

A NOVEL ENERGY FACTORIZATION APPROACH FOR THE DIFFUSE-INTERFACE MODEL WITH PENG-ROBINSON EQUATION OF STATE

JISHENG KOU*, SHUYU SUN[†], AND XIUHUA WANG[‡]

Abstract. The Peng-Robinson equation of state (PR-EoS) has become one of the most extensively applied equations of state in chemical engineering and petroleum industry due to its excellent accuracy in predicting the thermodynamic properties of a wide variety of materials, especially hydrocarbons. Although great efforts have been made to construct efficient numerical methods for the diffuse interface models with PR-EoS, there is still not a linear numerical scheme that can be proved to preserve the original energy dissipation law. In order to pursue such a numerical scheme, we propose a novel energy factorization (EF) approach, which first factorizes an energy function into a product of several factors and then treats the factors using their properties to obtain the semi-implicit linear schemes. We apply the EF approach to deal with the Helmholtz free energy density determined by PR-EoS, and then propose a linear semi-implicit numerical scheme that inherits the original energy dissipation law. Moreover, the proposed scheme is proved to satisfy the maximum principle in both the time semi-discrete form and the cell-centered finite difference fully discrete form under certain conditions. Numerical results are presented to demonstrate the stability and efficiency of the proposed scheme.

Key words. Diffuse interface model; Peng-Robinson equation of state; Energy stability; Maximum principle.

AMS subject classifications. 65N30, 65N50, 49S05.

1. Introduction. The Peng-Robinson equation of state (PR-EoS) [26] has become one of the most popular and useful tools for describing the thermodynamic properties of fluids in both academic and industrial fields, especially chemical engineering and petroleum industry [20]. Compared to the well-known Van der Waals equation of state, PR-EoS can provide more reasonable accuracy in predicting the properties of a wide variety of materials, such as N_2 , CO_2 and hydrocarbons. PR-EoS has been extensively applied for simulation of many important problems in petroleum and chemical engineering, for instance, phase equilibria calculations [7–9, 13, 21–23, 25] and prediction of surface tension between gas and liquid [7, 10, 11, 24]. Modeling and simulation of compressible multi-component two-phase flows with partial miscibility and realistic equations of state (e.g. PR-EoS) are intensively studied in recent years [6, 11, 12, 14, 15, 17–19, 27, 28]. On the basis of the thermodynamic fundamental laws and realistic equations of state (e.g. PR-EoS), general diffuse interface models for compressible multi-component two-phase flows have been proposed in [16] for isothermal fluids and [17, 18] for non-isothermal fluids.

This paper is primarily concerned with efficient numerical methods for an isothermal diffuse interface model with PR-EoS. There exist two primary challenging problems in designing numerical schemes for such model. One is that the Helmholtz free energy density determined by PR-EoS has the complicated structures and strong non-

*School of Civil Engineering, Shaoxing University, Shaoxing 312000, Zhejiang, China; School of Mathematics and Statistics, Hubei Engineering University, Xiaogan 432000, Hubei, China. Email: jishengkou@163.com.

[†]Corresponding author. Computational Transport Phenomena Laboratory, Division of Physical Science and Engineering, King Abdullah University of Science and Technology, Thuwal 23955-6900, Kingdom of Saudi Arabia. Email: shuyu.sun@kaust.edu.sa.

[‡]School of Mathematics and Statistics, Hubei Engineering University, Xiaogan 432000, Hubei, China.

linearity. The other is that the model follows the energy dissipation law, and thus numerical schemes should be constructed to preserve this feature at the discrete level. In this paper, we will resolve the above challenging problems and will develop a novel linear, energy stable numerical scheme.

We now provide the up-to-date review regarding the approaches in the literature employed to handle the bulk Helmholtz free energy density of PR-EoS and design energy stable numerical schemes. One approach is the convex splitting method [4,5] that has been extensively used for various phase-field models [2, 29, 33]. For the diffuse-interface models with PR-EoS, the convex splitting schemes inheriting the discrete energy dissipation law have been developed in a series of recent works [6, 13, 16, 17, 27, 28]. When the convex splitting approach is applied to the PR-EoS based Helmholtz free energy density, the ideal and repulsion terms are usually treated implicitly due to their convexity, while the attraction term with the concavity is treated explicitly. The convex splitting approach can produce unconditionally energy stable numerical schemes, but it results in the nonlinear discrete equations that demand the complicated implement of efficient nonlinear iterative solvers and also cost expensively in the computations. Approximating chemical potential by a difference of Helmholtz free energy density, a fully-implicit scheme has been developed in [12]. This scheme is proved to be unconditionally energy stable, but it still suffers from the nonlinearity of the resulted equations.

The invariant energy quadratization (IEQ) approach [34–36] is a novel and efficient method developed in recent years and has been intensively applied for various phase-field models. The essential idea of IEQ is to transform the bulk free energy into a quadratic form through introducing a set of new variables. The new variables can be updated with time steps via the semi-implicit linear schemes. A numerical scheme has been developed in [19] applying the IEQ approach to PR-EoS. As a modification of the IEQ approach, the scalar auxiliary variable (SAV) approach has been proposed in [30], which introduces a scalar auxiliary variable instead of the space-dependent new variables in the IEQ approach. It leads to unconditionally stable numerical schemes, which only need to solve the linear equations with constant coefficients at each time step. Recently, in [15], the SAV approach has been applied to construct unconditionally energy stable numerical schemes for the model proposed in [16], and moreover, a component-wise SAV approach has also been developed. The numerical schemes constructed by the IEQ and SAV approaches are linear, and consequently are easy-to-implement and very efficient in computations. However, while IEQ and SAV have become the very useful and successful tools applied for a variety of phase-field models, the produced schemes use the transformed free energies that are generally never equivalent to the original energies at the discrete level. Indeed, it has been indicated in [34] that the transformed free energies have the errors of the order of time step size against the original energies. Consequently, the schemes constructed by IEQ and SAV may not inherit the original energy dissipation law in theory although the dissipation of transformed energies can be proved. In numerical tests of [15], the original energy instability has been observed despite the transformed energies decrease with time steps. To our best knowledge, for the diffuse interface model with PR-EoS, there is not a linear semi-implicit numerical scheme inheriting the original energy dissipation law so far. In this paper, we will propose such a scheme.

In this paper, we will propose a novel energy factorization (EF) approach to construct an efficient numerical scheme for the diffuse interface model with PR-EoS. The key idea of EF is that we first factorize an energy function into a product of

several factors and then handle them by the use of their properties to obtain the semi-implicit linear schemes. The EF approach will not introduce any new independent energy variable, thus it can ensure the original energy dissipation law. Applying the EF approach to the model with PR-EoS, we will obtain a linear, efficient semi-implicit numerical scheme that inherits the original energy dissipation law.

We note that molar density is the primal variable in the models with PR-EoS. Moreover, for a realistic substance, molar density shall have the physical limits under given thermodynamical conditions, and as a result, the maximum principle is essential for numerical methods to ensure that numerical solutions are physically reasonable. However, there are no results concerning the maximum principle of numerical schemes for the models with PR-EoS so far. For the first time, the proposed numerical scheme will be proved to preserve the maximum principle under certain conditions. The proof will be provided for both the semi-discrete time matching scheme and the cell-centered finite difference fully discrete scheme.

The rest of this paper is organized as follows. In Section 2, we will provide a brief description of the diffuse interface model with PR-EoS. In Section 3, we will propose the energy factorization approach to deal with the bulk Helmholtz free energy density. In Section 4, we will present the semi-discrete time scheme and prove some theoretical results including the energy stability and the maximum principle. The fully discrete schemes will be developed and analyzed in Section 5. In Section 6, numerical results will be presented to validate the proposed numerical scheme. Finally, some concluding remarks are given in Section 7.

2. Model equations. We now give a brief description for the Helmholtz free energy density of a bulk fluid (denoted by f_b) determined by Peng-Robinson equation of state [26]. We denote by c molar density of a substance. For specified temperature T , f_b is a function of molar density c and can be expressed as a sum of three contributions

$$f_b(c) = f_b^{\text{ideal}}(c) + f_b^{\text{repulsion}}(c) + f_b^{\text{attraction}}(c), \quad (2.1)$$

where

$$f_b^{\text{ideal}}(c) = c\vartheta_0 + cRT \ln(c), \quad (2.2)$$

$$f_b^{\text{repulsion}}(c) = -cRT \ln(1 - \beta c), \quad (2.3)$$

$$f_b^{\text{attraction}}(c) = \frac{\alpha(T)c}{2\sqrt{2}\beta} \ln \left(\frac{1 + (1 - \sqrt{2})\beta c}{1 + (1 + \sqrt{2})\beta c} \right). \quad (2.4)$$

Here, R is the universal gas constant and ϑ_0 is an energy parameter that relies on the temperature and thermodynamical properties of a specific substance. The substance-specific parameters α and β can be determined from the critical properties and acentric factor of a specific substance

$$\alpha = 0.45724 \frac{R^2 T_c^2}{P_c} \left[1 + m(1 - \sqrt{T_r}) \right]^2, \quad \beta = 0.07780 \frac{RT_c}{P_c}, \quad (2.5)$$

where $T_r = T/T_c$ is the reduced temperature, T_c and P_c stands for the critical temperature and critical pressure respectively, and m is calculated from the acentric factor ω as follows

$$m = 0.37464 + 1.54226\omega - 0.26992\omega^2, \quad \omega \leq 0.49,$$

$$m = 0.379642 + 1.485030\omega - 0.164423\omega^2 + 0.016666\omega^3, \quad \omega > 0.49.$$

