + b)}{(1 + a^2 + x^2)(1 + b^2 + x^2)}. \tag{38}$$

From this, it is straightforward to show that $|h_2'(x)| \leq 1$ for all $x, a, b \in \mathbb{R}$. On the other hand, if $a = b$, we have

$$h_2'(x) = \frac{-4xa}{(1 + a^2 + x^2)^2}. \tag{39}$$

It easily follows that $|h_2'(x)| \leq 1$ for all $x, a \in \mathbb{R}$, and the result is proven. $\qquad\square$

We now proceed to prove the theorem.

Proof of Theorem 2.2. Let $\psi \in H_{\mathrm{per}}^4(\Omega)$ be the unique solution to (26) and define the iteration error at each stage via

$$e^k := \psi^k - \psi, \tag{40}$$

where $\psi^k \in H_{\mathrm{per}}^4(\Omega)$ is the kth iterate generated by the linear iteration scheme (25). Subtracting (26) from (25) yields

$$\mathcal{L}\left(e^{k+1}\right) = \nabla \cdot \begin{pmatrix} \mathsf{L}_x \left(\nabla \psi^k, \nabla \phi^n\right) - \mathsf{L}_x \left(\nabla \psi, \nabla \phi^n\right) \\ \mathsf{L}_y \left(\nabla \psi^k, \nabla \phi^n\right) - \mathsf{L}_y \left(\nabla \psi, \nabla \phi^n\right) \end{pmatrix}. \tag{41}$$

Taking the inner product with e^{k+1} leads to

$$\begin{aligned} \left(\mathcal{L}\left(e^{k+1}\right), e^{k+1}\right) &= \frac{1}{s} \left\|e^{k+1}\right\|^2 + \frac{A}{2} \left\|\nabla e^{k+1}\right\|^2 + \frac{3\epsilon^2}{4} \left\|\Delta e^{k+1}\right\|^2 \\ &= -\left(\mathsf{L}_x \left(\nabla \psi^k, \nabla \phi^n\right) - \mathsf{L}_x \left(\nabla \psi, \nabla \phi^n\right), \partial_x e^{k+1}\right) \\ &\quad - \left(\mathsf{L}_y \left(\nabla \psi^k, \nabla \phi^n\right) - \mathsf{L}_y \left(\nabla \psi, \nabla \phi^n\right), \partial_y e^{k+1}\right). \end{aligned} \tag{42}$$

Now, define

$$\begin{aligned} \mathcal{N}_1 &:= \frac{\mathsf{L}\left(\partial_x \psi^k, \partial_y \psi^k\right) - \mathsf{L}\left(\partial_x \phi^n, \partial_y \psi^k\right)}{\partial_x \psi^k - \partial_x \phi^n} - \frac{\mathsf{L}\left(\partial_x \psi, \partial_y \psi\right) - \mathsf{L}\left(\partial_x \phi^n, \partial_y \psi\right)}{\partial_x \psi - \partial_x \phi^n}, \\ \mathcal{N}_2 &:= \frac{\mathsf{L}\left(\partial_x \psi^k, \partial_y \phi^n\right) - \mathsf{L}\left(\partial_x \phi^n, \partial_y \phi^n\right)}{\partial_x \psi^k - \partial_x \phi^n} - \frac{\mathsf{L}\left(\partial_x \psi, \partial_y \phi^n\right) - \mathsf{L}\left(\partial_x \phi^n, \partial_y \phi^n\right)}{\partial_x \psi - \partial_x \phi^n}. \end{aligned} \tag{43}$$

Then, with L_x as in (9), we have

$$\mathsf{L}_x \left(\nabla \psi^k, \nabla \phi^n\right) - \mathsf{L}_x \left(\nabla \psi, \nabla \phi^n\right) = -\frac{1}{4} \left(\mathcal{N}_1 + \mathcal{N}_2\right). \tag{44}$$

The second nonlinear error term $\mathcal{N}_2$ can be represented as

$$\mathcal{N}_2 = h_1 \left(\partial_x \psi^k\right) - h_1 \left(\partial_x \psi\right), \quad \text{with} \quad a = \partial_y \phi^n, \quad b = \partial_x \phi^n. \tag{45}$$

8

An application of the mean value theorem and Lemma 2.3 yields point-wise estimate

$$\left|\mathcal{N}_2(\mathbf{x})\right| \leq \left|h_1'(\xi(\mathbf{x}))\right| \cdot \left|\partial_x \psi^k(\mathbf{x}) - \partial_x \psi(\mathbf{x})\right| \leq 2\left|\partial_x e^k(\mathbf{x})\right|, \quad \forall \mathbf{x} \in \overline{\Omega}, \tag{46}$$

where $\xi(\mathbf{x})$ is between $\partial_x \psi_x^k(\mathbf{x})$ and $\partial_x \psi(\mathbf{x})$.

Note that for the first nonlinear error term $\mathcal{N}_1$, neither Lemma 2.3 nor Lemma 2.4 can be applied directly. Instead, we perform the following, further decomposition: $\mathcal{N}_1 = \mathcal{N}_3 + \mathcal{N}_4$, with

$$\mathcal{N}_3 := \frac{\mathsf{L}\left(\partial_x \psi^k, \partial_y \psi\right) - \mathsf{L}\left(\partial_x \phi^n, \partial_y \psi\right)}{\partial_x \psi^k - \partial_x \phi^n} - \frac{\mathsf{L}\left(\partial_x \psi, \partial_y \psi\right) - \mathsf{L}\left(\partial_x \phi^n, \partial_y \psi\right)}{\partial_x \psi - \partial_x \phi^n},$$

$$\mathcal{N}_4 := \frac{\mathsf{L}\left(\partial_x \psi^k, \partial_y \psi^k\right) - \mathsf{L}\left(\partial_x \phi^n, \partial_y \psi^k\right)}{\partial_x \psi^k - \partial_x \phi^n} - \frac{\mathsf{L}\left(\partial_x \psi^k, \partial_y \psi\right) - \mathsf{L}\left(\partial_x \phi^n, \partial_y \psi\right)}{\partial_x \psi^k - \partial_x \phi^n}. \tag{47}$$

Observe that $\mathcal{N}_3$ and $\mathcal{N}_4$ can be represented as

$$\mathcal{N}_3 = h_1\left(\partial_x \psi^k\right) - h_1\left(\partial_x \psi\right), \quad \text{with} \quad a = \partial_y \psi, \quad b = \partial_x \phi^n, \tag{48}$$

$$\mathcal{N}_4 = h_2\left(\partial_y \psi^k\right) - h_2\left(\partial_y \psi\right), \quad \text{with} \quad a = \partial_x \psi^k, \quad b = \partial_x \phi^n. \tag{49}$$

Now, applications of the mean value theorem and Lemmas 2.3 and 2.4 result in the point-wise estimates

$$\left|\mathcal{N}_3(\mathbf{x})\right| \leq \left|h_1'\left(\xi_1(\mathbf{x})\right)\right| \cdot \left|\partial_x \psi^k(\mathbf{x}) - \partial_x \psi(\mathbf{x})\right| \leq 2\left|\partial_x e^k(\mathbf{x})\right|, \quad \forall \mathbf{x} \in \overline{\Omega}, \tag{50}$$

$$\left|\mathcal{N}_4(\mathbf{x})\right| \leq \left|h_2'\left(\xi_2(\mathbf{x})\right)\right| \cdot \left|\partial_y \psi^k(\mathbf{x}) - \partial_y \psi(\mathbf{x})\right| \leq \left|\partial_y e^k(\mathbf{x})\right|, \quad \forall \mathbf{x} \in \overline{\Omega}. \tag{51}$$

Substitution of (46), (50), and(51) into (44) yields the point-wise estimate

$$\left|\mathsf{L}_x\left(\nabla \psi^k, \nabla \phi^n\right) - \mathsf{L}_x\left(\nabla \psi, \nabla \phi^n\right)\right| \leq \left|\partial_x e^k\right| + \frac{1}{4}\left|\partial_y e^k\right|. \tag{52}$$

With an application of Cauchy's inequality, we arrive at

$$-\left(\mathsf{L}_x\left(\nabla \psi^k, \nabla \phi^n\right) - \mathsf{L}_x\left(\nabla \psi, \nabla \phi^n\right), \partial_x e^{k+1}\right) \leq \left(\left|\partial_x e^k\right|, \left|\partial_x e^{k+1}\right|\right) + \frac{1}{4}\left(\left|\partial_y e^k\right|, \left|\partial_x e^{k+1}\right|\right)$$
$$\leq \frac{1}{2}\left\|\partial_x e^k\right\|^2 + \frac{1}{8}\left\|\partial_y e^k\right\|^2 + \frac{5}{8}\left\|\partial_x e^{k+1}\right\|^2. \tag{53}$$

The second component of the nonlinear error term associated with L_y can be analyzed in exactly the same fashion. Specifically, one obtains the point-wise estimate

$$\left|\mathsf{L}_y\left(\nabla \psi^k, \nabla \phi^n\right) - \mathsf{L}_y\left(\nabla \psi, \nabla \phi^n\right)\right| \leq \frac{1}{4}\left|\partial_x e^k\right| + \left|\partial_y e^k\right|, \tag{54}$$

which, in turn, yields

$$-\left(\mathsf{L}_y\left(\nabla \psi^k, \nabla \phi^n\right) - \mathsf{L}_y\left(\nabla \psi, \nabla \phi^n\right), \partial_y e^{k+1}\right) \leq \frac{1}{8}\left\|\partial_x e^k\right\|^2 + \frac{1}{2}\left\|\partial_y e^k\right\|^2 + \frac{5}{8}\left\|\partial_y e^{k+1}\right\|^2. \tag{55}$$

Finally, a substitution of (53), (55) into (42) yields

$$\frac{1}{s}\left\|e^{k+1}\right\|^2 + \left(\frac{A}{2} - \frac{5}{8}\right)\left\|\nabla e^{k+1}\right\|^2 + \frac{3}{4}\epsilon^2\left\|\Delta e^{k+1}\right\|^2 \leq \frac{5}{8}\left\|\nabla e^k\right\|^2. \tag{56}$$

