



# Convex splitting method for the calculation of transition states of energy functional



Shuting Gu\*, Xiang Zhou<sup>1</sup>

Department of Mathematics, City University of Hong Kong, Tat Chee Ave, Kowloon, Hong Kong Special Administrative Region

## ARTICLE INFO

### Article history:

Received 20 October 2016

Received in revised form 8 September 2017

Accepted 19 October 2017

Available online 24 October 2017

### Keywords:

Transition state

Saddle point

Convex splitting method

Iterative minimization formulation

## ABSTRACT

Among numerical methods for partial differential equations arising from steepest descent dynamics of energy functionals (e.g., Allen–Cahn and Cahn–Hilliard equations), the convex splitting method is well-known to maintain unconditional energy stability for a large time step size. In this work, we show how to use the convex splitting idea to find transition states, i.e., index-1 saddle points of the same energy functionals. Based on the iterative minimization formulation (IMF) for saddle points [14], we introduce the convex splitting method to minimize the auxiliary functional at each cycle of the IMF. We present a general principle of constructing convex splitting forms for these auxiliary functionals and show how to avoid solving nonlinear equations. The new numerical scheme based on the convex splitting method allows for large time step sizes. The new methods are tested for the one dimensional Ginzburg–Landau energy functional in the search of the Allen–Cahn or Cahn–Hilliard types of transition states. We provide the numerical results of transition states for the two dimensional Landau–Brazovskii energy functional for diblock copolymers.

© 2017 Elsevier Inc. All rights reserved.

## 1. Introduction

For an energy functional, both its local minimizers and its unstable saddle points have important physical meanings for many problems in physics, chemistry, biology and material sciences. The local minimizers correspond to the stable configurations in physical models, and they manifest themselves as steady states of the gradient flow driven by the energy. For the spatially extended systems, these flows appear mathematically as the time-dependent partial differential equations (PDEs). These PDEs reflect the true physical dynamics of relaxations and they also serve as a convenient computational model to calculate the stable steady states. For instance, as the well-known phase separation and transition models, the Allen–Cahn (AC) [1] and Cahn–Hilliard (CH) [4] equations are the steepest descent dynamics of the Ginzburg–Landau energy functional under  $L^2$  and  $H^{-1}$  norms, respectively. Besides the local minima, the other critical points on an energy surface also play crucial roles for certain problems, such as the energy-barrier activated processes which escape from local minima by crossing saddle points. The infrequent hoppings between neighboring local minima, although randomly, occur in a quite certain style of travelling through transition states. These transition states, as the bottlenecks on the pathways of activated processes, belong to a class of saddle points with index one, i.e., the unstable critical point whose Hessian has only one negative eigenvalue.

\* Corresponding author.

E-mail addresses: shutinggu2-c@my.cityu.edu.hk (S. Gu), xiang.zhou@cityu.edu.hk (X. Zhou).

<sup>1</sup> The research of XZ was supported by the grants from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. CityU 11304314, 109113 and 11304715).

In this paper, we are interested in how to find these index-1 saddle points for a given smooth energy functional. The search for transition states, or index-1 saddle points, faces many challenges. A large number of numerical methods have been proposed and developed to address these challenges. There are also many applications of these algorithms in computational chemistry and material sciences. Refer to [33,41] for review of this topic. One class of numerical methods is to search the so-called minimum energy path (MEP). The points along an MEP with locally maximal energy value are then the index-1 saddle points. These path-finding methods include [8,9] and [23,20]. The other class of methods is to evolve a single state on the potential energy surface. The essential question is how to define some dynamics on the energy surface to converge to index-1 saddle points without knowing multiple local minima *a priori*. The intuitive idea of using the softest direction (corresponding to the minimal eigenvalue of the Hessian) to invert the force component along this min-mode direction proves very useful ([21,32]) and it was proposed probably as early as in 1970s in [7,6]. Many well-known algorithms and softwares such as the dimer method ([18,24,40]) or the activation–relaxation techniques ([31,30,5]) are based on this min-mode-following idea.

The underlying dynamics of these min-mode-following algorithms has been rigorously formulated and analyzed in [10]. This dynamics, with the name “gentlest ascent dynamics” (GAD), simultaneously evolves both a position variable and a direction variable. By analyzing the eigenvalue of the GAD, [10] proved the locally linear convergence to saddle point. To accelerate the convergence rate, a discrete iterative mapping, named the “iterative minimization formulation” (IMF), has been proposed in [14]. IMF has three advantageous features: (1) it has the quadratic convergence rate for non-degenerate saddle points; (2) it turns the problem of searching unstable saddle points into a series of minimization subproblems; (3) there is no restriction for numerical methods to solve minimization subproblems in the IMF. The only important issue in practice for better efficiency is to use the adaptive stopping rule in solving the subproblems. For a thorough discussion of the practical algorithms based on the IMF, the readers can refer to [15].

The advantages and flexibilities offered by the IMF immediately provide many new opportunities to explore the existing methods which were designed for searching local minima. The convex splitting method, originally proposed in [12], successfully gives unconditionally energy stable schemes to ensure a large time step size. The effectiveness of this method in resolving gradient dynamics as well as calculating the minimizers has been demonstrated by a vast number of applications, for example, the phase field model in [13], the phase field crystal model in [38], the thin film epitaxy model in [34], the binary fluid surfactant model in [16], as well as many others ([35,36,39,22]).

In this paper, our motivation is to test the performance of the convex splitting method if this strategy is used to locate the saddle point. As we mentioned earlier, the IMF solves the saddle point problem by solving a series of minimization subproblems, and these subproblems can be solved by running the steepest decent dynamics. Since the outer iteration of the IMF (referred as “cycle” in [14,15]) is of quadratic convergence rate and it usually only takes a few cycles in practice to reach the desired accuracy, then one can expect that a better method for the subproblem may gain a better speedup in efficiency. The benefit of the convex splitting idea here is that the subproblems can be solved by a large time step size. Therefore it takes less steps in each cycle to improve the overall efficiency of locating the saddle point in the IMF.

The idea of the convex splitting method is quite simple, but there are two important practical issues when applied to specific problems. The first is the construction of a convex splitting form for a given energy functional. In theory ([12]), there always exists a convex splitting form for any continuous functional. The explicit decomposition has to be sought for specific problems. The second is that one should try best to construct a *linear* time-implicit term in the convex splitting scheme since this can avoid solving a nonlinear system at each time step.

The contributions in our work of applying the convex splitting method to saddle point search problems include the following: (1) for any given convex splitting form of the original energy functional, we show how to obtain the corresponding convex splitting form of the auxiliary functional in the IMF. This means that we design an automatic procedure from the traditional convex splitting method for local minimizers to the convex splitting method for saddle points; (2) we shall see later that the auxiliary functional in the IMF consists of multiple terms involving the original energy functional. By adapting different convex splitting forms for different terms in the auxiliary functional, we can ensure the time-explicit discretization for nonlinear terms and obtain a linear system. We demonstrate how to achieve this by the examples of Ginzburg–Landau energy functional and the Landau–Brazovskii energy functional. The condition is that one need to know at least one convex splitting form with the time-explicit nonlinear term for the original energy functional.

The rest of the paper is organized as follows. In Section 2, we review the IMF for the saddle point search problem and the convex splitting method. In Section 3, we construct the convex splitting method for saddle point search problems and discuss numerical issues. Section 4 presents the detailed numerical schemes and substantial numerical results for the Ginzburg–Landau energy functional in the  $L^2$  and  $H^{-1}$  metric, respectively, subject to the Neumann or periodic boundary condition and the Landau–Brazovskii energy functional with periodic boundary condition. The conclusion is drawn in Section 5.

## 2. Review

In this section, we review two foundations of our method, which were born apparently from two different areas.

### 2.1. The IMF for the saddle point search

We first recall the iteration minimization formulation (IMF) in [14]. Let  $\mathcal{M}$  be a Hilbert space equipped with the norm  $\|\cdot\|$  and the inner product  $\langle \cdot, \cdot \rangle$ . Suppose that  $V(x) : \mathcal{M} \rightarrow \mathbb{R}$  is a sufficiently smooth potential function, then the IMF is the following iteration for the position variable  $x$  and the direction variable  $v$

$$\begin{cases} v^{(k+1)} = \operatorname{argmin}_{\|u\|=1} \langle u, H(x^{(k)})u \rangle, & \text{(a)} \\ x^{(k+1)} = \operatorname{argmin}_y L(y; x^{(k)}, v^{(k+1)}), & \text{(b)} \end{cases} \tag{2.1}$$

where

$$H(x^{(k)}) = \nabla^2 V(x^{(k)}),$$

and

$$\begin{aligned} L(y; x^{(k)}, v^{(k+1)}) = & (1 - \alpha)V(y) + \alpha V \left( y - (v^{(k+1)} \otimes v^{(k+1)})(y - x^{(k)}) \right) \\ & - \beta V \left( x^{(k)} + (v^{(k+1)} \otimes v^{(k+1)})(y - x^{(k)}) \right). \end{aligned} \tag{2.2}$$

$\alpha$  and  $\beta$  are two parameters and  $\alpha + \beta > 1$ . Two special choices for  $\alpha$  and  $\beta$  are: (i)  $(\alpha, \beta) = (2, 0)$ , then  $L(y; x, v) = -V(y) + 2V(y - (v \otimes v)(y - x))$ ; (ii)  $(\alpha, \beta) = (0, 2)$ , then  $L(y; x, v) = V(y) - 2V(x + (v \otimes v)(y - x))$ . So, the general form of  $L$  in (2.2) should be a linear combination of three terms involved in these two extreme cases and furthermore the coefficient of the first term  $V(y)$  is determined to be  $1 - \alpha$  by examining the Hessian  $\nabla_y^2 L(y = x; x, v)$ . The main properties of the auxiliary objective function  $L(y; x, v)$  when  $\alpha + \beta > 1$  are listed here for reference.

**Theorem 1** ([14]). *Suppose that  $x^*$  is a (non-degenerate) index-1 saddle point of a  $C^4$ -function  $V(x)$ , i.e., its all eigenvalues are  $\lambda_1 < 0 < \lambda_2 \leq \dots$ , and the auxiliary function  $L$  is defined by (2.2) with  $\alpha + \beta > 1$ , then*

- (1) *there exists a neighborhood  $\mathcal{U}$  of  $x^*$  such that for any  $x \in \mathcal{U}$ ,  $L(y; x, v)$  is strictly convex in  $y \in \mathcal{U}$  and thus has a unique minimum in  $\mathcal{U}$ ;*
- (2) *define the mapping  $\Phi : x \in \mathcal{U} \rightarrow \Phi(x) \in \mathcal{U}$  to be the unique minimizer of  $L$  in  $\mathcal{U}$  for any  $x \in \mathcal{U}$ . Then the mapping  $\Phi$  has only one fixed point  $x^*$ ;*
- (3) *the mapping  $x \rightarrow \Phi(x)$  has a quadratic convergence rate.*

The IMF includes two-level iterations. The top level is  $x \rightarrow \Phi(x)$ , referred as “cycle”. The  $k$ -th cycle means the step of  $x^{(k)} \rightarrow x^{(k+1)} = \Phi(x^{(k)})$ , which in practice consists of a second-level iterative procedure to solve (2.1)a for the min-mode (the so-called “rotation step”) and (2.1)b to update the position (“translation step”). The rotation step is a classical numerical eigenvector problem, for which many methods have been constructed such as the power method ([10]), the conjugate gradient method ([19,21]), the Lanczos algorithm ([31,30,5]) and the LOR in [26].

We are interested in spatially extended systems, i.e.,  $\mathcal{M}$  is a function space and  $V$  is actually a functional on  $\mathcal{M}$ . As a convention, we use  $F(\phi)$  rather than  $V(x)$  below to represent the functional of a spatial function  $\phi$ .

### 2.2. Convex splitting method

Let  $\phi(x, t) : [0, 1] \times \mathbb{R}^+ \rightarrow \mathbb{R}$  be the solution of the following PDE driven by the gradient flow

$$\frac{\partial \phi}{\partial t} = -\frac{\delta F}{\delta \phi}(\phi), \tag{2.3}$$

subject to certain boundary condition at  $x = 0$  and  $1$ .  $F$  is a sufficiently smooth free energy functional bounded from below.  $\frac{\delta F}{\delta \phi}$  is the first order variational derivative of  $F(\phi)$  with respect to  $\phi$ . A convex splitting form of  $F(\phi)$  means that two convex functionals exist, denoted by  $F_c$  and  $F_e$ , such that

$$F(\phi) = F_c(\phi) - F_e(\phi), \tag{2.4}$$

where “c” (“e”) refers to the contractive (expansive) part of the energy ([12]). Then the convex splitting scheme for (2.3) is

$$\frac{\phi^{n+1} - \phi^n}{\Delta t} = -\left( \frac{\delta F_c}{\delta \phi}(\phi^{n+1}) - \frac{\delta F_e}{\delta \phi}(\phi^n) \right), \tag{2.5}$$

where  $\phi^n \approx \phi(t_n)$  is the numerical solution at the  $n$ -th time level  $t_n = n\Delta t$  and  $\Delta t$  is the time step size. The time-discrete scheme (2.5) has the property of the so-called unconditional energy stability, as stated in the following theorem.