From the physical point of view, β is the effective volume of one mole of a substance. Let $v = \frac{1}{c}$ be molar volume, i.e., the average volume occupied by one mole of a substance, which includes the effective volume β and the space between molecules. Therefore, we know that generally $\beta \ll v$ for gas and liquid. For an ideal and simple example, if the molecules can be approximated as spherical particles, then we have

$$\beta c = \frac{\beta}{v} \leq \frac{\pi}{6}.$$

This fact suggests that βc has an upper bound for a specific substance under specified temperature. As a matter of fact, from the following form of Peng-Robinson equation of state (PR-EOS) [26]

$$P = \frac{cRT}{1 - \beta c} - \frac{\alpha(T)c^2}{1 + \beta c + \beta c(1 - \beta c)} > \frac{cRT}{1 - \beta c} - \alpha(T)c^2, \quad (2.7)$$

where P is the pressure, we can directly deduce

$$\beta c < 1 - \frac{cRT}{P + \alpha(T)c^2}.$$

To justify the boundedness of βc in realistic cases, we consider the species of n-butane. We can calculate $\beta = 7.2381 \times 10^{-5} \text{ m}^3/\text{mol}$ using (2.5) and the physical data of n-butane. Molar densities of gas and liquid of n-butane at the temperature 330 K and pressure 106.39 bar are $c^G = 249.1123 \text{ mol/m}^3$ and $c^L = 9526.8428 \text{ mol/m}^3$ respectively. Then we have

$$\beta c^G = 0.0180, \quad \beta c^L = 0.6896.$$

On the basis of the above physical observation, we assume that molar density c is always bounded as below

$$0 < c_m \leq c \leq c_M, \quad \beta c_M \leq \epsilon_0 < 1, \quad (2.8)$$

where c_m , c_M and ϵ_0 are determined by a specific substance under specified pressure and temperature. We remark that ϵ_0 usually does not take some very small value from physical property as stated above.

Since the diffuse interfaces always exist between gas and liquid phases. In addition to the bulk free energy density, the gradient free energy density accounting for the effect of the interfaces is defined as

$$f_{\nabla}(c) = \frac{1}{2}\kappa|\nabla c|^2, \quad (2.9)$$

where $\kappa > 0$ is the influence parameter that can be calculated as follows

$$\kappa = \alpha\beta^{2/3} [a_0(1 - T_r) + a_1],$$

$$a_0 = -\frac{10^{-16}}{1.2326 + 1.3757\omega}, \quad a_1 = \frac{10^{-16}}{0.9051 + 1.5410\omega}.$$

We denote by $f(c)$ the general Helmholtz free energy density

$$f(c) = f_b(c) + f_\nabla(c). \quad (2.10)$$

The chemical potential is defined as the variational derivative of $f(c)$

$$\mu(c) = \frac{\delta f(c)}{\delta c} = \mu_b(c) - \kappa \Delta c, \quad (2.11)$$

where $\mu_b(c) = f'_b(c)$ is the bulk chemical potential.

Let Ω be a connected and smooth space domain. We now state the model equation as follows [28]

$$\frac{\partial c}{\partial t} - \kappa \Delta c + \mu_b(c) = \mu_e, \quad (2.12a)$$

$$\int_{\Omega} c d\mathbf{x} = c_t, \quad (2.12b)$$

$$\nabla c \cdot \mathbf{n}_{\partial\Omega} = 0, \quad \mathbf{x} \in \partial\Omega, \quad c(\mathbf{x}, 0) = c_0(\mathbf{x}), \quad \mathbf{x} \in \Omega, \quad (2.12c)$$

where $\mathbf{n}_{\partial\Omega}$ denotes the normal unit outward vector to $\partial\Omega$, $c_t > 0$ is the total moles in Ω and μ_e is a Lagrange multiplier incorporated to enforce total moles conservation. We note that μ_e is constant in space but could vary with time. Moreover, as time goes on, the spatial distribution of molar density c will approach to an equilibrium state, and μ_e will also approach the chemical potential at the equilibrium state. Here, we consider the homogeneous Neumann boundary condition, but the proposed numerical schemes and theoretical analysis can be directly extended to various boundary conditions, for instance, Dirichlet boundary conditions.

We note that the model (2.12) obeys the energy dissipation law

$$\frac{\partial}{\partial t} \int_{\Omega} f(c(\mathbf{x}, t)) d\mathbf{x} = - \int_{\Omega} \left(\frac{\partial c}{\partial t} \right)^2 d\mathbf{x}. \quad (2.13)$$

3. Energy factorization approach. In this section, we propose a novel energy factorization (EF) approach to construct a linear, energy stable time matching scheme for the model (2.12). The basic idea of EF is to factorize an energy function or a term of the energy function into a product of several factors that can be separately treated by the use of their properties. Two different factorizations are proposed to deal with the ideal term and the repulsion term of the bulk Helmholtz free energy density respectively.

At the time discrete level, we denote the time step size by τ and set $t_n = n\tau$. Furthermore, we use c^n to denote the approximation of molar density c at the time t_n .

3.1. Factorization approach for the ideal term. The first energy factorization approach is proposed to deal with the ideal term. We define the function $H(c) = c \ln(c)$, which can be factorized into the product of a linear function c and a logarithm function $\ln(c)$. Apparently, $\ln(c)$ is a concave function, and thus, we have

$$\ln(c^{n+1}) \leq \ln(c^n) + \frac{1}{c^n} (c^{n+1} - c^n). \quad (3.1)$$

For $c^n > 0$ and $c^{n+1} > 0$, using (3.1), we can deduce that

$$\begin{aligned} H(c^{n+1}) - H(c^n) &= c^{n+1} \ln(c^{n+1}) - c^n \ln(c^n) \\ &= \ln(c^n) (c^{n+1} - c^n) + c^{n+1} (\ln(c^{n+1}) - \ln(c^n)) \\ &\leq \left(\ln(c^n) + \frac{c^{n+1}}{c^n} \right) (c^{n+1} - c^n). \end{aligned} \quad (3.2)$$

Then the ideal term $f_b^{\text{ideal}}(c)$ can be estimated as

$$f_b^{\text{ideal}}(c^{n+1}) - f_b^{\text{ideal}}(c^n) \leq \vartheta_0 (c^{n+1} - c^n) + RT \left(\ln(c^n) + \frac{c^{n+1}}{c^n} \right) (c^{n+1} - c^n), \quad (3.3)$$

which suggests us to define the ideal part of chemical potential at the $(n+1)$ -th time step as

$$\mu_{\text{ideal}}^{n+1} = \vartheta_0 + RT \ln(c^n) + RT \frac{c^{n+1}}{c^n}. \quad (3.4)$$

Therefore, from (3.3), we have

$$f_b^{\text{ideal}}(c^{n+1}) - f_b^{\text{ideal}}(c^n) \leq \mu_{\text{ideal}}^{n+1} (c^{n+1} - c^n). \quad (3.5)$$

We remark that the convex splitting approach for the ideal term $f_b^{\text{ideal}}(c)$ used in [28] leads to a highly nonlinear scheme, which demands the complicated implementations of efficient nonlinear iterative solvers and also costs expensively in computations. In contrast, μ_{ideal}^{n+1} is semi-implicit and linear with respect to c^{n+1} , thus it is easy-to-implement and can be solved at less computational costs. This approach can also be directly applied for the logarithmic Flory-Huggins potential [34, 37].

3.2. Factorization approach for the repulsion term. Since the function $-\ln(1 - \beta c)$ is a convex function, we would obtain a nonlinear chemical potential when the first energy factorization approach dealing with the ideal term is employed for the repulsion term $f_b^{\text{repulsion}}(c)$. In order to pursue a linear scheme, we introduce the second energy factorization approach to deal with the repulsion term $f_b^{\text{repulsion}}(c)$.

We define the modified repulsion energy function as

$$\widehat{f}_b^{\text{repulsion}}(c) = \lambda c + \frac{1}{RT} f_b^{\text{repulsion}}(c) = \lambda c - c \ln(1 - \beta c), \quad (3.6)$$

where λ is some positive constant relying on the specific substance. Apparently, $\widehat{f}_b^{\text{repulsion}}(c)$ is positive and bounded for molar density c satisfying (2.8). We further define the following intermediate energy function

$$G(c) = \sqrt{\lambda c - c \ln(1 - \beta c)}. \quad (3.7)$$

From this, we can factorize $\widehat{f}_b^{\text{repulsion}}(c)$ into the square of $G(c)$, i.e.,

$$\widehat{f}_b^{\text{repulsion}}(c) = G(c)^2. \quad (3.8)$$

For the function $G(c)$, we have the following key lemma regarding the choice of λ .

LEMMA 3.1. Assume that molar density c satisfies (2.8). If λ is taken such that

$$\lambda \geq \frac{\epsilon_0}{(1-\epsilon_0)^2} + \left(\frac{\epsilon_0^2}{(1-\epsilon_0)^4} - 2 \ln(1-\epsilon_0) \frac{\epsilon_0}{(1-\epsilon_0)^2} \right)^{1/2}, \quad (3.9)$$

where ϵ_0 is given in (2.8), then $G(c)$ is a concave function.