9

On the other hand, an application of Cauchy inequality shows that

$$\frac{1}{s}\left\|f\right\|^2 + \frac{3}{4}\epsilon^2\left\|\Delta f\right\|^2 \geq 2\sqrt{\frac{1}{s}\cdot\frac{3}{4}\epsilon^2}\left\|f\right\|\cdot\left\|\Delta f\right\| = \frac{\sqrt{3}\epsilon}{\sqrt{s}}\left\|f\right\|\cdot\left\|\Delta f\right\| \geq \frac{\sqrt{3}\epsilon}{\sqrt{s}}\left\|\nabla f\right\|^2, \qquad (57)$$

for any $f \in H^2_{\mathrm{per}}(\Omega)$. The last step comes from a simple estimate based on integration by parts: for all $f \in H^2_{\mathrm{per}}(\Omega)$,

$$\left\|\nabla f\right\|^2 = -(f, \Delta f) \leq \left\|f\right\|\cdot\left\|\Delta f\right\|. \qquad (58)$$

Now, going back to (56) and using (57), we get

$$\left(\frac{\sqrt{3}\epsilon}{\sqrt{s}} + \frac{A}{2} - \frac{5}{8}\right)\left\|\nabla e^{k+1}\right\|^2 \leq \frac{5}{8}\left\|\nabla e^k\right\|^2. \qquad (59)$$

As a result, the contraction mapping property is assured under the following condition

$$\alpha := \frac{\sqrt{3}\epsilon}{\sqrt{s}} + \frac{A}{2} > \frac{5}{4}. \qquad (60)$$

The result is proven. $\qquad\qquad\square$

3 A Fully Discrete Scheme

3.1 A Collocation Fourier Spectral Discretization of Space

So far, we have ignored the discretization of space, which is required for a fully practical method and implementation. Of course, Galerkin (spectral or finite element) methods will automatically inherit the properties described above. One can also use finite difference or collocation methods, the idea being to mimic the variational structure using summation-by-parts formulae and carefully constructed finite difference operators. In [18] we used a finite difference method and in [1] we used a spectral collocation method in precisely this way.

Motivated by the presence of the constant coefficient linear operator $\mathcal{L}$ appearing in the linear iteration scheme (25), together with the assumption of periodic boundary conditions, a natural choice here is to use collocation Fourier spectral differentiation in spatial discretization. Assume that $L_x = N_x \cdot h_x$ and $L_y = N_y \cdot h_y$, for some mesh sizes h_x, $h_y > 0$ and some positive integers N_x and N_y. For simplicity of presentation, we use a square domain, $i.e.$, $L_x = L_y = L$, and a uniform mesh: $h_x = h_y = h$, $N_x = N_y = N$. We will always assume that N is even. All the variables are evaluated/defined at the regular numerical grid vertices (p_i, p_j), $0 \leq i, j \leq N$, where $p_i = i \cdot h$.

For a periodic grid function $f : \{0, \ldots, N-1\} \times \{0, \ldots, N-1\} \to \mathbb{R}$, its discrete Fourier expansion is given by

$$f_{i,j} = \sum_{k,l=-N/2+1}^{N/2} \hat{f}_{k,l}\mathrm{e}^{\frac{2\pi\mathrm{i}}{L}(kp_i+lp_j)}. \qquad (61)$$

The $\hat{f}_{k,l}$ are the collocation Fourier coefficients and are different from the regular Fourier coefficients, in general, due to aliasing error. However, the two are equivalent if the continuous "version" of f is in $\mathcal{P}_N$, the span of the trigonometric polynomials of degree not greater than $N/2$. See, for example [19].

The collocation Fourier spectral approximations to the first and second order partial derivatives (in the x direction) of f are given by

$$(\mathcal{D}_{N,x}f)_{i,j} = \sum_{k,l=-N/2+1}^{N/2} \left(\frac{2k\pi i}{L}\right) \hat{f}_{k,l} e^{2\pi i(kp_i+lp_j)/L}, \tag{62}$$

$$\left(\mathcal{D}_{N,x}^2 f\right)_{i,j} = \sum_{k,l=-N/2+1}^{N/2} \left(\frac{-4\pi^2 k^2}{L^2}\right) \hat{f}_{k,l} e^{2\pi i(kp_i+lp_j)/L}. \tag{63}$$

The corresponding collocation spectral differentiations in the y direction can be defined in the same way. In turn, the discrete Laplacian, gradient and divergence operators become

$$\Delta_N f = \mathcal{D}_{N,x}^2 f + \mathcal{D}_{N,y}^2 f \;, \quad \nabla_N f = \begin{pmatrix} \mathcal{D}_{N,x} f \\ \mathcal{D}_{N,y} f \end{pmatrix} \;, \quad \nabla_N \cdot \begin{pmatrix} f_x \\ f_y \end{pmatrix} = \mathcal{D}_{N,x} f_x + \mathcal{D}_{N,y} f_y \;, \tag{64}$$

all at the point-wise level.

The fully discrete scheme is formulated as follows: given periodic grid functions ϕ^{n-1} and ϕ^n, find the periodic grid function ϕ^{n+1} that satisfies

$$
\begin{aligned}
\frac{\phi^{n+1} - \phi^n}{s} &= \nabla_N \cdot \begin{pmatrix} \mathsf{L}_x \left(\nabla_N \phi^{n+1}, \nabla_N \phi^n\right) \\ \mathsf{L}_y \left(\nabla_N \phi^{n+1}, \nabla_N \phi^n\right) \end{pmatrix} \\
&\quad + \frac{A}{2}\Delta_N \left(\widetilde{\Delta}_s \phi^n\right) - \epsilon^2 \Delta_N^2 \left(\frac{3}{4}\phi^{n+1} + \frac{1}{4}\phi^{n-1}\right).
\end{aligned} \tag{65}
$$

with the logarithmic flux terms L_x, L_y given in (9) and (10). The corresponding linear algorithm becomes

$$
\begin{aligned}
\mathcal{L}_N \left(\psi^{k+1}\right) &:= \left(\frac{1}{s} - \frac{A}{2}\Delta_N + \frac{3}{4}\epsilon^2 \Delta_N^2\right)\psi^{k+1} \\
&= \nabla_N \cdot \begin{pmatrix} \mathsf{L}_x \left(\nabla_N \psi^k, \nabla_N \phi^n\right) \\ \mathsf{L}_y \left(\nabla_N \psi^k, \nabla_N \phi^n\right) \end{pmatrix} - F_N \left(\phi^n, \phi^{n-1}\right),
\end{aligned} \tag{66}
$$

with the forcing term $F_N \left(\phi^n, \phi^{n-1}\right) := \frac{1}{s}\phi^n + \frac{A}{2}\Delta_N \left(-2\phi^n + \phi^{n-1}\right) - \frac{1}{4}\epsilon^2 \Delta_N^2 \phi^{n-1}$. The spatially discrete fixed point, ψ, satisfies

$$\mathcal{L}_N (\psi) = \nabla_N \cdot \begin{pmatrix} \mathsf{L}_x \left(\nabla_N \psi, \nabla_N \phi^n\right) \\ \mathsf{L}_y \left(\nabla_N \psi, \nabla_N \phi^n\right) \end{pmatrix} - F_N \left(\phi^n, \phi^{n-1}\right). \tag{67}$$

The unique solvability of the fully discrete linear iteration scheme (66) (at each iteration stage, k) is clear. For each discrete eigenfunction $e^{2\pi i(kp_i+lp_j)/L}$, the corresponding eigenvalue for the operator $\mathcal{L}_N$ is precisely

$$\lambda_{k,l} := \frac{1}{s} - \frac{A}{2}\left(\lambda_k + \lambda_l\right) + \frac{3}{4}\epsilon^2 \left(\lambda_k + \lambda_l\right)^2 > 0, \tag{68}$$

with $\lambda_k = -\frac{4\pi^2 k^2}{L^2}$, $\lambda_l = -\frac{4\pi^2 l^2}{L^2}$. This implies the unique unconditional solvability of the fully discrete algorithm (66). Naturally, the FFT can be very efficiently utilized to invert $\mathcal{L}_N$ and, therefore, to obtain numerical solutions.

3.2 Fully Discrete Energy Stability

Similar to an earlier work [1], we define a fully discrete analogue of the energy (1) and establish a discrete version of the global in time energy stability property, regardless of the time step size s and independent of the spatial resolution N. With any periodic grid functions f and g (over the 2D numerical grid described above), the discrete approximations to the L^2 norm and inner product are given as

$$\|f\|_2 = \sqrt{\langle f, f \rangle}, \quad \text{with} \quad \langle f, g \rangle = h^2 \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} f_{i,j} g_{i,j} \,. \tag{69}$$

A careful calculation shows that the following summation by parts formulas are valid:

$$\left\langle f, \nabla_N \cdot \begin{pmatrix} g_1 \\ g_2 \end{pmatrix} \right\rangle = -\left\langle \nabla_N f, \begin{pmatrix} g_1 \\ g_2 \end{pmatrix} \right\rangle, \tag{70}$$

$$\langle f, \Delta_N g \rangle = -\langle \nabla_N f, \nabla_N g \rangle, \qquad \langle f, \Delta_N^2 g \rangle = \langle \Delta_N f, \Delta_N g \rangle \,; \tag{71}$$

see the related derivations in [1]. The fully discrete energy is defined as

$$E_N(\phi) = E_{c,1,N}(\phi) + \frac{\epsilon^2}{2} \|\Delta_N \phi\|_2^2, \quad E_{c,1,N}(\phi) = h^2 \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \left(-\frac{1}{2} \ln \left(1 + |\nabla_N \phi|^2 \right)_{i,j} \right). \tag{72}$$

The proof of the following result is similar to the spatially continuous case and is, therefore, skipped to keep the presentation short.

Theorem 3.1. *For any $A \geq 0$ the scheme (65) is second order in time and spectrally accurate in space, i.e., its local truncation error is $O(h^m) + O(s^2)$, and is unconditionally strongly energy stable with respect to the discrete energy*

$$\mathcal{E}_N(\phi^n, \phi^{n-1}) := E_N(\phi^n) + \frac{A}{4} \left\| \nabla_N \left(\phi^n - \phi^{n-1} \right) \right\|_2^2 + \frac{B}{4} \left\| \Delta_N \left(\phi^n - \phi^{n-1} \right) \right\|_2^2, \tag{73}$$

i.e., $\mathcal{E}_N(\phi^{n+1}, \phi^n) \leq \mathcal{E}_N(\phi^n, \phi^{n-1})$, for any $n \geq 1$, and any $s > 0$. Furthermore, if $A \geq 1$ the fully discrete scheme is unconditionally uniquely solvable.