**Theorem 2 ([38]).** Suppose the free energy functional  $F(\phi)$  can be split into two parts  $F(\phi) = F_c(\phi) - F_e(\phi)$  as in (2.4). Then the time-discrete scheme (2.5) is unconditionally energy stable, meaning that for any time step size  $\Delta t > 0$ ,

$$F(\phi^{n+1}) \leq F(\phi^n), \quad n = 0, 1, 2, \dots$$

The convexity of  $F_c$  and  $F_e$  do not have to be valid in the whole configuration space. In practical applications, it only requires the convexity in bounded subsets of the configuration space, where the solution is known to be in.

### 3. Main method

Consider the 1-D spatial domain  $(a, b)$ . The high dimensional case follows exactly the same idea. Let  $\mathcal{M}$  be a function space on the interval  $[a, b]$ . For example,  $\mathcal{M}$  is the Hilbert space  $H^1([a, b])$  or some other subspaces of  $L^2([a, b])$ . Our goal is to find the transition states of the energy functional  $F$  on  $\mathcal{M}$ .

Assume that the second order variational derivative (Hessian) of  $F$  in  $\mathcal{M}$ , denoted by  $H(\phi) := \frac{\delta^2 F(\phi)}{\delta \phi^2}$ , exists. We rewrite the IMF in Section 2.1:

$$\begin{cases} v^{(k+1)} = \operatorname{argmin}_{\|v\|=1} \langle v, H(\phi^{(k)})v \rangle, & \text{(a)} \\ \phi^{(k+1)} = \operatorname{argmin}_{\phi} L(\phi; \phi^{(k)}, v^{(k+1)}), & \text{(b)} \end{cases} \tag{3.1}$$

where the auxiliary functional  $L$  is

$$L(\phi; \phi^{(k)}, v^{(k+1)}) = F(\phi) - \alpha F(\phi) + \alpha F(\phi - \tilde{\phi}) - \beta F(\phi^{(k)} + \tilde{\phi}), \tag{3.2}$$

where we define

$$\tilde{\phi} := (v^{(k+1)} \otimes v^{(k+1)})(\phi - \phi^{(k)}) \tag{3.3}$$

to ease the notation. The min-mode  $v^{(k+1)}$  belongs to  $\mathcal{T}_{\mathcal{M}}$ , the tangent space of  $\mathcal{M}$ .  $u \otimes v$  denotes the tensor defined by  $(u \otimes v)\phi = \langle v, \phi \rangle u$  for  $u, v \in \mathcal{T}_{\mathcal{M}}$ .

#### 3.1. Convex splitting method for minimizing the auxiliary functional $L$

We shall approximate the solution of the variational subproblem (3.1)b

$$\min_{\phi} L(\phi; \phi^{(k)}, v^{(k+1)})$$

at the  $k$ -th IMF cycle by the steady solution of the gradient flow associated with  $L$ ,

$$\frac{\partial \phi}{\partial t} = -\frac{\delta L}{\delta \phi}. \tag{3.4}$$

Note  $\phi^{(k)}$  and  $v^{(k+1)}$  are fixed here. The solution at infinite time is well-defined if the lowest eigenvalue of  $L$  at  $\phi = \phi^{(k)}$  is negative ([14,15]).

Next, we show how to construct the convex splitting scheme for equation (3.4). Our starting point is that a convex splitting form for  $F(\phi)$  has been given, say,

$$F(\phi) = F_c(\phi) - F_e(\phi). \tag{3.5}$$

By substituting (3.5) into (3.2), we find the following convex splitting form for  $L$

$$L(\phi; \phi^{(k)}, v^{(k+1)}) = L_c(\phi; \phi^{(k)}, v^{(k+1)}) - L_e(\phi; \phi^{(k)}, v^{(k+1)}), \tag{3.6}$$

where

$$L_c(\phi; \phi^{(k)}, v^{(k+1)}) = \begin{cases} F_c(\phi) + \alpha F_e(\phi) + \alpha F_c(\phi - \tilde{\phi}) + \beta F_e(\phi^{(k)} + \tilde{\phi}), & \text{if } \alpha \geq 0, \beta \geq 0; \\ F_c(\phi) + \alpha F_e(\phi) + \alpha F_c(\phi - \tilde{\phi}) - \beta F_c(\phi^{(k)} + \tilde{\phi}), & \text{if } \alpha \geq 0, \beta \leq 0; \\ F_c(\phi) - \alpha F_c(\phi) - \alpha F_e(\phi - \tilde{\phi}) + \beta F_e(\phi^{(k)} + \tilde{\phi}), & \text{if } \alpha \leq 0, \beta \geq 0, \end{cases} \tag{3.7}$$

and  $L_e(\phi; \phi^{(k)}, v^{(k+1)})$  is defined likewise by exchanging the subindices “e” and “c” in (3.7). Note that  $\tilde{\phi}$  has been defined in (3.3).

**Property 3.**  $L_c(\phi; \phi^{(k)}, v^{(k+1)})$  and  $L_e(\phi; \phi^{(k)}, v^{(k+1)})$  defined above are all convex with respect to  $\phi$  for any  $\phi^{(k)}$  and any  $v^{(k+1)}$ .

**Proof.** We only prove the case when  $\alpha \geq 0, \beta \geq 0$  since the other two cases can be proved similarly. The proofs of the convexity for  $L_c$  and  $L_e$  are the same, thus it suffices to only show that two special terms  $F_c(\phi - \tilde{\phi})$  and  $F_e(\phi^{(k)} + \tilde{\phi})$  in (3.7) are both convex in terms of  $\phi$ . Denote the Hessian of  $F_c(\phi)$  and  $F_e(\phi)$  by  $H_c$  and  $H_e$ , respectively, then both  $H_c$  and  $H_e$  are semi-positive definite by the property of convexity. By the definition of  $\tilde{\phi}$  in (3.3), the second order derivatives of  $F_c(\phi - \tilde{\phi})$  and  $F_e(\phi^{(k)} + \tilde{\phi})$  are  $[I - (v \otimes v)]H_c(\phi - \tilde{\phi})[I - (v \otimes v)]$  and  $(v \otimes v)H_e(\phi^{(k)} + \tilde{\phi})(v \otimes v)$ , respectively, where  $v = v^{(k+1)}$ . Since  $v$  is nonzero, these two (projected) Hessians are also semi-positive definite. This completes our proof.  $\square$

We now present our time-discrete numerical scheme for (3.4) based on the convex splitting idea in Section 2.2. Here we only consider the case  $\alpha \geq 0, \beta \geq 0$ . To simplify the notation, we drop out the parameters  $\phi^{(k)}$  and  $v^{(k+1)}$  in the expressions of  $L(\phi, \phi^{(k)}, v^{(k+1)})$ ,  $L_c(\phi, \phi^{(k)}, v^{(k+1)})$  and  $L_e(\phi, \phi^{(k)}, v^{(k+1)})$ . Let  $\phi^n$  be the numerical solution at the time level  $t_n$ . Our scheme is

$$\frac{\phi^{n+1} - \phi^n}{\Delta t} = - \left[ \frac{\delta L_c}{\delta \phi}(\phi^{n+1}) - \frac{\delta L_e}{\delta \phi}(\phi^n) \right], \tag{3.8}$$

with initial  $\phi^0 = \phi^{(k)}$ . The first order variational derivative of  $L_c$  is

$$\begin{aligned} \frac{\delta L_c}{\delta \phi}(\phi) &= \frac{\delta F_c}{\delta \phi}(\phi) + \alpha \frac{\delta F_e}{\delta \phi}(\phi) \\ &\quad + \alpha (I - (v^{(k+1)} \otimes v^{(k+1)})) \frac{\delta F_c}{\delta \phi}(\phi - (v^{(k+1)} \otimes v^{(k+1)})(\phi - \phi^{(k)})) \\ &\quad + \beta (v^{(k+1)} \otimes v^{(k+1)}) \frac{\delta F_e}{\delta \phi}(\phi^{(k)} + (v^{(k+1)} \otimes v^{(k+1)})(\phi - \phi^{(k)})), \end{aligned}$$

and  $\frac{\delta L_e}{\delta \phi}(\phi)$  has the similar form by switching subindices “c” and “e”.

For the same reason in Theorem 2, the following unconditional energy stability holds for the scheme (3.8). The proofs are straightforward and skipped.

**Lemma 4.** Suppose that  $\phi, \psi : [a, b] \times \mathbb{R}^+ \rightarrow \mathbb{R}$  are two periodic functions. If  $L(\phi)$  in (3.2) has the convex splitting form  $L(\phi) = L_c(\phi) - L_e(\phi)$  given in (3.6). Then

$$L(\phi) - L(\psi) \leq \langle \delta_\phi L_c(\phi) - \delta_\phi L_e(\psi), \phi - \psi \rangle_{L^2},$$

where  $\delta_\phi L_c$  and  $\delta_\phi L_e$  represent the first order variational derivatives of  $L_c$  and  $L_e$  with respect to  $\phi$ , respectively.

**Theorem 5.** If the energy functional  $F(\phi)$  has the convex splitting form  $F = F_c - F_e$ , then the time-discrete scheme (3.8) is unconditionally energy stable, meaning that for any time step size  $\Delta t > 0$ , we have

$$L(\phi^{n+1}) \leq L(\phi^n), \quad n = 0, 1, 2, \dots,$$

in each  $k$ -th cycle.

### 3.2. Avoid the time-implicit nonlinear term and construct the linear system

In many cases, several convex splitting forms for  $F$  may be found. Assume  $F(\phi)$  has two forms of convex splitting with the following properties:

$$F(\phi) = F_c^1(\phi) - F_e^n(\phi), \tag{3.9}$$

and

$$F(\phi) = \tilde{F}_c^n(\phi) - \tilde{F}_e^1(\phi), \tag{3.10}$$

where the superscripts “1” and “n” mean that the first order variational derivative is linear or nonlinear in  $\phi$ , respectively. Accordingly, two convex splitting schemes exist to solve the gradient flow  $\frac{\partial \phi}{\partial t} = -\frac{\delta F}{\delta \phi}$ :

$$\frac{\phi^{n+1} - \phi^n}{\Delta t} = -\frac{\delta F_c^1}{\delta \phi}(\phi^{n+1}) + \frac{\delta F_e^n}{\delta \phi}(\phi^n), \tag{3.11}$$

and

$$\frac{\phi^{n+1} - \phi^n}{\Delta t} = -\frac{\delta \tilde{F}_c^n}{\delta \phi}(\phi^{n+1}) + \frac{\delta \tilde{F}_e^1}{\delta \phi}(\phi^n). \tag{3.12}$$

Both schemes satisfy the unconditional energy stability. The difference is that the nonlinear terms are handled differently. The scheme (3.11) is time-explicit in nonlinear term and hence solves a linear system to generate  $\phi^{n+1}$  at the next time level. But the scheme (3.12) requires to solve a nonlinear equation for  $\phi^{n+1}$ . If one has a very efficient nonlinear solver, such as the multi-grid method in [25] for vector-valued Allen–Cahn equation, then the scheme (3.12) is also a good choice. In other cases, the linear scheme (3.11) is usually preferred.