Proof. We calculate the first and second derivatives of $G(c)$ as follows

$$G'(c) = \frac{1}{2}G(c)^{-1} \left(\lambda - \ln(1-\beta c) + \frac{\beta c}{1-\beta c} \right), \quad (3.10)$$

$$\begin{aligned} G''(c) &= -\frac{1}{4}G(c)^{-3} \left(\lambda - \ln(1-\beta c) + \frac{\beta c}{1-\beta c} \right)^2 \\ &\quad + \frac{1}{2}G(c)^{-1} \left(\frac{\beta}{1-\beta c} + \frac{\beta}{(1-\beta c)^2} \right) \\ &= -\frac{1}{4}G(c)^{-3} \left(\lambda^2 - 2\lambda \ln(1-\beta c) + (\ln(1-\beta c))^2 + \left(\frac{\beta c}{1-\beta c} \right)^2 \right. \\ &\quad \left. - 2\lambda \frac{\beta c}{(1-\beta c)^2} + 2 \ln(1-\beta c) \frac{\beta c}{(1-\beta c)^2} \right) \\ &\leq -\frac{1}{4}G(c)^{-3} \left(\lambda^2 - 2\lambda \frac{\beta c}{(1-\beta c)^2} + 2 \ln(1-\beta c) \frac{\beta c}{(1-\beta c)^2} \right) \\ &\leq -\frac{1}{4}G(c)^{-3} \left(\lambda^2 - 2\lambda \frac{\epsilon_0}{(1-\epsilon_0)^2} + 2 \ln(1-\epsilon_0) \frac{\epsilon_0}{(1-\epsilon_0)^2} \right). \end{aligned} \quad (3.11)$$

Applying the condition (3.9), we obtain $G''(c) \leq 0$, thus $G(c)$ is concave. \square

We can see from the proof of Lemma 3.1 that the condition (3.9) can be relaxed further substantially. Moreover, λ usually does not demand a large value in practice since ϵ_0 generally is not very small value; for instance, we calculate from (3.9) that $\lambda = 27.3656$ in numerical examples. We also note that λ is dimensionless.

The advantage of the factorization (3.8) is shown in the following lemma.

LEMMA 3.2. Assume that molar density satisfies (2.8) and λ is taken such that (3.9) holds. Then we have

$$G(c^{n+1})^2 - G(c^n)^2 \leq (G^{n+1} + G(c^n)) G'(c^n) (c^{n+1} - c^n), \quad (3.12)$$

where G^{n+1} is the linear approximation of $G(c^{n+1})$ as

$$G^{n+1} = G(c^n) + G'(c^n) (c^{n+1} - c^n). \quad (3.13)$$

Proof. The assumption implies the concavity of $G(c)$, so we have

$$G(c^{n+1}) \leq G(c^n) + G'(c^n) (c^{n+1} - c^n) = G^{n+1}. \quad (3.14)$$

Since $G(c) \geq 0$, we get $G^{n+1} \geq 0$ from (3.14) and consequently

$$G(c^{n+1})^2 \leq |G^{n+1}|^2. \quad (3.15)$$

From (3.13) and (3.15), we derive that

$$G(c^{n+1})^2 - G(c^n)^2 \leq |G^{n+1}|^2 - G(c^n)^2$$

$$\begin{aligned}
&= (G^{n+1} + G(c^n)) (G^{n+1} - G(c^n)) \\
&= (G^{n+1} + G(c^n)) G'(c^n) (c^{n+1} - c^n). \tag{3.16}
\end{aligned}$$

This ends the proof. \square

We are now ready to consider treatment of $f_b^{\text{repulsion}}(c)$ under the condition (3.9). Using Lemma 3.2, we can estimate the energy difference between two time steps as

$$\begin{aligned}
\widehat{f}_b^{\text{repulsion}}(c^{n+1}) - \widehat{f}_b^{\text{repulsion}}(c^n) &= G(c^{n+1})^2 - G(c^n)^2 \\
&\leq (G^{n+1} + G(c^n)) G'(c^n) (c^{n+1} - c^n). \tag{3.17}
\end{aligned}$$

It follows from (3.6) and (3.17) that

$$\begin{aligned}
f_b^{\text{repulsion}}(c^{n+1}) - f_b^{\text{repulsion}}(c^n) &= RT(\widehat{f}_b^{\text{repulsion}}(c^{n+1}) - \widehat{f}_b^{\text{repulsion}}(c^n)) \\
&\quad - RT\lambda (c^{n+1} - c^n) \\
&\leq RT((G^{n+1} + G(c^n)) G'(c^n) - \lambda) (c^{n+1} - c^n). \tag{3.18}
\end{aligned}$$

Thus, we can define the repulsion chemical potential at the $(n + 1)$ -th time step as

$$\mu_{\text{repulsion}}^{n+1} = RTG'(c^n) (G^{n+1} + G(c^n)) - \lambda RT. \tag{3.19}$$

Substituting (3.13) into (3.19), we rewrite

$$\mu_{\text{repulsion}}^{n+1} = RTG'(c^n) (2G(c^n) + G'(c^n) (c^{n+1} - c^n)) - \lambda RT, \tag{3.20}$$

which is a linear function of c^{n+1} . We note that the convex splitting approach for the repulsion term $f_b^{\text{repulsion}}(c)$ [28] results in a nonlinear scheme as well as the ideal term.

The following lemma is a direct consequence of the above analysis and the definition of $\mu_{\text{repulsion}}^{n+1}$.

LEMMA 3.3. *Assume that molar density satisfies (2.8) and λ is taken such that (3.9) holds. Then we have*

$$f_b^{\text{repulsion}}(c^{n+1}) - f_b^{\text{repulsion}}(c^n) \leq \mu_{\text{repulsion}}^{n+1} (c^{n+1} - c^n). \tag{3.21}$$

We remark that the proposed approach is different from the IEQ and SAV approaches. In the IEQ and SAV approaches, some new auxiliary energy variables are introduced, and from this, the original energy is transformed to a quadratic form. Nevertheless, the transformed energy is generally not equivalent to the original energy at the time discrete level [34]. In the proposed approach, we just use $G(c)$ as a function of c , but never introducing any new independent variable. This is a key feature of the proposed approach that allows us to apply the concavity of $G(c)$ to obtain the linear numerical scheme inheriting the dissipation law of the original energy at the discrete level.

4. Semi-implicit time-discrete scheme. In this section, we propose a semi-implicit time-discrete scheme based on the results presented in Section 3. The ideal and repulsion terms have been handled by the EF approach. Due to the concavity of

the attraction term [28], we treat it explicitly and define the corresponding chemical potential term as

$$\mu_{\text{attraction}}(c^n) = \frac{\alpha(T)}{2\sqrt{2}\beta} \ln \left(\frac{1 + (1 - \sqrt{2})\beta c^n}{1 + (1 + \sqrt{2})\beta c^n} \right) - \frac{\alpha(T)c^n}{1 + 2\beta c^n - (\beta c^n)^2}. \quad (4.1)$$

Let τ be the time step size and let c^0 be provided by the initial condition, we now state the semi-implicit linear time-discrete scheme as follows

$$\frac{c^{n+1} - c^n}{\tau} - \kappa \Delta c^{n+1} + \mu_{\text{ideal}}^{n+1} + \mu_{\text{repulsion}}^{n+1} + \mu_{\text{attraction}}(c^n) = \mu_e^{n+1}, \quad (4.2a)$$

$$\int_{\Omega} c^{n+1} d\mathbf{x} = c_t, \quad (4.2b)$$

$$\nabla c^{n+1} \cdot \mathbf{n}_{\partial\Omega} = 0, \quad \mathbf{x} \in \partial\Omega, \quad (4.2c)$$

where μ_{ideal}^{n+1} and $\mu_{\text{repulsion}}^{n+1}$ are defined in (3.4) and (3.20) respectively. For the convenience of theoretical analysis, we rewrite (4.2) as the following equivalent form

$$\frac{1}{\tau} (c^{n+1} - c^n) - \kappa \Delta c^{n+1} + \nu(c^n) c^{n+1} = s_r(c^n) + \mu_e^{n+1}, \quad (4.3a)$$

$$\int_{\Omega} c^{n+1} d\mathbf{x} = c_t, \quad (4.3b)$$

$$\nabla c^{n+1} \cdot \mathbf{n}_{\partial\Omega} = 0 \quad \text{on } \partial\Omega, \quad (4.3c)$$

where $\nu(c)$ and $s_r(c)$ are the functions of c defined as follows

$$\nu(c) = RT \left(\frac{1}{c} + G'(c)^2 \right), \quad (4.4)$$

$$s_r(c) = -\vartheta_0 - RT \ln(c) + RT (G'(c)^2 c - 2G(c)G'(c) + \lambda) - \mu_{\text{attraction}}(c). \quad (4.5)$$

In what follows, we use the traditional notations to denote the inner product of $L^2(\Omega)$ and $(L^2(\Omega))^d$ by (\cdot, \cdot) and the norm of $L^2(\Omega)$ and $(L^2(\Omega))^d$ by $\|\cdot\|$, where d is the spatial dimension.