In addition to the unconditional energy stability, the following proposition states a global in time bound for $\|\Delta \phi\|_2^2$ of the numerical solution.

Lemma 3.2. *Let $\Phi \in H^4_{\mathrm{per}}(\Omega)$. Suppose that $\phi^0_{i,j} := \Phi(p_i, p_j)$ and $\phi^{-1} \equiv \phi^0$. Then, for solutions of the fully discrete second order scheme (65), we have the global in time bounds*

$$E_N(\phi^n) \leq C_0, \quad \|\Delta_N \phi^n\|_2^2 \leq C_1, \tag{74}$$

for any $n \geq 1$, and any h and s, where $C_0, C_1 > 0$ depend upon ϵ, L and the data, but are independent of the step sizes h and s and of the final time T.

Proof. First, by the energy stability above and the definition of $\mathcal{E}_N$,

$$E_N(\phi^n) \leq \mathcal{E}_N(\phi^n, \phi^{n-1}) \leq \cdots \leq \mathcal{E}_N(\phi^0, \phi^{-1}) = E_N(\phi^0) \leq Ch^4 + E(\Phi) \leq C_0, \tag{75}$$

where a consistency argument for the collocation spectral approximation is applied in the last steps. For the second part, we need the point-wise estimate (see [1, 21])

$$F(\mathbf{y}) = -\frac{1}{2} \ln \left(1 + |\mathbf{y}|^2 \right) \geq -\frac{1}{2} \left(\beta |\mathbf{y}|^2 - \ln(\beta) + \beta - 1 \right), \quad \forall \, 0 < \beta \leq 1, \quad \forall \, \mathbf{y} \in \mathbb{R}^2, \tag{76}$$

and the following discrete elliptic regularity estimate in 2D: for all periodic grid functions ϕ,

$$\|\Delta_N \phi\|_2^2 \geq C_2 \|\nabla_N \phi\|_2^2, \quad C_2 = \frac{4\pi^2}{L^2}. \tag{77}$$

Then, with the choice of $\beta = \frac{\epsilon^2 C_2}{2}$, we obtain, for all periodic grid functions ϕ,

$$E_N(\phi) \geq \frac{L^2}{2}\left(\ln\left(\frac{\epsilon^2 C_2}{2}\right) - \frac{\epsilon^2 C_2}{2} + 1\right) + \frac{\epsilon^2}{4}\|\Delta_N \phi\|_2^2. \tag{78}$$

This, in turn, shows that

$$
\begin{aligned}
\|\Delta_N \phi^n\|_2^2 &\leq \frac{4}{\epsilon^2}\left(E_N(\phi^n) + \frac{L^2}{2}\left(-\ln\left(\frac{\epsilon^2 C_2}{2}\right) + \frac{\epsilon^2 C_1}{2} - 1\right)\right) \\
&\leq \frac{4}{\epsilon^2}\left(E_N(\phi^0) + \frac{L^2}{2}\left(-\ln\left(\frac{\epsilon^2 C_2}{2}\right) + \frac{\epsilon^2 C_2}{2} - 1\right)\right) \\
&\leq \frac{4}{\epsilon^2}\left(C_0 + \frac{L^2}{2}\left(-\ln\left(\frac{\epsilon^2 C_2}{2}\right) + \frac{\epsilon^2 C_2}{2} - 1\right)\right) =: C_1.
\end{aligned}
\tag{79}
$$

$\square$

Remark 3.3. *It is clear that the numerical solution of the fully discrete second order scheme (65) is mass-conserving at the discrete level, i.e., $\overline{\phi^{n+1}} = \overline{\phi^n} = \overline{\phi^0}$, where*

$$\overline{f} := \frac{1}{N^2}\sum_{i,j=0}^{N-1} f_{i,j}. \tag{80}$$

Without loss of generality, we may assume that $\overline{\phi^0} = 0$ so that $\overline{\phi^n} = 0$, for $n = 0, 1, 2, \ldots$, since only the gradient of ϕ is of consequence. Under this assumption, a discrete elliptic regularity can be applied so that we obtain a global in time H^2 bound for the numerical solution, at the discrete level:

$$\|\phi_N^n\|_{H^2} \leq C\|\Delta \phi_N^n\|_{L^2} \leq C\sqrt{C_1} =: C_3, \tag{81}$$

where ϕ_N^n is the continuous version of the discrete numerical solution ϕ^n obtained by interpolating into the space $\mathcal{P}_N$. Note that C_3 is also a global in time constant, only dependent upon ϵ, L, and the initial data.

Remark 3.4. *There are some alternate approaches to develop the energy stability for a second order accurate numerical scheme for the no-slope-selection model (4), in a modified way. For instance, in a very recent article [14], the authors found a certain PDE satisfied by the nonlinear integrand appearing in the nonlinear energy (1), and proposed a second order scheme to update such a nonlinear integrand. In turn, an alternate energy was defined and the non-increasing property was proved for the scheme with respect to the alternate energy. While the approach in [14] represents a clever way to achieve the desired numerical stability – since it avoids the complicated form of the nonlinear energy in the numerical scheme – it appears that such an alternate energy stability cannot assure an H^2 stability of the variable ϕ at the theoretical level. To the best of the authors' knowledge, the schemes presented in [18] and the linear iteration algorithm given by this paper are the only second order numerical schemes in which a global in time H^2 stability can be established for the height function ϕ.*

3.3 Contraction for the Fully Discrete Linear Iteration Method

For the fully discrete linear iteration scheme (66), all of the derivations and estimates in the proof of Theorem 2.2 can be extended to the spatially discrete case, with summation-by-parts replacing integration-by-parts in the arguments. The proof of the following is skipped for brevity of presentation.

Theorem 3.5. *The fully discrete linear iteration defined in (66) is a contraction mapping provided that* $\alpha := \frac{\sqrt{3}\epsilon}{\sqrt{s}} + \frac{A}{2} > \frac{5}{4}$.

4 Convergence Analysis for the Fully Discrete Scheme

We present a detailed convergence analysis for the fully discrete second order scheme (65) in this section. To simplify the presentation, a preliminary estimate for the nonlinear error term is given in the following lemma. For brevity, we state this preliminary result in the case with space kept continuous. It extension to the fully discrete error estimate is straightforward and we will cite this lemma in later analysis.

Lemma 4.1. *Let* $\Phi^n, \Phi^{n+1}, \phi^n, \phi^{n+1} \in C^1_{\mathrm{per}}(\Omega)$ *be arbitrary, and define* $e^k := \Phi^k - \phi^k$, $k = n, n+1$. *Then*

$$
\left| \mathsf{L}_x \left(\nabla \Phi^{n+1}, \nabla \Phi^n \right) - \mathsf{L}_x \left(\nabla \phi^{n+1}, \nabla \phi^n \right) \right| \leq \left(\left| \partial_x e^{n+1} \right| + \left| \partial_x e^n \right| \right) + \frac{1}{4} \left(\left| \partial_y e^{n+1} \right| + \left| \partial_y e^n \right| \right), \tag{82}
$$

$$
\left| \mathsf{L}_y \left(\nabla \Phi^{n+1}, \nabla \Phi^n \right) - \mathsf{L}_y \left(\nabla \phi^{n+1}, \nabla \phi^n \right) \right| \leq \frac{1}{4} \left(\left| \partial_x e^{n+1} \right| + \left| \partial_x e^n \right| \right) + \left(\left| \partial_y e^{n+1} \right| + \left| \partial_y e^n \right| \right). \tag{83}
$$

Proof. For the term L_x, we start from the following calculation:

$$
\begin{aligned}
& \mathsf{L}_x \left(\nabla \Phi^{n+1}, \nabla \Phi^n \right) - \mathsf{L}_x \left(\nabla \phi^{n+1}, \nabla \phi^n \right) \\
& = -\frac{1}{4} \left(\frac{\mathsf{L} \left(\partial_x \Phi^{n+1}, \partial_y \Phi^{n+1} \right) - \mathsf{L} \left(\partial_x \Phi^n, \partial_y \Phi^{n+1} \right)}{\partial_x \Phi^{n+1} - \partial_x \Phi^n} - \frac{\mathsf{L} \left(\partial_x \phi^{n+1}, \partial_y \phi^{n+1} \right) - \mathsf{L} \left(\partial_x \phi^n, \partial_y \phi^{n+1} \right)}{\partial_x \phi^{n+1} - \partial_x \phi^n} \right) \\
& \quad -\frac{1}{4} \left(\frac{\mathsf{L} \left(\partial_x \Phi^{n+1}, \partial_y \Phi^n \right) - \mathsf{L} \left(\partial_x \Phi^n, \partial_y \Phi^n \right)}{\partial_x \Phi^{n+1} - \partial_x \Phi^n} - \frac{\mathsf{L} \left(\partial_x \phi^{n+1}, \partial_y \phi^n \right) - \mathsf{L}(\partial_x \phi^n, \partial_y \phi^n)}{\partial_x \phi^{n+1} - \partial_x \phi^n} \right). \tag{84}
\end{aligned}
$$

