Available online at www.sciencedirect.com
ScienceDirect
Comput. Methods Appl. Mech. Engrg. 400 (2022) 115479
Computer methods in applied mechanics and engineering
www.elsevier.com/locate/cma

Efficient linear, fully-decoupled and energy stable numerical scheme for a variable density and viscosity, volume-conserved, hydrodynamically coupled phase-field elastic bending energy model of lipid vesicles

Xiaofeng Yang

Department of Mathematics, University of South Carolina, Columbia, SC 29208, USA

Received 20 April 2022; received in revised form 11 July 2022; accepted 27 July 2022
Available online xxxx

Abstract

We first establish a variable density and viscosity, volume-conserved, hydrodynamically coupled phase-field variable elastic bending energy model for lipid vesicles, and then construct an efficient time-discrete scheme for solving it. The numerical scheme combines the penalty method for solving the Navier–Stokes equation, the explicit-IEQ (invariant energy quadratization) method for the nonlinear potentials, and the operator-splitting method. Hence it is not only fully decoupled but also owns some desired properties of linearity and unconditional energy stability. The feature of full decoupling is achieved by introducing some auxiliary variables and designing additional ordinary differential equations which are used for discretizing the coupled and nonlinear terms. The solvability and the unconditional energy stability of the numerical scheme have been further rigorously and numerically proven. Several numerical examples are carried out on the sedimentation process of the vesicle cell to show the effectiveness of the model and scheme.

Keywords: Phase-field; Fully-decoupled; Elastic bending energy; Vesicle membrane; Variable density; Energy stability

1. Introduction

Certain amphiphilic molecules (such as surfactants, phospholipids, polymer amphiphiles) will spontaneously form an ordered combination of molecules with a closed bilayer structure when dispersed in water, called vesicles (or liposomes). The vesicle structure type is the basic functional unit, which provides the basis for drug transportation and related functional material design. The dynamic behavior of vesicles has very important applications in biology, medicine, and other fields. The phase-field (diffusive interface) approach, as an effective method for simulating multiphase flow, has been widely applied to investigate the dynamical deformation of lipid vesicle membranes, cf. [1–9]. The main idea of deriving the phase-field model of lipid vesicles is to minimize the total free energy in certain specific spaces, usually in the L^2 space or called as Allen–Cahn relaxation dynamics. By adopting a labeling function (phase-field variable) to label the two immiscible fluids isolated by the lipid vesicle membrane, and

E-mail address: xfyang@math.sc.edu.

https://doi.org/10.1016/j.cma.2022.115479

postulating the energy potential to approximate the mean curvature of the membrane surface, the phase-field vesicle model is obtained using the variational energetic approach. The obtained system is usually called the phase-field elastic bending energy model or phase-field vesicle model.

One of the important applications of the phase-field bending energy model is to investigate the dynamical deformations of lipid vesicles when an external field is applied (e.g., shear-driven deformation, gravity-driven sedimentation behavior, magnetic field-driven alignment, etc., see [10–22]). Assuming that the fluid inside and outside the lipid vesicles is incompressible, the flow-coupled vesicle model can be established by coupling the bending energy model with the Navier–Stokes equation, and the coupling of these two equations is achieved through the surface tension and advection terms. Regarding flow-coupled models, we find that most modeling and simulation-related works do not consider the impact of density change on vesicle dynamics. This may be reasonable for simulating some phenomena such as the shear flow of suspension. However, when considering the sedimentation dynamics of vesicles under the gravity force, see [18,19,23,24], if the density difference between the liquid in the vesicle and the surrounding liquid is large, the influence of the density difference cannot be ignored. *Therefore, the first goal of this paper is to couple variable density and viscosity into the flow-coupled bending energy model using the variational energetic approach.*

