SIAM J. SCI. COMPUT.
Vol. 42, No. 6, pp. B1350–B1377

EFFICIENT NUMERICAL METHODS FOR COMPUTING THE STATIONARY STATES OF PHASE FIELD CRYSTAL MODELS*

KAI JIANG†, WEI SI†, CHANG CHEN†, AND CHENGLONG BAO‡

Abstract. Finding the stationary states of a free energy functional is an important problem in phase field crystal (PFC) models. Many efforts have been devoted to designing numerical schemes with energy dissipation and mass conservation properties. However, most existing approaches are time-consuming due to the requirement of small effective step sizes. In this paper, we discretize the energy functional and propose efficient numerical algorithms for solving the constrained nonconvex minimization problem. A class of gradient-based approaches, which are the so-called adaptive accelerated Bregman proximal gradient (AA-BPG) methods, is proposed, and the convergence property is established without the global Lipschitz constant requirements. A practical Newton method is also designed to further accelerate the local convergence with convergence guarantee. One key feature of our algorithms is that the energy dissipation and mass conservation properties hold during the iteration process. Moreover, we develop a hybrid acceleration framework to accelerate the AA-BPG methods and most of the existing approaches through coupling with the practical Newton method. Extensive numerical experiments, including two three-dimensional periodic crystals in the Landau–Brazovskii (LB) model and a two-dimensional quasicrystal in the Lifshitz–Petrich (LP) model, demonstrate that our approaches have adaptive step sizes which lead to a significant acceleration over many existing methods when computing complex structures.

Key words. phase field crystal models, stationary states, adaptive accelerated Bregman proximal gradient methods, preconditioned conjugate gradient method, hybrid acceleration framework

AMS subject classifications. 35J60, 35Q74, 65N35

DOI. 10.1137/20M1321176

1. Introduction.

The phase field crystal (PFC) model is an important approach for describing many physical processes and material properties, such as the formation of ordered structures, the nucleation process, crystal growth, elastic and plastic deformations of the lattice, and dislocations [9, 31]. More concretely, letting the order parameter function be $\phi(\boldsymbol{r})$, the PFC model can be expressed by a free energy functional

$$E(\phi;\Theta) = G(\phi;\Theta) + F(\phi;\Theta), \tag{1.1}$$

where Θ are the physical parameters, $F[\phi]$ is the bulk energy with polynomial-type or log-type formulation, and $G[\phi]$ is the interaction energy that contains higher-order differential operators to form ordered structures [6, 26, 36]. A typical interaction

*Submitted to the journal's Computational Methods in Science and Engineering section February 24, 2020; accepted for publication (in revised form) August 7, 2020; published electronically November 9, 2020.

https://doi.org/10.1137/20M1321176

Funding: This work was supported in part by the National Natural Science Foundation of China (11771368, 11901338), the Hunan Science Foundation of China (2018JJ2376), the Youth Project (18B057), the Key Project (19A500) of the Education Department of the Hunan Province of China, and the Tsinghua University Initiative Scientific Research Program.

†School of Mathematics and Computational Science, Hunan Key Laboratory for Computation and Simulation in Science and Engineering, Xiangtan University, Xiangtan, Hunan, China, 411105 (kaijiang@xtu.edu.cn, 201610111098@smail.xtu.edu.cn, 2016750218@smail.xtu.edu.cn).

‡Corresponding author. Yau Mathematical Sciences Center, Tsinghua University, Beijing, China, 100084 (clbao@mail.tsinghua.edu.cn).

potential function for a domain Ω is

$$G(\phi) = \frac{1}{|\Omega|}\int_{\Omega}\left[\prod_{j=1}^{m}(\Delta + q_j^2)\phi\right]^2 d\boldsymbol{r}, \quad m \in \mathbb{N}, \tag{1.2}$$

which can be used to describe the pattern formation of periodic crystals, quasicrystals, and multipolynary crystals [26, 28]. In order to understand the theory of PFC models, as well as predict and guide experiments, it requires finding stationary states $\phi_s(\boldsymbol{r};\Theta)$ and constructing phase diagrams of the energy functional (1.1). Denote V as a feasible space; then the phase diagram is obtained via solving the minimization problem

$$\min_{\phi} E(\phi;\Theta) \quad \text{s.t.} \quad \phi \in V \tag{1.3}$$

with different physical parameters Θ, which brings the tremendous computational burden. Therefore, within an appropriate spatial discretization, the goal of this paper is to develop efficient and robust numerical methods for solving (1.3) with guaranteed convergence while keeping the desired dissipation and conservation properties during the iterative process.

Most existing numerical methods for computing the stationary states of PFC models can be classified into two categories. One is to solve the steady nonlinear Euler–Lagrange equations of (1.3) through different spatial discretization approaches. The other class aims at solving the nonlinear gradient flow equation by using the numerical PDE methods. In these approaches, there have been extensive works on energy stable numerical schemes for the time-dependent PFC model and its various extensions, such as the modified PFC (MPFC) [40, 22, 15] and square PFC (SPFC) models [11]. Typical energy stable schemes to gradient flows include convex splitting methods [41, 35], stabilized factor methods in both the first- and second-order temporal accuracy orders [33], the exponential time differencing schemes [13], and the recently developed invariant energy quadrature [46] and scalar auxiliary variable approaches [32] for a modified energy. It is noted that the gradient flow approach is able to describe the quasi-equilibrium behavior of PFC systems. Numerically, the gradient flow is discretized in both space and time domain via different discretization techniques and the stationary state is obtained with a proper choice of initialization. Many popular spatial approximations have been used, such as the finite difference method [41, 40, 17], the finite element method [14, 12], and the Fourier pseudospectral method [10, 20, 11].

Under an appropriate spatial discretization scheme, the infinite-dimensional problem (1.3) can be formulated as a minimization problem in a finite-dimensional space. Thus, there may exist alternative numerical methods that can converge to the steady states quickly by using modern optimization techniques. For example, similar ideas have shown success in computing steady states of the Bose–Einstein condensate [42] and the calculation of density functional theory [38, 27]. In this paper, in order to keep the mass conservation property, an additional constraint is imposed in (1.3) and the details will be given in the next section. Inspired by the recent advances in gradient-based methods, which have been successfully applied in image processing and machine learning, we propose an adaptive accelerated Bregman proximal gradient (AA-BPG) method for computing the stationary states of (1.3). In each iteration, the AA-BPG method updates the estimation of the order parameter function by solving linear equations which have closed form when using the pseudospectral discretization and chooses step sizes by using the line search algorithm initialized with the

Barzilai–Borwein (BB) method [1]. Meanwhile, a restart scheme is proposed such that the iterations satisfy the energy dissipation property and it is proved that the generated sequence converges to a stationary point of (1.3) without the assumption of the existence of the global Lipschitz constant of the bulk energy F. Moreover, a regularized Newton method is applied for further accelerating the local convergence. More specifically, a preconditioned conjugate gradient method is designed for solving the regularized Newton system efficiently. Extensive numerical experiments have demonstrated that our approach can quickly reach the vicinity of an optimal solution with moderately accuracy, even for very challenge cases.

The rest of this paper is organized as follows. In section 2, we present the PFC models considered in this paper and the projection method discretization. In section 3, we present the AA-BPG method for solving the constrained nonconvex optimization with proved convergence. In section 4, two choices of Bregman distance are proposed and applied for the PFC problems. In section 5, we design a practice Newton preconditioned conjugate gradient (Newton-PCG) method with gradient convergence guarantee. Then, a hybrid acceleration framework is proposed to further accelerate the calculation. Numerical results are reported in section 6 to illustrate the efficiency and accuracy of our algorithms. Finally, some concluding remarks are given in section 7.

1.1. Notations and definitions. Let C^k be the set of kth continuously differentiable functions on the whole space. The domain of a real-valued function f is defined as $\mathrm{dom} f := \{x : f(x) < +\infty\}$. We say f is proper if $f > -\infty$ and $\mathrm{dom} f \neq \emptyset$. For $\alpha \in \mathbb{R}$, let $[f \leq \alpha] := \{x : f(x) \leq \alpha\}$ be the α-(sub)level set of f. We say that f is level bounded if $[f \leq \alpha]$ is bounded for all $\alpha \in \mathbb{R}$. f is lower semicontinuous if all level sets of f are closed. For a proper function f, the subgradient [8] of f at $x \in \mathrm{dom} f$ is defined as $\partial f(x) = \{u : f(y) - f(x) - \langle u, y - x\rangle \geq 0 \ \forall\, y \in \mathrm{dom} f\}$. A point x is called a stationary point of f if $0 \in \partial f(x)$.

2. Problem formulation.

2.1. Physical models. Two classes of PFC models are considered in the paper. The first one is the Landau–Brazovskii (LB) model, which can characterize the phase and phase transitions of periodic crystals [6]. It has been discovered in many different scientific fields, e.g., polymeric materials [34]. In particular, the energy functional of LB model is

$$E_{LB}(\phi) = \frac{1}{|\Omega|}\int_{\Omega} \left\{ \underbrace{\frac{\xi^2}{2}[(\Delta+1)\phi]^2}_{G(\phi)} + \underbrace{\frac{\tau}{2!}\phi^2 - \frac{\gamma}{3!}\phi^3 + \frac{1}{4!}\phi^4}_{F(\phi)} \right\} d\boldsymbol{r}, \tag{2.1}$$

where $\phi(\boldsymbol{r})$ is a real-valued function which measures the order of system in terms of order parameter. Ω is the bounded domain of the system, ξ is the bare correlation length, τ is the dimensionless reduced temperature, and γ is the phenomenological coefficient. Compared with double-well bulk energy [36], the cubic term in the LB functional helps us study the first-order phase transition.

The second one is the Lifshitz–Petrich (LP) model, which can simulate quasiperiodic structures, such as the bifrequency excited Faraday wave [26], and explain the stability of soft-matter quasicrystals [25, 18]. Since quasiperiodic structures are space-filling without decay, it is necessary to define the average spacial integral over the whole space as $⨍ = \lim_{R\to\infty} \frac{1}{|B_R|}\int_{B_R}$, where $B_R \subset \mathbb{R}^d$ is the ball centered at the

origin with radii R. The energy functional of the LP model is given by

$$E_{LP}(\phi) = \fint \left\{ \underbrace{\frac{c}{2}[(\Delta + q_1^2)(\Delta + q_2^2)\phi]^2}_{G(\phi)} + \underbrace{\frac{\varepsilon}{2}\phi^2 - \frac{\kappa}{3}\phi^3 + \frac{1}{4}\phi^4}_{F(\phi)} \right\} d\boldsymbol{r}, \tag{2.2}$$

where c is the energy penalty, and ε and κ are phenomenological coefficients.

Furthermore, we impose the following mean zero condition of order parameter on the LB and LP systems, respectively, to ensure the mass conservation:

$$\frac{1}{|\Omega|}\int_\Omega \phi(\boldsymbol{r})d\boldsymbol{r} = 0 \quad \text{or} \quad \fint \phi(\boldsymbol{r})d\boldsymbol{r} = 0. \tag{2.3}$$

The equality constraint condition is from the definition of the order parameter which is the deviation from average density.

2.2. Projection method discretization. In this section, we introduce the projection method [20], a high-dimensional interpretation approach which can avoid the Diophantine approximation error in computing quasiperiodic systems, to discretize the LB and LP energy functionals. It is noted that the stationary state in the LB model is periodic, and thus it can be discretized by the Fourier pseudospectral method, which is a special case of the projection method. Therefore, we only consider the projection method discretization of the LP model (2.2). We immediately have the following orthonormal property in the average spacial integral sense:

$$\fint e^{i\boldsymbol{k}\cdot\boldsymbol{r}} e^{-i\boldsymbol{k}'\cdot\boldsymbol{r}}\, d\boldsymbol{r} = \delta_{\boldsymbol{k}\boldsymbol{k}'} \quad \forall \boldsymbol{k}, \boldsymbol{k}' \in \mathbb{R}^d. \tag{2.4}$$

For a quasiperiodic function, we can define the Bohr–Fourier transformation as [21]

$$\hat{\phi}(\boldsymbol{k}) = \fint \phi(\boldsymbol{r}) e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}\, d\boldsymbol{r}, \quad \boldsymbol{k} \in \mathbb{R}^d. \tag{2.5}$$

In this paper, we carry out the above computation in a higher dimension using the projection method, which is based on the fact that a d-dimensional quasicrystal can be embedded into an n-dimensional periodic structure ($n \geqslant d$) [16]. The dimension n is the number of linearly independent numbers over the rational number field. Using the projection method, the order parameter $\phi(\boldsymbol{r})$ can be expressed as

$$\phi(\boldsymbol{r}) = \sum_{\boldsymbol{h}\in\mathbb{Z}^n} \hat{\phi}(\boldsymbol{h}) e^{i[(\mathcal{P}\cdot\mathbf{B}\boldsymbol{h})^\top\cdot\boldsymbol{r}]}, \quad \boldsymbol{r} \in \mathbb{R}^d, \tag{2.6}$$

where $\mathbf{B} \in \mathbb{R}^{n\times n}$ is invertible, related to the n-dimensional primitive reciprocal lattice. The corresponding computational domain in physical space is $2\pi\mathbf{B}^{-T}\tau$, $\tau \in [0,1)^n$. The projection matrix $\mathcal{P} \in \mathbb{R}^{d\times n}$ depends on the property of quasicrystals, such as rotational symmetry [16]. If we consider periodic crystals, the projection matrix becomes the d-order identity matrix, and then the projection reduces to the common Fourier spectral method. The Fourier coefficient $\hat{\phi}(\boldsymbol{h})$ satisfies

$$X := \left\{ (\hat{\phi}(\boldsymbol{h}))_{\boldsymbol{h}\in\mathbb{Z}^n} : \hat{\phi}(\boldsymbol{h}) \in \mathbb{C}, \ \sum_{\boldsymbol{h}\in\mathbb{Z}^n} |\hat{\phi}(\boldsymbol{h})| < \infty \right\}. \tag{2.7}$$