The first term appearing above can be decomposed as follows:

$$
\begin{aligned}
& \frac{\mathsf{L}(\partial_x \Phi^{n+1}, \partial_y \Phi^{n+1}) - \mathsf{L}(\partial_x \Phi^n, \partial_y \Phi^{n+1})}{\partial_x \Phi^{n+1} - \partial_x \Phi^n} - \frac{\mathsf{L}(\partial_x \phi^{n+1}, \partial_y \phi^{n+1}) - \mathsf{L}(\partial_x \phi^n, \partial_y \phi^{n+1})}{\partial_x \phi^{n+1} - \partial_x \phi^n} \\
& \qquad\qquad\qquad\qquad = \mathcal{N}_5 + \mathcal{N}_6 + \mathcal{N}_7, \tag{85}
\end{aligned}
$$

where

$$
\mathcal{N}_5 := \frac{\mathsf{L} \left(\partial_x \Phi^{n+1}, \partial_y \Phi^{n+1} \right) - \mathsf{L} \left(\partial_x \Phi^n, \partial_y \Phi^{n+1} \right)}{\partial_x \Phi^{n+1} - \partial_x \Phi^n} - \frac{\mathsf{L} \left(\partial_x \phi^{n+1}, \partial_y \Phi^{n+1} \right) - \mathsf{L} \left(\partial_x \Phi^n, \partial_y \Phi^{n+1} \right)}{\partial_x \phi^{n+1} - \partial_x \Phi^n}, \tag{86}
$$

$$
\mathcal{N}_6 := \frac{\mathsf{L} \left(\partial_x \Phi^n, \partial_y \Phi^{n+1} \right) - \mathsf{L} \left(\partial_x \phi^{n+1}, \partial_y \Phi^{n+1} \right)}{\partial_x \Phi^n - \partial_x \phi^{n+1}} - \frac{\mathsf{L} \left(\partial_x \phi^n, \partial_y \Phi^{n+1} \right) - \mathsf{L} \left(\partial_x \phi^{n+1}, \partial_y \Phi^{n+1} \right)}{\partial_x \phi^n - \partial_x \phi^{n+1}}, \tag{87}
$$

$$
\mathcal{N}_7 := \frac{\mathsf{L} \left(\partial_x \phi^{n+1}, \partial_y \Phi^{n+1} \right) - \mathsf{L} \left(\partial_x \phi^n, \partial_y \Phi^{n+1} \right)}{\partial_x \phi^{n+1} - \partial_x \phi^n} - \frac{\mathsf{L} \left(\partial_x \phi^{n+1}, \partial_y \phi^{n+1} \right) - \mathsf{L} \left(\partial_x \phi^n, \partial_y \phi^{n+1} \right)}{\partial_x \phi^{n+1} - \partial_x \phi^n}. \tag{88}
$$

14

The following representations are clear:

$$\mathcal{N}_5 = h_1\left(\partial_x \Phi^{n+1}\right) - h_1\left(\partial_x \phi^{n+1}\right), \quad a = \partial_y \Phi^{n+1}, \quad b = \partial_x \Phi^n, \tag{89}$$

$$\mathcal{N}_6 = h_1\left(\partial_x \Phi^n\right) \quad - h_1\left(\partial_x \phi^n\right), \quad\; a = \partial_y \Phi^{n+1}, \quad b = \partial_x \phi^{n+1}, \tag{90}$$

$$\mathcal{N}_7 = h_2\left(\partial_y \Phi^{n+1}\right) - h_2\left(\partial_y \phi^{n+1}\right), \quad a = \partial_x \phi^{n+1}, \quad b = \partial_x \phi^n. \tag{91}$$

Similar to (46), (50), (51), applications of the mean value theorem and Lemmas 2.3 and 2.4 result in the point-wise estimates

$$|\mathcal{N}_5(\mathbf{x})| \le \left|h_1'(\xi_1(\mathbf{x}))\right| \cdot \left|\partial_x \Phi^{n+1}(\mathbf{x}) - \partial_x \phi^{n+1}(\mathbf{x})\right| \le 2\left|\partial_x e^{n+1}(\mathbf{x})\right|, \quad \forall \mathbf{x} \in \overline{\Omega}, \tag{92}$$

$$|\mathcal{N}_6(\mathbf{x})| \le \left|h_1'(\xi_2(\mathbf{x}))\right| \cdot \left|\partial_x \Phi^n(\mathbf{x}) - \partial_x \phi^n(\mathbf{x})\right| \quad\; \le 2\left|\partial_x e^n(\mathbf{x})\right|, \qquad \forall \mathbf{x} \in \overline{\Omega}, \tag{93}$$

$$|\mathcal{N}_7(\mathbf{x})| \le \left|h_2'(\xi_3(\mathbf{x}))\right| \cdot \left|\partial_y \Phi^{n+1}(\mathbf{x}) - \partial_y(\mathbf{x})\phi^{n+1}\right| \le \left|\partial_y e^{n+1}(\mathbf{x})\right|, \qquad \forall \mathbf{x} \in \overline{\Omega}. \tag{94}$$

Then we arrive at

$$\left| \frac{\mathsf{L}\left(\partial_x \Phi^{n+1}, \partial_y \Phi^{n+1}\right) - \mathsf{L}\left(\partial_x \Phi^n, \partial_y \Phi^{n+1}\right)}{\partial_x \Phi^{n+1} - \partial_x \Phi^n} - \frac{\mathsf{L}\left(\partial_x \phi^{n+1}, \partial_y \phi^{n+1}\right) - \mathsf{L}\left(\partial_x \phi^n, \partial_y \phi^{n+1}\right)}{\partial_x \phi^{n+1} - \partial_x \phi^n} \right|$$
$$\le 2\left(\left|\partial_x e^{n+1}\right| + \left|\partial_x e^n\right|\right) + \left|\partial_y e^{n+1}\right|. \tag{95}$$

The second term in (84) can be analyzed in an analogous manner; the details of the following are omitted for the sake of brevity:

$$\left| \frac{\mathsf{L}\left(\partial_x \Phi^{n+1}, \partial_y \Phi^n\right) - \mathsf{L}\left(\partial_x \Phi^n, \partial_y \Phi^n\right)}{\partial_x \Phi^{n+1} - \partial_x \Phi^n} - \frac{\mathsf{L}\left(\partial_x \phi^{n+1}, \partial_y \phi^n\right) - \mathsf{L}\left(\partial_x \phi^n, \partial_y \phi^n\right)}{\partial_x \phi^{n+1} - \partial_x \phi^n} \right|$$
$$\le 2\left(\left|\partial_x e^{n+1}\right| + \left|\partial_x e^n\right|\right) + \left|\partial_y e^n\right|. \tag{96}$$

A combination of (84), (95) and (96) results in (82). The estimate (83) involving L_y is derived by a similar approach. The result is proven. $\qquad\square$

The convergence theorem for the fully discrete second order scheme (65) is stated below.

Theorem 4.2. *Denote by Φ the exact smooth, periodic solution for the no-slope-selection model (4). Set $\phi_{i,j}^0 := \Phi(p_i, p_j, 0)$ and $\phi^{-1} \equiv \phi^0$. Define $e_{i,j}^n := \Phi(x_i, y_j, s \cdot n) - \phi_{i,j}^n$, with ϕ the fully discrete solution for the second order scheme (65). Then, provided s is sufficiently small, we have the following error estimate in a discrete $L^\infty(0, T; H^1) \cap L^2(0, T; H^3)$ norm: for any k, with $1 \le k \le T/s$,*

$$\left\|\nabla_N e^k\right\|_2 + \left(\frac{s\epsilon^2}{4} \sum_{l=1}^{k} \left\|\nabla_N\left(\Delta_N e^l\right)\right\|_2^2\right)^{1/2} \le C\left(s^2 + h^m\right), \tag{97}$$

where $C > 0$ is a constant that depends upon ϵ, L, the final time T, and the exact solution Φ, but is independent of the step sizes h and s.

Proof. A consistency analysis shows that

$$\frac{\Phi^{n+1} - \Phi^n}{s} = \nabla_N \cdot \left(\begin{array}{c} \mathsf{L}_x\left(\nabla_N \Phi^{n+1}, \nabla_N \Phi^n\right) \\ \mathsf{L}_y\left(\nabla_N \Phi^{n+1}, \nabla_N \Phi^n\right) \end{array} \right) + \frac{A}{2}\Delta_N\left(\Phi^{n+1} - 2\Phi^n + \Phi^{n-1}\right)$$
$$- \epsilon^2 \Delta_N^2 \left(\frac{3}{4}\Phi^{n+1} + \frac{1}{4}\Phi^{n-1}\right) + \tau^{n+1/2}, \tag{98}$$

where the local truncation error satisfies the estimate

$$\left\|\tau^{n+1/2}\right\|_2 \leq C\left(s^2 + h^m\right). \tag{99}$$

The functions L_x, L_y are given in (9) and (10). The last estimate is based on a Fourier spectral differentiation analysis in space and Taylor expansions in the time dimension. The details are left to interested readers.

Subtracting the numerical scheme (65) from (98), we get the equation for the numerical error function:

$$\frac{e^{n+1} - e^n}{s} - \frac{A}{2}\Delta_N\left(e^{n+1} - 2e^n + e^{n-1}\right) + \epsilon^2\Delta_N^2\left(\frac{3}{4}e^{n+1} + \frac{1}{4}e^{n-1}\right)$$
$$= \nabla_N \cdot \left(\begin{array}{c} \mathsf{L}_x\left(\nabla_N\Phi^{n+1}, \nabla_N\Phi^n\right) - \mathsf{L}_x\left(\nabla_N\phi^{n+1}, \nabla_N\phi^n\right) \\ \mathsf{L}_y\left(\nabla_N\Phi^{n+1}, \nabla_N\Phi^n\right) - \mathsf{L}_y\left(\nabla_N\phi^{n+1}, \nabla_N\phi^n\right) \end{array}\right) + \tau^{n+1/2}. \tag{100}$$

Taking the discrete L^2 inner product with $-2\Delta_N e^{n+1}$ gives

$$\left\|\nabla_N e^{n+1}\right\|_2^2 - \left\|\nabla_N e^n\right\|_2^2 + \left\|\nabla_N(e^{n+1} - e^n)\right\|_2^2 + As\left\langle\Delta_N e^{n+1}, \Delta_N(e^{n+1} - 2e^n + e^{n-1})\right\rangle$$
$$-\epsilon^2 s\left\langle\Delta_N e^{n+1}, \Delta_N^2\left(\frac{3}{2}e^{n+1} + \frac{1}{2}e^{n-1}\right)\right\rangle + 2s\left\langle\tau^{n+1/2}, \Delta_N e^{n+1}\right\rangle$$
$$= -s\left\langle\left(\begin{array}{c} \mathsf{L}_x\left(\nabla_N\Phi^{n+1}, \nabla_N\Phi^n\right) - \mathsf{L}_x\left(\nabla_N\phi^{n+1}, \nabla_N\phi^n\right) \\ \mathsf{L}_y\left(\nabla_N\Phi^{n+1}, \nabla_N\Phi^n\right) - \mathsf{L}_y\left(\nabla_N\phi^{n+1}, \nabla_N\phi^n\right) \end{array}\right), 2\nabla_N\Delta_N e^{n+1}\right\rangle, \tag{101}$$

where the summation-by-parts formulae (70) and (71) were applied. The term associated with the truncation error can be controlled by Cauchy's inequality:

$$-2\left\langle\tau^{n+1/2}, \Delta_N e^{n+1}\right\rangle \leq 2\left\|\tau^{n+1/2}\right\|_2 \cdot \left\|\Delta_N e^{n+1}\right\|_2 \leq \left\|\tau^{n+1/2}\right\|_2^2 + \left\|\Delta_N e^{n+1}\right\|_2^2. \tag{102}$$