4.1. Well-posedness of the solution. We now show the existence and uniqueness of the solution of the semi-discrete scheme (4.3) as follows.

THEOREM 4.1. *Assume that c^n satisfies the condition (2.8) and λ is taken to satisfy (3.9). There exists a unique c^{n+1} to solve (4.3) weakly in $H^1(\Omega)$.*

Proof. Suppose that c is a solution of the following homogeneous problem

$$\frac{1}{\tau} c - \kappa \Delta c + \nu(c^n) c = \mu_e, \quad (4.6a)$$

$$\int_{\Omega} c d\mathbf{x} = 0, \quad (4.6b)$$

$$\nabla c \cdot \mathbf{n}_{\partial\Omega} = 0 \quad \text{on } \partial\Omega. \quad (4.6c)$$

By Fredholm alternative theorem, it suffices to prove $c \equiv 0$. Multiplying the equation (4.6a) by c and integrating it over Ω , we obtain

$$\frac{1}{\tau} \|c\|^2 + \kappa \|\nabla c\|^2 + (\nu(c^n), c^2) = (\mu_e, c). \quad (4.7)$$

The right-hand side term vanishes due to (4.6b). Furthermore, $\nu(c^n) > 0$ holds for $c_m \leq c^n \leq c_M$. The equation (4.7) can be reduced into

$$\frac{1}{\tau} \|c\|^2 + \kappa \|\nabla c\|^2 \leq 0. \quad (4.8)$$

This means that $c \equiv 0$ almost everywhere in Ω as well as $\mu_e = 0$ from (4.6a). \square

4.2. Maximum principle. We now prove that the time scheme given in (4.2) follows the maximum principle, which can ensure that the schemes presented in Section 3 are always well defined.

THEOREM 4.2. *Assume that c^n satisfies the condition (2.8) and λ is taken to satisfy (3.9). If*

$$\max_{c_m \leq c \leq c_M} (c_m \nu(c) - s_r(c)) \leq \mu_e^{n+1} \leq \min_{c_m \leq c \leq c_M} (c_M \nu(c) - s_r(c)), \quad (4.9)$$

then we have $c_m \leq c^{n+1} \leq c_M$ almost everywhere.

Proof. Suppose that $c_m \leq c^n \leq c_M$ holds for $n \geq 0$. We first prove that $c^{n+1} \geq c_m$ almost everywhere. Let $c_-^{n+1} = \min(c^{n+1} - c_m, 0)$ and apparently $c_-^{n+1} \leq 0$. Multiplying the equation (4.3a) by c_-^{n+1} and then integrating it over Ω , we have

$$\begin{aligned} \frac{1}{\tau} (c^{n+1} - c^n, c_-^{n+1}) - \kappa (\Delta c^{n+1}, c_-^{n+1}) + (\nu(c^n)(c^{n+1} - c_m), c_-^{n+1}) \\ = (\mu_e^{n+1}, c_-^{n+1}) + (s_r(c^n) - c_m \nu(c^n), c_-^{n+1}). \end{aligned} \quad (4.10)$$

For the first term on the left-hand side of (4.10), thanks to $c^n \geq c_m$, we deduce

$$(c^{n+1} - c^n, c_-^{n+1}) = \|c_-^{n+1}\|^2 - (c^n - c_m, c_-^{n+1}) \geq \|c_-^{n+1}\|^2. \quad (4.11)$$

By using the boundary condition, the second term on the left-hand side of (4.10) becomes

$$-\kappa (\Delta c^{n+1}, c_-^{n+1}) = \kappa \|\nabla c_-^{n+1}\|^2. \quad (4.12)$$

We observe that $\nu(c)$ is a strictly monotonically decreasing positive function over the interval $[c_m, c_M]$. The third term on the left-hand side of (4.10) is bounded below

$$(\nu(c^n)(c^{n+1} - c_m), c_-^{n+1}) = (\nu(c^n), |c_-^{n+1}|^2) \geq \nu(c_M) \|c_-^{n+1}\|^2. \quad (4.13)$$

Using the condition (4.9) and taking into account $c_-^{n+1} \leq 0$, we estimate the first term on the right-hand side of (4.10) as

$$\begin{aligned} (\mu_e^{n+1}, c_-^{n+1}) &\leq \left(\max_{c_m \leq c \leq c_M} (c_m \nu(c) - s_r(c)), c_-^{n+1} \right) \\ &\leq (c_m \nu(c^n) - s_r(c^n), c_-^{n+1}). \end{aligned} \quad (4.14)$$

Combining (4.11)-(4.14), we derive from (4.10) that

$$\frac{1}{\tau} \|c_-^{n+1}\|^2 + \kappa \|\nabla c_-^{n+1}\|^2 + \nu(c_M) \|c_-^{n+1}\|^2 \leq 0. \quad (4.15)$$

It follows from (4.15) that $\|c_-^{n+1}\|^2 = 0$ and $\|\nabla c_-^{n+1}\|^2 = 0$. Consequently, $c^{n+1} \geq c_m$ almost everywhere.

We turn to prove $c^{n+1} \leq c_M$ almost everywhere. We define $c_+^{n+1} = \max(c^{n+1} - c_M, 0)$ and apparently $c_+^{n+1} \geq 0$. Similar to (4.10), we can get

$$\begin{aligned} & \frac{1}{\tau} (c^{n+1} - c^n, c_+^{n+1}) - \kappa (\Delta c^{n+1}, c_+^{n+1}) + (\nu(c^n)(c^{n+1} - c_M), c_+^{n+1}) \\ & = (\mu_e^{n+1}, c_+^{n+1}) + (s_r(c^n) - \nu(c^n)c_M, c_+^{n+1}). \end{aligned} \quad (4.16)$$

Taking into account $c^n \leq c_M$, we deduce

$$(c^{n+1} - c^n, c_+^{n+1}) = \|c_+^{n+1}\|^2 - (c^n - c_M, c_+^{n+1}) \geq \|c_+^{n+1}\|^2. \quad (4.17)$$

Using the condition (4.9) and $c_+^{n+1} \geq 0$, we derive

$$\begin{aligned} (\mu_e^{n+1}, c_+^{n+1}) & \leq \left(\min_{c_m \leq c \leq c_M} (c_M \nu(c) - s_r(c)), c_+^{n+1} \right) \\ & \leq (c_M \nu(c^n) - s_r(c^n), c_+^{n+1}). \end{aligned} \quad (4.18)$$

Using the similar routines as in the derivations of (4.15), we can obtain

$$\frac{1}{\tau} \|c_+^{n+1}\|^2 + \kappa \|\nabla c_+^{n+1}\|^2 + \nu(c_M) \|c_+^{n+1}\|^2 \leq 0. \quad (4.19)$$

This implies that $\|c_+^{n+1}\|^2 = 0$ and $\|\nabla c_+^{n+1}\|^2 = 0$, and thus, $c^{n+1} \leq c_M$ almost everywhere. \square

We remark that it is reasonable to assume the condition (4.9), as it is likely to be required in a typical physical setting. As a matter of fact, we can derive its a priori bounds just assuming $c^{n+1} > 0$. Integrating the equation (4.3a) over Ω and using the constraint (4.3b), we get

$$\mu_e^{n+1} = \frac{1}{|\Omega|} \int_{\Omega} (\nu(c^n)c^{n+1} - s_r(c^n)) d\mathbf{x}, \quad (4.20)$$

where $|\Omega|$ is the measure of Ω . Using the constraint (4.3b) again, we can estimate

$$\int_{\Omega} \nu(c^n)c^{n+1} d\mathbf{x} \leq \max(\nu(c^n)) \int_{\Omega} c^{n+1} d\mathbf{x} \leq \nu(c_m)c_t, \quad (4.21)$$

$$\int_{\Omega} \nu(c^n)c^{n+1} d\mathbf{x} \geq \min(\nu(c^n)) \int_{\Omega} c^{n+1} d\mathbf{x} \geq \nu(c_M)c_t. \quad (4.22)$$

Let us denote $s_r^m = \min_{c \in [c_m, c_M]}(s_r(c))$ and $s_r^M = \max_{c \in [c_m, c_M]}(s_r(c))$. Thus, μ_e^{n+1} is bounded as

$$\nu(c_M) \frac{c_t}{|\Omega|} - s_r^M \leq \mu_e^{n+1} \leq \nu(c_m) \frac{c_t}{|\Omega|} - s_r^m. \quad (4.23)$$

Numerical results will also be presented to verify (4.9).

4.3. Energy stability. We define the total free energy as

$$F(c^n) = (f_b(c^n), 1) + \frac{1}{2}\kappa\|\nabla c^n\|^2. \quad (4.24)$$

The following theorem shows that the proposed scheme inherits the dissipation law of the original energy at the discrete level.