In practice, let $\boldsymbol{N} = (N_1, N_2, \ldots, N_n) \in \mathbb{N}^n$, and let

$$X_{\boldsymbol{N}} := \{\hat{\phi}(\boldsymbol{h}) \in X : \hat{\phi}(\boldsymbol{h}) = 0 \ \forall |h_j| > N_j/2,\ j = 1, 2, \ldots, n\}. \tag{2.8}$$

The number of elements in the set is $N = (N_1 + 1)(N_2 + 1)\cdots(N_n + 1)$. Together with (2.4) and (2.6), the discretized energy function (2.2) is

$$E_{\boldsymbol{h}}(\hat{\Phi}) = G_{\boldsymbol{h}}(\hat{\Phi}) + F_{\boldsymbol{h}}(\hat{\Phi}), \tag{2.9}$$

where G_h and F_h are the discretized interaction and bulk energies:

$$\begin{aligned} G_{\boldsymbol{h}}(\hat{\Phi}) &= \frac{c}{2} \sum_{\boldsymbol{h}_1+\boldsymbol{h}_2=0} \left[q_1^2 - (\mathcal{P}\mathbf{B}\boldsymbol{h})^\top(\mathcal{P}\mathbf{B}\boldsymbol{h})\right]^2 \left[q_2^2 - (\mathcal{P}\mathbf{B}\boldsymbol{h})^\top(\mathcal{P}\mathbf{B}\boldsymbol{h})\right]^2 \hat{\phi}(\boldsymbol{h}_1)\hat{\phi}(\boldsymbol{h}_2), \\ F_{\boldsymbol{h}}(\hat{\Phi}) &= \frac{\varepsilon}{2} \sum_{\boldsymbol{h}_1+\boldsymbol{h}_2=\mathbf{0}} \hat{\phi}(\boldsymbol{h}_1)\hat{\phi}(\boldsymbol{h}_2) - \frac{\kappa}{3} \sum_{\boldsymbol{h}_1+\boldsymbol{h}_2+\boldsymbol{h}_3=\mathbf{0}} \hat{\phi}(\boldsymbol{h}_1)\hat{\phi}(\boldsymbol{h}_2)\hat{\phi}(\boldsymbol{h}_3) \\ &\quad + \frac{1}{4} \sum_{\boldsymbol{h}_1+\boldsymbol{h}_2+\boldsymbol{h}_3+\boldsymbol{h}_4=\mathbf{0}} \hat{\phi}(\boldsymbol{h}_1)\hat{\phi}(\boldsymbol{h}_2)\hat{\phi}(\boldsymbol{h}_3)\hat{\phi}(\boldsymbol{h}_4), \end{aligned} \tag{2.10}$$

and $\boldsymbol{h}_j \in \mathbb{Z}^n$, $\hat{\phi}_j \in X_{\boldsymbol{N}}$, $j = 1, 2, \ldots, 4$, $\hat{\Phi} = (\hat{\phi}_1, \hat{\phi}_2, \ldots, \hat{\phi}_N) \in \mathbb{C}^N$. It is clear that the nonlinear terms in F_h are n-dimensional convolutions in the reciprocal space. A direct evaluation of these convolution terms is extremely expensive. Instead, these terms are simple multiplication in the n-dimensional physical space. Similar to the pseudospectral approach, these convolutions can be efficiently calculated through the FFT. Moreover, the mass conservation constraint (2.3) is discretized as

$$e_1^\top \hat{\Phi} = 0, \tag{2.11}$$

where $e_1 = (1, 0, \ldots, 0)^\top \in \mathbb{R}^N$. Therefore, we obtain the following finite-dimensional minimization problem:

$$\min_{\hat{\Phi}\in\mathbb{C}^N} E_{\boldsymbol{h}}(\hat{\Phi}) = G_{\boldsymbol{h}}(\hat{\Phi}) + F_{\boldsymbol{h}}(\hat{\Phi}) \quad \text{s.t.} \quad e_1^\top \hat{\Phi} = 0. \tag{2.12}$$

For simplicity, we omit the subscription in $G_{\boldsymbol{h}}$ and $F_{\boldsymbol{h}}$ in the following context. According to (2.10), denoting $\mathcal{F}_N \in \mathbb{C}^{N\times N}$ as the discretized Fourier transformation matrix, we have

$$\nabla G(\hat{\Phi}) = D\hat{\Phi}, \quad \nabla F(\hat{\Phi}) = \mathcal{F}_N^{-1}\Lambda\mathcal{F}_N\hat{\Phi}, \tag{2.13}$$

$$\nabla^2 G(\hat{\Phi}) = D, \quad \nabla^2 F(\hat{\Phi}) = \mathcal{F}_N^{-1}\Lambda^{(')}\mathcal{F}_N, \tag{2.14}$$

where D is a diagonal matrix with nonnegative entries $c\left[q_1^2 - (\mathcal{P}\mathbf{B}\boldsymbol{h})^\top(\mathcal{P}\mathbf{B}\boldsymbol{h})\right]^2 \times \left[q_2^2 - (\mathcal{P}\mathbf{B}\boldsymbol{h})^\top(\mathcal{P}\mathbf{B}\boldsymbol{h})\right]^2$ and $\Lambda, \Lambda^{(')} \in \mathbb{R}^{N\times N}$ are also diagonal matrices but related to $\hat{\Phi}$. In the next section, we propose the adaptive accelerated Bregman proximal gradient (AA-BPG) method for solving the constrained minimization problem (2.12).

3. The AA-BPG method. Consider the minimization problem that has the form

$$\min_x E(x) = f(x) + g(x), \tag{3.1}$$

where $f \in C^2$ is proper but nonconvex and g is proper, lower semicontinuous, and convex. Let the domain of E be $\mathrm{dom}E = \{x \mid E(x) < +\infty\}$; then we make the following assumptions.

Assumption 3.1. E is bounded below, and for any $x^0 \in \mathrm{dom}E$, the sublevel set $\mathcal{M}(x^0) := \{x|E(x) \le E(x^0)\}$ is compact.

Let h be a strongly convex function such that $\mathrm{dom}h \subset \mathrm{dom}f$ and $\mathrm{dom}g \cap \mathrm{intdom}\, h \neq \emptyset$. Then, it induces the *Bregman divergence* [7] defined as

$$D_h(x,y) = h(x) - h(y) - \langle \nabla h(y), x - y\rangle \quad \forall (x,y) \in \mathrm{dom}h \times \mathrm{intdom}\, h. \tag{3.2}$$

It is noted that $D_h(x,y) \ge 0$ and $D_h(x,y) = 0$ if and only if $x = y$ due to the strong convexity of h. Furthermore, $D_h(x,\bar{x}) \to 0$ as $x \to \bar{x}$. In recent years, Bregman distance–based proximal methods [2, 5] have been proposed and applied for solving (3.1) in a general nonconvex setting [24]. Basically, given the current estimation $x^k \in \mathrm{intdom}\, h$ and step size $\alpha_k > 0$, it updates x^{k+1} via

$$x^{k+1} = \operatorname*{argmin}_x \left\{ g(x) + \langle x - x^k, \nabla f(x^k)\rangle + \frac{1}{\alpha_k} D_h(x, x^k) \right\}. \tag{3.3}$$

Under suitable assumptions, it is proved in [24] that the iterate $\{x^k\}$ has a convergence property similar to that of the traditional proximal gradient method [3], while iteration (3.3) does not require the Lipschitz condition on ∇f. Motivated by the Nesterov acceleration technique [37, 3], we add an extrapolation step before (3.3), and thus the iterate becomes

$$\begin{aligned} y^k &= x^k + w_k(x^k - x^{k-1}), \\ x^{k+1} &= \operatorname*{argmin}_x \left\{ g(x) + \langle x - y^k, \nabla f(y^k)\rangle + \frac{1}{\alpha_k} D_h(x, y^k) \right\}, \end{aligned} \tag{3.4}$$

where $w_k \in [0, \bar{w}]$. It is noted that the minimization problems in (3.3) and (3.4) are well defined and single valued as g is convex and h is strongly convex. Although the extrapolation step accelerates the convergence in some cases, it may generate the oscillating phenomenon of the objective value $E(x)$ that slows down the convergence [30]. Therefore, we propose a restart algorithm that leads to a convergent algorithm for solving (3.1) with the energy dissipation property. Given $\alpha_k > 0$, define

$$z^k = \operatorname*{argmin}_x \left\{ g(x) + \langle x - y^k, \nabla f(y^k)\rangle + \frac{1}{\alpha_k} D_h(x, y^k) \right\}, \tag{3.5}$$

and reset $w_k = 0$ if the following does not hold:

$$E(x^k) - E(z^k) \ge c\|x^k - x^{k+1}\|^2 \tag{3.6}$$

for some constant $c > 0$. In the next section, we will show that (3.6) holds when $w_k = 0$. Overall, the AA-BPG algorithm is presented in Algorithm 3.1.

Step size estimation. In each step, α_k is chosen adaptively by the backtracking linear search method, which is initialized by the BB step [1] estimation, i.e.,

$$\alpha_k = \frac{\langle s_k, s_k\rangle}{\langle s_k, v_k\rangle} \quad \text{or} \quad \frac{\langle v_k, s_k\rangle}{\langle v_k, v_k\rangle}, \tag{3.7}$$

where $s_k = x^k - x^{k-1}$ and $v_k = \nabla f(x^k) - \nabla f(x^{k-1})$. Let $\eta > 0$ be a small constant and z^k be obtained from (3.5); then we adopt the step size α_k whenever the following inequality holds:

$$E(y^k) - E(z^k) \ge \eta\|y^k - z^k\|^2. \tag{3.8}$$

The detailed estimation method is presented in Algorithm 3.2.

Algorithm 3.1. AA-BPG algorithm.

Require: $x^1 = x^0$, $\alpha_0 > 0$, $w_0 = 0$, $\rho \in (0,1)$, $\eta, c, \bar{w} > 0$, and $k = 1$.

while the stop criterion is not satisfied **do**
  Update $y^k = x^k + w^k(x^k - x^{k-1})$
  Estimate α_k by Algorithm 3.2
  Calculate z^k via (3.5)
  if (3.6) holds **then**
    $x^{k+1} = z^k$ and update $w_{k+1} \in [0, \bar{w}]$.
  else
    $x^{k+1} = x^k$ and reset $w_{k+1} = 0$.
  end if
  $k = k + 1$.
end while

Algorithm 3.2. Estimation of α_k at y^k.

Require: x^k, y^k, $\eta > 0$ and $\rho \in (0,1)$ and $\alpha_{\min}, \alpha_{\max} > 0$

Initialize α_k by BB step (3.7).
for $j = 1, 2 \ldots$ **do**
  Calculate z^k via (3.5)
  if (3.8) holds or $\alpha_k < \alpha_{\min}$ **then**
    break
  else
    $\alpha_k = \rho\alpha_k$
  end if
end for
Output $\alpha_k = \max(\min(\alpha_k, \alpha_{\max}), \alpha_{\min})$.

3.1. Convergence analysis. In this section, we focus on the convergence analysis of the proposed AA-BPG method. Before proceeding, we introduce a significant definition used in analysis.

DEFINITION 3.1. *A function $f \in C^2$ is R_f-relative smooth if there exists a strongly convex function $h \in C^2$ such that*

$$R_f \nabla^2 h(x) - \nabla^2 f(x) \succeq 0 \quad \forall x \in \operatorname{intdom} h. \tag{3.9}$$

Throughout this section, we impose the next assumption on f.

Remark 3.2. If $h = \|\cdot\|^2/2$, the relative smoothness becomes the Lipschitz smoothness.

Assumption 3.3. There exists $R_f > 0$ such that f is R_f-relative smooth with respect to a strongly convex function $h \in C^2$.

Remark 3.4. In the LB model (2.1) and LP model (2.2), their bulk energies are fourth degree polynomials and their gradients are not Lipschitz continuous. However, we will show that relative smoothness constant R_f can be $O(1)$ through appropriately choosing the strongly convex function h in Lemma 4.4.

3.2. Convergence property. In this subsection, we will prove the convergence property of Algorithm 3.1. The outline of the proof is given in Figure 1. Under Assumption 3.3, we have the following useful lemma as stated in [2].

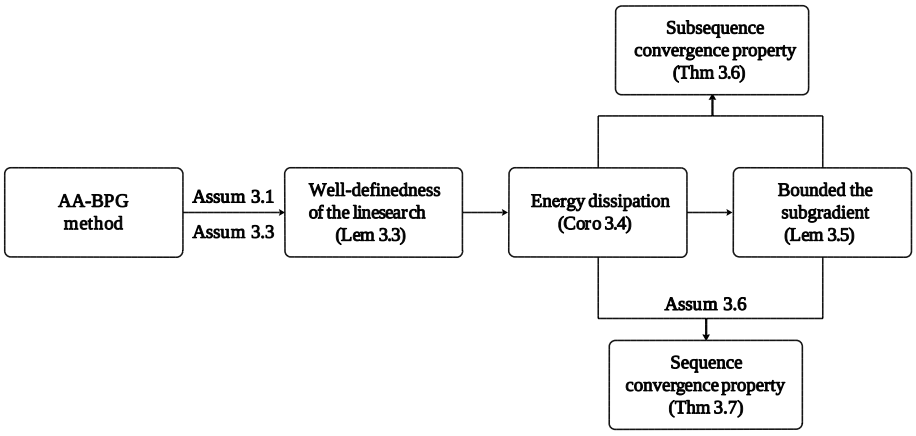

FIG. 1. *The flow chart of the convergence proof of Algorithm* 3.1.

LEMMA 3.2 (see [2]). *If f is R_f-relative smooth with respect to h, then*

$$f(x) - f(y) - \langle \nabla f(y), x - y \rangle \le R_f D_h(x, y) \quad \forall x, y \in \operatorname{intdom} h. \tag{3.10}$$

Based on the above lemma, the descent property of the iteration generated by Bregman proximal operator (3.5) is established as follows.