In the context of the IMF, our proposed convex splitting scheme for the saddle point problem faces the same difficulty of possible emergence of nonlinear time-implicit term. If only one convex splitting form is available, then the convex splitting form of the auxiliary functional  $L$  specified by (3.7) inevitably runs into this trouble. Either  $L_c$  or  $L_e$  includes both  $F_c$  and  $F_e$ , which leads to the appearance of at least one nonlinear term at the implicit time level. However, if one has both (3.9) and (3.10), then this difficulty can be circumvented by combining them together. Take  $\alpha, \beta \geq 0$  as an example again. Substituting (3.9) into the first and the third terms on the right hand side of (3.2) and substituting (3.10) into the second and the fourth terms on the right hand side of (3.2), then we have the following decomposition of  $L$ :

$$\begin{aligned} L(\phi) &= F_c^1(\phi) - F_e^n(\phi) - \alpha[\tilde{F}_c^n(\phi) - \tilde{F}_e^1(\phi)] \\ &\quad + \alpha[F_c^1(\phi - \tilde{\phi}) - F_e^n(\phi - \tilde{\phi})] - \beta[\tilde{F}_c^n(\phi^{(k)} + \tilde{\phi}) - \tilde{F}_e^1(\phi^{(k)} + \tilde{\phi})] \\ &= [F_c^1(\phi) + \alpha\tilde{F}_e^1(\phi) + \alpha F_c^1(\phi - \tilde{\phi}) + \beta\tilde{F}_e^1(\phi^{(k)} + \tilde{\phi})] \\ &\quad - [F_e^n(\phi) + \alpha\tilde{F}_c^n(\phi) + \alpha F_e^n(\phi - \tilde{\phi}) + \beta\tilde{F}_c^n(\phi^{(k)} + \tilde{\phi})] \\ &=: L_c(\phi) - L_e(\phi), \end{aligned}$$

where  $\tilde{\phi} = (v^{(k+1)} \otimes v^{(k+1)})(\phi - \phi^{(k)})$ . It is easy to see that both  $L_c$  and  $L_e$  are convex with respect to  $\phi$ . The proof is exactly the same as that for Property 3. Theorem 5 in Section 3.1 also holds true in this case.

In constructing two convex splitting forms satisfying the above conditions, we do not require the convexity holds for the whole configuration space. In fact, for many examples, the global convexity for all four functions in (3.9) and (3.10) is not possible. We only need the convexity properties valid locally for a certain range of the values of  $\phi$  in which the solution of interest does not violate.

In summary, the semi-discrete scheme for (3.4) by using the two forms of convex splitting of  $F(\phi)$ , (3.9) and (3.10), has the following expression:

$$\frac{\phi^{n+1} - \phi^n}{\Delta t} = -\left[\frac{\delta L_c}{\delta \phi}\right]^{n+1} + \left[\frac{\delta L_e}{\delta \phi}\right]^n, \tag{3.13}$$

where

$$\begin{aligned} \left[\frac{\delta L_c}{\delta \phi}\right]^{n+1} &= \frac{\delta F_c^1}{\delta \phi}(\phi^{n+1}) + \alpha \frac{\delta \tilde{F}_e^1}{\delta \phi}(\phi^{n+1}) \\ &\quad + \beta(v^{(k+1)} \otimes v^{(k+1)}) \frac{\delta \tilde{F}_e^1}{\delta \phi}(\phi^{(k)} + (v^{(k+1)} \otimes v^{(k+1)})(\phi^{n+1} - \phi^{(k)})) \\ &\quad + \alpha(I - (v^{(k+1)} \otimes v^{(k+1)})) \frac{\delta F_c^1}{\delta \phi}(\phi^{n+1} - (v^{(k+1)} \otimes v^{(k+1)})(\phi^{n+1} - \phi_k)), \\ \left[\frac{\delta L_e}{\delta \phi}\right]^n &= \frac{\delta F_e^n}{\delta \phi}(\phi^n) + \alpha \frac{\delta \tilde{F}_c^n}{\delta \phi}(\phi^n) \\ &\quad + \beta(v^{(k+1)} \otimes v^{(k+1)}) \frac{\delta \tilde{F}_c^n}{\delta \phi}(\phi^{(k)} + (v^{(k+1)} \otimes v^{(k+1)})(\phi^n - \phi^{(k)})) \\ &\quad + \alpha(I - (v^{(k+1)} \otimes v^{(k+1)})) \frac{\delta F_e^n}{\delta \phi}(\phi^n - (v^{(k+1)} \otimes v^{(k+1)})(\phi^n - \phi^{(k)})), \end{aligned}$$

and the initial value  $\phi^0 = \phi^{(k)}$ .  $\alpha \geq 0, \beta \geq 0$  and  $\alpha + \beta > 1$ . It is easy to see that  $\left[\frac{\delta L_c}{\delta \phi}\right]^{n+1}$  is linear and  $\left[\frac{\delta L_e}{\delta \phi}\right]^n$  is nonlinear. So the scheme (3.13) indeed corresponds to a linear system.

### 4. Applications

#### 4.1. 1D example: Ginzburg–Landau free energy

We apply our method to the Ginzburg–Landau free energy on  $[0, 1]$ ,

$$F(\phi) = \int_0^1 \left[ \frac{\kappa^2}{2} \left( \frac{\partial \phi}{\partial x} \right)^2 + f(\phi) \right] dx, \tag{4.1}$$

where  $\phi(x)$  is an order parameter representing for example the concentration of one of the component in a binary alloy. The mobility parameter  $\kappa > 0$ .  $f(\phi) = (\phi^2 - 1)^2/4$ . If we consider the gradient flow of  $F$  in the  $L^2([0, 1])$  space with the standard  $L^2$  inner product  $\langle \cdot, \cdot \rangle_{L^2}$ , then we obtain the (non-conserved) Allen–Cahn (AC) equation

$$\frac{\partial \phi}{\partial t} = - \frac{\delta F}{\delta \phi}(\phi) = \kappa^2 \Delta \phi - f'(\phi) = \kappa^2 \Delta \phi - (\phi^3 - \phi), \tag{4.2}$$

where  $\Delta = \partial_{xx}$ . If the gradient flow is defined in the  $H^{-1}$  metric  $\langle \cdot, \cdot \rangle_{H^{-1}}$ , then we have the (conserved) Cahn–Hilliard (CH) equation

$$\frac{\partial \phi}{\partial t} = \Delta \frac{\delta F}{\delta \phi} = -\kappa^2 \Delta^2 \phi + \Delta(\phi^3 - \phi). \tag{4.3}$$

We are interested in the unstable index-1 saddle point of the Ginzburg–Landau free energy (4.1). These saddle points correspond to the “spike-like” stationary solutions, or “canonical nuclei” discussed in [3]. Similarly to the AC and CH equations, which arise in  $L^2$  metric and  $H^{-1}$  metric, respectively, we search for the saddle points of  $F$  both in  $L^2$  and in  $H^{-1}$  metrics. The calculations of transition states and transition rates for the CH equation have already been done in [42,28] by using the string method ([8]) and the GAD ([10]).

We consider both the Neumann and periodic boundary condition. The Neumann boundary condition is  $\partial_x \phi(0) = \partial_x \phi(1) = 0$  for the AC equation and  $\partial_x \phi(0) = \partial_x \phi(1) = \partial_x^3 \phi(0) = \partial_x^3 \phi(1) = 0$  for the CH equation. The periodic boundary condition simply means that  $\phi(x) = \phi(x + 1)$ ,  $\forall x \in [0, 1]$ , which induces a degeneracy at any stationary solution corresponding to the translation invariance in the spatial variable  $\phi(x) \rightarrow \phi(x + c)$ . This means that the second smallest eigenvalue of the index-1 saddle points is zero. The degeneracy from the periodic boundary condition does not affect the quadratic convergence rate of the IMF ([17]). It is also noted that the mass  $\int_0^1 \phi(x) dx$  is conservative in  $H^{-1}$  metric and any stationary solution is still stationary if an arbitrary constant is added. This degeneracy can be eliminated by restricting the solutions in the space where the mass is chosen beforehand; for the same reason, any eigenvectors or perturbations should be restricted to having zero mass. We refer the reader to [2,3] and references therein on the existence of index-1 saddle point for sufficient small  $\kappa$ . We restrict our calculation to the case of not too large domain, i.e., the parameter  $\kappa$  in (4.1) is not too small, but small enough to posses saddle points.

Now we discuss some details of our method. In the IMF, the auxiliary functional  $L$  given by (3.2) (setting  $\alpha = 0, \beta = 2$ ) is

$$L(\phi; \phi^{(k)}, v^{(k+1)}) = F(\phi) - 2F(\hat{\phi}) \tag{4.4}$$

with

$$\hat{\phi} := \phi^{(k)} + \tilde{\phi} = \phi^{(k)} + (v^{(k+1)} \otimes v^{(k+1)})(\phi - \phi^{(k)}), \tag{4.5}$$

where  $(v \otimes v)u = \langle v, u \rangle v, \forall u, v$ , is associated with either  $L^2$  metric or  $H^{-1}$  metric. So the formal notation  $\langle \cdot, \cdot \rangle$  means either  $\langle \cdot, \cdot \rangle_{L^2}$  (AC-type) or  $\langle \cdot, \cdot \rangle_{H^{-1}}$  (CH-type). Next, we give two convex splitting forms of  $F(\phi)$  as discussed in (3.9) and (3.10). The convex splitting form  $F(\phi) = F_c^n - F_e^n$  can be taken as

$$F_c^n(\phi) = \int_0^1 \left[ \frac{\kappa^2}{2} \left( \frac{\partial \phi}{\partial x} \right)^2 + \phi^2 + \frac{1}{4} \right] dx, \quad F_e^n(\phi) = \int_0^1 -\frac{1}{4} \phi^4 + \frac{3}{2} \phi^2 dx, \tag{4.6}$$

and the convex splitting form  $F(\phi) = \tilde{F}_c^n - \tilde{F}_e^n$  is chosen as

$$\tilde{F}_c^n(\phi) = \int_0^1 \left[ \frac{\kappa^2}{2} \left( \frac{\partial \phi}{\partial x} \right)^2 + \frac{1}{4} \phi^4 + \frac{1}{4} \right] dx, \quad \tilde{F}_e^n(\phi) = \int_0^1 \frac{1}{2} \phi^2 dx. \tag{4.7}$$

**Remark 1.** The functionals in (4.7) are always convex but the second functional  $F_e^n$  in (4.6) are convex only in the region  $\phi \in [-1, 1]$ . In general, when the global convexity is not available, to obtain the locally contractive  $F_c$  and expansive  $F_e$ , one usually has to introduce a sufficiently large positive constant  $C$ . For instance the quadratic term  $-\phi^2$  in (4.1) can be written as  $C\phi^2 - (1/2 + C)\phi^2$ , then the convex region of  $F_e^n$  is  $\phi^2 \leq (2C + 1)/3$ . The form (4.6) corresponds to  $C = 1$ .

The  $L_2$  gradient flow of the Ginzburg–Landau functional ensures that the solution  $\phi(t)$  always remains in  $[-1, 1]$  by the maximum principle if the initial  $|\phi(0)| \leq 1$ ; however, the  $H^{-1}$  Cahn–Hilliard flow may not always satisfy this condition. It is not a trivial work to obtain such  $L_\infty$  bounds *a priori* and to derive an implementable optimal choice of  $C$ . Refer to [11] for



the theoretic investigation on a convex splitting scheme for a phase field crystal model. In our numerical simulations, we choose the minimal  $C$  such that the local convexity holds for the initial and it happens that we did not observe unstable phenomena from our empirical results of two applications tested here. However, it should be noted that this choice has no theoretic foundation and is not guaranteed to work in any situation.

4.1.1. Saddle points in  $L^2$  metric

The second-order variational operator  $\mathbf{H}(\phi)$  of the energy functional  $F$ , evaluated at  $\phi$ , is  $\mathbf{H}(\phi)\psi = \delta_\phi^2 F \psi = -\kappa^2 \Delta \psi + f''(\phi)\psi, \forall \psi \in H^2([0, 1])$ , where  $H^2$  is the standard Sobolev space. The eigenvalue problem for this operator is defined by

$$\mathbf{H}(\phi)\psi = -\kappa^2 \Delta \psi + f''(\phi)\psi = \lambda \psi, \tag{4.8}$$

subject to boundary conditions, where  $\lambda$  is the eigenvalue. By the result of the Rayleigh quotient, the eigen-pair for the min-mode,  $\{\lambda_1, \psi_1\}$ , solves the variational problem

$$\min_{\psi \in H^1([0,1])} \mathcal{R}(\psi) := \frac{\langle \psi, \mathbf{H}\psi \rangle_{L^2}}{\|\psi\|_{L^2}^2} = \frac{\int_0^1 \kappa^2 |\nabla \psi|^2 + f''(\phi)\psi^2 dx}{\int_0^1 |\psi|^2 dx}.$$

After the min-mode is obtained, the subproblem of minimizing the auxiliary functional (4.4) is then solved by evolving the gradient flow

$$\begin{aligned} \frac{\partial \phi}{\partial t} &= -\frac{\delta L}{\delta \phi}(\phi) = -\frac{\delta F}{\delta \phi}(\phi) + 2(v \otimes v) \frac{\delta F}{\delta \phi}(\hat{\phi}), \\ &= \kappa^2 \Delta \phi - \phi^3 + \phi - 2(v \otimes v)(\kappa^2 \Delta \hat{\phi} - \hat{\phi}^3 + \hat{\phi}). \end{aligned} \tag{4.9}$$

$v = v^{(k+1)}$  is the min-mode of  $\mathbf{H}(\phi^{(k)})$ .  $\hat{\phi}$  is defined in (4.5).