The term associated with the artificial diffusion can be handled by

$$-\left\langle\Delta_N e^{n+1}, \Delta_N(e^{n+1} - 2e^n + e^{n-1})\right\rangle$$
$$= -\left\|\Delta_N e^{n+1}\right\|_2^2 + 2\left\langle\Delta_N e^{n+1}, \Delta_N e^n\right\rangle - \left\langle\Delta_N e^{n+1}, \Delta_N e^{n-1}\right\rangle$$
$$\leq -\left\|\Delta_N e^{n+1}\right\|_2^2 + \left\|\Delta_N e^{n+1}\right\|_2^2 + \left\|\Delta_N e^n\right\|_2^2 + \frac{1}{2}\left(\left\|\Delta_N e^{n+1}\right\|_2^2 + \left\|\Delta_N e^{n-1}\right\|_2^2\right)$$
$$= \left\|\Delta_N e^n\right\|_2^2 + \frac{1}{2}\left(\left\|\Delta_N e^{n+1}\right\|_2^2 + \left\|\Delta_N e^{n-1}\right\|_2^2\right). \tag{103}$$

The term associated with the surface diffusion can be analyzed as follows:

$$-\left\langle\Delta_N e^{n+1}, \Delta_N^2\left(\frac{3}{2}e^{n+1} + \frac{1}{2}e^{n-1}\right)\right\rangle = \left\langle\nabla_N\Delta_N e^{n+1}, \nabla_N\Delta_N\left(\frac{3}{2}e^{n+1} + \frac{1}{2}e^{n-1}\right)\right\rangle$$
$$= \frac{3}{2}\left\|\nabla_N\Delta_N e^{n+1}\right\|_2^2 + \frac{1}{2}\left\langle\nabla_N\Delta_N e^{n+1}, \nabla_N\Delta_N e^{n-1}\right\rangle$$
$$\geq \frac{3}{2}\left\|\nabla_N\Delta_N e^{n+1}\right\|_2^2 - \frac{1}{4}\left\|\nabla_N\Delta_N e^{n+1}\right\|_2^2 - \frac{1}{4}\left\|\nabla_N\Delta_N e^{n-1}\right\|_2^2$$
$$= \frac{5}{4}\left\|\nabla_N\Delta_N e^{n+1}\right\|_2^2 - \frac{1}{4}\left\|\nabla_N\Delta_N e^{n-1}\right\|_2^2. \tag{104}$$

For the nonlinear term, an extension of Lemma 4.1 to the space discrete case indicates that

$$\left|\mathsf{L}_x\left(\nabla_N\Phi^{n+1}, \nabla_N\Phi^n\right) - \mathsf{L}_x\left(\nabla_N\phi^{n+1}, \nabla_N\phi^n\right)\right|$$
$$\leq \left(\left|\mathcal{D}_{Nx}e^{n+1}\right| + \left|\mathcal{D}_{Nx}e^n\right|\right) + \frac{1}{4}\left(\left|\mathcal{D}_{Ny}e^{n+1}\right| + \left|\partial_y e^n\right|\right), \tag{105}$$
$$\left|\mathsf{L}_y\left(\nabla_N\Phi^{n+1}, \nabla_N\Phi^n\right) - \mathsf{L}_y\left(\nabla_N\phi^{n+1}, \nabla_N\phi^n\right)\right|$$
$$\leq \left(\left|\mathcal{D}_{Ny}e^{n+1}\right| + \left|\mathcal{D}_{Ny}e^n\right|\right) + \frac{1}{4}\left(\left|\mathcal{D}_{Nx}e^{n+1}\right| + \left|\mathcal{D}_{Nx}e^n\right|\right). \tag{106}$$

Therefore, we arrive at

$$-\left\langle \begin{pmatrix} \mathsf{L}_x\left(\nabla_N\Phi^{n+1},\nabla_N\Phi^n\right) - \mathsf{L}_x\left(\nabla_N\phi^{n+1},\nabla_N\phi^n\right) \\ \mathsf{L}_y\left(\nabla_N\Phi^{n+1},\nabla_N\Phi^n\right) - \mathsf{L}_y\left(\nabla_N\phi^{n+1},\nabla_N\phi^n\right) \end{pmatrix}, 2\nabla_N\Delta_N e^{n+1}\right\rangle$$

$$\leq \quad \frac{\epsilon^2}{4}\left\|\nabla_N\Delta_N e^{n+1}\right\|_2^2 + \frac{4}{\epsilon^2}\left(\frac{17}{4}\left\|\nabla_N e^{n+1}\right\|_2^2 + \frac{17}{4}\left\|\nabla_N e^n\right\|_2^2\right). \tag{107}$$

Subsequently, a substitution of estimates (102) – (107) into (101) yields

$$\left\|\nabla_N e^{n+1}\right\|_2^2 - \left\|\nabla_N e^n\right\|_2^2 + \epsilon^2 s\left\|\nabla_N\Delta_N e^{n+1}\right\|_2^2$$

$$\leq \quad \frac{1}{4}\epsilon^2 s\left\|\nabla_N\Delta_N e^{n-1}\right\|_2^2 + \frac{17}{\epsilon^2}s\left(\left\|\nabla_N e^{n+1}\right\|_2^2 + \left\|\nabla_N e^n\right\|_2^2\right) + s\left\|\tau^{n+1/2}\right\|_2^2$$

$$+ \left(\frac{A}{2}+1\right)s\left\|\Delta_N e^{n+1}\right\|_2^2 + As\left\|\Delta_N e^n\right\|_2^2 + \frac{A}{2}s\left\|\Delta_N e^{n-1}\right\|_2^2. \tag{108}$$

Meanwhile, the following estimate for $\left\|\Delta_N e^k\right\|_2^2$ is valid:

$$\left\|\Delta_N e^k\right\|_2^2 = -\left\langle\nabla_N e^k, \nabla_N\Delta_N e^k\right\rangle \leq \beta\epsilon^2\left\|\nabla_N\Delta_N e^k\right\|_2^2 + \frac{1}{4\beta\epsilon^2}\left\|\nabla_N e^k\right\|_2^2, \tag{109}$$

for all $\beta > 0$. Using the last estimate, with a careful choice of β, we obtain

$$\left\|\nabla_N e^{n+1}\right\|_2^2 - \left\|\nabla_N e^n\right\|_2^2 + \frac{3}{4}\epsilon^2 s\left\|\nabla_N\Delta_N e^{n+1}\right\|_2^2$$

$$\leq \quad \frac{3}{8}\epsilon^2 s\left\|\nabla_N\Delta_N e^{n-1}\right\|_2^2 + \frac{1}{8}\epsilon^2 s\left\|\nabla_N\Delta_N e^n\right\|_2^2$$

$$+ \frac{C}{\epsilon^2}s\left(\left\|\nabla_N e^{n+1}\right\|_2^2 + \left\|\nabla_N e^n\right\|_2^2 + \left\|\nabla_N e^{n-1}\right\|_2^2\right) + s\left\|\tau^{n+1/2}\right\|_2^2. \tag{110}$$

Finally, if s is small enough($s < \frac{\epsilon^2}{C}$), summing in time and applying a discrete Gronwall inequality, we arrive at the desired discrete $L^\infty(0,T;H^1)\cap L^2(0,T;H^3)$ error estimate for solutions of the fully discrete scheme (65). $\qquad\square$

Remark 4.3. *A detailed consistency analysis shows that a regularity of $\Phi \in W^{3,\infty}(0,T;H^m) \cap W^{2,\infty}(0,T;H^{m+4})$ is needed for the exact solution Φ to make the local truncation error estimate (99) valid at every time step. Moreover, a more careful analysis indicates that a reduced regularity assumption $\Phi \in H^3(0,T;H^m) \cap H^2(0,T;H^{m+4})$ can be made for Φ, if we only need the local truncation error estimate (99) satisfied in $\ell^2(0,T)$, at the discrete time level. That could also lead to the desired convergence analysis result as in (97).*

In this paper, we made an assumption that "Φ is the exact smooth solution" in Thm. 4.2, for simplicity of presentation.

Remark 4.4. *We note that the initialization $\phi^{-1} \equiv \phi^0$ is only first order accurate. However, this "one-step" degradation of accuracy does not affect the overall numerical accuracy. In more detail, we observe that, although truncation error at the first time step is only of $O(s)$, the numerical solution ϕ^1 updated from the fully discrete scheme (65) is still an $O(s^2)$ approximation to Φ^1, because of its detailed expansions. In other words, we have an exact approximation ϕ^0 and an $O(s^2)$ approximation ϕ^1. Subsequently, after the first time step, all the local truncation errors are of second order accurate in time. In turn, both the theoretical analysis presented in this paper and the numerical experiments have indicated a full second order accuracy in time.*
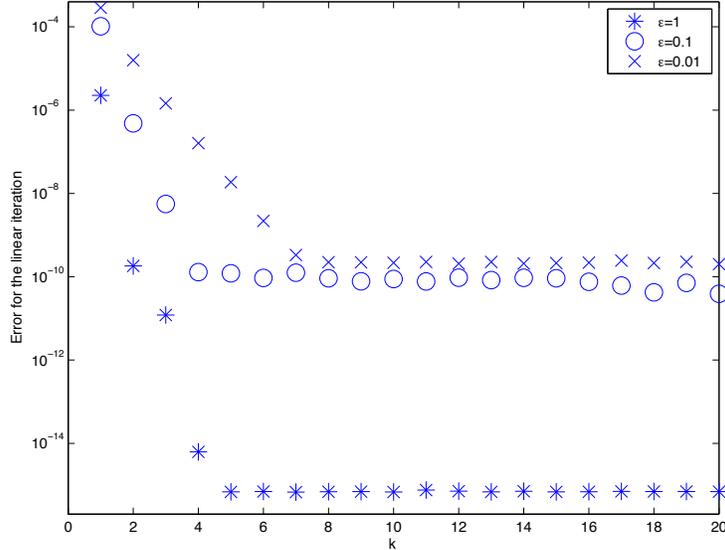
Figure 1: Dependence of the convergence rate of the linear iteration method on the surface diffusion parameter ϵ. Here we plot the L^2 norm of the error for the linear iteration versus the iteration stage k, with time step $s = 0.01$ and artificial diffusion parameter $A = 2$.