THEOREM 4.3. *Assume that c^n satisfies (2.8) and λ is taken to satisfy (3.9). Under the condition (4.9), for any time step size τ , we have*

$$F(c^{n+1}) \leq F(c^n). \quad (4.25)$$

Proof. It follows from the concavity of $f_b^{\text{attraction}}(c)$ that

$$(f_b^{\text{attraction}}(c^{n+1}) - f_b^{\text{attraction}}(c^n), 1) \leq (\mu_{\text{attraction}}(c^n), c^{n+1} - c^n). \quad (4.26)$$

Applying (3.5), (3.21) and (4.26), we deduce that

$$(f_b(c^{n+1}) - f_b(c^n), 1) \leq (\mu_{\text{ideal}}^{n+1} + \mu_{\text{repulsion}}^{n+1} + \mu_{\text{attraction}}(c^n), c^{n+1} - c^n). \quad (4.27)$$

For the gradient contribution to the free energy, we can derive that

$$\begin{aligned} \frac{1}{2} (\|\nabla c^{n+1}\|^2 - \|\nabla c^n\|^2) &= \frac{1}{2} \int_{\Omega} (|\nabla c^{n+1}|^2 - |\nabla c^n|^2) d\mathbf{x} \\ &= \int_{\Omega} \nabla c^{n+1} \cdot \nabla (c^{n+1} - c^n) d\mathbf{x} - \frac{1}{2} \int_{\Omega} |\nabla (c^{n+1} - c^n)|^2 d\mathbf{x} \\ &\leq \int_{\Omega} \nabla c^{n+1} \cdot \nabla (c^{n+1} - c^n) d\mathbf{x} \\ &= - \int_{\Omega} (c^{n+1} - c^n) \Delta c^{n+1} d\mathbf{x}. \end{aligned} \quad (4.28)$$

Taking into account the equation (4.2a) and the mass constraint (4.2b), we deduce from (4.27) and (4.28) that

$$\begin{aligned} F(c^{n+1}) - F(c^n) &= (f_b(c^{n+1}) - f_b(c^n), 1) + \frac{1}{2}\kappa (\|\nabla c^{n+1}\|^2 - \|\nabla c^n\|^2) \\ &\leq (\mu_{\text{ideal}}^{n+1} + \mu_{\text{repulsion}}^{n+1} + \mu_{\text{attraction}}(c^n) - \kappa\Delta c^{n+1}, c^{n+1} - c^n) \\ &= \left(\mu_e^{n+1} - \frac{c^{n+1} - c^n}{\tau}, c^{n+1} - c^n \right) \\ &= -\frac{1}{\tau} \|c^{n+1} - c^n\|^2. \end{aligned} \quad (4.29)$$

Thus, the energy inequality (4.25) is proved. \square

5. Fully discrete scheme. In this section, we consider the fully discrete scheme. The cell-centered finite difference (CCFD) method [31] is employed as the spatial discretization method. We note that the CCFD method is equivalent to a mixed finite element method with quadrature rules [1]. Here, we present the numerical scheme in two-dimensional case only, but it is straightforward to extend it to the three-dimensional case.

We consider a rectangular domain as $\Omega = [0, l_x] \times [0, l_y]$, where $l_x > 0$ and $l_y > 0$. For simplicity, a uniform mesh of Ω is used as $0 = x_0 < x_1 < \dots < x_N = l_x$ and $0 = y_0 < y_1 < \dots < y_M = l_y$, where N and M are integers. We also introduce the intermediate points $x_{i+\frac{1}{2}} = \frac{x_i + x_{i+1}}{2}$ and $y_{j+\frac{1}{2}} = \frac{y_j + y_{j+1}}{2}$. The mesh size is denoted by $h = x_{i+1} - x_i = y_{j+1} - y_j$.

5.1. Notations and fully discrete scheme. To formulate the fully discrete scheme, we define the following discrete function spaces:

$$\mathcal{V}_c = \{c : (x_{i+\frac{1}{2}}, y_{j+\frac{1}{2}}) \mapsto \mathbb{R}, \quad 0 \leq i \leq N-1, \quad 0 \leq j \leq M-1\},$$

$$\mathcal{V}_u = \{u : (x_i, y_{j+\frac{1}{2}}) \mapsto \mathbb{R}, \quad 0 \leq i \leq N, \quad 0 \leq j \leq M-1\},$$

$$\mathcal{V}_v = \{v : (x_{i+\frac{1}{2}}, y_j) \mapsto \mathbb{R}, \quad 0 \leq i \leq N-1, \quad 0 \leq j \leq M\}.$$

For components of discrete functions in the above spaces, we denote $c_{i+\frac{1}{2}, j+\frac{1}{2}} = c(x_{i+\frac{1}{2}}, y_{j+\frac{1}{2}})$ for $c \in \mathcal{V}_c$, $u_{i, j+\frac{1}{2}} = u(x_i, y_{j+\frac{1}{2}})$ for $u \in \mathcal{V}_u$ and $v_{i+\frac{1}{2}, j} = v(x_{i+\frac{1}{2}}, y_j)$ for $v \in \mathcal{V}_v$.

For $c \in \mathcal{V}_c$, we define the difference operators $\delta_x^c[c] \in \mathcal{V}_u$ and $\delta_y^c[c] \in \mathcal{V}_v$ as follows

$$\delta_x^c[c]_{i, j+\frac{1}{2}} = \frac{c_{i+\frac{1}{2}, j+\frac{1}{2}} - c_{i-\frac{1}{2}, j+\frac{1}{2}}}{h}, \quad 1 \leq i \leq N-1, \quad 0 \leq j \leq M-1, \quad (5.1a)$$

$$\delta_y^c[c]_{i+\frac{1}{2}, j} = \frac{c_{i+\frac{1}{2}, j+\frac{1}{2}} - c_{i+\frac{1}{2}, j-\frac{1}{2}}}{h}, \quad 0 \leq i \leq N-1, \quad 1 \leq j \leq M-1. \quad (5.1b)$$

On the boundary, applying the homogeneous Neumann boundary condition, we take the difference operators as

$$\delta_x^c[c]_{i, j+\frac{1}{2}} = 0, \quad i \in \{0, N\}, \quad 0 \leq j \leq M-1, \quad (5.2a)$$

$$\delta_y^c[c]_{i+\frac{1}{2}, j} = 0, \quad j \in \{0, M\}, \quad 0 \leq i \leq N-1. \quad (5.2b)$$

We introduce the subsets of \mathcal{V}_u and \mathcal{V}_v involving the boundary condition as

$$\mathcal{V}_u^0 = \{u \in \mathcal{V}_u \mid u_{0, j+\frac{1}{2}} = u_{N, j+\frac{1}{2}} = 0, \quad 0 \leq j \leq M-1\}, \quad (5.3)$$

$$\mathcal{V}_v^0 = \{v \in \mathcal{V}_v \mid v_{i+\frac{1}{2}, 0} = v_{i+\frac{1}{2}, M} = 0, \quad 0 \leq i \leq N-1\}. \quad (5.4)$$

Apparently, $\delta_x^c[c] \in \mathcal{V}_u^0$ and $\delta_y^c[c] \in \mathcal{V}_v^0$.

The difference operators for $u \in \mathcal{V}_u$ and $v \in \mathcal{V}_v$ are defined as

$$\delta_x^u[u]_{i+\frac{1}{2}, j+\frac{1}{2}} = \frac{u_{i+1, j+\frac{1}{2}} - u_{i, j+\frac{1}{2}}}{h}, \quad (5.5)$$

$$\delta_y^v[v]_{i+\frac{1}{2}, j+\frac{1}{2}} = \frac{v_{i+\frac{1}{2}, j+1} - v_{i+\frac{1}{2}, j}}{h}, \quad (5.6)$$

where $0 \leq i \leq N-1$, $0 \leq j \leq M-1$.

We define the following discrete inner-products:

$$\langle c, c' \rangle = h^2 \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} c_{i+\frac{1}{2}, j+\frac{1}{2}} c'_{i+\frac{1}{2}, j+\frac{1}{2}}, \quad c, c' \in \mathcal{V}_c,$$

$$\langle u, u' \rangle = h^2 \sum_{i=1}^{N-1} \sum_{j=0}^{M-1} u_{i, j+\frac{1}{2}} u'_{i, j+\frac{1}{2}}, \quad u, u' \in \mathcal{V}_u^0,$$

$$\langle v, v' \rangle = h^2 \sum_{i=0}^{N-1} \sum_{j=1}^{M-1} v_{i+\frac{1}{2}, j} v'_{i+\frac{1}{2}, j}, \quad v, v' \in \mathcal{V}_v^0.$$

The discrete norms for $c \in \mathcal{V}_c$, $u \in \mathcal{V}_u^0$ and $v \in \mathcal{V}_v^0$ are denoted as

$$\|c\| = \langle c, c \rangle, \quad \|u\| = \langle u, u \rangle, \quad \|v\| = \langle v, v \rangle.$$

The semi-implicit fully discrete scheme is stated: given $c^n \in \mathcal{V}_c$, find $c^{n+1} \in \mathcal{V}_c$ such that

$$\frac{c^{n+1} - c^n}{\tau} - \kappa (\delta_x^u [\delta_x^c [c^{n+1}]] + \delta_y^u [\delta_y^c [c^{n+1}]]) + \nu(c^n) c^{n+1} = s_r(c^n) + \mu_e^{n+1} \quad (5.7a)$$

$$\langle c^{n+1}, 1 \rangle = c_t, \quad (5.7b)$$

where the formulations of functions $\nu(c)$ and $s_r(c)$ are given in (4.4) and (4.5) respectively. We note that the boundary condition has already been considered in the definitions of the operators δ_x^c and δ_y^c in (5.2).