*Convex splitting scheme.* We apply (4.6) and (4.7) to the convex splitting of  $L(\phi)$ , then  $L(\phi) = [F_c^1(\phi) + 2\tilde{F}_e^1(\hat{\phi})] - [F_e^n(\phi) + 2\tilde{F}_c^n(\hat{\phi})] =: L_c(\phi) - L_e(\phi)$ , where

$$\begin{aligned} L_c(\phi) &= \int_0^1 \left[ \frac{\kappa^2}{2} (\phi_x)^2 + \phi^2 + \frac{1}{4} + \hat{\phi}^2 \right] dx, \\ L_e(\phi) &= \int_0^1 \left[ -\frac{1}{4} \phi^4 + \frac{3}{2} \phi^2 + \kappa^2 (\hat{\phi}_x)^2 + \frac{1}{2} \hat{\phi}^4 + \frac{1}{2} \right] dx. \end{aligned}$$

The first order variational derivatives of  $L_c(\phi)$  and  $L_e(\phi)$  are, respectively,

$$\begin{aligned} \delta_\phi L_c(\phi) &= -\kappa^2 \Delta \phi + 2\phi + 2(v \otimes v)\hat{\phi}, \\ \delta_\phi L_e(\phi) &= -\phi^3 + 3\phi + 2(v \otimes v) \left( -\kappa^2 \Delta \hat{\phi} + \hat{\phi}^3 \right). \end{aligned}$$

Therefore, the convex splitting scheme for (4.9) is written as

$$\begin{aligned} \frac{\phi^{n+1} - \phi^n}{\Delta t} &= \left[ \kappa^2 \Delta \phi - 2\phi - 2(v \otimes v)\phi \right]^{n+1} \\ &\quad + \left[ -\phi^3 + 3\phi + 2(v \otimes v)(-\kappa^2 \Delta \hat{\phi} + \hat{\phi}^3) \right]^n, \end{aligned} \tag{4.10}$$

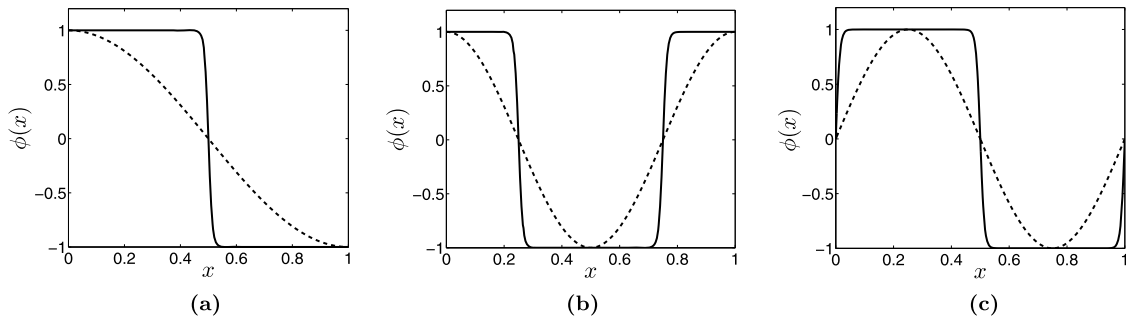
after using  $\langle v, v \rangle_{L^2} = 1$  and  $(v \otimes v)\hat{\phi} = (v \otimes v)\phi$ .

*Non-convex-splitting scheme.* We need a traditional semi-implicit scheme, which is not derived from the convex splitting idea, as the benchmark for comparison. The idea is to take out the linearized part of the nonlinear term in  $f'$  and make it implicit together with the Laplace operator. Linearizing the cubic term  $\phi^3$  in (4.9),  $(\phi^{n+1})^3 \approx (\phi^n)^3 + 3(\phi^n)^2(\phi^{n+1} - \phi^n)$ , we then have the following scheme for (4.9)

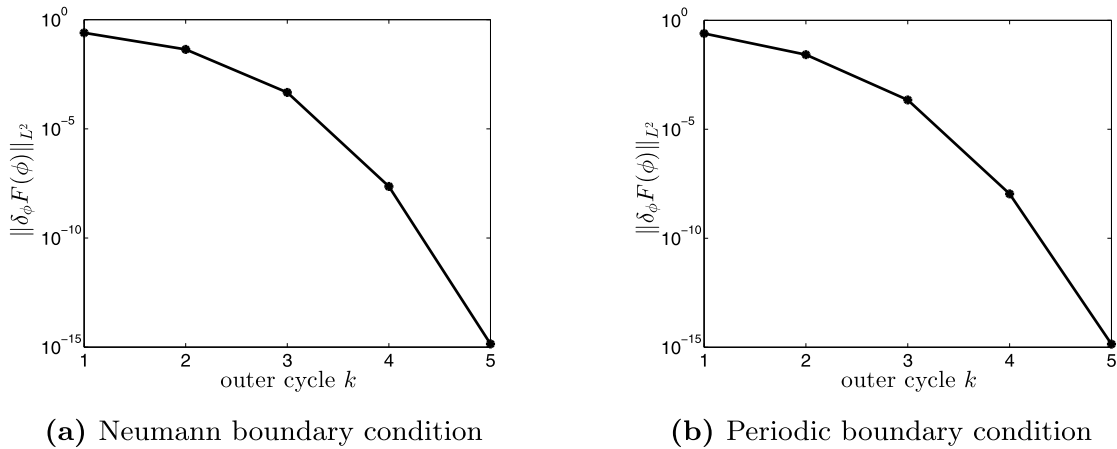
$$\begin{aligned} \frac{\phi^{n+1} - \phi^n}{\Delta t} &= \kappa^2 \Delta \phi^{n+1} - \left[ (\phi^n)^3 + 3(\phi^n)^2(\phi^{n+1} - \phi^n) \right] + \phi^{n+1} \\ &\quad - 2 \left[ \langle v, \kappa^2 \Delta \hat{\phi} - \hat{\phi}^3 + \hat{\phi} \rangle v \right]^n. \end{aligned} \tag{4.11}$$

Set the parameter  $\kappa = 0.01$ . The finite difference method is used for spatial discretization with the mesh grid  $\{x_i = ih, i = 0, 1, 2, \dots, N\}$ .  $h = 1/N$ .  $N = 200$ . A finer mesh with  $N = 1000$  is also used to verify all numerical results. There are only





**Fig. 1.** Profiles of some saddle points (solid lines) of  $F(\phi)$  in  $L^2$  metric computed from various initial states (dashed lines). (a)–(b): the Neumann boundary condition; (c): the periodic boundary condition. The free energy  $F$  for these three plotted states from left to right are 0.0094, 0.0188 and 0.0188, respectively.  $\kappa = 0.01$ .



**Fig. 2.** The validation of the quadratic convergence rate of the IMF mapping  $\Phi$  by plotting the decay of the error, measured by the force  $\|\delta_\phi F(\phi^{(k)})\|_{L^2}$  at each cycle  $k$ .

two locally stable states in  $L^2$  metric. They are the two homogeneous constant states:  $\phi_+(x) \equiv 1$  and  $\phi_-(x) \equiv -1$ . For the Neumann boundary condition, Fig. 1a and Fig. 1b show the transition states calculated from the convex splitting scheme (4.10) with the initial conditions  $\phi_0(x) = \cos \pi x$  and  $\phi_0(x) = \cos 2\pi x$ , respectively. For the periodic boundary condition, Fig. 1c shows the transition state obtained from the initial condition  $\phi_0(x) = \sin 2\pi x$ , which looks almost identical to the one in Fig. 1b after a simple spatial translation.

Next we present the numerical evidence on the performance of our numerical method. We first validate the quadratic convergence rate of the IMF for both boundary conditions. In this validation, which only cares about the rate of the mapping  $\phi \rightarrow \Phi(\phi)$ , we actually solve each subproblem with an extremely high precision. The numerical results are presented in Fig. 2 and are consistent with Theorem 1.

Now we compare the performance of the convex splitting (“CS”) scheme (4.10) against the non-convex splitting (“nCS”) scheme (4.11). Firstly, we examine their performance for the subproblem, i.e., within a fixed cycle where only the inner iteration is running. Take the first cycle for example, which is to solve  $\min_\phi L(\phi; \phi^{(0)}, v^{(1)})$  with the initial  $\phi(t = 0) = \phi^{(0)}$ . We measure the error by the gradient force  $err := \|\delta_\phi L(\phi^n)\|_{L^2}$  and calculate the iteration number required to attain the given error tolerance for the CS scheme (4.10) and the nCS scheme (4.11). The tolerances we tested are the following three values:  $err \leq 1.0 \times 10^{-4}$ ,  $err \leq 1.0 \times 10^{-6}$  and  $err \leq 1.0 \times 10^{-8}$ . We can observe from Table 1 for this simple subproblem that (1) the CS scheme obviously have much better stability than the nCS scheme when the time step size is large ( $\infty$  in this table means that the numerical results diverge); (2) For the small  $\Delta t$  such that both schemes converge, both schemes have the similar required number of iterations (i.e., the time steps) to decrease  $\delta_\phi L$  to certain accuracy. This is expected since both solves the same steepest descent flow for  $L$  with the same step size.

Secondly, we show the comparison of the overall efficiency in locating the saddle point. To be more transparent, we fix the number of iterations in each cycle and count the required number of the outer cycles to reach some prescribed tolerance for the error which is defined as  $\|\delta_\phi F(\phi^{(k)})\|_{L^2}$ . The results are summarized in Table 2. The total iteration number, which is the indicator of the total computational cost, is therefore equal to the number of cycles multiplied by the “iter#” specified in the corresponding columns of Table 2. The key conclusions from this table, for both Neumann and periodic boundary conditions, are the following: (1) for the CS scheme, the larger the time step size is, the smaller the total computational

**Table 1**

The comparison of the CS scheme (4.10) and nCS scheme (4.11) for the sub-problem within the first cycle  $\phi^{(0)} \rightarrow \phi^{(1)} = \Phi(\phi^{(0)})$ . The integers shown in the table are the required number of iterations for the CS scheme (4.10) and the nCS (4.11) to achieve the three choices of the prescribed error tolerance  $\|\delta_\phi L(\phi^n)\|_{L^2} \leq 10^{-4}, 10^{-6}$  and  $10^{-8}$ .

$\Delta t$	The required number of iterations					
	$err = 10^{-4}$		$err = 10^{-6}$		$err = 10^{-8}$	
	(4.10)	(4.11)	(4.10)	(4.11)	(4.10)	(4.11)
0.01	777	763	1092	1074	1407	1386
0.1	92	78	129	111	166	144
5.0	16	$\infty$	23	$\infty$	30	$\infty$
10	14	$\infty$	21	$\infty$	28	$\infty$

(a) Neumann boundary condition. The initial state is  $\phi^{(0)} = \cos \pi x$ .

$\Delta t$	The required number of iterations					
	$err = 10^{-4}$		$err = 10^{-6}$		$err = 10^{-8}$	
	(4.10)	(4.11)	(4.10)	(4.11)	(4.10)	(4.11)
0.01	728	722	1043	1034	1358	1347
0.1	86	74	123	108	160	141
5.0	14	$\infty$	22	$\infty$	29	$\infty$
10	13	$\infty$	20	$\infty$	27	$\infty$

(b) Periodic boundary condition. The initial state is  $\phi^{(0)} = \sin 2\pi x$ .

**Table 2**

The comparison of the required number of outer cycles for the CS scheme (4.10) and the nCS scheme (4.11) to attain the given error tolerance  $\|\delta_\phi F(\phi^{(k)})\|_{L^2} \leq 10^{-8}$  when the inner iteration number is fixed as 50, 80 and 100, respectively, for different choices of the time step size.

$\Delta t$	The number of cycles					
	iter# = 50		iter# = 80		iter# = 100	
	(4.10)	(4.11)	(4.10)	(4.11)	(4.10)	(4.11)
0.01	29	28	18	18	15	14
0.1	4	3	3	2	2	2
5.0	1	$\infty$	1	$\infty$	1	$\infty$

(a) Neumann boundary condition. The initial state is  $\phi^{(0)} = \cos \pi x$ .

$\Delta t$	The number of cycles					
	iter# = 50		iter# = 80		iter# = 100	
	(4.10)	(4.11)	(4.10)	(4.11)	(4.10)	(4.11)
0.01	28	27	17	17	14	14
0.1	4	3	3	2	2	2
5.0	1	$\infty$	1	$\infty$	1	$\infty$

(b) Periodic boundary condition. The initial state is  $\phi^{(0)} = \sin 2\pi x$ .