LEMMA 3.3. *Let $\alpha > 0$, and suppose Assumption* 3.3 *holds. If*

$$z = \operatorname*{argmin}_x \left\{ g(x) + \langle x - y, \nabla f(y) \rangle + \frac{1}{\alpha} D_h(x, y) \right\}, \tag{3.11}$$

then there exists some $\sigma > 0$ such that

$$E(y) - E(z) \ge \left(\frac{1}{\alpha} - R_f \right) \frac{\sigma}{2} \|z - y\|^2. \tag{3.12}$$

Proof. Since h is strongly convex, there exists some constant $\sigma > 0$ such that $h(x) - \sigma\|x\|^2/2$ is convex. Then, $\nabla^2 h(x) - \sigma I \succeq 0$ and we have

$$D_h(z, x) = h(z) - h(y) - \langle \nabla h(y), z - y \rangle \ge \frac{\sigma}{2} \|z - y\|^2. \tag{3.13}$$

From the optimal condition of (3.11), we have

$$\begin{aligned} E(y) = f(y) + g(y) &= \left[f(y) + \langle \nabla f(y), x - y \rangle + \frac{1}{\alpha} D_h(x, y) + g(x) \right]_{x=y} \\ &\ge f(y) + \langle \nabla f(y), z - y \rangle + \frac{1}{\alpha} D_h(z, y) + g(z) \\ &\ge f(z) - R_f D_h(z, y) + \frac{1}{\alpha} D_h(z, y) + g(z) \\ &= E(z) + \left(\frac{1}{\alpha} - R_f \right) D_h(z, y) \ge E(z) + \left(\frac{1}{\alpha} - R_f \right) \frac{\sigma}{2} \|z - y\|^2, \end{aligned}$$

where the second inequality follows from (3.10) and the last inequality follows from (3.13). ∎

Remark 3.5. Lemma 3.3 shows that the nonrestart condition (3.6) and the linear search condition (3.8) are satisfied when

$$0 < \alpha < \bar{\alpha} := \min\left(\frac{1}{2c/\sigma + R_f}, \frac{1}{2\eta/\sigma + R_f}\right) \quad \text{and} \quad 0 < \alpha_{\min} \le \bar{\alpha}. \tag{3.14}$$

Therefore, the line search in Algorithm 3.2 stops in finite iterations, and thus Algorithm 3.1 is well defined.

In the following analysis, we always assume that the parameter $\alpha_{\min}$ satisfies (3.14) for simplicity. Therefore, we can obtain the sufficient decrease property of the sequence generated by Algorithm 3.1.

COROLLARY 3.4. *Suppose Assumptions* 3.1 *and* 3.3 *hold. Let* $\{x^k\}$ *be the sequence generated by Algorithm* 3.1. *Then,* $\{x^k\} \subset \mathcal{M}(x^0)$ *and*

$$E(x^k) - E(x^{k+1}) \ge c_0 \|x^k - x^{k+1}\|^2, \tag{3.15}$$

where $c_0 = \min(c, \eta)$.

The proof of Corollary 3.4 is a straightforward result, as the AA-BPG algorithm is well defined and condition (3.6) or (3.8) holds at each iteration. Let $\mathcal{B}(x^0)$ be the closed ball that contains $\mathcal{M}(x^0)$. Since $h, F \in C^2$, there exist $\rho_h, \rho_f > 0$ such that

$$\rho_h = \sup_{x \in \mathcal{B}(x^0)} \|\nabla^2 h(x)\|, \quad \rho_f = \sup_{x \in \mathcal{B}(x^0)} \|\nabla^2 f(x)\|. \tag{3.16}$$

Thus, we can show that the subgradient of each step generated by Algorithm 3.1 is bounded by the movement of x^k.

LEMMA 3.5 (bounded by the subgradient). *Suppose Assumptions* 3.1 *and* 3.3 *hold. Let* $\{x^k\}$ *be the sequence generated by Algorithm* 3.1. *Then, there exists* $c_1 = \rho_f + \rho_h/\alpha_{\min} > 0$ *such that*

$$\operatorname{dist}(\mathbf{0}, \partial E(x^{k+1})) \le c_1(\|x^{k+1} - x^k\| + \bar{w}\|x^k - x^{k-1}\|), \tag{3.17}$$

where $\operatorname{dist}(\mathbf{0}, \partial E(x^{k+1})) = \inf\{\|y\| : y \in \partial E(x^{k+1})\}$, ρ_h, ρ_f *are defined in* (3.16), *and* $\bar{w}, \alpha_{\min}$ *are constants defined in Algorithms* 3.1 *and* 3.2, *respectively.*

Proof. By the first-order optimality condition of (3.4), we get

$$\begin{aligned} &\mathbf{0} \in \nabla f(y^k) + \frac{1}{\alpha_k}\left(\nabla h(x^{k+1}) - \nabla h(y^k)\right) + \partial g(x^{k+1}) \\ \iff &-\nabla f(y^k) - \frac{1}{\alpha_k}\left(\nabla h(x^{k+1}) - \nabla h(y^k)\right) \in \partial g(x^{k+1}). \end{aligned}$$

Since $f \in C^2$, we know [39, Theorem 5.38] that

$$\partial E(x) = \nabla f(x) + \partial g(x). \tag{3.18}$$

From Lemma 3.3, we have $x^k, x^{k-1} \in \mathcal{M}(x^0)$; then $y^k = (1+w_k)x^k - w_k x^{k-1} \in \mathcal{B}(x^0)$.

Together with (3.18), we have

$$\begin{aligned}\operatorname{dist}(\mathbf{0},\partial E(x^{k+1})) &= \inf_{y\in\partial g(x^{k+1})}\|\nabla f(x^{k+1})+y\|\\ &\le \left\|\nabla f(x^{k+1})-\nabla f(y^k)-\frac{1}{\alpha_k}\left(\nabla h(x^{k+1})-\nabla h(y^k)\right)\right\|\\ &\le \|\nabla f(x^{k+1})-\nabla f(y^k)\|+\frac{1}{\alpha_k}\|\nabla h(x^{k+1})-\nabla h(y^k)\|\\ &\le \left(\rho_f+\frac{\rho_h}{\alpha_k}\right)\|x^{k+1}-y^k\|\\ &\le c_1(\|x^{k+1}-x^k\|+\bar{w}\|x^k-x^{k-1}\|),\end{aligned}$$

where the last inequality is from $y^k = x^k + w_k(x^k - x^{k-1})$ and $w^k \in [0, \bar{w}]$. ∎

Now, we are ready to establish the subconvergence property of Algorithm 3.1.

THEOREM 3.6. *Suppose Assumptions* 3.1 *and* 3.3 *hold. Let* $\{x^k\}$ *be the sequence generated by Algorithm* 3.1. *Then, for any limit point* x^* *of* $\{x^k\}$, *we have* $\mathbf{0} \in \partial E(x^*)$.

Proof. From Corollary 3.4, we know that $\{x^k\} \subset \mathcal{M}(x^0) \subset \mathcal{B}(x^0)$ and thus bounded. Then, the set of limit points of $\{x^k\}$ is nonempty. For any limit point x^*, there exists a subsequence $\{x^{k_j}\}$ such that $x^{k_j} \to x^*$ as $j \to \infty$. We know that $\{E(x^k)\}$ is a decreasing sequence. Together with the fact that E is bounded below, there exists some $\bar{E}$ such that $E(x^k) \to \bar{E}$ as $k \to \infty$. Moreover, it has

$$E(x^0) - \bar{E} = \lim_{K\to\infty}\sum_{j=0}^{K}\left(E(x^j) - E(x^{j+1})\right) \ge c_0 \lim_{K\to\infty}\sum_{j=0}^{K}\|x^j - x^{j+1}\|^2 \tag{3.19}$$

and implies that $\|x^k - x^{k-1}\| \to 0$ as $k \to \infty$. As a result,

$$\lim_{k\to\infty}\|x^k - y^{k-1}\| \le \lim_{k\to\infty}(\|x^k - x^{k-1}\| + \bar{\omega}\|x^{k-1} - x^{k-2}\|) = 0.$$

Together with (3.17), it implies that there exists $u^{k_j} \in \partial g(x^{k_j})$ such that

$$\lim_{j\to\infty}\|\nabla f(x^{k_j}) + u^{k_j}\| = 0 \Rightarrow \lim_{j\to\infty} u^{k_j} = -\nabla f(x^*), \tag{3.20}$$

as ∇f is continuous and $x^{k_j} \to x^*$ when $j \to \infty$.

Next, we prove $\lim_{j\to\infty} g(x^{k_j}) = g(x^*)$. It is easy to know that $\lim_{j\to\infty} x^{k_j - p} = x^*$ for finite $p \ge 0$ since $\lim_{k\to\infty}\|x^k - x^{k-1}\| = 0$. Thus, we have $y^{k_j-1} = x^{k_j-1} + w_{k_j-1}(x^{k_j-1} - x^{k_j-2}) \to x^*$ as $j \to \infty$. From (3.4), we know that

$$\begin{aligned} & g(x^{k_j}) + \langle x^{k_j} - y^{k_j-1}, \nabla f(y^{k_j-1})\rangle + \frac{1}{\alpha_k}D(x^{k_j}, y^{k_j-1})\\ \le\ & g(x) + \langle x - y^{k_j-1}, \nabla f(y^{k_j-1})\rangle + \frac{1}{\alpha_k}D(x, y^{k_j-1}) \quad \forall x.\end{aligned} \tag{3.21}$$

Let $x = x^*$ and $j \to \infty$; then we get $\limsup_{j\to\infty} g(x^{k_j}) \le g(x^*)$. By the fact that $g(x)$ is lower semicontinuous, it has $\lim_{j\to\infty} g(x^{k_j}) = g(x^*)$.

Thus, by the convexity of g, we have

$$g(x) \ge g(x^{k_j}) + \langle u^{k_j}, x - x^{k_j}\rangle \quad \forall x \in \operatorname{dom} g. \tag{3.22}$$

Let $j \to \infty$ in (3.22); then, using the fact that $x^{k_j} \to x^*$, $g(x^{k_j}) \to g(x^*)$ as $j \to \infty$, and (3.20), we have $-\nabla f(x^*) \in \partial g(x^*)$, and thus $0 \in \partial E(x^*)$. ∎

Furthermore, the subsequence convergence can be strengthened by imposing the next assumption on E, which is known as the Kurdyka–Lojasiewicz (KL) property [4].

Assumption 3.6. $E(x)$ is the KL function; i.e., for all $\bar{x} \in \mathrm{dom}\partial E := \{x : \partial E(x) \neq \emptyset\}$, there exist $\eta > 0$, a neighborhood U of $\bar{x}$, and $\psi \in \Psi_\eta := \{\psi \in C[0,\eta) \cap C^1(0,\eta),$ where ψ is concave, $\psi(0) = 0$, and $\psi^{'} > 0$ on $(0,\eta)\}$ such that for all $x \in U \cap \{x : E(\bar{x}) < E(x) < E(\bar{x}) + \eta\}$, the following inequality holds:

$$\psi^{'}(E(x) - E(\bar{x}))\,\mathrm{dist}(\mathbf{0}, \partial E(x)) \geq 1. \tag{3.23}$$

THEOREM 3.7. *Suppose Assumptions* 3.1, 3.3, *and* 3.6 *hold. Let* $\{x^k\}$ *be the sequence generated by Algorithm* 3.1. *Then, there exists a point* $x^* \in \mathcal{B}(x^0)$ *such that*

$$\lim_{k\to+\infty} x^k = x^*, \quad \mathbf{0} \in \partial E(x^*). \tag{3.24}$$

Proof. The proof is in Appendix A. ∎

It is known from [4] that many functions satisfy Assumption 3.6, including the energy function in PFC models. In the following context, we apply the AA-BPG method for solving the PFC models (2.12) by introducing two Bregman distances.

4. AA-BPG method for solving PFC models. The problem (2.12) can be reduced to (3.1) by setting

$$f(\hat{\Phi}) = F(\hat{\Phi}), \quad g(\hat{\Phi}) = G(\hat{\Phi}) + \delta_{\mathcal{S}}(\hat{\Phi}), \tag{4.1}$$

where $\mathcal{S} = \{\hat{\Phi} : e_1^\top \hat{\Phi} = 0\}$ and $\delta_{\mathcal{S}}(\hat{\Phi}) = 0$ if $\hat{\Phi} \in \mathcal{S}$ and $+\infty$ otherwise. The main difficulty of applying Algorithm 3.1 is solving the subproblem (3.5) efficiently. In this section, two different strongly convex functions h are chosen as

$$h(x) = \frac{1}{2}\|x\|^2 \ \ \text{(P2)} \qquad \text{and} \qquad h(x) = \frac{a}{4}\|x\|^4 + \frac{b}{2}\|x\|^2 + 1 \ \ \text{(P4)}, \tag{4.2}$$

where $a, b > 0$ and (P2) and (P4) represent the highest order of the ℓ^2 norm.

Case (P2). The Bregman distance of D_h is reduced to the Euclidean distance, i.e.,

$$D_h(x, y) = \frac{1}{2}\|x - y\|^2. \tag{4.3}$$

The subproblem (3.5) is reduced to

$$\min_{\hat{\Phi}} \ G(\hat{\Phi}) + \langle \nabla F(\hat{\Psi}^k), \hat{\Phi} - \hat{\Psi}^k\rangle + \frac{1}{2\alpha_k}\|\hat{\Phi} - \hat{\Psi}^k\|^2 \quad \text{s.t.} \quad e_1^\top \hat{\Phi} = 0, \tag{4.4}$$

where $\hat{\Psi}^k = \hat{\Phi}^k + w_k(\hat{\Phi}^k - \hat{\Phi}^{k-1})$. Although (4.4) is a constrained minimization problem, it has a closed form solution based on our discretization which leads to a fast computation.