The following summation-by-parts formulas are derived by direct calculations [3, 32, 33]

$$\langle u, \delta_x^c [c] \rangle = -\langle \delta_x^u [u], c \rangle, \quad u \in \mathcal{V}_u^0, \quad c \in \mathcal{V}_c, \quad (5.8)$$

$$\langle v, \delta_y^c [c] \rangle = -\langle \delta_y^v [v], c \rangle, \quad v \in \mathcal{V}_v^0, \quad c \in \mathcal{V}_c. \quad (5.9)$$

5.2. Unique solvability. We first demonstrate the unique solvability of the linear system (5.7).

THEOREM 5.1. *Assume that $c^n \in \mathcal{V}_c$ satisfies (2.8) and λ is taken to satisfy (3.9). There exists a unique $c^{n+1} \in \mathcal{V}_c$ such that (5.7) holds.*

Proof. It suffices to prove the following homogeneous problem has a unique zero solution in \mathcal{V}_c

$$\frac{c}{\tau} - \kappa (\delta_x^u [\delta_x^c [c]] + \delta_y^u [\delta_y^c [c]]) + \nu(c^n) c = \mu_e, \quad (5.10a)$$

$$\langle c, 1 \rangle = 0. \quad (5.10b)$$

Suppose that there exists a nonzero solution $c \in \mathcal{V}_c$ satisfying (5.10). We take the inner product of (5.10) with c

$$\frac{1}{\tau} \|c\|^2 - \kappa \langle \delta_x^u [\delta_x^c [c]], c \rangle - \kappa \langle \delta_y^u [\delta_y^c [c]], c \rangle + \langle \nu(c^n) c, c \rangle = \langle \mu_e, c \rangle, \quad (5.11)$$

which can be further reduced using (5.8), (5.9) and (5.10b) into

$$\frac{1}{\tau} \|c\|^2 + \kappa (\|\delta_x^c[c]\|^2 + \|\delta_y^c[c]\|^2) + \nu(c_M) \|c\|^2 \leq 0. \quad (5.12)$$

Since $\nu(c_M) > 0$, we concluded from (5.12) that $c \equiv 0$. This yields a contradiction, and thus, (5.7) is uniquely solvable. \square

5.3. Discrete maximum principle. We are going to prove the discrete maximum principle of molar density. The following lemma is an essential ingredient of the proof.

LEMMA 5.1. *Let $c^- = \min(c - c_m, 0)$ and $c^+ = \max(c - c_M, 0)$, where $c \in \mathcal{V}_c$ and $c_m < c_M$. Then we have*

$$\langle \delta_x^c[c^-], \delta_x^c[c^-] \rangle \leq -\langle \delta_x^u[\delta_x^c[c]], c^- \rangle, \quad (5.13)$$

$$\langle \delta_x^c[c^+], \delta_x^c[c^+] \rangle \leq -\langle \delta_x^u[\delta_x^c[c]], c^+ \rangle, \quad (5.14)$$

$$\langle \delta_y^c[c^-], \delta_y^c[c^-] \rangle \leq -\langle \delta_y^v[\delta_y^c[c]], c^- \rangle, \quad (5.15)$$

$$\langle \delta_y^c[c^+], \delta_y^c[c^+] \rangle \leq -\langle \delta_y^v[\delta_y^c[c]], c^+ \rangle. \quad (5.16)$$

Proof. Let a and b be two real scalar numbers. We further define $a^- = \min(a - c_m, 0)$ and $b^- = \min(b - c_m, 0)$. It is apparent that

$$(a - c_m) a^- = |a^-|^2, \quad (b - c_m) b^- = |b^-|^2. \quad (5.17)$$

Since $a - c_m \geq a^-$ and $b^- \leq 0$, we get

$$(a - c_m) b^- \leq a^- b^-. \quad (5.18)$$

Applying (5.8), (5.17) and (5.18), we can derive the inequality (5.13) as follows

$$\begin{aligned} -\langle \delta_x^u[\delta_x^c[c]], c^- \rangle &= \langle \delta_x^c[c], \delta_x^c[c^-] \rangle \\ &= h^2 \sum_{i=1}^{N-1} \sum_{j=0}^{M-1} \delta_x^c[c]_{i,j+\frac{1}{2}} \delta_x^c[c^-]_{i,j+\frac{1}{2}} \\ &= \sum_{i=1}^{N-1} \sum_{j=0}^{M-1} \left(c_{i+\frac{1}{2},j+\frac{1}{2}} - c_{i-\frac{1}{2},j+\frac{1}{2}} \right) \left(c_{i+\frac{1}{2},j+\frac{1}{2}}^- - c_{i-\frac{1}{2},j+\frac{1}{2}}^- \right) \\ &\geq \sum_{i=1}^{N-1} \sum_{j=0}^{M-1} \left(|c_{i+\frac{1}{2},j+\frac{1}{2}}^-|^2 + |c_{i-\frac{1}{2},j+\frac{1}{2}}^-|^2 - 2c_{i+\frac{1}{2},j+\frac{1}{2}}^- c_{i-\frac{1}{2},j+\frac{1}{2}}^- \right) \\ &= \sum_{i=1}^{N-1} \sum_{j=0}^{M-1} \left(c_{i+\frac{1}{2},j+\frac{1}{2}}^- - c_{i-\frac{1}{2},j+\frac{1}{2}}^- \right)^2 \\ &= \langle \delta_x^c[c^-], \delta_x^c[c^-] \rangle. \end{aligned} \quad (5.19)$$

We now prove the inequality (5.14). Let $a^+ = \max(a - c_M, 0)$ and $b^+ = \max(b - c_M, 0)$, then we have

$$(a - c_M) a^+ = |a^+|^2, \quad (b - c_M) b^+ = |b^+|^2. \quad (5.20)$$

Since $a - c_M \leq a^+$ and $b^+ \geq 0$, we get

$$(a - c_M)b^+ \leq a^+b^+. \quad (5.21)$$

Applying (5.8), (5.20) and (5.21), we deduce the inequality (5.14) as

$$\begin{aligned} -\langle \delta_x^u[\delta_x^c[c]], c^+ \rangle &= \langle \delta_x^c[c], \delta_x^c[c^+] \rangle \\ &= h^2 \sum_{i=1}^{N-1} \sum_{j=0}^{M-1} \delta_x^c[c]_{i,j+\frac{1}{2}} \delta_x^c[c^+]_{i,j+\frac{1}{2}} \\ &= \sum_{i=1}^{N-1} \sum_{j=0}^{M-1} \left(c_{i+\frac{1}{2},j+\frac{1}{2}} - c_{i-\frac{1}{2},j+\frac{1}{2}} \right) \left(c_{i+\frac{1}{2},j+\frac{1}{2}}^+ - c_{i-\frac{1}{2},j+\frac{1}{2}}^+ \right) \\ &\geq \sum_{i=1}^{N-1} \sum_{j=0}^{M-1} \left(|c_{i+\frac{1}{2},j+\frac{1}{2}}^+|^2 + |c_{i-\frac{1}{2},j+\frac{1}{2}}^+|^2 - 2c_{i+\frac{1}{2},j+\frac{1}{2}}^+ c_{i-\frac{1}{2},j+\frac{1}{2}}^+ \right) \\ &= \langle \delta_x^c[c^+], \delta_x^c[c^+] \rangle. \end{aligned} \quad (5.22)$$

The rest inequalities (5.15) and (5.16) can be proved by the similar approaches. \square

We are now ready to prove the maximum principle of the fully discrete scheme.

THEOREM 5.2. *Assume that $c^n \in \mathcal{V}_c$ satisfies (2.8) and λ is taken to satisfy (3.9). Under the condition (4.9), we have $c_m \leq c^{n+1} \leq c_M$.*

Proof. It can be proved using the similar routines as in the proof of Theorem 4.2. The major difference is that we use Lemma 5.1 to treat the discrete form of Laplace operator. So we just give a brief proof.

Assuming that $c_m \leq c^n \leq c_M$ for $n \geq 0$, we will prove that $c_m \leq c^{n+1} \leq c_M$. Let $c_-^{n+1} = \min(c^{n+1} - c_m, 0)$. It can be derived from (5.7a) that

$$\begin{aligned} \frac{1}{\tau} \langle c^{n+1} - c^n, c_-^{n+1} \rangle - \kappa \langle \delta_x^u[\delta_x^c[c^{n+1}]] + \delta_y^u[\delta_y^c[c^{n+1}]], c_-^{n+1} \rangle \\ + \langle \nu(c^n)(c^{n+1} - c_m), c_-^{n+1} \rangle = \langle \mu_e^{n+1}, c_-^{n+1} \rangle + \langle s_r(c^n) - c_m \nu(c^n), c_-^{n+1} \rangle. \end{aligned} \quad (5.23)$$

The second term on the left-hand side of (5.23) can be estimated using Lemma 5.1

$$-\kappa \langle \delta_x^u[\delta_x^c[c^{n+1}]] + \delta_y^u[\delta_y^c[c^{n+1}]], c_-^{n+1} \rangle \geq \kappa (\|\delta_x^c[c_-^{n+1}]\|^2 + \|\delta_y^c[c_-^{n+1}]\|^2). \quad (5.24)$$