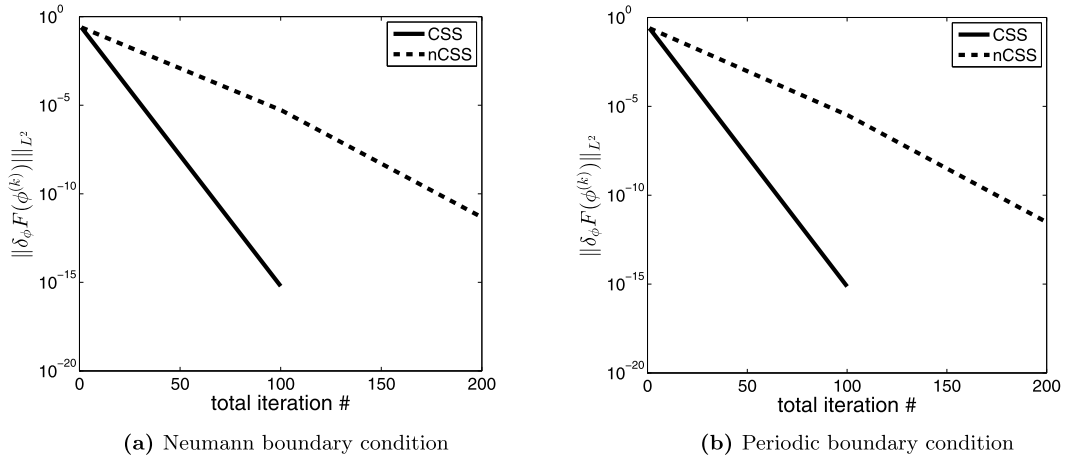
cost is; (2) for the nCS scheme, it is divergent when a large time step size is applied; (3) for the very small time step size such as 0.01, there is no much difference in the computational cost of the CS and nCS schemes. The last observation is consistent with the known empirical experience of the convex splitting schemes: when both the CS and nCS converge, the decay rate of the objective function in the CS may not be better.

To better visualize the improvement of the CS scheme over the nCS scheme, we draw the decay of the error  $\|\delta_\phi F(\phi)\|_{L^2}$  with respect to the total iteration number in Fig. 3. This plot illustrates how much accuracy one can obtain (the vertical axis) with the available total computational cost (the horizontal axis), in which the solid lines are from the CS scheme and the dashed lines are from the nCS scheme. Note that the step sizes for the CS and nCS are different in order to obtain this efficiency gain.

4.1.2. Saddle points in  $H^{-1}$  metric

In this part, we study the transition state of  $F(\phi)$  in  $H^{-1}$  metric. Note that the inner product and the norm in  $H^{-1}$  metric can be transformed to those in  $L^2$  metric as follows

$$\|\phi\|_{H^{-1}}^2 = \langle (-\Delta)^{-1}\phi, \phi \rangle_{L^2}, \quad \langle \phi, \psi \rangle_{H^{-1}} = \langle (-\Delta)^{-1}\phi, \psi \rangle_{L^2}, \tag{4.12}$$



**Fig. 3.** The decay of the error measured by the gradient  $\|\delta_\phi F(\phi^{(k)})\|_{L^2}$  with the total iteration number (i.e., the cost) for the CS scheme (4.10) (solid line) and the nCS scheme (4.11) (dashed line). The inner iteration number in each cycle is fixed as 100 and the time step sizes are  $\Delta t = 5.0$  and  $0.1$  for the CS and nCS schemes, respectively.

where  $(-\Delta)^{-1}$ , a bounded positive self-adjoint linear operator, is the inverse of  $-\Delta$  subject to certain boundary condition. Then the variational derivatives between the  $L^2$  metric and the  $H^{-1}$  metric can be linked as follows:

$$\delta_\phi F|_{H^{-1}} = -\Delta \delta_\phi F, \quad \tilde{\mathbf{H}} := \delta_\phi^2 F|_{H^{-1}} = -\Delta \delta_\phi^2 F,$$

where  $\delta_\phi F$  and  $\delta_\phi^2 F$  represent respectively the first and the second order variational derivatives of  $F(\phi)$  in  $L^2$  metric.

The gradient flow  $\partial_t \phi = -(-\Delta \delta_\phi F) = -\kappa^2 \Delta^2 \phi + \Delta(\phi^3 - \phi)$  is the CH equation (4.3). It is known that the solution  $\phi(x, t)$  of the CH equation (4.3) satisfies  $\int_0^1 \phi(x, t) dx \equiv \int_0^1 \phi(x, 0) dx, \forall t > 0$ . In our problem of finding the saddle point in the  $H^{-1}$  metric, we choose a fixed mass  $m = 0.6$  beforehand and we are interested in the saddle points satisfying  $\int_0^1 \phi(x) dx = m$ . We also require any eigenvectors or perturbations to belong to the subspace  $\{\psi : \int_0^1 \psi(x) dx = 0\}$ . The eigenvalue problem of  $\tilde{\mathbf{H}}$  is

$$\begin{cases} \tilde{\mathbf{H}}(\phi)\psi = -\Delta(-\kappa^2 \Delta \psi + f''(\phi)\psi) = \lambda \psi, \\ \int_0^1 \psi(x) dx = 0, \end{cases} \tag{4.13}$$

subject to the Neumann or periodic boundary condition. For nonzero eigenvalue, the eigenvector  $\psi$  automatically satisfies  $\int_0^1 \psi dx = 0$ ; For zero eigenvalue, the condition  $\int_0^1 \psi dx = 0$  needs to be imposed additionally. If we introduce the projection  $\mathbf{P}u := u - \int_0^1 u(x) dx$ , then the eigen-problem (4.13) is equivalent to  $\mathbf{P}\tilde{\mathbf{H}}\mathbf{P}u = \lambda u$  for any  $u$  without the zero-mass constraint. The min-mode is then equal to  $u$  if the eigenvalue is nonzero and equal to  $\mathbf{P}u$  if the eigenvalue is zero.

For the periodic boundary condition, there exists one degenerate ( $\lambda = 0$ ) direction  $\psi_0$  for the Hessian  $\tilde{\mathbf{H}}(\phi)$ , which is the spatial derivative of  $\phi$ , i.e.,

$$\psi_0(x) = \phi_x(x).$$

To see this, we just need verify (4.13) by using the periodic boundary condition:  $\int_0^1 \phi_x(x) dx = \phi(1) - \phi(0) = 0$  and

$$\int_0^1 -\Delta(-\kappa^2 \Delta \phi_x + f''(\phi)\phi_x) dx = \int_0^1 \kappa^2 \partial_x^5 \phi(x) - \partial_x^2(\partial_x(f'(\phi(x)))) dx = 0.$$

The Rayleigh quotient with respect to  $H^{-1}$  metric is

$$\tilde{\mathcal{R}}(\psi) = \frac{(\psi, \tilde{\mathbf{H}}\psi)_{H^{-1}}}{\|\psi\|_{H^{-1}}^2} = \frac{\int_0^1 \kappa^2 |\nabla \psi|^2 + f''(\phi)\psi^2 dx}{\int_0^1 \psi \Delta^{-1} \psi dx},$$

and thus the min-mode is the minimizer of the problem

$$\operatorname{argmin}_\psi \left\{ \tilde{\mathcal{R}}(\psi) : \int_0^1 \psi dx = 0, \|\psi\|_{H^{-1}} = 1 \right\}. \tag{4.14}$$

For the IMF in this  $H^{-1}$  case, the subproblem of minimizing the auxiliary functional  $L$  for a given  $\phi^{(k)}$  at cycle  $k$  is

$$\phi^{(k+1)} = \underset{\int_0^1 \phi(x) dx = m}{\operatorname{argmin}} L(\phi; \phi^{(k)}, v^{(k+1)}), \tag{4.15}$$

where  $m = \int_0^1 \phi^{(k)} dx$ . This means that each IMF cycle  $\phi^{(k)} \rightarrow \phi^{(k+1)}$  should also conserve the mass. The expression of  $L(\phi)$  is defined as (4.4), with the modification of  $\hat{\phi}$  as follows

$$\begin{aligned} \hat{\phi} &:= \phi^{(k)} + \left\langle v, \phi - \phi^{(k)} \right\rangle_{H^{-1}} v = \phi^{(k)} + \left\langle -\Delta^{-1} v, \phi - \phi^{(k)} \right\rangle_{L^2} v \\ &= \phi^{(k)} + \left\langle w, \phi - \phi^{(k)} \right\rangle_{L^2} v, \end{aligned} \tag{4.16}$$

where

$$w := -\Delta^{-1} v$$

is the unique solution satisfying the equation  $-\Delta w = v$  and  $\int_0^1 w dx = 0$ . Then,  $\delta_\phi L(\phi) = \delta_\phi F(\phi) - 2w \left\langle \delta_\phi F(\hat{\phi}), v \right\rangle_{L^2}$ , and the  $H^{-1}$  gradient flow of  $L$  is

$$\begin{aligned} \frac{\partial \phi}{\partial t} &= \Delta \frac{\delta L}{\delta \phi}(\phi) = \Delta \left( \frac{\delta F}{\delta \phi}(\phi) \right) + 2v \left\langle \frac{\delta F}{\delta \phi}(\hat{\phi}), v \right\rangle_{L^2}, \\ &= -\kappa^2 \Delta^2 \phi + \Delta(\phi^3 - \phi) + 2v \left\langle v, -\kappa^2 \Delta \hat{\phi} + \hat{\phi}^3 - \hat{\phi} \right\rangle_{L^2}, \end{aligned} \tag{4.17}$$

where  $v = v^{(k+1)}$  refers to the min-mode of (4.13) at  $\phi^{(k)}$  (normalized under  $H^{-1}$  metric, i.e.,  $\|v\|_{H^{-1}} = 1$ ). Note that the scalar in the second line of (4.17) is the  $L^2$  inner product due to the cancelation of  $\Delta$  and  $\Delta^{-1}$  in the calculation, but the variable  $\hat{\phi}$  involves the  $H^{-1}$  metric via  $w$ . This suggests one more computational cost of solving  $w = -\Delta^{-1} v$  during the IMF subproblem than the GAD scheme where  $\hat{\phi}$  is actually just  $\phi^{(k)}$  ([15]). Our result (4.17) is consistent with the equation (3.4) in [28] which was written for the GAD in the finite dimension, i.e.,  $\hat{\phi} = \phi^{(k)}$ .

**Remark 2.** We show that the flow (4.17) conserves the initial mass  $\int_0^1 \phi(x) dx$ , sharing exactly the same property as the CH equation (4.3). So the constraint in (4.15) holds automatically. This result immediately implies that the IMF mapping  $\phi^{(k)} \rightarrow \phi^{(k+1)}$  does not change the mass at each cycle  $k$ . To prove our conclusion, after integrating the two sides of (4.17) and using the boundary conditions (either Neumann or periodic), one remains to show the following condition for the eigenvector  $v$ :  $\int_0^1 v(x) dx = 0$ . This is exactly the condition that the min-mode  $v$  satisfies in (4.13).