LEMMA 4.1. *Given* $\alpha_k > 0$, *if* $e_1^\top \hat{\Psi}^k = 0$, *the minimizer of* (4.4), *denoted by* $\hat{\Phi}^{k+1}$, *is given by*

$$\hat{\Phi}^{k+1} = (I + \alpha_k D)^{-1}\left(\hat{\Psi}^k - \alpha_k \mathcal{P}_1 \nabla F(\hat{\Psi}^k)\right), \tag{4.5}$$

where D *is defined in* (2.13) *and* $\mathcal{P}_1 = I - e_1 e_1^\top$ *is the projection into the set* $\mathcal{S}$.

Proof. The KKT conditions for this subproblem (3.5) can be written as

$$\nabla G(\hat{\Phi}^{k+1}) + \nabla F(\hat{\Psi}^k) + \frac{1}{\alpha_k}\left(\hat{\Phi}^{k+1} - \hat{\Psi}^k\right) - \xi_k e_1 = 0, \tag{4.6}$$

$$e_1^\top \hat{\Phi}^{k+1} = 0, \tag{4.7}$$

where ξ_k is the Lagrange multiplier. Taking the inner product with e_1 in (4.6), we obtain

$$\xi_k = e_1^\top \left(\nabla G(\hat{\Phi}^{k+1}) + \nabla F(\hat{\Phi}^k) - \frac{1}{\alpha_k}\hat{\Psi}^k\right).$$

Using (4.7) and (2.13), we know that

$$e_1^\top \nabla G(\hat{\Phi}^{k+1}) = e_1^\top (D\hat{\Phi}^{k+1}) = 0.$$

Together with $e_1^\top \hat{\Psi}^k = 0$, we have $\xi_k = e_1^\top \nabla F(\hat{\Psi}^k)$. Substituting it into (4.6), it follows that

$$\hat{\Phi}^{k+1} = (\alpha_k D + I)^{-1}\left(\hat{\Psi}^k - \alpha_k \mathcal{P}_1 \nabla F(\hat{\Psi}^k)\right). \qquad \square$$

It is noted that from the proof of Lemma 4.1, the feasibility assumption $e_1^\top \hat{\Psi}^k = 0$ holds as long as $e_1^\top \hat{\Phi}^0 = 0$, which can be set in the initialization. The detailed algorithm is given in Algorithm 4.1 with $K = 2$.

Case (P4). In this case, the subproblem (3.5) is reduced to

$$\min_{\hat{\Phi}} G(\hat{\Phi}) + \langle \nabla F(\hat{\Psi}^k), \hat{\Phi} - \hat{\Psi}^k\rangle + D_h(\hat{\Phi}, \hat{\Psi}^k) \quad \text{s.t.} \quad e_1^\top \hat{\Phi} = 0, \tag{4.8}$$

where $\hat{\Psi}^k = \hat{\Phi}^k + w_k(\hat{\Phi}^k - \hat{\Phi}^{k-1})$. The next lemma shows the optimal condition of minimizing (4.8).

LEMMA 4.2. *Given $\alpha^k > 0$, if $e_1^\top \hat{\Psi}^k = 0$, the minimizer of* (4.8)*, denoted by $\hat{\Phi}^{k+1}$, is given by*

$$\hat{\Phi}^{k+1} = [\alpha_k D + (ap^* + b)I]^{-1}(\nabla h(\hat{\Psi}^k) - \alpha_k \mathcal{P}_1 \nabla F(\hat{\Psi}^k)), \tag{4.9}$$

where D is given in (2.13) *and p^* is a fixed point of $p = \|\hat{\Phi}^{k+1}\|^2 := r(p)$.*

Proof. The KKT conditions of (4.8) imply that there exists a Lagrange multiplier ξ_k such that $(\hat{\Phi}^{k+1}, \xi_k)$ satisfies

$$\alpha_k \nabla G(\hat{\Phi}^{k+1}) + \alpha_k \nabla F(\hat{\Psi}^k) + \nabla h(\hat{\Phi}^{k+1}) - \nabla h(\hat{\Psi}^k) - \xi_k e_1 = 0, \tag{4.10}$$

$$e_1^\top \hat{\Phi}^{k+1} = 0. \tag{4.11}$$

Since $e_1^\top \hat{\Phi}^k = 0$ and $\nabla h(x) = (a\|x\|^2 + b)x$, (4.11) and (2.13) imply

$$e_1^\top \nabla G(\hat{\Phi}^{k+1}) = e_1^\top (D\hat{\Phi}^{k+1}) = 0, \quad e_1^\top \nabla h(\hat{\Psi}^k) = (a\|\hat{\Psi}^k\|^2 + b) e_1^\top \hat{\Psi}^k = 0,$$

where D is defined in (2.13). Substituting the above equalities into (4.10) implies $\xi_k = \alpha_k e_1^\top \nabla F(\hat{\Psi}^k)$. Denote

$$p := \|\hat{\Phi}^{k+1}\|^2 \geq 0, \quad \beta := \nabla h(\hat{\Psi}^k) - \alpha_k \nabla F(\hat{\Psi}^k) + \xi e_1 = \nabla h(\hat{\Psi}^k) - \alpha_k \mathcal{P}_1 \nabla F(\hat{\Psi}^k).$$

From (4.10), we obtain a fixed point problem with respect to p:

$$p = \|\hat{\Phi}^{k+1}\|^2 = \|[D + (ap + b)I]^{-1}\beta\|^2 := r(p). \tag{4.12}$$

Let $R(p) = r(p) - p$. Then $R(0) = \|(D + bI)^{-1}\beta\|^2 \geq 0$, $R(p) \to -\infty$ as $p \to \infty$, and

$$R'(p) = -2a \sum_{i=1}^{n} \frac{\beta_i^2}{(D_{ii} + ap + b)^3} - 1 < 0 \quad \forall p \geq 0,$$

and there is a unique zero $p^* \geq 0$ of $R(p)$, i.e., $p^* = r(p^*)$. Thus,

$$\hat{\Phi}^{k+1} = [\alpha_k D + (ap^* + b)I]^{-1}(\nabla h(\hat{\Psi}^k) - \alpha_k \mathcal{P}_1 \nabla F(\hat{\Psi}^k)). \qquad \square$$

It is noted that the fixed point equation (4.12) is a nonlinear scalar equation which can be efficiently solved by many existing solvers. The detailed algorithm is given in Algorithm 4.1 with $K = 4$.

Algorithm 4.1. AA-BPG-K method for PFC model.

Require: $\hat{\Phi}^1 = \hat{\Phi}^0$, $\alpha_0 > 0$, $w_0 \in [0, 1]$, $\rho \in (0, 1)$, $\eta, c, \bar{w} > 0$, and $k = 1$.

while stop criterion is not satisfied **do**
  Update $\hat{\Psi}^k = \hat{\Phi}^k - w_k(\hat{\Phi}^k - \hat{\Phi}^{k-1})$
  Estimate α_k by Algorithm 3.2
  if $K = 2$ **then**
    Calculate $z^k = (\alpha_k D + I)^{-1}\left(\hat{\Psi}^k - \alpha_k \mathcal{P}_1 \nabla F(\hat{\Psi}^k)\right)$
  else if $K = 4$ **then**
    Calculate the fixed point of (4.12).
    Calculate $z^k = [\alpha_k D + (ap^* + b)I]^{-1}(\nabla h(\hat{\Psi}^k) - \alpha_k \mathcal{P}_1 \nabla F(\hat{\Psi}^k))$
  end if
  if $E(\hat{\Phi}^k) - E(z^k) \geq c\|\hat{\Phi}^k - z^k\|^2$ **then**
    $\Phi^{k+1} = z^k$ and update $w_{k+1} \in [0, \bar{w}]$.
  else
    $\Phi^{k+1} = \Phi^k$ and reset $w_{k+1} = 0$.
  end if
  $k = k + 1$.
end while

4.1. Convergence analysis for Algorithm 4.1. The convergence analysis can be directly applied for Algorithm 4.1 if the assumptions required in Theorem 3.6 hold. We first show that the energy function E in PFC model satisfies Assumptions 3.1 and 3.6. Then, Assumption 3.3 is analyzed for Cases (P2) and (P4) independently.

LEMMA 4.3. *Let $E_0 = F(\hat{\Phi}) + G(\hat{\Phi})$, and let $E(\hat{\Phi}) = E_0(\hat{\Phi}) + \delta_{\mathcal{S}}(\hat{\Phi})$ be the energy functional which is defined in* (4.1)*. Then, it satisfies the following:*

1. *E is bounded below, and the sublevel set $\mathcal{M}(\hat{\Phi}^0)$ is compact for any $\hat{\Phi}^0 \in \mathcal{S}$.*
2. *E is a KL function and thus satisfies Assumption* 3.6*.*

Proof. From the continuity and the coercive property of F, i.e., $F(\hat{\Phi}) \to +\infty$ as $\hat{\Phi} \to \infty$, the sublevel set $\mathcal{S}_0 := \{\hat{\Phi} : E_0(\hat{\Phi}) \leq E_0(\hat{\Phi}^0)\}$ is compact for any $\hat{\Phi}^0$. Together with $\mathcal{S}$ being closed, it follows that $\mathcal{M}(\hat{\Phi}^0) = \mathcal{S} \cap \mathcal{S}_0$ is compact for any $\hat{\Phi}^0$.

Moreover, according to Example 2 in [4], it is easy to know that $E(\hat{\Phi})$ is a semi-algebraic function, and then it is a KL function by Theorem 2 in [4]. $\square$

LEMMA 4.4. *Let $F(\hat{\Phi})$ be defined in* (2.12)*. Then, we have the following:*

1. *If h is chosen as Case* (P2) *in* (4.2)*, then F is relative smooth with respect to h in $\mathcal{M}$ for any compact set $\mathcal{M}$.*
2. *If h is chosen as Case* (P4) *in* (4.2)*, then F is relative smooth with respect to h.*

Proof. Denote $\hat{\Phi}^{\otimes k} := \hat{\Phi} \otimes \hat{\Phi} \otimes \cdots \otimes \hat{\Phi}$, where $\otimes$ is the tensor product. Then, $F(\hat{\Phi})$ is the 4th-degree polynomial, i.e., $F(\hat{\Phi}) = \sum_{k=2}^{4} \langle \mathcal{A}_k, \hat{\Phi}^{\otimes k} \rangle$, where the kth-degree monomials are arranged as a kth-order tensor $\mathcal{A}_k$. For any compact set $\mathcal{M}$, $\nabla^2 F$ is bounded, and thus F is relative smooth with respect to any polynomial function in $\mathcal{M}$ which includes Case (P2). When h is chosen as Case (P4), according to Lemma 2.1 in [24], there exists $R_F > 0$ such that $F(\hat{\Phi})$ is R_F-relative smooth with respect to $h(x)$. ∎

Combining Lemmas 4.3 and 4.4 with Theorem 3.7, we can directly establish the convergence of Algorithm 4.1.

THEOREM 4.5. *Let $E(\hat{\Phi}) = F(\hat{\Phi}) + G(\hat{\Phi}) + \delta_{\mathcal{S}}(\hat{\Phi})$ be the energy function which is defined in* (4.1)*. The following results hold:*

1. *Let $\{\hat{\Phi}^k\}$ be the sequence generated by Algorithm* 4.1 *with $K = 2$. If $\{\hat{\Phi}^k\}$ is bounded, then $\{\hat{\Phi}^k\}$ converges to some $\hat{\Phi}^*$ and $0 \in \partial E(\hat{\Phi}^*)$.*
2. *Let $\{\hat{\Phi}^k\}$ be the sequence generated by Algorithm* 4.1 *with $K = 4$. Then, $\{\hat{\Phi}^k\}$ converges to some $\hat{\Phi}^*$ and $0 \in \partial E(\hat{\Phi}^*)$.*

It is noted that when h is chosen as Case (P2), we cannot bound the growth of F, as F is a fourth-order polynomial. Thus, the boundedness assumption of $\{\hat{\Phi}^k\}$ is imposed, which is similar to the requirement in the semi-implicit scheme [33].

5. Newton-PCG method. Despite the fast initial convergence speed of the gradient-based methods, the tail convergence speed becomes slow. Therefore, it can be further locally accelerated by the feature of Hessian-based methods. In this section, we design a practical Newton method to solve the PFC models (2.12) and provide a hybrid accelerated framework.

5.1. Our method. Define $Z := [0, I_{N-1}]^\top$. Any vector $\hat{\Phi}$ that satisfies the constraint $e_1^\top \hat{\Phi} = 0$ has the form of $\hat{\Phi} = ZU$ with $U \in \mathbb{C}^{N-1}$. Since $Z^\top Z = I_{N-1}$, we can also obtain U from $\hat{\Phi}$ by $U = Z^\top \hat{\Phi}$. Therefore, the problem (2.12) is equivalent to

$$\min_{U \in \mathbb{C}^{N-1}} E(ZU) = G(ZU) + F(ZU). \tag{5.1}$$

Let $\tilde{E}(U) := E(ZU)$, $\tilde{G}(U) := E(ZU)$, $\tilde{F}(U) := F(ZU)$; then we have the following facts:

$$\begin{aligned} \tilde{g} &:= \nabla \tilde{E}(U) = Z^\top \nabla E(ZU) = Z^\top g, \\ \tilde{\mathcal{J}} &:= \nabla^2 \tilde{E}(U) = Z^\top \nabla^2 E(ZU) Z = Z^\top \mathcal{J} Z, \end{aligned} \tag{5.2}$$

where $g = \nabla E(ZU)$ and $\mathcal{J} = \nabla^2 E(ZU)$. Therefore, finding the steady states of PFC models is equivalent to solving the nonlinear equations

$$\nabla \tilde{E}(U) = \mathbf{0}. \tag{5.3}$$

Due to the nonconvexity of $\tilde{E}(U)$, the Hessian matrix $\tilde{\mathcal{J}}$ may not be positive definite, and thus a regularized Newton method is applied.