5 Numerical Simulation Results

5.1 Convergence of the Linear Iteration Scheme

In this subsection we present some tests, the results of which support the theoretical convergence for the proposed linear iteration algorithm (25). It is clear that the convergence rate of the linear iteration scheme will dependent on the values of surface diffusion coefficient, ϵ, the artificial diffusion coefficient, A, and time step size s, as well as others, like the value of N. We will vary ϵ, A, and s and compare the convergence rates. We take the following exact profile for the phase variable:

$$\psi(x, y) = \sin(2\pi x)\cos(2\pi y) \tag{111}$$

over the domain $\Omega = (0, 1)^2$. Making this the exact solution requires that we manufacture appropriate values for ϕ^n and F in (26). In Figs. 1, 2, 3, we plot the iteration error $\left\|e^k\right\|_2$ versus k, where $e^k := \psi^k - \psi$, as in Theorem 2.2. For the tests, we fix $N = 64$, and do not explore the convergence rate dependence on this parameter here. Of course, (111) will not be the solution to the fully discrete equation (67). We expect to and, in fact, do see a finite saturation of the iteration error in Figs. 1, 2, 3. And, as expected, the saturation levels differ for different values of the parameters. Naturally, a larger value of N will allow for smaller saturation levels in each case, but at the cost of more computation.

For the first test, the results for which are reported in Fig. 1, we fix $s = 0.01$ and $A = 2$ and vary ϵ: $\epsilon = 1$, $\epsilon = 0.1$ and $\epsilon = 0.01$. It is clear that the linear iteration error reaches a saturation after a few ($k \leq 8$) iteration stages. From Fig. 1, we observe that the convergence rate for the linear iteration increases with an increasing value of ϵ. This implies that numerical implementation of the linear iteration algorithm (25) becomes more challenging with a smaller surface diffusion coefficient. This result matches with our theoretical analysis in proof of Theorem 2.2. Also note, that each iteration of the linear iteration method reduces the iteration by roughly a constant amount, which is not surprising since we have a pure contraction of the error.
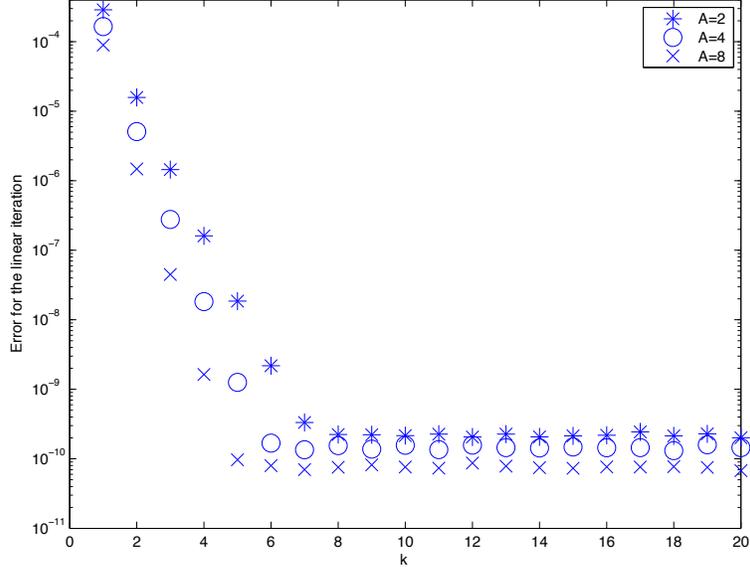
18

Figure 2: Dependence of the convergence rate of the linear iteration method on the artificial diffusion parameter A. Here we plot the L^2 norm of the error for the linear iteration versus the iteration stage k, with time step $s = 0.01$ and the surface diffusion coefficient $\epsilon = 0.01$.

For the second test, the results for which are reported in Fig. 2, we fix $s = 0.01$ and $\epsilon = 0.01$ and we vary A: $A = 2$, $A = 4$ and $A = 8$. Again, the linear iteration error reaches a saturation after a few (≤ 8) iteration stages. Moreover, the convergence rate for the linear iteration increases with an increasing value of A. For the third test, the results for which are reported in Fig. 3, we fix $\epsilon = 0.01$ and $A = 2$ and we vary s: $s = 10^{-2}$, $s = 10^{-3}$ and $s = 10^{-4}$. The linear iteration error reaches a saturation after a few iteration stages. The convergence rate for the linear iteration increases with a decreasing value of s. These results likewise match with our theoretical analysis in the proof of Theorem 2.2. As before, the iteration error is decreased by roughly the same amount for each iteration stage.

5.2 Convergence of the Convex Splitting Scheme

In this subsection we perform a numerical accuracy check for the energy stable, fully discrete second order scheme (65). Similar to the last example, the computational domain is set to be $\Omega = (0, 1)^2$, and the exact profile for the phase variable is set to be

$$\Phi(x, y, t) = \sin(2\pi x)\cos(2\pi y)\cos(t). \tag{112}$$

As with the last test, to make Φ satisfy the original PDE (4), we have to add an artificial, time-dependent forcing term, which we do. The proposed second order scheme (65) (with Fourier spectral differentiation in space) can be implemented to (4), with the linear iteration (25) applied to solve the nonlinear system. We compute solutions with grid sizes $N = 64$ to $N = 192$ in increments of 16, and we solve up to time $T = 1$. The errors are reported at this final time. Two parameters for the surface diffusion are used: $\epsilon = 0.5$ and $\epsilon = 0.05$. The time step s is determined by the linear refinement path $s = 0.5h$, where h is the spatial grid size. Figs. 4 and 5 show the discrete L^1, L^2 and L^∞ norm s of the errors between the numerical and exact solutions. A clear second order accuracy is observed in all cases.

19

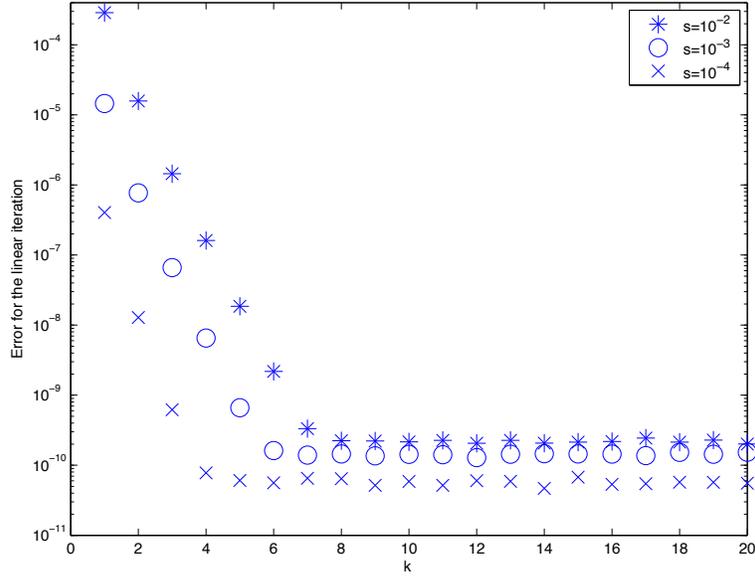Figure 3: Dependence of the convergence rate of the linear iteration method on the time step s. Here we plot the L^2 norm of the error for the linear iteration versus the iteration stage k, with artificial diffusion parameter $A = 2$ and the surface diffusion coefficient $\epsilon = 0.01$.
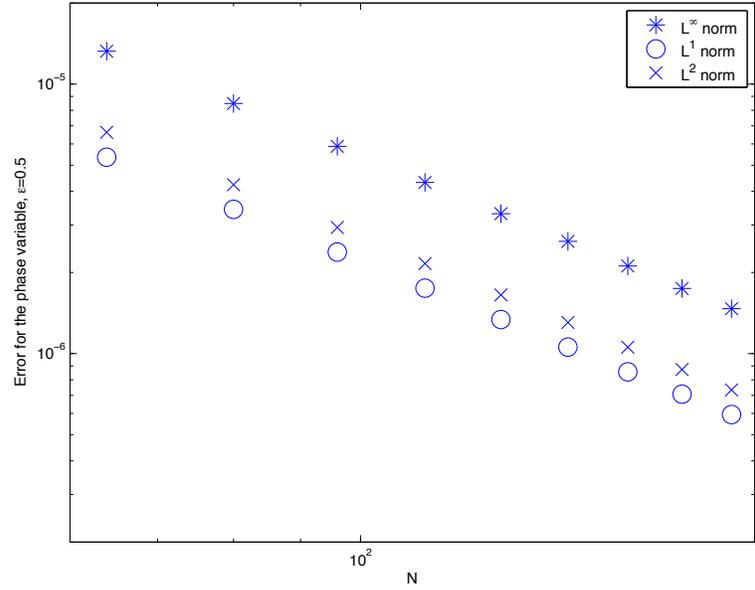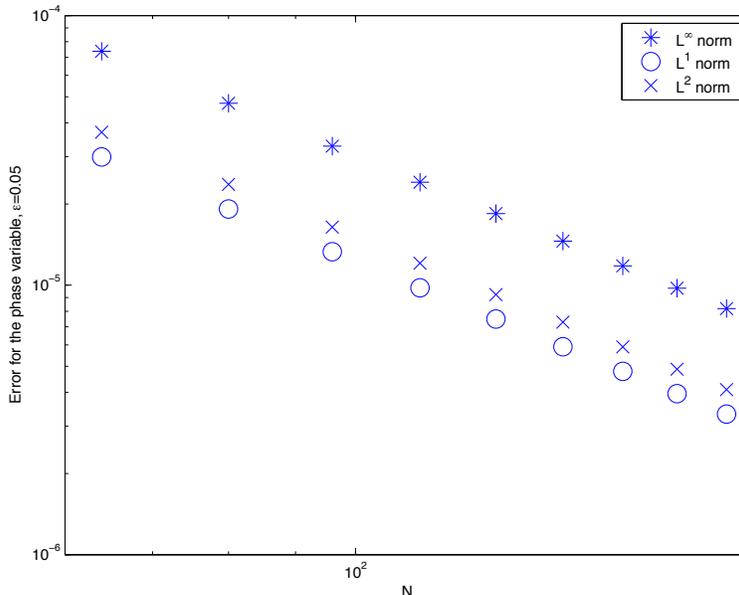

Figure 4: L^1, L^2 and L^∞ numerical errors at $T = 1.0$ plotted versus N for the fully discrete second order scheme (65), with the linear iteration algorithm (25) applied. The surface diffusion parameter is taken to be $\epsilon = 0.5$ and the time step size is $s = 0.5h$. The data lie roughly on curves CN^{-2}, for appropriate choices of C, confirming the full second-order accuracy of the scheme.