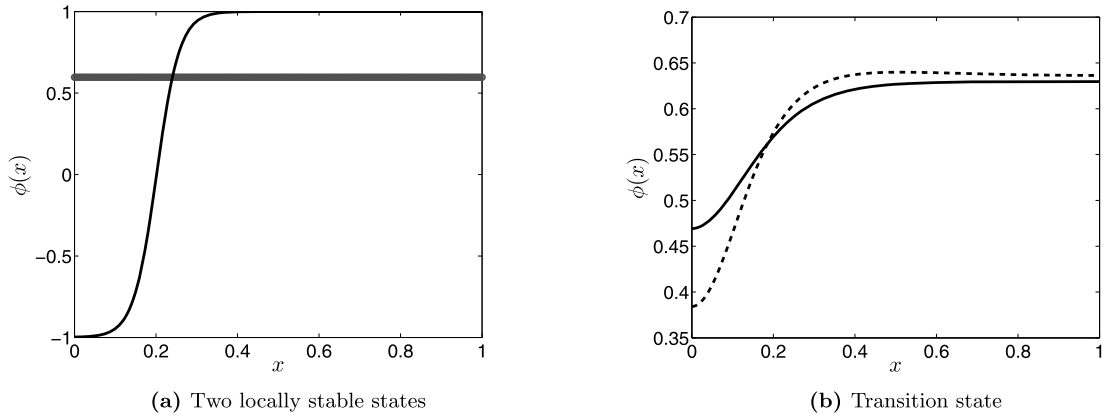
*Convex splitting scheme.* We now test the convex splitting form of (4.6), (4.7), which corresponds to  $L(\phi) = L_c(\phi) - L_e(\phi)$ , where  $L_c(\phi) = F_c^l(\phi) + 2\tilde{F}_e^l(\hat{\phi})$ ,  $L_e(\phi) = F_e^n(\phi) + 2\tilde{F}_c^n(\hat{\phi})$ , and  $\hat{\phi}$  is defined in (4.16). Since  $\delta_\phi L_c(\phi) = -\kappa^2 \Delta \phi + 2\phi + 2 \left\langle v, \hat{\phi} \right\rangle_{L^2} w$ , and  $\delta_\phi L_e(\phi) = -\phi^3 + 3\phi + 2 \left\langle -\kappa^2 \Delta \hat{\phi} + \hat{\phi}^3, v \right\rangle_{L^2} w$ , then the convex splitting scheme for (4.17) is for  $n \geq 0$

$$\begin{aligned} \frac{\phi^{n+1} - \phi^n}{\Delta t} &= \left[ -\kappa^2 \Delta^2 \phi^{n+1} + 2\Delta \phi^{n+1} - 2 \left\langle w, \phi^{n+1} \right\rangle_{L^2} \left\langle v, v \right\rangle_{L^2} v \right] \\ &\quad - 2 \left\langle v, \phi^{(k)} \right\rangle_{L^2} v + 2 \left\langle w, \phi^{(k)} \right\rangle_{L^2} \left\langle v, v \right\rangle_{L^2} v \\ &\quad + \left[ \Delta(\phi^n)^3 - 3\Delta \phi^n + 2 \left\langle v, -\kappa^2 \Delta \hat{\phi}^n + (\hat{\phi}^n)^3 \right\rangle_{L^2} v \right], \end{aligned} \tag{4.18}$$

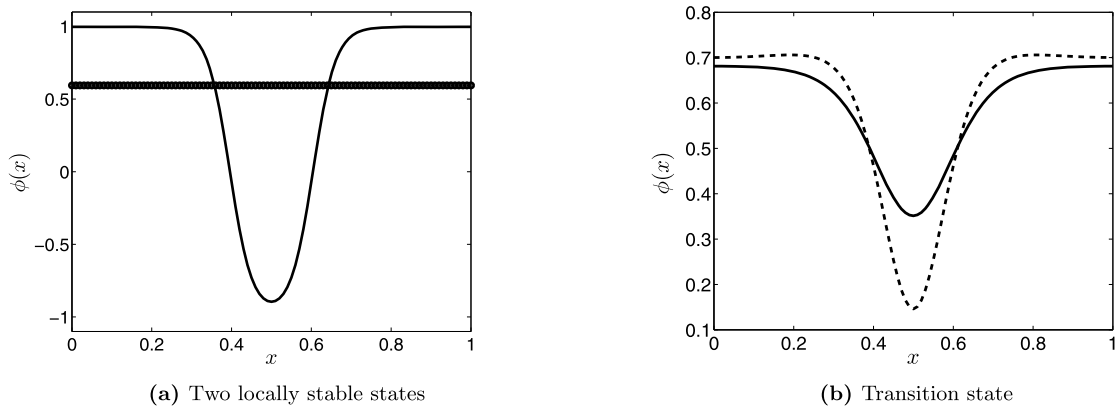
where  $v = v^{(k+1)}$ ,  $w = -\Delta^{-1} v^{(k+1)}$  and  $\hat{\phi}^n$  is from (4.16) by letting  $\phi = \phi^n$ . On the right hand side of this CS scheme (4.18), the first line is linear in the unknown  $\phi^{n+1}$ , the second line is independent of  $n$ , the third line is the nonlinear term in  $\phi^n$ .

*Non-convex-splitting scheme.* The same linearization trick used in (4.11) for the previous AC type problem is applied here again to construct the following non-convex splitting scheme for CH type equation for comparison:

$$\begin{aligned} \frac{\phi^{n+1} - \phi^n}{\Delta t} &= -\kappa^2 \Delta^2 \phi^{n+1} - \Delta \phi^{n+1} + \Delta \left[ (\phi^n)^3 + 3(\phi^n)^2(\phi^{n+1} - \phi^n) \right] \\ &\quad - 2 \left[ \left\langle v, \kappa^2 \Delta \hat{\phi} - \hat{\phi}^3 + \hat{\phi} \right\rangle v \right]^n. \end{aligned} \tag{4.19}$$



**Fig. 4.** (Neumann boundary condition.  $H^{-1}$  metric.) (a): the two stable stationary states of  $F(\phi)$  with the mass  $\int_0^1 \phi dx = 0.6$ .  $F = 0.10240$  for the trivial constant state (the thick line) and  $F = 0.03772$  for the transition layer state (the thin line). (b): one of the transition states (solid line) with the free energy 0.10241 whose first 3 eigenvalues are  $\lambda = -3.41, 3.91$  and 18.14, calculated from the initial condition (dashed line) whose first 3 eigenvalues are  $\lambda = -10.97, 3.45$  and 17.48. Note that the vertical axes in subfigures are in different scales.



**Fig. 5.** (Periodic boundary condition.  $H^{-1}$  metric.) (a): the two stable stationary states of  $F(\phi)$  in  $H^{-1}$  metric with the given mass  $\int_0^1 \phi dx = 0.6$ .  $F = 0.10240$  for the constant state and  $F = 0.07510$  for the other non-constant state. (b): one of the transition states (solid line) with the free energy 0.10285 whose first three smallest (numerical) eigenvalues are  $-12.75, -1.86 \times 10^{-7}$  (corresponding to theoretical zero eigenvalue) and 46.15. The initial state is shown as the dashed line in which the first 3 smallest eigenvalues are  $\lambda = -27.13, -6.81$  and 44.98.

The parameter  $\kappa = 0.04$ . We set the mesh grid  $\{x_i = ih, i = 0, 1, 2, \dots, N\}$ .  $h = 1/N$ .  $N = 200$ . The function  $\phi(x)$  is represented by  $\boldsymbol{\phi} = (\phi_0, \phi_1, \phi_2, \dots, \phi_N)^T$  for the Neumann boundary condition and  $\boldsymbol{\phi} = (\phi_0, \phi_1, \phi_2, \dots, \phi_{N-1})^T$  for the periodic boundary condition;  $\phi_i \approx \phi(x_i)$ . The matrix form of the Hessian  $\tilde{\mathbf{H}}$  at the state  $\boldsymbol{\phi}$  is  $\tilde{\mathbf{H}}(\boldsymbol{\phi}) = A \nabla^2 F_h(\boldsymbol{\phi})$ , where  $\nabla^2 F_h(\boldsymbol{\phi}) = \frac{\partial^2 F_h}{\partial \phi_i \partial \phi_j}$ .  $A$  and  $F_h(\boldsymbol{\phi})$  denote the discretized forms of the operator  $-\Delta$  and the potential energy  $F(\phi)$ , respectively.

The min-mode  $v$  of the Hessian matrix  $\tilde{\mathbf{H}}$  can be calculated according to the Rayleigh quotient (4.14).

Fig. 4 and Fig. 5 show some of the stationary states of  $F$  in  $H^{-1}$  metric for the Neumann and periodic boundary conditions, respectively. The stationary points identified by us agree with the result in [42]. Next we compare the performance of the CS scheme (4.18) and the nCS scheme (4.19). We also start from their performance for the subproblem in the first cycle as before. Table 3 compares the number of inner iterations to attain the three error tolerances:  $\|\Delta \delta_\phi L\|_{H^{-1}} \leq 10^{-4}, 10^{-5}$  or  $10^{-6}$ . Basically we have the same observations as for the  $L_2$ -metric case. The CS scheme allows the step size as large as 0.1 but the nCS is not able in this example. This table also shows that the nCS scheme, if it converges, requires a less number of iterations to reach the tolerance, in particular for the periodic boundary condition. This is possible because a scheme with a better stability does not guarantee a faster convergence rate toward the minimum.

Table 4 shows the same quantities of the overall performance of the convex splitting scheme in this  $H^{-1}$  metric as measured before in Table 2 in the  $L_2$  metric. The conclusions we draw from this table are qualitatively the same as from Table 2. In the end, based on the experiments corresponding to Table 4, we plot the evolution of the error measured by the force  $\|\Delta \delta_\phi F(\phi^{(k)})\|_{H^{-1}}$  against the total iteration number for the CS scheme and the nCS scheme by using their own optimal time step sizes respectively. Refer to Fig. 6. Here for the CH-type problem, in order to illustrate that our results are robust with respect to the initial guess, we added the random perturbation to generate multiple initial states so that multiple lines are plotted for different initial guesses.

**Table 3**

The comparison of the CS scheme (4.18) and nCS scheme (4.19) for the subproblem  $\phi^{(0)} \rightarrow \phi^{(1)} = \Phi(\phi^{(0)})$ . The integers in the table are the required number of iterations to achieve the three prescribed tolerances  $\|\Delta\delta_\phi L(\phi^n)\|_{H^{-1}} \leq 10^{-4}, 10^{-5}$  and  $10^{-6}$ .

$\Delta t$	The required number of iterations					
	$err = 10^{-4}$		$err = 10^{-5}$		$err = 10^{-6}$	
	(4.18)	(4.19)	(4.18)	(4.19)	(4.18)	(4.19)
$10^{-3}$	2956	2647	3332	2993	3709	3339
$10^{-2}$	550	273	617	308	685	343
$10^{-1}$	283	$\infty$	321	$\infty$	360	$\infty$
1.0	14	$\infty$	262	$\infty$	298	$\infty$

(a) Neumann boundary condition.

$\Delta t$	The required number of iterations					
	$err = 10^{-4}$		$err = 10^{-5}$		$err = 10^{-6}$	
	(4.18)	(4.19)	(4.18)	(4.19)	(4.18)	(4.19)
$10^{-3}$	1022	820	1163	939	1305	1057
$10^{-2}$	281	77	316	88	351	99
$10^{-1}$	208	$\infty$	233	$\infty$	257	$\infty$

(b) Periodic boundary condition.

**Table 4**

The comparison of the number of outer cycles required for the CS scheme (4.18) and the nCS scheme (4.19) to attain the given error tolerance  $\|\Delta\delta_\phi F(\phi^{(k)})\|_{H^{-1}} \leq 10^{-8}$ , when the inner iteration number is fixed for the different choices of the time step size  $\Delta t = 10^{-3}, 10^{-2}$  and  $10^{-1}$ . The corresponding initial states are specified in Fig. 4b and Fig. 5b, respectively.

$\Delta t$	The number of cycles					
	iter# = 40		iter# = 50		iter# = 60	
	(4.18)	(4.19)	(4.18)	(4.19)	(4.18)	(4.19)
$10^{-3}$	118	102	95	82	79	69
$10^{-2}$	28	$\infty$	22	$\infty$	16	$\infty$
$10^{-1}$	16	$\infty$	17	$\infty$	16	$\infty$

(a) Neumann boundary condition.

$\Delta t$	The number of cycles					
	iter# = 50		iter# = 80		iter# = 100	
	(4.18)	(4.19)	(4.18)	(4.19)	(4.18)	(4.19)
$10^{-3}$	34	25	22	16	18	13
$10^{-2}$	13	$\infty$	7	$\infty$	8	$\infty$
$10^{-1}$	11	$\infty$	8	$\infty$	7	$\infty$

(b) Periodic boundary condition.

4.2. 2D example: Landau–Brazovskii free energy

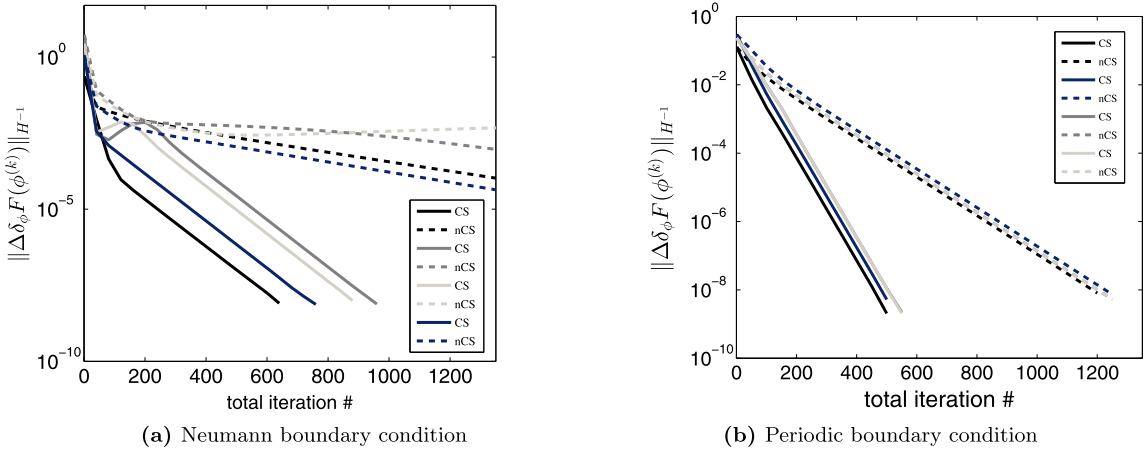
In this section, we study the nucleation problem of phase transition in diblock copolymers ([29,37]). The model is described by the Landau–Brazovskii energy functional of the order parameter  $\phi$

$$F(\phi) = \int_{\Omega} \frac{\xi^2}{2} [(\Delta + 1)\phi(\mathbf{r})]^2 + \Phi(\phi) \, d\mathbf{r}, \tag{4.20}$$

where  $\Phi(\phi) = \frac{\tau}{2}\phi^2 - \frac{\gamma}{3!}\phi^3 + \frac{1}{4!}\phi^4$ . The parameters are  $\tau = -0.15, \xi = 1.0, \gamma = 0.25$ . We compute the transition state of this  $F$  in  $H^{-1}$  metric.