Computing the Newton direction. Denote $\tilde{\mathcal{J}}_k := \nabla^2 \tilde{E}(U^k)$ and $\tilde{g}_k := \nabla \tilde{E}(U^k)$; then we find the approximated Newton direction d_k by solving

$$(\tilde{\mathcal{J}}_k + \mu_k I) d_k = -\tilde{g}_k, \tag{5.4}$$

where regularized parameter μ_k is chosen as

$$-c_1 \min\{0, \lambda_{\min}(\tilde{\mathcal{J}}_k)\} + c_2 \|\tilde{g}_k\| \le \mu_k \le \bar{\mu} < +\infty \ (c_1 \ge 1, c_2 > 0). \tag{5.5}$$

Thus, (5.4) is a symmetric, positive definite linear system. To accelerate the convergence, a preconditioned conjugate gradient (PCG) method is adopted. More specifically, in the kth step, we terminate the PCG iterates whenever $\|(\tilde{\mathcal{J}}_k + \mu_k I) d_k + \tilde{g}_k\| \le \eta_k$, in which η_k is set as

$$\eta_k = \tau \min\{1, \|\tilde{g}_k\|\}, \quad 0 < \tau < 1, \tag{5.6}$$

and the preconditioner M_k is adaptively obtained by setting

$$M_k = Z(H_k + \mu_k I)^{-1} Z^\top \quad \text{with} \quad H_k = D + \delta_k I, \tag{5.7}$$

where D is from (2.13) and some $\delta_k > 0$. Let $A = \tilde{\mathcal{J}}_k + \mu_k I$, $b = -\tilde{g}_k$, and $M = M_k$; then the PCG method is given in Algorithm 5.1, where $\|x\|_A := \langle x, Ax \rangle$.

Algorithm 5.1. PCG(η) for solving $Ax = b$.

Require: A, b, η, k_{max}, preconditioner M.

Set $x^0 = 0$, $r_0 = Ax^0 - b = -b$, $p_0 = -M^{-1} r_0$, $i = 0$.
while $\|r_i\| > \eta$ or $i < k_{max}$ **do**
  $\alpha_{i+1} = \dfrac{\|r_i\|^2_{M^{-1}}}{\|p_i\|^2_A}$
  $x^{i+1} = x^i + \alpha_{i+1} p_i$
  $r_{i+1} = r_i + \alpha_{i+1} A p_i$
  $\beta_{i+1} = \dfrac{\|r_{i+1}\|^2_{M^{-1}}}{\|r_i\|^2_{M^{-1}}}$
  $p_{i+1} = -M^{-1} r_{i+1} + \beta_{i+1} p_i$
  $i = i + 1$
end while

Computing the step size t_k. Once the Newton direction d_k is obtained, the line search technique is applied for finding an appropriate step size t_k that satisfies the following inequality:

$$\tilde{E}(U^k + t_k d_k) \le \tilde{E}(U^k) + \nu t_k \langle \tilde{g}_k, d_k \rangle, \quad 0 < \nu < 1. \tag{5.8}$$

The existence of $t_k > 0$ that satisfies (5.8) is given in Lemma 5.3. Then, U^{k+1} is updated by $U^{k+1} = U^k + t_k d_k$. Our proposed algorithm is summarized in Algorithm 5.2.

5.2. Convergence analysis for Algorithm 5.2. We first establish several properties related to the direction d_k computed by the PCG method.

LEMMA 5.1. *Consider a linear system $Ax = b$, where A is symmetric and positive definite. Let $\{x^i\}$ be the sequence generated by Algorithm* 5.1*; then it satisfies*

$$\frac{1}{\lambda_{\max}(A)} \le \frac{\langle x^i, b \rangle}{\|b\|^2} \le \frac{1}{\lambda_{\min}(A)} \quad \forall i = 1, 2, \ldots. \tag{5.9}$$

Algorithm 5.2. Newton-PCG method.

Require: $U^0, \varepsilon, \bar{\mu}, c_1 \geq 1$, $c_2 > 0$, $0 < \nu, \rho, \tau < 1$;
$k = 0, \tilde{g}_0 = \nabla \tilde{E}(U^0)$;
while stop criterion is not satisfied **do**
  Choose $-c_1 \min\{0, \lambda_{\min}(\tilde{\mathcal{J}}_k)\} + c_2 \|\tilde{g}_k\| \leq \mu_k \leq \bar{\mu}$;
  Update $\eta_k = \tau \min(1, \|\tilde{g}_k\|)$.
  Find direction d_k by solving (5.4) via PCG(η_k) using Algorithm 5.1;
  for $n = 0, 1, 2, \ldots$ **do**
    $t_k = \rho^n$;
    if $\tilde{E}(U^k + t_k d_k) \leq \tilde{E}(U^k) + \nu t_k \langle \tilde{g}_k, d_k \rangle$ **then**
      Break;
    end if
  end for
  $U^{k+1} = U^k + t_k d_k$;
  $k = k + 1$;
end while

Proof. The proof is in Appendix B. ∎

Then, we know that the d_k is a descent direction from the next lemma.

LEMMA 5.2 (descent direction). *Let d_k be generated by the PCG(η_k) method (Algorithm 5.1). If $\|\tilde{g}_k\| > 0$, then we have*

$$-\langle d_k, \tilde{g}_k \rangle \geq l_k := \frac{\|\tilde{g}_k\|^2}{\lambda_{\max}(\tilde{\mathcal{J}}_k + \mu_k I)} \quad \text{and} \quad \|d_k\| \leq \bar{d} := \frac{\tau + 1}{c_2}, \tag{5.10}$$

where τ, K, and c_1, c_2 are defined in (5.6), (5.15), *and* (5.5), *respectively.*

Proof. The first inequality is a direct consequence of Lemma 5.1. Moreover, let $r_k = (\tilde{\mathcal{J}}_k + \mu_k I) d_k + \tilde{g}_k$. By Algorithm 5.1 and (5.6), we have $\|r_k\| \leq \eta_k \leq \tau \|\tilde{g}_k\|$. Then,

$$\|d_k\| = \|(\tilde{\mathcal{J}}_k + \mu_k I)^{-1}(r_k - \tilde{g}_k)\| \leq \frac{\|r_k - \tilde{g}_k\|}{\lambda_{\min}(\tilde{\mathcal{J}}_k + \mu_k I)} \leq \frac{\|r_k\| + \|g_k\|}{c_2 \|g_k\|} \leq \frac{\tau + 1}{c_2},$$

where the second inequality is from (5.5). ∎

LEMMA 5.3 (lower bound of t_k). *Let d_k be generated by the PCG(η_k) method (Algorithm 5.1). If $\|\tilde{g}_k\| \geq \varepsilon > 0$, then for any $\nu \in (0, 1)$, there exists $M_k > 0$ and*

$$t_{\max}^k := \min\left\{ \frac{2(1-\nu) l_k}{M_k \bar{d}^2}, 1 \right\}, \tag{5.11}$$

such that the inequality (5.8) *holds for $t_k \in (0, t_{\max}^k]$, where l_k is defined in* (5.10).

Proof. By the Taylor expansion, we have

$$\tilde{E}(U^k + t d_k) = \tilde{E}(U^k) + t \langle \tilde{g}_k, d_k \rangle + \frac{t^2}{2} \langle d_k, \nabla^2 \tilde{E}(\xi^t) d_k \rangle, \tag{5.12}$$

where $\xi^t \in \mathcal{V}_k = \{V | V = U^k + t d_k, t \in [0, 1]\}$. As $\tilde{E} \in C^2$, there exists $M_k > 0$ such that $M_k = \sup\{\|\nabla^2 \tilde{E}(V)\| | V \in \mathcal{V}_k\}$. Then, (5.12) and (5.10) imply that

$$\tilde{E}(U^k + t d_k) \leq \tilde{E}(U^k) + \nu t \langle \tilde{g}_k, d_k \rangle - (1 - \nu) l_k t + \frac{M_k \bar{d}^2}{2} t^2. \tag{5.13}$$

Define $Q(t) = (1-\nu)l_k t - \frac{M_k \bar{d}^2}{2}t^2$; then we know that $Q(t) \geq 0$ for all $t \in [0, \frac{2(1-\nu)l_k}{M_k \bar{d}^2}]$, which implies that (5.8) holds for all $t \in (0, t^k_{\max}]$. ∎

THEOREM 5.4. *Let $\tilde{E}$ be defined in* (5.1)*, and let $\{U^k\}$ be the infinite sequence generated by Algorithm* 5.2*. Then, $\{U^k\}$ is bounded and has the following property:*

$$\lim_{k\to+\infty} \|\tilde{g}_k\| = 0. \tag{5.14}$$

Proof. Due to the continuity of $\tilde{F}$, $\tilde{G}$ in (5.1) and the coercive property of $\tilde{F}$, the sublevel set $\mathcal{M}_0 = \{U : \tilde{E}(U) \leq \tilde{E}(U^0)\}$ is compact for any U^0. By the inequality (5.8), it is easy to know that $\{\tilde{E}(U^k)\}$ is a decreasing sequence, and thus $\{U^k\} \subset \mathcal{M}_0$ and there exists some $\bar{E}$ such that $\tilde{E}(U^k) \to \bar{E}$ as $k \to \infty$. Moreover, from (5.10) and $t_k \in (0, 1]$, we know that there exists a compact set $\mathcal{B}_0$ such that $\{U^k + td^k | t \in (0, 1]\} \subset \mathcal{B}_0$, and thus there exists $M > 0$ such that

$$\|\nabla^2 \tilde{E}(U)\| \leq M \quad \forall\, U \in \mathcal{B}_0. \tag{5.15}$$

From the proof of Lemma 5.3, we know that $M_k \leq M$ for all k. Moreover, there exists some $\bar{\lambda} > 0$ such that $\lambda_{\max}(\tilde{\mathcal{J}}_k + \mu_k I) \leq \bar{\lambda}$ for all k. We prove (5.14) by contradiction. Assume $\limsup_{k\to+\infty} \|\tilde{g}_k\| = \varepsilon > 0$, and define the index set

$$\mathcal{I} = \bigcup_{k=1}^{\infty} \mathcal{I}_k := \{j \in \mathbb{N} : j \leq k, \|\tilde{g}_j\| \geq \varepsilon/2\}. \tag{5.16}$$

Then, we know that $|\mathcal{I}| = \infty$, where $|\mathcal{I}|$ denotes the number of the elements of $\mathcal{I}$. Moreover, for all $j \in \mathcal{I}$, we know that

$$l_j \geq \varepsilon/2\bar{\lambda} \quad \text{and} \quad t^j_{\max} \geq \bar{t} = \min\left\{\frac{(1-\nu)\varepsilon}{M\bar{\lambda}\bar{d}^2}, 1\right\}. \tag{5.17}$$

Thus, $\bar{t}$ is a uniform lower bound for the step size t at U^j for $j \in \mathcal{I}$, i.e., $t_j \geq t$ for all $j \in \mathcal{I}$, and we have

$$\tilde{E}(U^0) - \tilde{E}(U^{k+1}) = \sum_{j=0}^{k}(\tilde{E}(U^j) - \tilde{E}(U^{j+1})) \geq \sum_{j\in\mathcal{I}_k}(\tilde{E}(U^j) - \tilde{E}(U^{j+1})) \tag{5.18}$$

$$\geq \sum_{j\in\mathcal{I}_k} -\nu t_j \langle \tilde{g}_j, d_j\rangle \geq \frac{\nu\bar{t}\varepsilon}{2\bar{\lambda}}|\mathcal{I}_k|. \tag{5.19}$$

Let $k \to \infty$ in (5.18); then we know that $\tilde{E}(U^0) - \bar{E} \geq +\infty$, which leads to a contradiction. ∎

5.3. Hybrid acceleration framework. Many gradient-based methods have a good convergent performance at the beginning but often show slow tail convergence near the stationary states. In this case, the Newton-like method is a natural choice and has a better convergence speed when the iteration is near the stationary states. It is noted that the Hessian-based method is sensitive to the initial point. A key step of mixing two methods is designing a proper criterion to determine when to launch the Hessian-based method. It is difficult to develop a perfect strategy for all kinds of PFC models. In our experiments, we switch to the Newton-PCG algorithm when one of the following criteria is met:

$$|E(\hat{\Phi}^k) - E(\hat{\Phi}^{k-1})| < \varepsilon_1 \quad \text{or} \quad \|g_k - g_{k-1}\| < \varepsilon_2, \tag{5.20}$$

where $\varepsilon_1, \varepsilon_2 > 0$. Our proposed hybrid accelerated framework is summarized in Algorithm 5.3. The M method stands for a certain existing method, such as our AA-BPG method.

Algorithm 5.3. Hybrid acceleration framework (N-M method).

Require: $\Phi^0, \varepsilon_1, \varepsilon_2$, and $k = 0$.
1: **while** stop criterion is not satisfied **do**
2: **if** switching condition is satisfied **then**
3: Perform Newton-PCG method (Algorithm 5.2);
4: **else**
5: Perform M method;
6: **end if**
7: $k = k + 1$;
8: **end while**

Remark 5.1. The idea of a hybrid method provides a general framework for local acceleration. Our Newton-PCG methods cannot only combine with the AA-BPG methods but also with many existing methods. It is worth noting that directly using the Newton-PCG method may converge to a bad stationary point or lead to slow convergence since the initial point is not good.

6. Numerical results. In this section, we present several numerical examples for our proposed methods and compare the efficiency and accuracy with existing methods. Our approaches contain AA-BPG-2 and AA-BPG-4 (see Algorithm 4.1) and a hybrid method (see Algorithm 5.2), and the comparison methods [45, 33, 46, 32] include the first-order temporal accuracy semi-implicit scheme (SIS), the first-order temporal accuracy stabilized semi-implicit scheme (SSIS1), the second-order temporal accuracy stabilized semi-implicit scheme (SSIS2), the invariant energy quadrature scheme (IEQ) and scalar auxiliary variable (SAV) approaches. All methods are employed to calculate the stationary states of finite-dimensional PFC models, including the LB model for periodic crystals and the LP model for quasicrystals. Note that these methods all guarantee mass conservation. The step sizes α_k in our approaches are obtained adaptively by the linear search technique, while the fixed step sizes α of others are chosen to guarantee the best performance on the premise of energy dissipation. In efficient implementation of the Newton-PCG method, the parameters in (5.6) and (5.7) are set with $\tau = 0.01$, $\delta_k = 0.7 \max \Lambda_k^{(')}$, and μ_k is chosen as [43]. To show the energy tendency obviously, we calculate a reference energy E_s by choosing the invariant energy value as the grid size converges to 0. From our numerical tests, the reference energy has 14 significant decimal digits. All experiments were performed on a workstation with a 3.20 GHz CPU (i7-8700, 12 processors). All code was written in MATLAB language without parallel implementation.