Figure 5: L^1, L^2 and L^∞ numerical errors at $T = 1.0$ plotted versus N for the fully discrete second order scheme (65), with the linear iteration algorithm (25) applied. The surface diffusion parameter is taken to be $\epsilon = 0.05$ and the time step size is $s = 0.5h$. The data lie roughly on curves CN^{-2}, for appropriate choices of C, confirming the full second-order accuracy of the scheme.

5.3 Coarsening and Energy Dissipation

Typically one is interested in the how properties associated with the solutions to (4) and (6) scale with time, where it is assumed that $\epsilon \ll \min\{L_x, L_y\}$. The physically interesting quantities that may be obtained from the solutions of these equations are the surface roughness, defined as

$$w(t) = \sqrt{\frac{1}{|\Omega|} \int_\Omega \left|\phi(\mathbf{x}, t) - \bar\phi(t)\right|^2 \mathrm{d}\mathbf{x}}, \quad \text{with } \bar\phi(t) = \frac{1}{|\Omega|} \int_\Omega \phi(\mathbf{x}, t)\mathrm{d}\mathbf{x}, \tag{113}$$

the characteristic pyramid/mound size, denoted $\lambda(t)$, and the energy. For the no-slope-selection equation (4), one obtains $w \sim O\left(t^{1/2}\right)$, $\lambda(t) \sim O\left(t^{1/4}\right)$, and $E \sim O\left(-\ln(t)\right)$. (See [11, 12] and references therein.) Observe that for the no-slope-selection equation (4), the average mound height (measured by the roughness) grows faster then the mound width (measured by λ), which is expected because there is no preferred slope of the height function ϕ. Also, note that in the rigorous setting, for example [8, 9, 12], one can only (at best) obtain lower bounds for the energy dissipation and, conversely, upper bounds for the roughness growth. However, the rates quoted as the upper or lower bounds are typically observed for the averaged values of the quantities of interest. Predicting these scaling laws numerically is quite challenging, since doing so requires very long simulation times. To adequately capture the full range of coarsening behaviors, numerical simulations for the coarsening process require short- and long-time accuracy and stability, in addition to high spatial accuracy for small values of ϵ.

Here we show a numerical simulation result using the proposed second order scheme (65) combined with the linear iteration algorithm (25) for the no-slope-selection equation (4) to compare our computed solutions against the predicted coarsening rates. This test is a repeat of those given in our previous papers [1, 21]. The surface diffusion coefficient parameter is taken to be $\epsilon = 0.02$. For the domain we take $L = L_x = L_y = 12.8$ and $h = L/N$, where h is the uniform spatial step size.
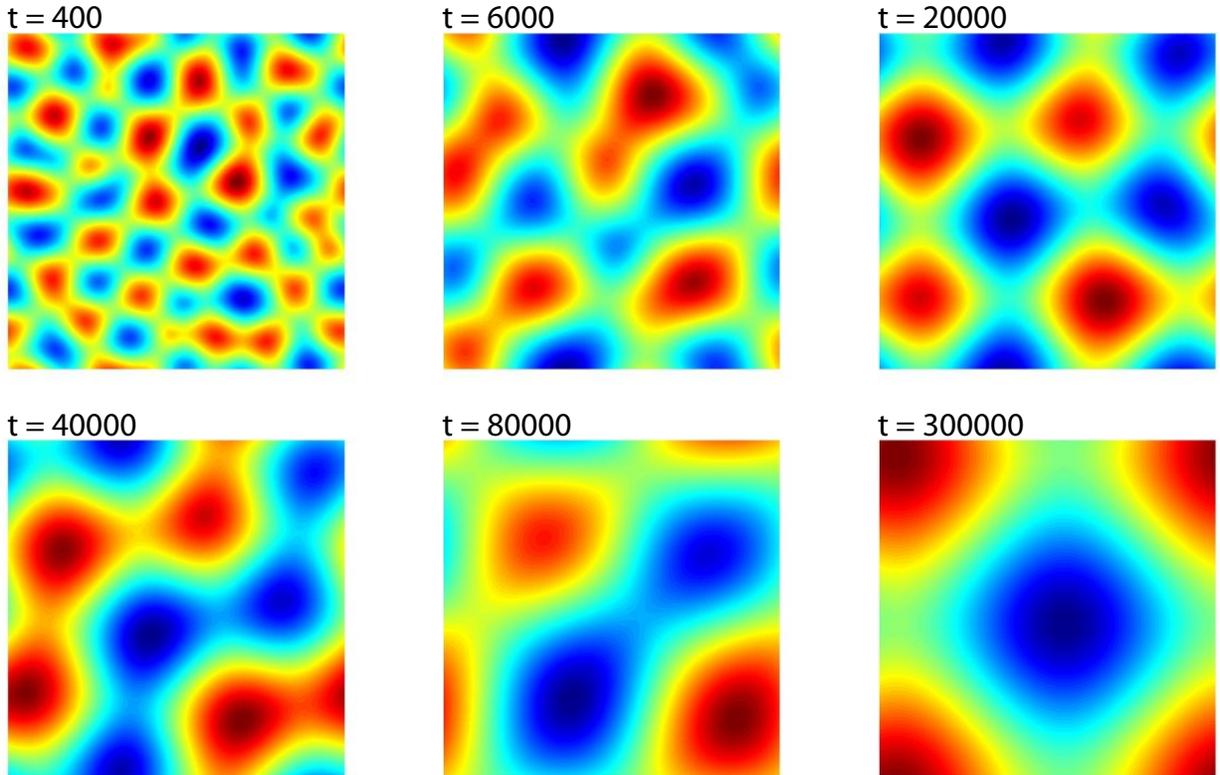
Figure 6: (Color online.) Snapshots of the computed height function ϕ at the indicated times for the parameters $L = 12.8$, $\epsilon = 0.02$. Note that the color scale changes with time. The hills (red) at early times are not as high as time at later times, and similarly with the valley (blue). To see how the average height/depth changes with time, see Fig. 8.

For such a value of ϵ, our previous numerical experiments have shown that $N = 512$ is adequate to resolve the small structures in the solution. For the temporal step size s, we use increasing values of s, namely, $s = 0.004$ on the time interval $[0, 400]$, $s = 0.04$ on the time interval $[400, 6000]$, $s = 0.16$ on the time interval $[6000, 10^5]$, and $s = 0.32$ for $t > 10^5$. Whenever a new time step size is applied, we initiate the two-step numerical scheme by taking $\phi^{-1} = \phi^0$, with the initial data ϕ^0 given by the final time output of the last time period. Both the energy stability and second order numerical accuracy are assured by our arguments in Sections 3, 4. Figure 6 presents time snapshots of the film height ϕ with $\epsilon = 0.02$. Significant coarsening in the system is evident. At early times many small hills (red) and valleys (blue) are present. At the final time, $t = 300000$, a one-hill-one-valley structure emerges, and further coarsening is not possible.

The long time characteristics of the solution, especially the energy decay rate and surface roughness growth rate, are of interest to surface scientists. Recall that, at the space-discrete level, the energy is defined via (72). The space-continuous surface roughness is defined in (113), and an analogous fully discrete version is also available. At the PDE level, the lower bound for the energy decay rate is of the order of $-\ln(t)$, and the upper bound for the standard deviation growth rate is of the order of $t^{1/2}$, as established for the no-slope-selection equation (4) in Li and Liu's work [12]. Figures 7 and 8 present the semi-log plots for the energy versus time and log-log plots for the roughness versus time, respectively, with the given physical parameter $\epsilon = 0.02$. The detailed scaling "exponents" are obtained using least squares fits of the computed data up to time $t = 400$.
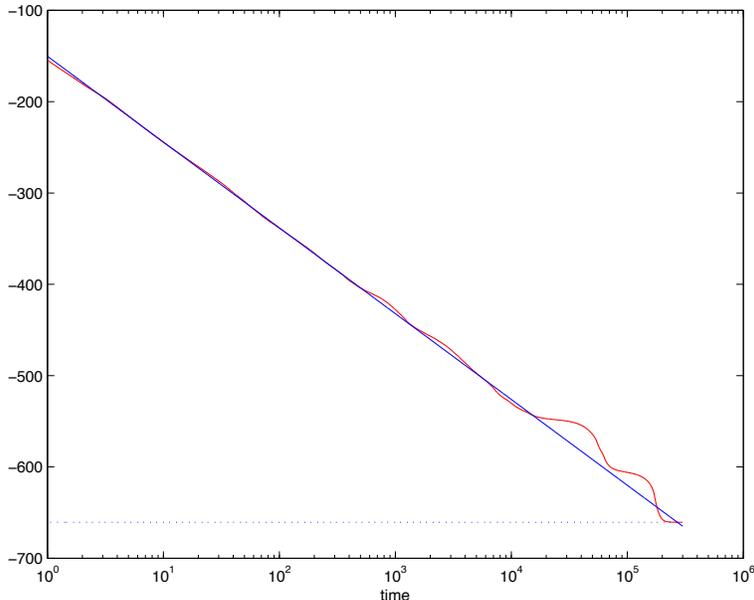
Figure 7: Semi-log plot of the temporal evolution the energy E_N for $\epsilon = 0.02$. The energy decreases like $-\ln(t)$ until saturation. The dotted lines correspond to the minimum energy reached by the numerical simulation. The red lines represent the energy plot obtained by the simulations, while the straight lines are obtained by least squares approximations to the energy data. The least squares fit is only taken for the linear part of the calculated data, only up to about time $t = 400$. The fitted line has the form $m_e \ln(t) + b_e$, with $m_e = -40.68$, $b_e = -152.73$.