Following the similar routines used to deduce (4.15), we obtain

$$\frac{1}{\tau} \|c_-^{n+1}\|^2 + \kappa (\|\delta_x^c[c_-^{n+1}]\|^2 + \|\delta_y^c[c_-^{n+1}]\|^2) + \nu(c_M) \|c_-^{n+1}\|^2 \leq 0, \quad (5.25)$$

which yields $c_-^{n+1} = 0$, and thus, $c^{n+1} \geq c_m$. It is similar to prove $c^{n+1} \leq c_M$. \square

5.4. Energy stability. We denote the discrete total free energy as

$$F_h(c^n) = \langle f_b(c^n), 1 \rangle + \frac{1}{2} (\|\delta_x^c[c^n]\|^2 + \|\delta_y^c[c^n]\|^2). \quad (5.26)$$

THEOREM 5.3. *Assume that c^n satisfies (2.8) and λ is taken to satisfy (3.9). Under the condition (4.9), for any time step size τ , we have*

$$F_h(c^{n+1}) \leq F_h(c^n). \quad (5.27)$$

Table 6.1: Physical parameters of nC₄

$P_c(\text{bar})$	$T_c(\text{K})$	ω
38.0	425.2	0.199

Proof. Using (5.8), we can derive that

$$\begin{aligned}
\frac{1}{2} (\|\delta_x^c[c^{n+1}]\|^2 - \|\delta_x^c[c^n]\|^2) &= \frac{1}{2} (\langle \delta_x^c[c^{n+1}], \delta_x^c[c^{n+1}] \rangle - \langle \delta_x^c[c^n], \delta_x^c[c^n] \rangle) \\
&= \langle \delta_x^c[c^{n+1}], \delta_x^c[c^{n+1}] - \delta_x^c[c^n] \rangle - \frac{1}{2} \|\delta_x^c[c^{n+1}] - \delta_x^c[c^n]\|^2 \\
&\leq \langle \delta_x^c[c^{n+1}], \delta_x^c[c^{n+1}] - \delta_x^c[c^n] \rangle \\
&= -\langle \delta_x^u[\delta_x^c[c^{n+1}]], c^{n+1} - c^n \rangle. \tag{5.28}
\end{aligned}$$

It is similar to deduce that

$$\frac{1}{2} (\|\delta_y^c[c^{n+1}]\|^2 - \|\delta_y^c[c^n]\|^2) \leq -\langle \delta_y^u[\delta_y^c[c^{n+1}]], c^{n+1} - c^n \rangle. \tag{5.29}$$

Using similar arguments in (4.27) and the estimates (5.28) and (5.28), we obtain

$$\begin{aligned}
\langle f_b(c^{n+1}) - f_b(c^n), 1 \rangle + \frac{1}{2} (\|\delta_x^c[c^{n+1}]\|^2 - \|\delta_x^c[c^n]\|^2) \\
+ \frac{1}{2} (\|\delta_y^c[c^{n+1}]\|^2 - \|\delta_y^c[c^n]\|^2) \leq -\frac{1}{\tau} \|c^{n+1} - c^n\|^2, \tag{5.30}
\end{aligned}$$

which yields (5.27). \square

6. Numerical results. In this section, we present some numerical results to show the performance of the proposed method and verify the theoretical analysis. We consider a hydrocarbon substance, n-butane (nC₄). The related physical data is provided in Table 6.1. The temperature is fixed at 330 K. The value of ϑ_0 has no effects for the isothermal fluids, so we take $\vartheta_0 = 0$. In numerical tests, we use the gas molar density $c^G = 249.1123 \text{ mol/m}^3$ and the liquid molar density $c^L = 9526.8428 \text{ mol/m}^3$ to initialize the distribution of molar density. We note that mass transfer between two phases may take place in dynamical process; that is, the gas phase may be condensed into the liquid phase, while the liquid phase may also be vaporized into the gas phase. As a result, we take $c_m = 0.9c^G$ and $c_M = 1.1c^L$ in (2.8) and calculate $\epsilon_0 = \beta c_M = 0.7585$. To satisfy the condition (3.9), we use the value of ϵ_0 to calculate λ as follows

$$\lambda = \frac{\epsilon_0}{(1 - \epsilon_0)^2} + \left(\frac{\epsilon_0^2}{(1 - \epsilon_0)^4} - 2 \ln(1 - \epsilon_0) \frac{\epsilon_0}{(1 - \epsilon_0)^2} \right)^{1/2} = 27.3656. \tag{6.1}$$

In all numerical tests, the spatial domain is taken as $\Omega = [-L, L]^2$, where $L = 15 \text{ nm}$, and a uniform rectangular mesh with 100×100 elements is employed. The proposed method admits the use of a very time step size, so we take $\tau = 10^{10} \text{ s}$.

In this example, we simulate the droplet shrinking problem for 200 time steps. In Figure 6.1, we show molar density distributions after different time steps, which demonstrate that the square droplet is changing to a circle shape. In Figures 6.2, we also illustrate the distributions of the bulk free energy density and gradient free energy density after 200 time steps, respectively.

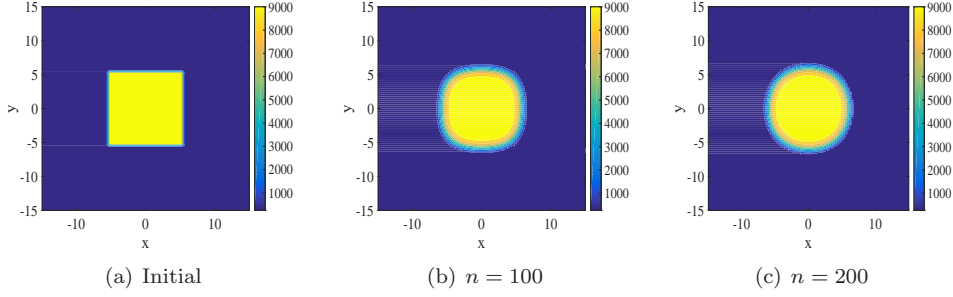


Fig. 6.1: Example 1: molar density distributions.

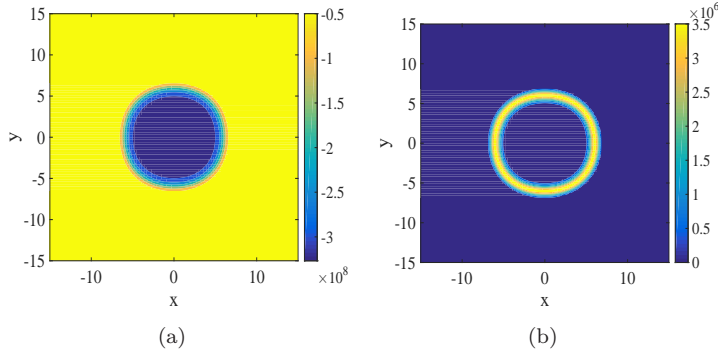


Fig. 6.2: Example 1: (a) the bulk free energy density after 200 time steps; (b) the gradient free energy after 200 time steps.

In Figure 6.3(a), the lower bound $\underline{\mu}$ and the upper bound $\bar{\mu}$ are calculated as

$$\underline{\mu} = \max_{c_m \leq c \leq c_M} (c_m \nu(c) - s_r(c)), \quad \bar{\mu} = \min_{c_m \leq c \leq c_M} (c_M \nu(c) - s_r(c)).$$

Figure 6.3(a) shows that μ_e^n always varies between $\underline{\mu}$ and $\bar{\mu}$, thus the condition (4.9) is invariably true. Figures 6.3(b) and 6.3(c) depict the minimum and maximum values of molar density c^n . Due to the effects of liquid vaporization and gas condensation, molar density changes a lot at the beginning period and then tends to the steady values in the later period, but we observe that c^n always fluctuates between c_m and c_M , thus the maximum principle is verified.

Figure 6.4(a) illustrates that the total energies are dissipated with time steps, while Figure 6.4(b) plots total energies at the last twenty time steps, which are still decreasing. Therefore, the proposed scheme can preserve the energy dissipation law.

7. Conclusions. A novel energy factorization (EF) approach has been proposed to construct the linear energy stable numerical scheme for the diffuse interface model with the Peng-Robinson equation of state that is one of the most useful and prominent tools in chemical engineering and petroleum industry. Compared with the convex-splitting schemes, the semi-implicit numerical scheme constructed by the EF approach is linear and easy-to-implement. Compared with the IEQ/SAV schemes, the proposed scheme inherits the original energy dissipation law. Moreover, we prove that the maximum principle holds for both the time semi-discrete form and the cell-centered finite difference fully discrete form under certain conditions. Numerical results validate the stability and efficiency of the proposed scheme. In the future work, we will study the

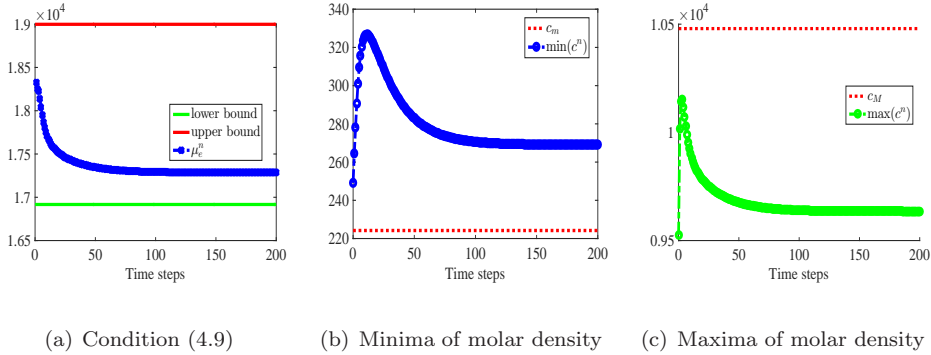


Fig. 6.3: Example 1: verification of the condition (4.9) and the maximum principle.