6.1. AA-BPG method.

6.1.1. Periodic crystals. For the LB model, we use three-dimensional periodic crystals of the double gyroid and the sigma phase, as shown in Figure 2, to demonstrate the performance of our approach. In the hybrid method of Algorithm 5.2, we choose the gradient difference $\|g_k - g_{k-1}\| < 10^{-3}$ as the measurement to launch the Newton-PCG algorithm.

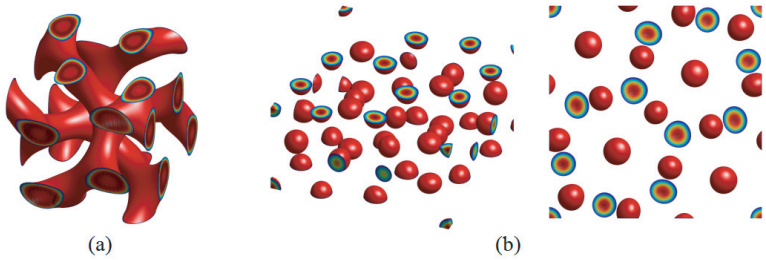

FIG. 2. *The stationary periodic crystals in the LB model.* (a) *Double gyroid phase with* $\xi = 0.1$, $\tau = -2.0$, $\gamma = 2.0$. (b) *Sigma phase with* $\xi = 0.1$, $\tau = 0.01$, $\gamma = -2.0$ *from two perspectives.*

Double gyroid. The double gyroid phase is a continuous network periodic phase whose initial values can be chosen as

$$\phi(\boldsymbol{r}) = \sum_{\boldsymbol{h}\in\Lambda_0^{DG}} \hat{\phi}(\boldsymbol{h}) e^{i(\mathbf{B}\boldsymbol{h})^\top \cdot \boldsymbol{r}}, \tag{6.1}$$

where initial lattice points set $\Lambda_0^{DG} \subset \mathbb{Z}^3$ only on which of the Fourier coefficients located are nonzero. The corresponding Λ_0^{DG} of the double gyroid phase can be found in Table 1 in [19]. The double gyroid structure belongs to the cubic crystal system; therefore, the 3-order invertible matrix can be chosen as $\mathbf{B} = (1/\sqrt{6})\mathbf{I}_3$. Correspondingly, the computational domain in physical space is $\Omega = [0, 2\sqrt{6}\pi)^3$. The parameters in LB model (2.1) are set as $\xi = 0.1$, $\tau = -2.0$, and $\gamma = 2.0$. 128^3 wavefunctions are used in these simulations. Figure 2(a) shows the stationary solution of the double gyroid profile.

Figure 3 gives the iteration process of the above-mentioned approaches, including the relative energy difference and the gradient changes with iterations, and the CPU time cost. The reference energy value $E_s = -12.94291551898271$ is the finally convergent value. As is evident from these results, our AA-BPG methods are most efficient among these approaches under the premise of ensuring energy dissipation. The AA-BPG-4 and AA-BPG-2 approaches have nearly the same numerical behaviors; however, the AA-BPG-4 method spends a little more CPU time than the AA-BPG-2 scheme does. The reason is attributed to the cost of solving the subproblem (3.5) at each step. For AA-BPG-2 scheme, (3.5) can be solved analytically, while for the AA-BPG-4 method, (3.5) is required to numerically solve a nonlinear system. In Figure 4, we give the step sizes of the AA-BPG-2/4 scheme.

The SIS, SAV, and IEQ approaches have almost the same iterations. Theoretically, the convergence of SIS is based on the assumption of global Lipschitz constants, while the SAV method always has a modified energy dissipation through adding an arbitrary scalar auxiliary parameter C which guarantees the boundedness of the bulk energy term. The original energy dissipation property of the SAV method depends on the selection of C. For computing the double gyroid phase, we find that when C is smaller than 10^6, the SAV scheme cannot keep the original energy dissipation property even if we adopt a small step size 0.001. Further increasing C to 10^8, we can use a large step size $\alpha = 0.2$ to obtain the original energy dissipation feature. Note that there exists a gap between the modified energy and the original energy no matter what the auxiliary parameters are. Like in the SAV method, similar results and phenomena have been also found in the IEQ approach. Among the three methods, SIS spends the fewest CPU times. The reason is that the SAV and IEQ methods

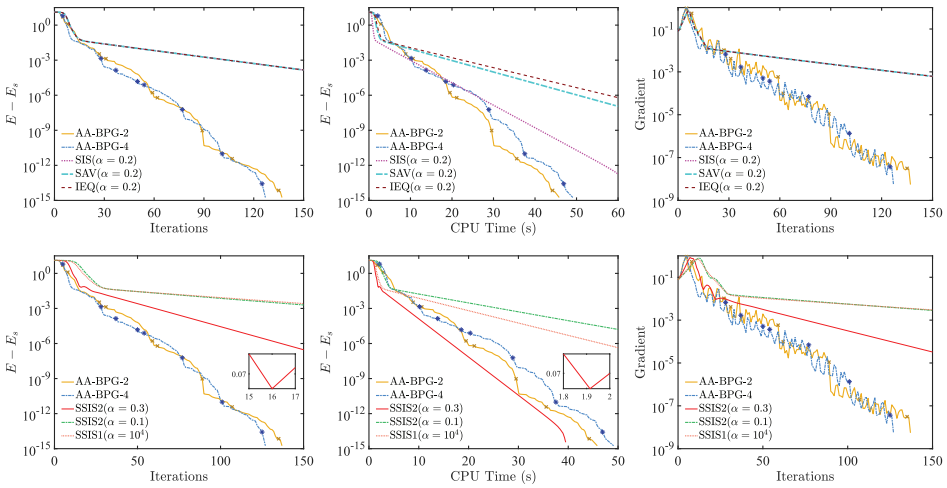

FIG. 3. *Double gyroid phase: comparisons of numerical behaviors of the AA-BPG-2/4 approaches with the following: First row: SIS, SAV, and IEQ. Second row: SSIS1 and SSIS2. Left column: Relative energy over iterations. Middle column: Relative energy over CPU times. Right column: Gradient over iterations. The blue and yellow ×s mark where restarts occurred.*

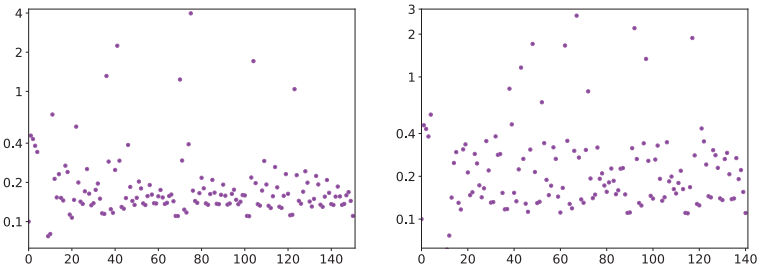

FIG. 4. *Double gyroid phase: the step sizes of the following: Left: AA-BPG-2. Right: the AA-BPG-4 approach.*

require one to solve a subsystem at each step while SIS does not.

SSIS1 is an unconditionally stable scheme through imposing a stabilized term on SIS. Its energy law holds under the assumption of the stabilizing parameter being greater than half of the global Lipschitz constant. The step size α can be arbitrarily large, while the effective step size has a limit. From the numerical results, SSIS1 with $\alpha = 10^4$ shows a slower convergent rate than SIS with $\alpha = 0.2$ does. An interesting scheme is the conditionally stable SSIS2 proposed in [33] that introduces a center difference stabilizing term to guarantee the second-order temporal accuracy. From the point of continuity, SSIS2 actually adds an inertia term onto the original gradient flow system. The inertia term can accelerate the convergent speed but often accompanied with some oscillations if the step size is large. As Figure 3 shows, when $\alpha = 0.1$, SSIS2 has almost the same convergent speed as SSIS1 and holds the energy dissipation property. If increasing α, such as 0.3, SSIS2 obtains an accelerated speed but with oscillations.

Sigma phase. The second periodic structure considered here is the sigma phase, which is a spherical packed phase recently discovered in a block copolymer experiment [23], and in the self-consistent mean-field simulation [44]. The sigma phase has a larger, much more complicated tetragonal unit cell with 30 atoms. For such a pattern, we implement our algorithm on bounded computational domain $\Omega = [0, 27.7884) \times [0, 27.7884) \times [0, 14.1514)$. Correspondingly, the initial values can be found in [44]. When computing the sigma phase, the parameters are set as $\xi = 1.0$, $\tau = 0.01$, and $\gamma = 2.0$, and $256 \times 256 \times 128$ wavefunctions are used to discretize the LB energy functional. The stationary morphology is shown in Figure 2(b). As far as we know, it is the first time such a complicated sigma phase in such a simple PFC model has been found.

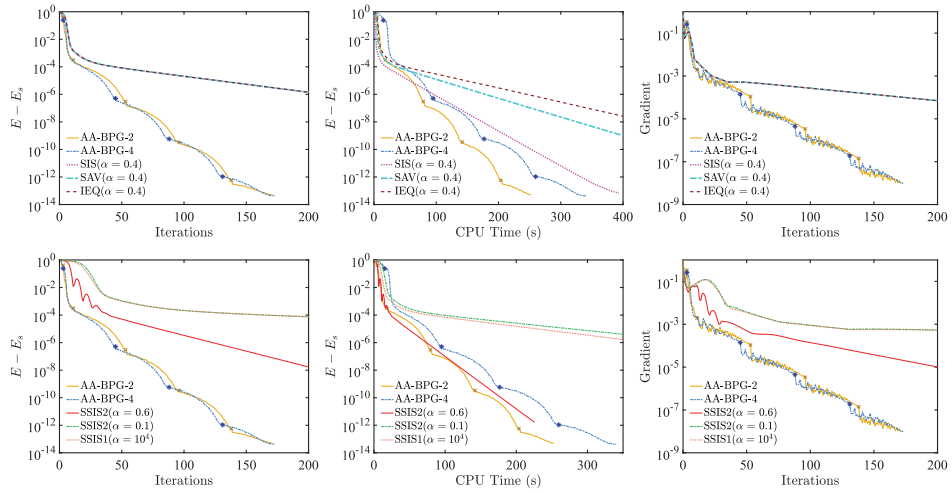

FIG. 5. *Sigma phase: Comparisons of numerical behaviors of the AA-BPG-2/4 approaches with other numerical methods. The information in these plots is the same as in Figure 3.*

Figure 5 compares our proposed methods with other numerical schemes. We still use the reference energy value $E_s = -0.93081648457086$ as the baseline to observe the relative energy changes of various numerical approaches. Again, as shown in these results, on the premise of energy dissipation, the new developed gradient-based approaches demonstrate a better performance over the existing methods in computing the sigma phase. Among these methods, the AA-BPG-2 method is the most efficient.

6.1.2. Quasicrystals. For the LP free energy (2.2), we take the two-dimensional dodecagonal quasicrystal as an example of examining the performance of our proposed approach. For dodecagonal quasicrystals, two length scales q_1 and q_2 are equal to 1 and $2\cos(\pi/12)$, respectively. Two-dimensional dodecagonal quasicrystals can be embedded into four-dimensional periodic structures; therefore, the projection method is carried out in four-dimensional space. The 4-order invertible matrix $\mathbf{B}$ associated with four-dimensional periodic structure is chosen as $\mathbf{I}_4$. The corresponding computational domain in real space is $[0, 2\pi)^4$. The projection matrix $\mathcal{P}$ in (2.6) of the dodecagonal quasicrystals is

$$\mathcal{P} = \begin{pmatrix} 1 & \cos(\pi/6) & \cos(\pi/3) & 0 \\ 0 & \sin(\pi/6) & \sin(\pi/3) & 1 \end{pmatrix}. \tag{6.2}$$

The initial solution is

$$\phi(\boldsymbol{r}) = \sum_{\boldsymbol{h}\in\Lambda_0^{QC}} \hat{\phi}(\boldsymbol{h}) e^{i[(\mathcal{P}\cdot\mathbf{B}\boldsymbol{h})^\top\cdot\boldsymbol{r}]}, \quad \boldsymbol{r}\in\mathbb{R}^2, \tag{6.3}$$

where initial lattice points set $\Lambda_0^{QC}\subset\mathbb{Z}^4$ can be found in Table 3 in [20] on which of the Fourier coefficients $\hat{\phi}(\boldsymbol{h})$ located are nonzero.

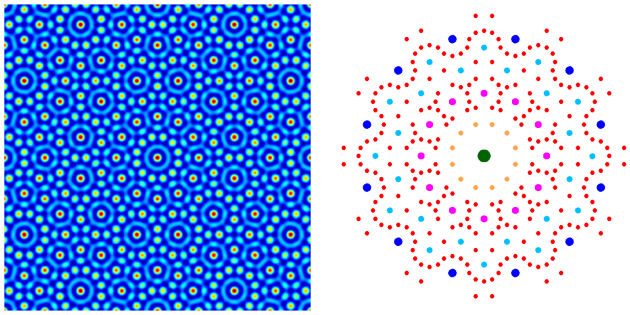

FIG. 6. *The stationary dodecagonal quasicrystal phase in the LP model with* $c = 24$, $\varepsilon = -6$, $\kappa = 6$. *Left: physical morphology. Right: Fourier spectral points whose coefficient intensity is larger than* 0.001.