A clear observation of the $-\ln(t)$ and $t^{1/2}$ scaling laws can be made, with different coefficients dependent upon ϵ, or, equivalently, the domain size, L.

Now we recall that a lower bound for the energy (1), assuming $\Omega = (0, L) \times (0, L)$, which has been derived in our earlier article [21] and polished in a more recent one [1]:

$$E(\phi) \geq \frac{L^2}{2} \left(\ln \left(\frac{4\epsilon^2 \pi^2}{L^2} \right) - \frac{4\epsilon^2 \pi^2}{L^2} + 1 \right) =: \gamma \, . \tag{114}$$

Obviously, since the energy is bounded below it cannot keep decreasing at the rate $-\ln(t)$. This fact manifests itself in the calculated data as the rate of decrease of the energy, for example, begins to wildly deviate from the predicted $-\ln(t)$ curve. Sometimes the rate of decrease increases, and sometimes it slows as the systems "feels" the periodic boundary conditions. Interestedly, regardless of this later-time deviation from the accepted rates, the time at which the system saturates (*i.e.*, the time when the energy abruptly and essentially stops decreasing) is roughly that predicted by extending the blue lines in Fig. 7 to the predicted minimum energy (114).

Remark 5.1. *In this numerical simulation, the time step sizes are taken as 4 times larger as the ones taken in the first order linear splitting scheme presented in our earlier work [1], at different time range. Meanwhile, the computational cost at each time step is about 3 to 5 times as that of the first order scheme, due to the presence of linear iteration algorithm. Thus, in the final analysis, the total computational cost is at a comparable level as that of the first order scheme.*

For the long time simulation, both the first order and second order schemes have produced similar evolutionary curves in terms of energy and the standard deviation, as presented in Figs. 7 and 8. A more detailed calculation shows that long time asymptotic growth rate of the standard deviation
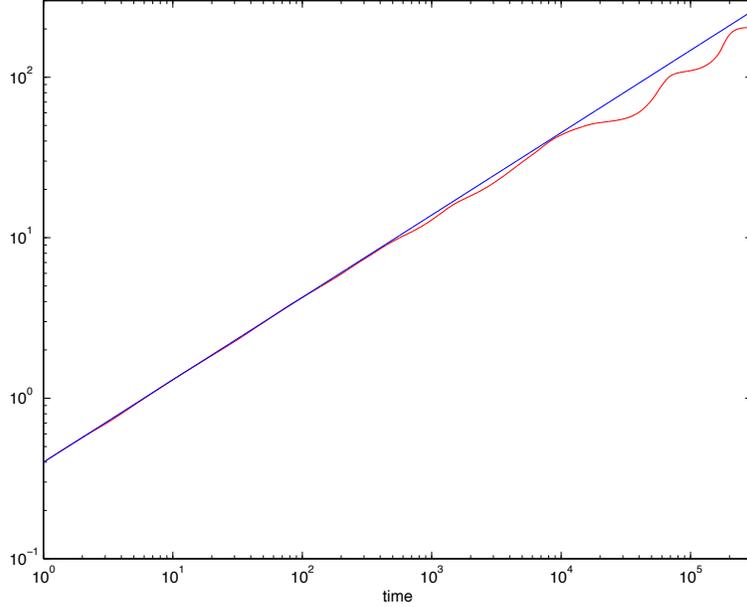
23

Figure 8: The log-log plot of the standard deviation (or roughness) of ϕ, denoted $w(t)$ for $\epsilon = 0.02$. For the no slope selection model, $w(t)$ grows like $t^{1/2}$. The red lines represent the plot obtained by the numerical simulations, while the straight lines are linear least squares approximations to the $t^{1/2}$ growth. The least squares fit is only taken for the linear part of the calculated data, only up to about time $t = 400$. The (blue) fitting line has the form $b_r t^{m_r}$, with $b_r = 0.40$, $m_r = 0.51$.

given by the second order numerical simulation is closer to $t^{1/2}$ than that by the first order scheme. Here we found $m_r = 0.51$, as in recorded in Fig. 8, while in [1] this exponent was found to be $m_r = 0.52$. This gives more evidence that the second order scheme is able to produce more accurate long time numerical simulation results than the first order schemes, even if its time step size is 4 times larger than the later (so that they have comparable computational costs).

6   Summary and Remarks

In this paper we have presented a linear iteration algorithm to implement an unconditionally energy stable second-order convex splitting scheme for thin film epitaxy without slope selection, *i.e*, the no-slope-selection equation (4). The second order convex splitting was given by a recent article [18]. Here an $O(s^2)$ artificial diffusion term, a Douglas-Dupont-type regularization, is added to assure a contraction mapping property of our proposed linear iteration. The addition of this regularization does not affect the unconditional unique solvability and unconditional energy stability of the scheme. Moreover, a global in time H^2 bound for the numerical solution is obtained at the discrete level and the convergence for the numerical scheme in a discrete $L^\infty(0, T; H^1) \cap L^2(0, T; H^3)$ norm is proved. This convergence if made possible by the available bounds of the nonlinear terms involved in the numerical scheme.

   The work here demonstrates an important tool to implement a highly nonlinear scheme, namely a linear iteration. We envision that this technique will have more applications in many other nonlinear convex splitting schemes for gradient equations. Here, the nonlinear system can be decomposed as an iteration of purely linear solvers, which can be very efficiently implemented with the help of FFT in a collocation Fourier spectral setting. The numerical simulation experiments

showed that the second order scheme, combined with the linear iteration algorithm, is able to produce a more accurate long time numerical results than the first order schemes reported in [1, 21], with a comparable computational cost. This is remarkable when one notes that the scheme in [1] is linear.

Acknowledgment

References

[1] Chen, W.; Conde, S.; Wang, C.; Wang, X. and Wise, S.M.; *A linear energy stable scheme for a thin film model without slope selection*, J. Sci. Comput., 52, (2012) 546-562.

[2] Chen, W.; and Wang, Y.; *A mixed finite element method for thin film epitaxy*, Numer. Math., 122, (2012) 771-793.

[3] Ehrlich, G. and Hudda, F.G.; *Atomic view of surface diffusion: Tungsten on tungsten*, J. Chem. Phys., 44, (1966) 1036-1099.

[4] Evans, J.W.; Thiel, P.A.; and Bartelt, M.C.; *Morphological evolution during epitaxial thin film growth: Formation of 2D islands and 3D mounds*, Surface Science Reports 61 (2006) 1-128.

[5] Eyre, D.J.; *Unconditionally gradient stable time marching the Cahn-Hilliard equation*, in Computational and Mathematical Models of Microstructural Evolution, J. W. Bullard, R. Kalia, M. Stoneham, and L.Q. Chen, eds., Materials Research Society, Warrendale, PA, 53 (1998) 1686-1712.

[6] Hu, Z.; Wise, S.M.; Wang, C.; and Lowengrub, J.S.; *Stable and efficient finite-difference nonlinear-multigrid schemes for the Phase Field Crystal equation*, J. Comput. Phys., 228 (2009) 5323-5339.

[7] Johnson, M.D.; Orme, C.; Hunt, A.W.; Graff, D.; Sudijono, J.; Sander, L.M.; and Orr, B.G.; *Stable and unstable growth in molecular beam epitaxy*, Phys. Rev. Lett., 72 (1994) 116-119.

[8] Kohn, R.V.; *Energy-driven pattern formation*, in Proceedings of the International Congress of Mathematicians, M. Sanz-Sole, J. Soria, J.L. Varona, and J. Verdera, eds., European Mathematical Society Publishing House, Madrid, 1 (2007) 359-384.

[9] Kohn, R.V. and Yan, X.; *Upper bound on the coarsening rate for an epitaxial growth model*, Comm. Pure Appl. Math., 56 (2003) 1549-1564.

[10] Li, B.; *High-order surface relaxation versus the Ehrlich-Schwoebel effect*, Nonlinearity, 19 (2006), 25812603.

[11] Li, B. and Liu, J.-G.; *Thin film epitaxy with or without slope selection*, Euro. J. Appl. Math., 14 (2003) 713-743.

[12] Li, B. and Liu, J.-G.; *Epitaxial growth without slope selection: energetics, coarsening, and dynamic scaling*, J. Nonlinear Sci., 14 (2004) 429-451.

[13] Moldovan, D. and Golubovic L.; *Interfacial coarsening dynamics in epitaxial growth with slope selection*, Phys. Rev. E, 61 (2000) 6190-6214.

[14] Qiao, Z.; Sun Z.; and Zhang Z.; *Stability and convergence of second-order schemes for the nonlinear epitaxial growth model without slope selection*, Math. Comp., (submitted and in review).

[15] Qiao, Z.; Zhang Z.; and Tang, T.; *An adaptive time-stepping strategy for the molecular beam epitaxy models*, SIAM J Sci. Comput., 33 (2011) 1395-1414.

[16] Qiao, Z.; Sun Z.; and Zhang Z.; *The stability and convergence of two linearized finite difference schemes for the nonlinear epitaxial growth model*, Numer. Methods Partial Diff. Eq., 28 (2012) 1893-1915.

[17] Schwoebel, R.L.; *Step motion on crystal surfaces: II*, J. Appl. Phys., 40 (1969) 614-618.

[18] Shen, J.; Wang, C.; Wang, X. and Wise, S.M.; *Second-order convex splitting schemes for gradient flows with Ehrlich-Schwoebel type energy: Application to thin film epitaxy*, SIAM J. Numer. Anal., 50, (2012) 105-125.

[19] Trefethen, L.N.; *Spectral Methods in Matlab*, SIAM, Philadelphia, PA, 2000.

[20] Wise, S.M.; Wang, C. and Lowengrub, J.S.; *An energy stable and convergent finite-difference scheme for the Phase Field Crystal equation*, SIAM J. Numer. Anal., 47, (2009) 2269-2288.

[21] Wang, C.; Wang, X. and Wise, S.M.; *Unconditionally stable schemes for equations of thin film epitaxy*, Discrete and Continuous Dynamical Systems-Series A, 28, (2010) 405-423.

[22] Xu, C. and Tang, T.; *Stability analysis of large time-stepping methods for epitaxial growth models*, SIAM J. Numer. Anal., 44, (2006) 1759-1779.