We consider the two dimensional domain  $\Omega = [0, L_x] \times [0, L_y]$  and the periodic boundary condition.  $\phi$  satisfies the mass conservation  $\int_{\Omega} \phi(\mathbf{r}) \, d\mathbf{r} = 0$ . The eigenvalue problem in the IMF for this case then reads

$$\begin{cases} \tilde{\mathbf{H}}(\phi)\psi = -\Delta[\xi^2(\Delta + 1)^2 + \Phi''(\phi)]\psi = \lambda\psi, \\ \int_{\Omega} \psi(\mathbf{r}) \, d\mathbf{r} = 0. \end{cases}$$



**Fig. 6.** The evolution of the errors measured by  $\|\Delta\delta_\phi F(\phi)\|_{H^{-1}}$  w.r.t. the total iteration number for the CS scheme (4.18) (solid lines) and the nCS scheme (4.19) (dashed lines). Different small perturbations around the initial condition used in Table 4 are added as the new initial conditions here to produce multiple lines. For the Neumann boundary condition in (a), the inner iteration number is 40 and the time step sizes are  $\Delta t = 10^{-1}$  and  $10^{-3}$  for the CS and nCS schemes, respectively. For the periodic boundary condition in (b), the inner iteration number is 50 and the time step sizes are the same as in (a). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

The min-mode  $v$  is the eigenvector of  $\tilde{\mathbf{H}}$  corresponding to the smallest eigenvalue. The gradient flow associated with the minimization subproblem for  $L(\phi; \phi^{(k)}, v^{(k+1)}) = F(\phi) - 2F(\hat{\phi})$  is

$$\frac{\partial \phi}{\partial t} = \Delta \frac{\delta L}{\delta \phi} = \xi^2 \Delta (\Delta + 1)^2 \phi + \Delta \Phi_1(\phi) + 2 \left\langle v, \xi^2 (\Delta + 1)^2 \hat{\phi} + \Phi_1(\hat{\phi}) \right\rangle_{L^2} v, \quad (4.21)$$

where  $\Phi_1(\phi) = \Phi'(\phi) = \tau\phi - \gamma\phi^2/2! + \phi^3/3!$  and  $\hat{\phi}$  is defined the same as (4.16).

*Convex-splitting scheme.* We give two convex splitting forms of  $F$  as follows

$$F_c^1(\phi) = \int_{\Omega} \frac{\xi^2}{2} [(\Delta + 1)\phi]^2 + \frac{\tau + 22.75}{2} \phi^2 \, d\mathbf{r}, \quad (4.22)$$

$$F_e^n(\phi) = \int_{\Omega} \left[ -\frac{1}{4!} \phi^4 + \frac{\gamma}{3!} \phi^3 + \frac{22.75}{2} \phi^2 \right] d\mathbf{r}, \quad (4.23)$$

and

$$\tilde{F}_c^n(\phi) = \int_{\Omega} \frac{\xi^2}{2} |\Delta\phi|^2 + \left( \frac{1}{4!} \phi^4 - \frac{\gamma}{3!} \phi^3 + \frac{\xi^2 + \tau}{2} \phi^2 \right) d\mathbf{r}, \quad (4.24)$$

$$\tilde{F}_e^1(\phi) = \int_{\Omega} \xi^2 |\nabla\phi|^2 \, d\mathbf{r}. \quad (4.25)$$

By (4.22), (4.23) and (4.25), (4.25), and the convex splitting form of  $L(\phi) = L_c(\phi) - L_e(\phi)$  with  $L_c(\phi) = F_c^1(\phi) + 2\tilde{F}_e^1(\hat{\phi})$  and  $L_e(\phi) = F_e^n(\phi) + 2\tilde{F}_c^n(\hat{\phi})$ , we have the convex splitting scheme for (4.21) as follows

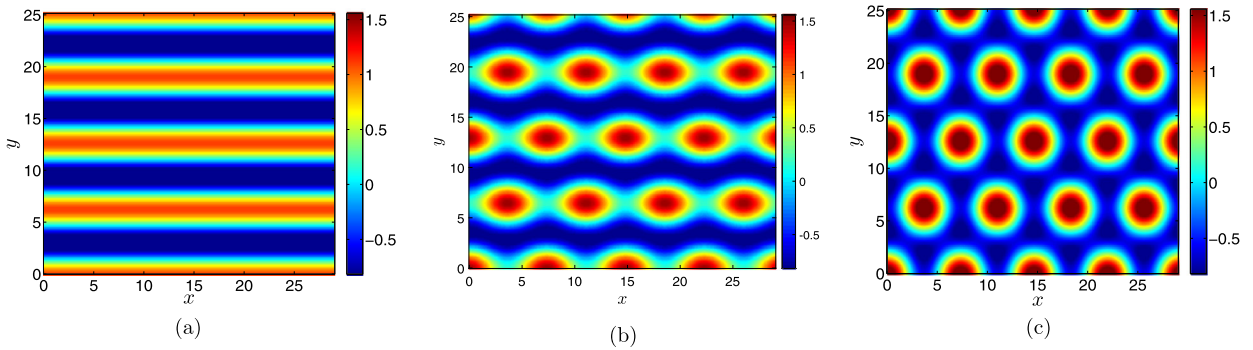
$$\frac{\phi^{n+1} - \phi^n}{\Delta t} = \Delta [\delta_\phi L_c(\phi)]^{n+1} - \Delta [\delta_\phi L_e(\phi)]^n,$$

i.e.,

$$\begin{aligned} \frac{\phi^{n+1} - \phi^n}{\Delta t} &= \left[ \xi^2 \Delta (\Delta + 1)^2 \phi + (\tau + 22.75) \Delta \phi \right]^{n+1} + 4\xi^2 \left[ \langle \Delta v, v \rangle \langle w, \phi \rangle v \right]^{n+1} \\ &\quad + 4\xi^2 \left[ \langle \Delta \phi^{(k)}, v \rangle - \langle w, \phi^{(k)} \rangle \langle \Delta v, v \rangle \right] v - \Delta \left[ -\frac{1}{3!} \phi^3 + \frac{\gamma}{2} \phi^2 + 22.75 \phi \right]^n \\ &\quad + 2 \left\langle v, \xi^2 \Delta^2 \hat{\phi} + (\xi^2 + \tau) \hat{\phi} - \frac{\gamma}{2} \hat{\phi}^2 + \frac{1}{3!} \hat{\phi}^3 \right\rangle v \right]^n, \end{aligned} \quad (4.26)$$

where  $\langle \cdot, \cdot \rangle$  means  $\langle \cdot, \cdot \rangle_{L^2}$ .





**Fig. 7.** Two stable stationary states (a) and (c) and the transition state (b) of the 2-D Landau-Brazovskii energy in  $H^{-1}$  metric. Their free energies are  $-16.486$ ,  $-16.447$  and  $-17.290$ , from left to right. Their smallest eigenvalues are  $3.50 \times 10^{-6}$ ,  $-4.32 \times 10^{-6}$ , and  $3.32 \times 10^{-6}$ , respectively.

**Remark 3.** In (4.23), we choose the constant  $C = 22.75$  in Remark 1, then the convex region for (4.23) is  $[-6.5, 6.5]$ . One can see later that the local minimum and the saddle point lie between  $-1.5$  and  $1.5$ . But our initial guess spans the interval  $[-6.5, 6.5]$ , so we use a large  $C = 22.75$ . Fortunately, we find that  $\phi$  always locates in  $[-6.5, 6.5]$  by tracking the numerical value of  $\phi$ . Note that (4.22), (4.24) and (4.25) are globally convex.

*Non-convex-splitting scheme.* This scheme is constructed by the same idea as before by applying the Taylor expansion of the nonlinear term  $\Phi_1(\phi^{n+1})$  around the solution at the old time step  $\phi^n$ ,

$$\Phi_1(\phi^{n+1}) \approx \Phi_1(\phi^n) + \Phi_1'(\phi^n)(\phi^{n+1} - \phi^n).$$

After simplification, we get the non-convex splitting scheme as follows:

$$\begin{aligned} \frac{\phi^{n+1} - \phi^n}{\Delta t} &= \xi^2 \Delta (\Delta + 1)^2 \phi^{n+1} + \Delta \left( \tau - \gamma \phi^n + \frac{1}{2} [\phi^n]^2 \right) \phi^{n+1} \\ &+ \Delta \left( \frac{\gamma}{2} [\phi^n]^2 - \frac{1}{3} [\phi^n]^3 \right) + 2 \left[ \langle v, \xi^2 (\Delta + 1)^2 \hat{\phi} + \Phi_1(\hat{\phi}) \rangle v \right]^n. \end{aligned} \tag{4.27}$$

In the numerical simulation, we choose the domain  $\Omega = [0, \frac{16}{\sqrt{3}}\pi] \times [0, 8\pi]$  and set the mesh grid  $\{x_i = ih_x, i = 0, 1, 2, \dots, N_x\}$  and  $\{y_j = jh_y, j = 0, 1, 2, \dots, N_y\}$ .  $h_x = 1/N_x, h_y = 1/N_y, N_x = N_y = 100$ . We first consider the gradient flow  $\partial_t \phi = \Delta \delta_\phi F = \Delta [\xi^2 (\Delta + 1)^2 \phi + \Phi_1(\phi)]$  to find the steady states of  $F(\phi)$ . The initial conditions are chosen as equation (2.21) in [37]. For this steepest descent flow of  $F$ , we tested the convex splitting method based on (4.22) (4.23) (obtained by setting  $v = 0$  in (4.26)), and the semi-implicit scheme based on the Taylor expansion for the nonlinear term (this scheme is obtained by directly setting  $v = 0$  in (4.27)). The numerical results show that all schemes are stable with the time step sizes as large as 50.

Fig. 7(a) and Fig. 7(c) are the two metastable states of the lamellar phase and the cylindrical phase respectively. Fig. 7(b) is the transition state obtained by the IMF. The initial guess for the IMF in our test has the gradient  $\|\Delta \delta_\phi F(\phi^0)\|_{H^{-1}} = 10^{-2}$  and the minimal eigenvalue  $-2.0 \times 10^{-12}$ .

Table 5 shows the number of inner iterations in the first cycle and the number of outer cycles in the whole process for the CS scheme (4.26) and the nCS scheme (4.27). For this particular example, we find that both schemes perform quite well for large time step sizes and their performances are almost identically. So, in this 2D example, the both schemes are quite successful. However, for the convex splitting scheme, we have offered an automatic procedure from the decomposition of  $F$  to the decomposition of  $L$ , so our CS scheme developed here inherits the advantage of convex splitting method to allow large time step sizes. Note that the linearization idea of constructing semi-implicit scheme has shown the unconditionally stability in search of local minimizers for this Landau-Brazovskii example. So for other schemes which are not based on the convex splitting idea but have proven to work for  $F$  with unconditional stability, one may also be able to construct some resemblant schemes for  $L$ , such as the scheme (4.27) we derived above. Then what is not obvious is the theoretic question on the stability of these new schemes for  $L$ . This could be left as a future project to generalize our idea of testing the convex splitting method in this article.

### 4.3. Discussion

The choice of the initial guess is an important practical issue for all existing numerical methods of calculating the saddle points. The GAD or the IMF only has the local convergence and thus in extreme cases, one can easily construct a special initial guess which does not have convergence. In addition, the subproblem of minimizing the auxiliary functional is well-defined only when the minimal eigenvalue of the original Hessian is negative; otherwise, it should not be solved

**Table 5**

(a) The comparison of the CS (4.26) and nCS scheme (4.27) for the subproblem of the first cycle. The integers shown are the required number of iterations to achieve the three prescribed tolerances  $\|\Delta\delta_\phi L(\phi^n)\|_{H^{-1}} \leq 10^{-2}$ ,  $10^{-3}$  and  $2 \times 10^{-4}$ ; (b) The number of outer cycles required for (4.26) and (4.27) to attain the given tolerance  $\|\Delta\delta_\phi F(\phi^{(k)})\|_{H^{-1}} \leq 10^{-6}$ .