The parameters in LP models are set as $c = 24$, $\varepsilon = -6$, and $\kappa = 6$, and 38^4 wavefunctions are used to discretize the LP energy functional. The convergent stationary quasicrystal is given in Figure 6, including its order parameter distribution and Fourier spectrum. The numerical behavior of different approaches can be found in Figure 7. To better observe the change tendency, we use the convergent energy value $E_s = -15.97486323815640$ as a baseline to show the relative energy changes versus iterations. We find again that our proposed approaches are more efficient than others.

6.2. Local acceleration. The motivation of the hybrid method is providing a framework to locally accelerate the existing methods. Certainly, the Newton-PCG method is suitable for all alternative methods mentioned above. In Figure 8, we give a detailed comparison of our Newton-PCG method applied to alternative methods. For method M, the acceleration ratio is defined as

$$\text{Acceleration ratio} := \frac{\text{CPU times of original method M}}{\text{CPU times of hybrid method N-M}}. \tag{6.4}$$

All numerical parameters, such as step size, of all alternative approaches are kept the same as the former to guarantee the best performance. To launch the Newton-PCG method, we choose the gradient difference $\|g_k - g_{k-1}\| < 10^{-3}$ in computing the crystal and energy difference $|E(\hat{\Phi}^k) - E(\hat{\Phi}^{k-1})| < 10^{-4}$ in computing the quasicrystal as the measurement. As shown in our numerical results, our Hessian-based methods can accelerate all the existing methods with the acceleration ratio ranging from 2–14. After using the proposed local acceleration, we observe that all the compared approaches have similar performance in terms of the CPU time. Moreover, it is noted that the acceleration ratio for the AA-BPG-2 method is the smallest one, as it shows the best performance without coupling the Newton-PCG method.

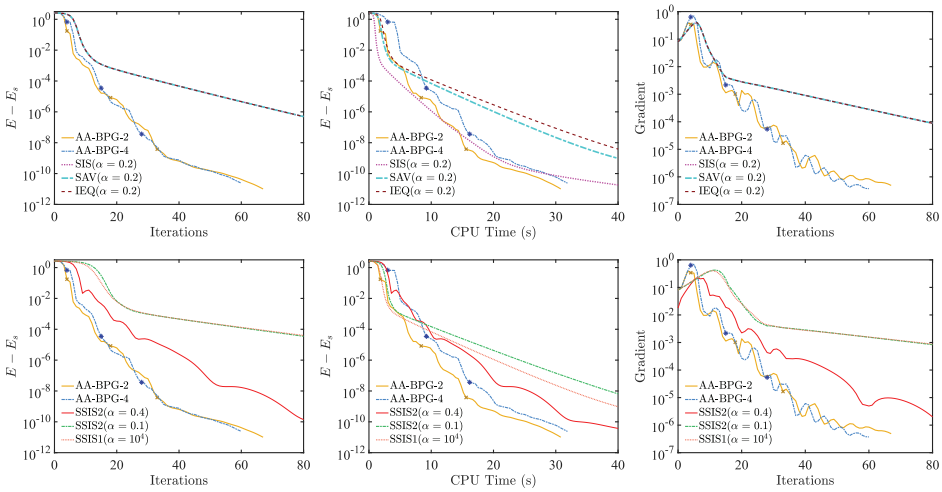

FIG. 7. *Dodecagonal quasicrystal: Comparisons of numerical behaviors of the AA-BPG-2/4 approaches with other numerical methods. The details of these images are the same as in Figure 3.*

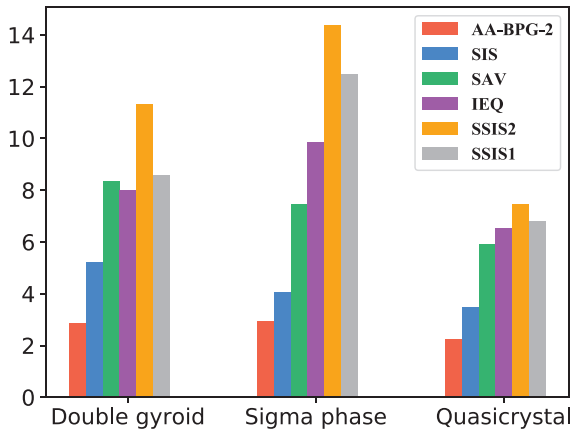

FIG. 8. *The acceleration ratio of applying the Newton-PCG algorithm to existing methods compared with original ones for computing periodic crystals and quasicrystals.*

7. Conclusion. In this paper, efficient and robust computational approaches have been proposed to find the stationary states of PFC models. Instead of formulating the energy minimization as a gradient flow, we applied the modern optimization methods directly on the discretized energy with mass conservation and energy dissipation. Moreover, the AA-BPG methods with a suitable choice of h overcome the global Lipschitz constant requirement in theoretical analysis and the step sizes are adaptively obtained by the line search technique. We also propose a practical Newton-PCG method and introduce a hybrid framework to further accelerate the local convergence of gradient-based methods. Extensive results in computing periodic crystals and quasicrystals show their advantages in terms of computation efficiency. Thus, it motivates us to continue finding the deep relationship between the gradient

flow and the optimization, applying our methods to many related problems, such as SPFC and MPFC models, and extending to more spatial discretization methods.

Appendix A. Proof of Theorem 3.6. Before we prove the convergent property, we first present a useful lemma for our analysis.

LEMMA A.1 (uniformized KL property [4]). *Let Ω be a compact set and E be a constant on Ω. Then, there exist $\epsilon > 0$, $\eta > 0$, and $\psi \in \Psi_\eta$ such that for all $\bar{u} \in \Omega$ and all $u \in \Gamma_\eta(\bar{u}, \epsilon)$, one has*

$$\psi^{'}(E(u) - E(\bar{u}))\mathrm{dist}(\mathbf{0}, \partial E(u)) \geq 1, \tag{A.1}$$

where $\Psi_\eta = \{\psi \in C[0,\eta) \cap C^1(0,\eta)$, ψ is concave, $\psi(0) = 0$, $\psi^{'} > 0$ on $(0,\eta)\}$ and $\Gamma_\eta(x,\epsilon) = \{y \| \|x - y\| \leq \epsilon,\ E(x) < E(y) < E(x) + \eta\}$.

Now, we show the proof of Theorem 3.6, which is similar to the framework in [2].

Proof. Let $S(x^0)$ be the set of limiting points of the sequence $\{x^k\}_{k=0}^{\infty}$ starting from x^0. By the boundedness of $\{x^k\}_{k=0}^{\infty}$ and the fact that $S(x^0) = \cap_{q\in\mathbb{N}}\overline{\cup_{k\geq q}\{x^k\}}$, it follows that $S(x^0)$ is a nonempty and compact set. Moreover, from (3.15), we know that $E(x)$ is a constant on $S(x^0)$, denoted by E^*. If there exists some k_0 such that $E(x^{k_0}) = E^*$, then we have $E(x^k) = E^*$ for all $k \geq k_0$, which is from (3.15). In the following proof, we assume that $E(x^k) > E^*$ for all k. Therefore, for all $\epsilon, \eta > 0$, there exists some $\ell > 0$ such that for all $k > \ell$, we have $\mathrm{dist}(S(x^0), x^k) \leq \epsilon$ and $E^* < E(x^k) < E^* + \eta$, i.e.,

$$x \in \Gamma_\eta(x^*, \epsilon) \quad \forall x^* \in S(x^0). \tag{A.2}$$

Applying Lemma A.1 for all $k > \ell$, we have

$$\psi^{'}(E(x^k) - E^*)\mathrm{dist}(\mathbf{0}, E(x^k)) \geq 1.$$

From (3.17), it implies that

$$\psi^{'}(E(x^k) - E^*) \geq \frac{1}{c_1(\|x^k - x^{k-1}\| + \bar{w}\|x^{k-1} - x^{k-2}\|)}. \tag{A.3}$$

By the convexity of ψ, we have

$$\psi(E(x^k) - E^*) - \psi(E(x^{k+1}) - E^*) \geq \psi^{'}(E(x^k) - E^*)(E(x^k) - E(x^{k+1})). \tag{A.4}$$

Define $\Delta_{p,q} = \psi(E(x^p) - E^*) - \psi(E(x^q) - E^*)$ and $C = (1 + \bar{w})c_1/c_0 > 0$. Together with (A.3), (A.4), and (3.15), we have for all $k > \ell$

$$\Delta_{k,k+1} \geq \frac{c_0\|x^{k+1} - x^k\|^2}{c_1(\|x^k - x^{k-1}\| + \bar{w}\|x^{k-1} - x^{k-2}\|)} \geq \frac{\|x^{k+1} - x^k\|^2}{C(\|x^k - x^{k-1}\| + \|x^{k-1} - x^{k-2}\|)}. \tag{A.5}$$

Therefore,

$$2\|x^{k+1} - x^k\| \leq \frac{1}{2}(\|x^k - x^{k-1}\| + \|x^{k-1} - x^{k-2}\|) + 2C\Delta_{k,k+1}, \tag{A.6}$$

which is from the geometric inequality. For any $k > \ell$, summing up (A.6) for $i = \ell + 1, \ldots, k$, it implies that

$$2\sum_{i=\ell+1}^{k}\|x^{i+1} - x^i\| \leq \frac{1}{2}\sum_{i=\ell+1}^{k}(\|x^i - x^{i-1}\| + \|x^{i-1} - x^{i-2}\|) + 2C\sum_{i=\ell+1}^{k}\Delta_{i,i+1}$$

$$\leq \sum_{i=\ell+1}^{k}\|x^{i+1} - x^i\| + \|x^{\ell+1} - x^\ell\| + \|x^\ell - x^{\ell-1}\| + 2C\Delta_{\ell+1,k+1},$$

where the last inequality is from the fact that $\Delta_{p,q} + \Delta_{q,r} = \Delta_{p,r}$ for all $p, q, r \in \mathbb{N}$. For any $k \geq \ell$, since $\psi \geq 0$, we have

$$\sum_{i=\ell+1}^{k} \|x^{i+1} - x^i\| \leq \|x^{\ell+1} - x^\ell\| + \|x^\ell - x^{\ell-1}\| + 2C\psi(E(x^{\ell+1}) - E^*). \tag{A.7}$$

This easily implies that $\sum_{k=1}^{\infty} \|x^{k+1} - x^k\| < \infty$. Together with Theorem 3.6, we obtain

$$\lim_{k\to+\infty} x^k = x^*, \quad \mathbf{0} \in \partial E(x^*) = 0. \qquad \square$$

Appendix B. Proof of Lemma 5.1. The proof is similar to the framework in [47]. Let x^* be the exact solution, and let $e_i = x^* - x^i$ for all i. We first prove some important properties of Algorithm 5.1.

Property I: $r_i = Ax^i - b$. From step 4 of Algorithm 5.1, we have $\alpha_i A p_{i-1} = Ax^i - Ax^{i-1}$. Then,

$$\begin{aligned} r_i &= r_{i-1} + \alpha_i A p_{i-1} = r_0 + \sum_{j=1}^{i} \alpha_j A p_{j-1} = -b + \sum_{j=1}^{i} \alpha_j A p_{j-1} \\ &= -b + \sum_{j=1}^{i} (Ax^j - Ax^{j-1}) = -b + Ax^i - Ax^0 = Ax^i - b. \end{aligned}$$

Property II: $\langle p_i, b\rangle = \|r_i\|^2_{M^{-1}}$ $(i = 0, 1, 2, \dots)$. By formula (5.40) in [29], we know that $\langle r_i, r_j\rangle_{M^{-1}} = 0$ $(i \neq j)$. Together with the definition of β_i and p_i in Algorithm 5.1, we get

$$\begin{aligned} &\langle p_0, b\rangle = \langle p_0, -r_0\rangle = \langle M^{-1} r_0, r_0\rangle = \|r_0\|^2_{M^{-1}}, \\ &\langle p_i, b\rangle = \langle p_i, -r_0\rangle = \langle M^{-1} r_i, r_0\rangle + \beta_i \langle p_{i-1}, -r_0\rangle = \beta_i \langle p_{i-1}, -r_0\rangle = \left(\prod_{j=1}^{i} \beta_i\right) \langle p_0, -r_0\rangle \\ &\quad = \left(\prod_{j=1}^{i} \beta_i\right) \|r_0\|^2_{M^{-1}} = \left(\prod_{j=2}^{i} \beta_i\right) \|r_1\|^2_{M^{-1}} = \|r_i\|^2_{M^{-1}} \quad \forall i = 1, 2, \dots. \end{aligned} \tag{B.1}$$

Property III: $\|e_i\|_A \geq \|e_{i+1}\|_A$. According to the iteration of p_i, one has

$$\begin{aligned} \langle p_i, -r_{i+1}\rangle &= \langle -M^{-1} r_i + \beta_i p_{i-1}, -r_{i+1}\rangle = 0 + \beta_i \langle p_{i-1}, -r_{i+1}\rangle \\ &= \left(\prod_{j=1}^{i} \beta_j\right) \langle p_0, -r_{i+1}\rangle = \left(\prod_{j=1}^{i} \beta_j\right) \langle M^{-1} r_0, r_{i+1}\rangle = 0. \end{aligned} \tag{B.2}$$

By Property I, we have $Ae_{i+1} = A(x^* - x^{i+1}) = b - Ax^{i+1} = -r_{i+1}$, which implies that $\langle p_i, Ae_{i+1}\rangle = 0$. Using the fact that $e_i = e_{i+1} + x^{i+1} - x^i = e_{i+1} + \alpha_{i+1} p_i$, the following equation holds for all $i \geq 0$:

$$\begin{aligned} \|e_i\|^2_A &= \|e_{i+1} + \alpha_{i+1} p_i\|^2_A = \|e_{i+1}\|^2_A + 2\alpha_{i+1} \langle p_i, Ae_{i+1}\rangle + \|\alpha_{i+1} p_i\|^2_A \\ &= \|e_{i+1}\|^2_A + \alpha^2_{i+1} \|p_i\|^2_A \geq \|e_{i+1}\|^2_A. \end{aligned} \tag{B.3}$$

Property IV: $\langle \boldsymbol{x^i, b}\rangle \geq \langle \boldsymbol{x^{i-1}, b}\rangle$. The definition of α_j gives

$$\|r_{j-1}\|^2_{M^{-1}} = \alpha_j \|p_{j-1}\|^2_A.$$

Together with (B.1) and (B.3), we have

$$\begin{aligned}\langle x^i, b\rangle &= \langle x^{i-1}, b\rangle + \langle \alpha_i p_{i-1}, b\rangle = \langle x^0, b\rangle + \sum_{j=1}^{i} \langle \alpha_j p_{j-1}, b\rangle = \sum_{j=1}^{i} \alpha_j \|r_{j-1}\|^2_{M^{-1}} \\ &= \sum_{j=1}^{i} \alpha_j^2 \|p_{j-1}\|^2_A = \sum_{j=1}^{i} (\|e_{j-1}\|^2_A - \|e_j\|^2_A) = \|e_0\|^2_A - \|e_i\|^2_A,\end{aligned} \tag{B.4}$$

which implies that $\langle x^i, b\rangle \geq \langle x^{i-1}, b\rangle$ by the monotonicity of $\|e_i\|^2_A$.