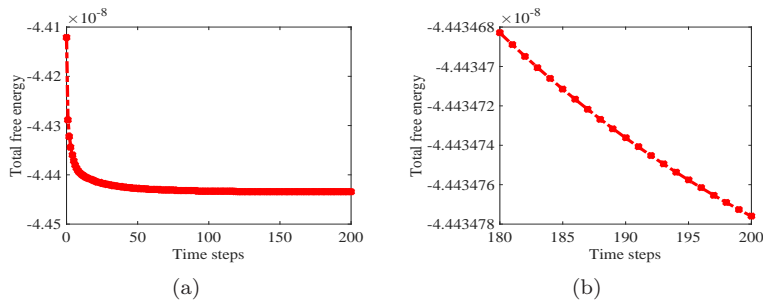


Fig. 6.4: Example 1: total energy profiles with time steps.

applications of the EF approach to multi-component fluids and phase-field models.

REFERENCES

- [1] T. Arbogast, M.F. Wheeler, I. Yotov. Mixed finite elements for elliptic problems with tensor coefficients as cell-centered finite differences. *SIAM Journal on Numerical Analysis*, 34(2): 828–852, 1997.
- [2] A. Baskaran, J. Lowengrub, C. Wang, S. Wise. Convergence analysis of a second order convex splitting scheme for the modified phase field crystal equation. *SIAM Journal on Numerical Analysis*, 51(5): 2851–2873, 2013.
- [3] Y. Chen, J. Shen. Efficient, adaptive energy stable schemes for the incompressible Cahn-Hilliard Navier-Stokes phase-field models. *Journal of Computational Physics*, 308: 40–56, 2016.
- [4] C. M. Elliott, A. M. Stuart, The global dynamics of discrete semilinear parabolic equations, *SIAM Journal on Numerical Analysis*, 30: 1622–1663, 1993.
- [5] D. J. Eyre. Unconditionally gradient stable time marching the Cahn-Hilliard equation. *Computational and mathematical models of microstructural evolution (San Francisco, CA, 1998)*, Mater. Res. Soc. Sympos. Proc., 529: 39–46. MRS, Warrendale, PA, 1998.
- [6] X. Fan, J. Kou, Z. Qiao, S. Sun. A Componentwise Convex Splitting Scheme for Diffuse Interface Models with Van der Waals and Peng-Robinson Equations of State. *SIAM Journal on Scientific Computing*, 39(1): B1–B28, 2017.
- [7] A. Firoozabadi. *Thermodynamics of hydrocarbon reservoirs*. McGraw-Hill New York, 1999.
- [8] T. Jindrová, J. Mikyška. Fast and robust algorithm for calculation of two-phase equilibria at given volume, temperature, and moles. *Fluid Phase Equilibria*, 353:101–114, 2013.
- [9] T. Jindrová, J. Mikyška. General algorithm for multiphase equilibria calculation at given volume, temperature, and moles. *Fluid Phase Equilibria*, 393:7–25, 2015.
- [10] J. Kou, S. Sun, X. Wang. Efficient numerical methods for simulating surface tension of multi-component mixtures with the gradient theory of fluid interfaces. *Computer Methods in Applied Mechanics and Engineering*, 292: 92–106, 2015.

- [11] J. Kou, S. Sun. Numerical methods for a multi-component two-phase interface model with geometric mean influence parameters. *SIAM Journal on Scientific Computing*, 37(4): B543–B569, 2015.
- [12] J. Kou, S. Sun. Unconditionally stable methods for simulating multi-component two-phase interface models with Peng-Robinson equation of state and various boundary conditions. *Journal of Computational and Applied Mathematics*, 291(1): 158–182, 2016.
- [13] J. Kou, S. Sun. A stable algorithm for calculating phase equilibria with capillarity at specified moles, volume and temperature using a dynamic model. *Fluid Phase Equilibria*, 456: 7–24, 2018.
- [14] J. Kou, S. Sun. Multi-scale diffuse interface modeling of multi-component two-phase flow with partial miscibility. *Journal of Computational Physics*, 318: 349–372, 2016.
- [15] J. Kou, S. Sun, X. Wang. Linearly decoupled energy-stable numerical methods for multicomponent two-phase compressible flow. *SIAM Journal on Numerical Analysis*, 56(6): 3219–3248, 2018.
- [16] J. Kou, S. Sun. Thermodynamically consistent modeling and simulation of multi-component two-phase flow with partial miscibility. *Computer Methods in Applied Mechanics and Engineering*, 331: 623–649, 2018.
- [17] J. Kou, S. Sun. Thermodynamically consistent simulation of nonisothermal diffuse-interface two-phase flow with Peng-Robinson equation of state, *Journal of Computational Physics*, 371: 581–605, 2018.
- [18] J. Kou, S. Sun. Entropy stable modeling of non-isothermal multi-component diffuse-interface two-phase flows with realistic equations of state, *Computer Methods in Applied Mechanics and Engineering*, 341: 221–248, 2018.
- [19] H. Li, L. Ju, C. Zhang, Q. Peng. Unconditionally energy stable linear schemes for the diffuse interface model with Peng-Robinson equation of state. *Journal of Scientific Computing*, 75(2): 993–1015, 2018.
- [20] J. S. Lopez-Echeverry, S. Reif-Acherman, E. Araujo-Lopez. Peng-Robinson equation of state: 40 years through cubics. *Fluid Phase Equilibria*, 447: 39–71, 2017.
- [21] M. L. Michelsen. State function based flash specifications. *Fluid Phase Equilibria*, 158–160: 617–626, 1999.
- [22] J. Mikyška, A. Firoozabadi. A new thermodynamic function for phase-splitting at constant temperature, moles, and volume. *AIChE Journal*, 57(7): 1897–1904, 2011.
- [23] J. Mikyška. A Collection of Analytical Solutions for the Flash Equilibrium Calculation Problem. *Transport in Porous Media*, doi.org/10.1007/s11242-018-1160-9, 2018.
- [24] C. Miqueu, B. Mendiboure, C. Graciaa, J. Lachaise. Modelling of the surface tension of binary and ternary mixtures with the gradient theory of fluid interfaces. *Fluid Phase Equilibria*, 218: 189–203, 2004.
- [25] N.R. Nagarajan, A.S. Cullick. New strategy for phase equilibrium and critical point calculations by thermodynamic energy analysis. Part I. Stability analysis and flash. *Fluid Phase Equilibria*, 62(3): 191–210, 1991.
- [26] D. Peng, D.B. Robinson. A new two-constant equation of state. *Industrial and Engineering Chemistry Fundamentals*, 15(1): 59–64, 1976.
- [27] Q. Peng. A convex-splitting scheme for a diffuse interface model with Peng-Robinson equation of state. *Advances in Applied Mathematics and Mechanics*, 9(5): 1162–1188, 2017.
- [28] Z. Qiao, S. Sun. Two-phase fluid simulation using a diffuse interface model with Peng-Robinson equation of state. *SIAM Journal on Scientific Computing*, 36(4): B708–B728, 2014.
- [29] J. Shen, X. Yang. Decoupled, energy stable schemes for phase-field models of two-phase incompressible flows. *SIAM Journal on Numerical Analysis*, 53(1): 279–296, 2015.
- [30] J. Shen, J. Xu, J. Yang. The scalar auxiliary variable (SAV) approach for gradient flows. *Journal of Computational Physics*, 353: 407–416, 2018.
- [31] G. Tryggvason, R. Scardovelli, S. Zaleski. *Direct Numerical Simulations of Gas-Liquid Multiphase Flows*. Cambridge University Press, New York, 2011.
- [32] S. M. Wise. Unconditionally Stable Finite Difference, Nonlinear Multigrid Simulation of the Cahn-Hilliard-Hele-Shaw System of Equations. *Journal of Scientific Computing*, 44: 38–68, 2010.
- [33] S. M. Wise, C. Wang, J. S. Lowengrub. An energy-stable and convergent finite-difference scheme for the phase field crystal equation. *SIAM Journal on Numerical Analysis*, 47(3): 2269–2288, 2009.

- [34] X. Yang, J. Zhao. On Linear and Unconditionally Energy Stable Algorithms for Variable Mobility Cahn-Hilliard Type Equation with Logarithmic Flory-Huggins Potential. *Communications in Computational Physics*, 25(3): 703–728, 2019.
- [35] X. Yang, L. Ju. Efficient linear schemes with unconditionally energy stability for the phase field elastic bending energy model. *Computer Methods in Applied Mechanics and Engineering*, 315: 691–712, 2017.
- [36] X. Yang, J. Zhao, Q. Wang. Numerical approximations for the molecular beam epitaxial growth model based on the invariant energy quadratization method. *Journal of Computational Physics*, 333: 104–127, 2017.
- [37] G. Zhu, J. Kou, S. Sun, J. Yao, A. Li. Decoupled, energy stable schemes for a phase-field surfactant model. *Computer Physics Communications*, 233: 67–77, 2018.