$\Delta t$	The required number of iterations					
	$err = 10^{-2}$		$err = 10^{-3}$		$err = 2 \times 10^{-4}$	
	(4.26)	(4.27)	(4.26)	(4.27)	(4.26)	(4.27)
0.1	127	126	1930	1928	10651	10645
1.0	14	14	192	190	1061	1055
10	3	3	18	16	105	99

(a) Inner iteration comparison.

$\Delta t$	The number of cycles					
	iter# = 500		iter# = 800		iter# = 1000	
	(4.26)	(4.27)	(4.26)	(4.27)	(4.26)	(4.27)
0.1	191	191	381	380	239	238
1.0	40	39	25	25	21	20
10	6	5	4	4	4	3

(b) Outer cycle comparison.

thoroughly, but limited to a fixed few number of iterations, such as in the GAD. For the readers particularly interested in the practical convergence issue, refer to the discussion in [15]. Most numerical results we reported here are for one typical choice of initial conditions (except for Fig. 6), but our unreported numerical experiments by trying various initial guesses, still strongly support our main conclusions.

Finally, we emphasize another critical implementation issue of selecting the correct “min-mode”. In theory, the calculation of the min-mode is straight forward by minimizing the corresponding Rayleigh quotient. But if one starts from a local minimizer of  $F$ , say  $\bar{\phi}$ , then the eigenvalues at this locally stable state are  $\{0 = \lambda_1 < \lambda_2 < \dots\}$  for the periodic boundary condition and the min-mode is then the zero eigenvector  $v_1 = \partial_x \bar{\phi}$ . However, taking  $\partial_x \bar{\phi}$  as the min-mode is a very bad choice since it will not push the state away but only translate the state back and forth in space. This pathological case could also appear for some special initial states in the convex region of  $F$ , for example, when the minimal eigenvalue  $\lambda_1$  crosses over zero from positive to negative (from the convex region to the non-convex region), i.e., near the so-called branching point. The remedy to avoid this pathological situation in practice is simple: at a state  $\phi$ , whenever the angle between  $v_1$  and  $\partial_x \phi$  is close to  $0^\circ$  or  $180^\circ$  (determined by a prescribed threshold), a constraint is added to make sure that the min-mode  $v$  in use for the auxiliary functional  $L$  must be orthogonal to  $\partial_x \phi$ . In this way, a strictly positive  $\lambda_2$  is selected and accordingly,  $v_2$  is selected as the “min-mode”. As long as  $\lambda_1$  starts to take a negative value, the angle defined above automatically becomes  $90^\circ$  and there is no interference between the min-mode  $v_1$  and the translation direction  $\partial_x \phi$ . After taking care of this issue, we found that for many initial guesses we tried, we did observe the convergence of the algorithm.

## 5. Conclusion

We have demonstrated how the convex splitting method can improve the efficiency of the transition-state calculation by allowing for the preferred large time step size. For the 1-D Ginzburg–Landau energy, this new method has been applied to find index-1 saddle points of the Allen–Cahn and Cahn–Hilliard types, i.e., under the  $L_2$  and  $H^{-1}$  metrics, respectively. Besides, we also test this method for the 2-D Landau–Brazovskii energy functional in  $H^{-1}$  metric. The main advantage of using the convex splitting scheme is to avoid the instability when the time step size is large. And it is also very inspiring that our extensive numerical studies in this paper have shown the significant improvement of the computational efficiency. Therefore, for spatially extended systems driven by an energy functional such as the phase field models or the Kohn–Sham density functional ([27]), we have reasons to speculate that many matured and excellent numerical methods for the traditional gradient dynamics may be able to exhibit their new vitalities for saddle point calculation, if they are correctly wrapped by the iterative minimization formulation.

## References

- [1] Samuel M. Allen, John W. Cahn, A microscopic theory for antiphase boundary motion and its application to antiphase domain coarsening, *Acta Metall.* 27 (1979) 1085–1095.
- [2] Peter W. Bates, Paul C. Fife, Spectral comparison principles for the Cahn–Hilliard and phase-field equations and time scales for coarsening, *Phys. D* 43 (2–3) (1990) 335–348.
- [3] Peter W. Bates, Paul C. Fife, The dynamics of nucleation for the Cahn–Hilliard equation, *SIAM J. Appl. Math.* 53 (4) (1993) 990–1008.
- [4] John W. Cahn, John E. Hilliard, Free energy of a nonuniform system. I. Interfacial free energy, *J. Chem. Phys.* 28 (2) (1958) 258–267.
- [5] E. Cancès, F. Legoll, M.-C. Marinica, K. Minoukadeh, F. Willaime, Some improvements of the activation–relaxation technique method for finding transition pathways on potential energy surfaces, *J. Chem. Phys.* 130 (11) (2009) 114711.

- [6] C.J. Cerjan, W.H. Miller, On finding transition states, *J. Chem. Phys.* 75 (6) (1981) 2800–2806.
- [7] G.M. Crippen, H.A. Scheraga, Minimization of polypeptide energy: XI. The method of gentlest ascent, *Arch. Biochem. Biophys.* 144 (2) (1971) 462–466.
- [8] W. E. W. Ren, E. Vanden-Eijnden, String method for the study of rare events, *Phys. Rev. B* 66 (2002) 052301.
- [9] W. E. W. Ren, E. Vanden-Eijnden, Simplified and improved string method for computing the minimum energy paths in barrier-crossing events, *J. Chem. Phys.* 126 (2007) 164103.
- [10] W. E. X. Zhou, The gentlest ascent dynamics, *Nonlinearity* 24 (6) (2011) 1831.
- [11] Matt Elsey, Benedikt Wirth, A simple and efficient scheme for phase field crystal simulation, *ESAIM: M2AN* 47 (5) (2013) 1413–1432.
- [12] David J. Eyre, An unconditionally stable one-step scheme for gradient systems, manuscript, <http://www.math.utah.edu/~eyre/research/methods/papers.html>, 1998.
- [13] David J. Eyre, Unconditionally gradient stable time marching the Cahn–Hilliard equation, in: Jeffrey W. Bullard, et al. (Eds.), *Computational and Mathematical Models of Microstructural Evolution*, vol. 529, MRS, Warrendale, PA, 1998, pp. 39–46.
- [14] W. Gao, J. Leng, X. Zhou, An iterative minimization formulation for saddle point search, *SIAM J. Numer. Anal.* 53 (4) (2015) 1786–1805.
- [15] W. Gao, J. Leng, X. Zhou, Iterative minimization algorithm for efficient calculations of transition states, *J. Comput. Phys.* 309 (2016) 69–87.
- [16] S. Gu, H. Zhang, Z. Zhang, An energy-stable finite-difference scheme for the binary fluid-surfactant system, *J. Comput. Phys.* 270 (2014) 416–431.
- [17] Shuting Gu, On the Calculation of Transition States, PhD thesis, City University of Hong Kong, 2017.
- [18] G. Henkelman, H. Jónsson, A dimer method for finding saddle points on high dimensional potential surfaces using only first derivatives, *J. Chem. Phys.* 111 (15) (1999) 7010–7022.
- [19] G. Henkelman, H. Jónsson, Theoretical calculations of dissociative adsorption of  $\text{CH}_4$  on an Ir(111) surface, *Phys. Rev. Lett.* 86 (4) (2001) 664–667.
- [20] G. Henkelman, B.P. Uberuaga, H. Jónsson, A climbing image nudged elastic band method for finding saddle points and minimum energy paths, *J. Chem. Phys.* 113 (22) (2000) 9901–9904.
- [21] A. Heyden, A.T. Bell, F.J. Keil, Efficient methods for finding transition states in chemical reactions: comparison of improved dimer method and partitioned rational function optimization method, *J. Chem. Phys.* 123 (2005) 224101.
- [22] Z. Hu, S.M. Wise, C. Wang, J.S. Lowengrub, Stable and efficient finite-difference nonlinear-multigrid schemes for the phase field crystal equation, *J. Comput. Phys.* 228 (15) (2009) 5323–5339.
- [23] H. Jónsson, G. Mills, K.W. Jacobsen, Nudged elastic band method for finding minimum energy paths of transitions, in: B.J. Berne, G. Ciccotti, D.F. Coker (Eds.), *Classical and Quantum Dynamics in Condensed Phase Simulations, Proceedings of the International School of Physics, LERICI, Villa Marigola, World Scientific, New Jersey, 1998*, p. 385.
- [24] J. Kästner, P. Sherwood, Superlinearly converging dimer method for transition state search, *J. Chem. Phys.* 128 (2008) 014106.
- [25] Ralf Kornhuber, Rolf Krause, Robust multigrid methods for vector-valued Allen–Cahn equations with logarithmic free energy, *Comput. Vis. Sci.* 9 (2) (Jun. 2006) 103–116.
- [26] J. Leng, W. Gao, C. Shang, Z.-P. Liu, Efficient softest mode finding in transition states calculations, *J. Chem. Phys.* 138 (9) (2013) 094110.
- [27] Chen Li, Jianfeng Lu, Weitao Yang, Gentlest ascent dynamics for calculating first excited state and exploring energy landscape of Kohn–Sham density functionals, *J. Chem. Phys.* 143 (22) (2015) 224110.
- [28] T. Li, P. Zhang, W. Zhang, Nucleation rate calculation for the phase transition of diblock copolymers under stochastic Cahn–Hilliard dynamics, *Multiscale Model. Simul.* 11 (1) (2013) 385–409.
- [29] Tiejun Li, Pingwen Zhang, Wei Zhang, Nucleation rate calculation for the phase transition of diblock copolymers under stochastic Cahn–Hilliard dynamics, *Multiscale Model. Simul.* 11 (2013) 385–409.
- [30] R. Malek, N. Mousseau, Dynamics of Lennard–Jones clusters: a characterization of the activation–relaxation technique, *Phys. Rev. E* 22 (6) (2000) 7723–7728.
- [31] N. Mousseau, G.T. Barkema, Traveling through potential energy surfaces of disordered materials: the activation–relaxation technique, *Phys. Rev. E* 57 (1998) 2419.
- [32] R.A. Olsen, G.J. Kroes, G. Henkelman, A. Arnaldsson, H. Jónsson, Comparison of methods for finding saddle points without knowledge of the final states, *J. Chem. Phys.* 121 (20) (2004) 9776–9792.
- [33] H. Bernhard Schlegel, Exploring potential energy surfaces for chemical reactions: an overview of some practical methods, *J. Comput. Chem.* 24 (12) (2003) 1514–1527.
- [34] J. Shen, C. Wang, X.M. Wang, 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) (2012) 105–125.
- [35] J. Shen, X.F. Yang, Decoupled energy stable schemes for phase-field model of two-phase complex fluids, *SIAM J. Sci. Comput.* 36 (1) (2014) B122–B145.
- [36] C. Wang, S.M. Wise, An energy stable and convergent finite-difference scheme for the modified phase field crystal equation, *SIAM J. Numer. Anal.* 49 (3) (2011) 945–969.
- [37] Robert A. Wickham, An-Chang Shi, Zhen-Gang Wang, Nucleation of stable cylinders from a metastable lamellar phase in a diblock copolymer melt, *J. Chem. Phys.* 118 (2003) 10293–10305.
- [38] S.M. Wise, C. Wang, J.S. Lowengrub, An energy-stable and convergent finite-difference scheme for the phase field crystal equation, *SIAM J. Numer. Anal.* 47 (3) (2009) 2269–2288.
- [39] S.M. Wise, Unconditionally stable finite difference, nonlinear multigrid simulation of the Cahn–Hilliard–Hele–Shaw system of equations, *J. Sci. Comput.* 44 (1) (2010) 38–68.
- [40] J. Zhang, Q. Du, Shrinking dimer dynamics and its applications to saddle point search, *SIAM J. Numer. Anal.* 50 (2012) 1899–1921.
- [41] L. Zhang, W. Ren, A. Samanta, Q. Du, Recent developments in computational modelling of nucleation in phase transformations, *npj Comput. Mater.* 2 (2016) 16003.
- [42] W. Zhang, T. Li, P. Zhang, Numerical study for the nucleation of one-dimensional stochastic Cahn–Hilliard dynamics, *Commun. Math. Sci.* 10 (2012) 1105–1132.