Now, we can prove the main result. By using the definition of p_0 and α_1, we obtain

$$\begin{aligned}\frac{\langle x^i, b\rangle}{\|b\|^2} &\geq \frac{\langle x^1, b\rangle}{\|b\|^2} = \frac{\langle x^0 + \alpha_1 p_0, b\rangle}{\|b\|^2} = \alpha_1 \frac{\langle p_0, b\rangle}{\|b\|^2} = \frac{\langle r_0, p_0\rangle}{\langle p_0, Ap_0\rangle} \frac{\langle M^{-1}b, b\rangle}{\|b\|^2} \\ &= \frac{\langle Mp_0, p_0\rangle}{\langle p_0, Ap_0\rangle} \frac{\langle M^{-1}b, b\rangle}{\|b\|^2} \geq \frac{\langle Mp_0, p_0\rangle}{\langle p_0, Ap_0\rangle} \frac{1}{\lambda_{\max}(M)}.\end{aligned} \tag{B.5}$$

Since M is positive, we know that $M = M^{1/2}M^{1/2}$, where $M^{1/2}$ is still positive. As a result,

$$\|M\| = \lambda_{max}(M) = \lambda_{\max}(M^{1/2}M^{1/2}) = \lambda^2_{\max}(M^{1/2}) = \|M^{1/2}\|^2. \tag{B.6}$$

Let $y = M^{1/2}p_0$; then we get

$$\begin{aligned}\frac{\langle Mp_0, p_0\rangle}{\langle p_0, Ap_0\rangle} &= \frac{\langle y, y\rangle}{\langle y, M^{-1/2}AM^{-1/2}y\rangle} \geq \frac{1}{\lambda_{\max}(M^{-1/2}AM^{-1/2})} = \frac{1}{\|M^{-1/2}AM^{-1/2}\|} \\ &\geq \frac{1}{\|M^{-1/2}\| \cdot \|A\| \cdot \|M^{-1/2}\|} = \frac{\|M\|}{\|A\|} = \frac{\lambda_{\max}(M)}{\lambda_{\max}(A)},\end{aligned} \tag{B.7}$$

where the second inequality takes the fact that $\|AB\| \leq \|A\| \cdot \|B\|$. Together with (B.5), we get

$$\frac{\langle x^i, b\rangle}{\|b\|^2} \geq \frac{\langle Mp_0, p_0\rangle}{\langle p_0, Ap_0\rangle} \frac{1}{\lambda_{\max}(M)} \geq \frac{\lambda_{\max}(M)}{\lambda_{\max}(A)} \frac{1}{\lambda_{\max}(M)} = \frac{1}{\lambda_{\max}(A)}. \tag{B.8}$$

To verify another inequality, we use (B.4) and the fact that $e_0 = x^* - x^0 = -A^{-1}b$:

$$\frac{\langle x^i, b\rangle}{\|b\|^2} = \frac{\|e_0\|^2_A - \|e_i\|^2_A}{\|b\|^2} \leq \frac{\|e_0\|^2_A}{\|b\|^2} = \frac{\|A^{-1}b\|^2_A}{\|b\|^2} = \frac{\langle b, A^{-1}b\rangle}{\|b\|^2} \leq \frac{1}{\lambda_{\min}(A)}.$$

Acknowledgments. We would like to thank the anonymous reviewers and the associate editor for their valuable comments and suggestions, which substantially improved this paper. In addition, C. Bao would like to thank Lili Ju for fruitful discussions.

REFERENCES

[1] J. Barzilai and J. M. Borwein, *Two-point step size gradient methods*, IMA J. Numer. Anal., 8 (1988), pp. 141–148.

[2] H. H. Bauschke, J. Bolte, and M. Teboulle, *A descent lemma beyond Lipschitz gradient continuity: First-order methods revisited and applications*, Math. Oper. Res., 42 (2017), pp. 330–348.

[3] A. Beck and M. Teboulle, *A fast iterative shrinkage-thresholding algorithm for linear inverse problems*, SIAM J. Imaging Sci., 2 (2009), pp. 183–202, https://doi.org/10.1137/080716542.

[4] J. Bolte, S. Sabach, and M. Teboulle, *Proximal alternating linearized minimization for nonconvex and nonsmooth problems*, Math. Program., 146 (2014), pp. 459–494.

[5] J. Bolte, S. Sabach, M. Teboulle, and Y. Vaisbourd, *First order methods beyond convexity and Lipschitz gradient continuity with applications to quadratic inverse problems*, SIAM J. Optim., 28 (2018), pp. 2131–2151, https://doi.org/10.1137/17M1138558.

[6] S. Brazovskii, *Phase transition of an isotropic system to a nonuniform state*, J. Exp. Theor. Phys., 41 (1975), pp. 85–89.

[7] L. M. Bregman, *The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming*, USSR Comput. Math. Math. Phys., 7 (1967), pp. 200–217.

[8] H. Brezis, *Functional Analysis, Sobolev Spaces and Partial Differential Equations*, Springer, New York, 2011.

[9] L. Chen, *Phase-field models for microstructure evolution*, Annu. Rev. Mater. Res., 32 (2002), pp. 113–140.

[10] L. Q. Chen and J. Shen, *Applications of semi-implicit Fourier-spectral method to phase field equations*, Comput. Phys. Commun., 108 (1998), pp. 147–158.

[11] K. Cheng, C. Wang, and S. M. Wise, *An energy stable BDF2 Fourier pseudo-spectral numerical scheme for the square phase field crystal equation*, Commun. Comput. Phys., 26 (2019), pp. 1335–1364.

[12] Q. Du and J. Zhang, *Adaptive finite element method for a phase field bending elasticity model of vesicle membrane deformations*, SIAM J. Sci. Comput., 30 (2008), pp. 1634–1657, https://doi.org/10.1137/060656449.

[13] Q. Du and W.-x. Zhu, *Stability analysis and application of the exponential time differencing schemes*, J. Comput. Math., 22 (2004), pp. 200–209.

[14] X. Feng, Y. He, and C. Liu, *Analysis of finite element approximations of a phase field model for two-phase fluids*, Math. Comp., 76 (2007), pp. 539–571.

[15] R. Guo and Y. Xu, *A high order adaptive time-stepping strategy and local discontinuous Galerkin method for the modified phase field crystal equation*, Commun. Comput. Phys., 24 (2018), pp. 123–151.

[16] H. Hiller, *The crystallographic restriction in higher dimensions*, Acta Crystallogr. Sect. A, 41 (1985), pp. 541–544.

[17] Z. Hu, S. M. Wise, C. Wang, and J. S. Lowengrub, *Stable and efficient finite-difference nonlinear-multigrid schemes for the phase field crystal equation*, J. Comput. Phys., 228 (2009), pp. 5323–5339.

[18] K. Jiang, J. Tong, P. Zhang, and A.-C. Shi, *Stability of two-dimensional soft quasicrystals in systems with two length scales*, Phys. Rev. E (3), 92 (2015), 042159.

[19] K. Jiang, C. Wang, Y. Huang, and P. Zhang, *Discovery of new metastable patterns in diblock copolymers*, Commun. Comput. Phys., 14 (2013), pp. 443–460.

[20] K. Jiang and P. Zhang, *Numerical methods for quasicrystals*, J. Comput. Phys., 256 (2014), pp. 428–440.

[21] Y. Katznelson, *An Introduction to Harmonic Analysis*, Cambridge University Press, Cambridge, UK, 2004.

[22] H. G. Lee, J. Shin, and J.-Y. Lee, *First- and second-order energy stable methods for the modified phase field crystal equation*, Comput. Methods Appl. Mech. Engrg., 321 (2017), pp. 1–17.

[23] S. Lee, M. Bluemle, and F. Bates, *Discovery of a Frank-Kasper σ phase in sphere-forming block copolymer melts*, Science, 330 (2010), pp. 349–353.

[24] Q. Li, Z. Zhu, G. Tang, and M. B. Wakin, *Provable Bregman-Divergence Based Methods for Nonconvex and Non-Lipschitz Problems*, preprint, https://arxiv.org/abs/1904.09712, 2019.

[25] R. Lifshitz and H. Diamant, *Soft quasicrystals—why are they stable?*, Philos. Mag., 87 (2007), pp. 3021–3030.

[26] R. Lifshitz and D. Petrich, *Theoretical model for Faraday waves with multiple-frequency*

forcing, Phys. Rev. Lett., 79 (1997), pp. 1261–1264.
[27] X. Liu, Z. Wen, X. Wang, M. Ulbrich, and Y. Yuan, *On the analysis of the discretized Kohn–Sham density functional theory*, SIAM J. Numer. Anal., 53 (2015), pp. 1758–1785, https://doi.org/10.1137/140957962.
[28] S. K. Mkhonta, K. R. Elder, and Z.-F. Huang, *Exploring the complex world of two-dimensional ordering with three modes*, Phys. Rev. Lett., 111 (2013), 035501, https://doi.org/10.1103/PhysRevLett.111.035501.
[29] J. Nocedal and S. J. Wright, *Numerical Optimization*, Springer, New York, 2006.
[30] B. O'donoghue and E. Candes, *Adaptive restart for accelerated gradient schemes*, Found. Comput. Math., 15 (2015), pp. 715–732.
[31] N. Provatas and K. Elder, *Phase-Field Methods in Materials Science and Engineering*, Wiley-VCH, Weinheim, Germany, 2010.
[32] J. Shen, J. Xu, and J. Yang, *A new class of efficient and robust energy stable schemes for gradient flows*, SIAM Rev., 61 (2019), pp. 474–506, https://doi.org/10.1137/17M1150153.
[33] J. Shen and X. Yang, *Numerical approximations of Allen-Cahn and Cahn-Hilliard equations*, Discrete Contin. Dyn. Syst., 28 (2010), pp. 1669–1691.
[34] A.-C. Shi, J. Noolandi, and R. C. Desai, *Theory of anisotropic fluctuations in ordered block copolymer phases*, Macromolecules, 29 (1996), pp. 6487–6504.
[35] J. Shin, H. G. Lee, and J.-Y. Lee, *First and second order numerical methods based on a new convex splitting for phase-field crystal equation*, J. Comput. Phys., 327 (2016), pp. 519–542.
[36] J. Swift and P. C. Hohenberg, *Hydrodynamic fluctuations at the convective instability*, Phys. Rev. A, 15 (1977), pp. 319–328, https://doi.org/10.1103/PhysRevA.15.319.
[37] P. Tseng, *On Accelerated Proximal Gradient Methods for Convex-Concave Optimization*, manuscript, 2008.
[38] M. Ulbrich, Z. Wen, C. Yang, D. Klöckner, and Z. Lu, *A proximal gradient method for ensemble density functional theory*, SIAM J. Sci. Comput., 37 (2015), pp. A1975–A2002, https://doi.org/10.1137/14098973X.
[39] J. Van Tiel, *Convex Analysis: An Introductory Text*, manuscript, 1984.
[40] C. Wang and S. M. Wise, *An energy stable and convergent finite-difference scheme for the modified phase field crystal equation*, SIAM J. Numer. Anal., 49 (2011), pp. 945–969, https://doi.org/10.1137/090752675.
[41] S. M. Wise, C. Wang, and J. S. Lowengrub, *An energy-stable and convergent finite-difference scheme for the phase field crystal equation*, SIAM J. Numer. Anal., 47 (2009), pp. 2269–2288, https://doi.org/10.1137/080738143.
[42] X. Wu, Z. Wen, and W. Bao, *A regularized Newton method for computing ground states of Bose-Eeinstein condensates*, J. Sci. Comput., 73 (2017), pp. 303–329.
[43] X. Xiao, Y. Li, Z. Wen, and L. Zhang, *A regularized semi-smooth Newton method with projection steps for composite convex programs*, J. Sci. Comput., 76 (2018), pp. 364–389.
[44] N. Xie, W. Li, F. Qiu, and A.-C. Shi, *σ phase formed in conformationally asymmetric AB-type block copolymers*, ACS Macro Lett., 3 (2014), pp. 906–910.
[45] C. Xu and T. Tang, *Stability analysis of large time-stepping methods for epitaxial growth models*, SIAM J. Numer. Anal., 44 (2006), pp. 1759–1779, https://doi.org/10.1137/050628143.
[46] X. Yang, *Linear, first and second-order, unconditionally energy stable numerical schemes for the phase field model of homopolymer blends*, J. Comput. Phys., 327 (2016), pp. 294–316.
[47] X.-Y. Zhao, D. Sun, and K.-C. Toh, *A Newton-CG augmented Lagrangian method for semi-definite programming*, SIAM J. Optim., 20 (2010), pp. 1737–1765, https://doi.org/10.1137/080718206.