Considering that the classic phase-field bending energy model [7–9] can only approximately conserve the volume, in this article, we directly apply the so-called conserved Allen–Cahn method developed in [25] to achieve accurate volume conservation. In fact, there are many different available methods to achieve the volume conservation, for example, various versions of the conservative Allen–Cahn equation (see [26–31]), Lagrangian multiplier method (see [1,32]), and even the Cahn–Hilliard dynamics that can be used when deriving the model (see [33,34]). Among these methods, the way of adding a nonlocal Lagrange multiplier term into the time-dependent PDE system to achieve exact volume conservation is more suitable for developing the numerical scheme with energy stability because the modified system can not only have the same spatial order as the classical phase-field bending energy model but also guarantee the energy law of dissipation at the same time. After obtaining such a highly nonlinear, nonlocal, flow-coupled, variable-density/viscosity, phase-field elastic bending energy model for lipid vesicles, *the second goal of this article is to design some efficient and energy-stable numerical algorithms to solve it, especially linear and fully decoupled type schemes*.

So far, many effective numerical schemes related to the phase-field vesicle model have been developed, especially for the case of no flow-field, for example, the linear stabilization method [35], the nonlinear functional derivative method [33], the Exponential Time Differencing (ETD) method [32], the Invariant Energy Quadratization (IEQ) method [36] and its variant version of Scalar Auxiliary Variable (SAV) method [37–39], etc. As for the Navier–Stokes equation, we all know that there are many effective numerical approaches, for example, the Gauge/penalty/projection methods (cf. [40–46]), etc. Therefore, a natural question arises, is it possible to obtain an effective numerical method for the hydrodynamically coupled model by simply combining numerical methods for these two equations? The answer to this question can be explored by classifying the available numerical schemes of the flow-coupled phase-field model because the particular vesicle model studied in this article is also derived from the phase-field approach.

If the existing time-marching numerical schemes of the flow-coupled phase-field model are categorized based on the way of the time discretization for those nonlinearly coupled terms, there are usually four different methods that are capable of following the energy stability unconditionally, including the fully-implicit method (cf. [47,48]), the semi-implicit method (cf. [49–51]), the explicit-stabilization method (cf. [52–55]), and the explicit-auxiliary variable method [38,39]. In the fully-implicit method, all terms are discretized implicitly, resulting in a nonlinearly coupled algorithm and leading to high computational cost. After using the semi-implicit method, a linear algorithm can be obtained, and even a time-marching algorithm with the second-order temporal accuracy (for constant-density case), but the disadvantage is that the obtained scheme is still fully coupled. For comparison, the use of the explicit-stabilization method [33,52–55] and the explicit-auxiliary variable method [38,39] has great advantages in obtaining a complete decoupling scheme. These two methods are currently the only two methods that can achieve a fully decoupling structure while maintaining unconditional energy stability, but there are still some shortcomings. In the explicit-stabilization method [33,52–55], the key to realizing the decoupling structure is to introduce some stabilization terms to the advection term in the phase-field equation. This method inevitably leads to solving the variable-coefficient phase-field equation at every time step, which requires higher computational costs than equations with only constant coefficients. In addition, because the flow coupling model studied in this paper is

nonlocal, the direct application of the explicit-stabilization method needs to solve a nonlocal system with many variable coefficients, which can cause more additional computational cost in practice. The disadvantage of the explicit-auxiliary variable approach developed in [38,39] is that this method only considers the case of constant density/viscosity, so it cannot be directly applied to the vesicle model with variable density/viscosity studied in this article. Therefore, we hope to make up for the shortcomings of these methods, that is, to construct a full decoupling scheme for the variable-density phase-field vesicle model. It is expected that the developed scheme can solve as many linear systems with constant coefficients as possible while verifying the energy stability unconditionally, thereby reducing the computational cost due to dealing with variable-coefficient systems as much as possible.

To achieve the above purposes, inspired by the explicit-auxiliary variable method developed in [38,39] for the constant density case, we develop a novel explicit-IEQ method to deal with the nonlinear potentials and nonlinear couplings. The key idea of the new time-marching explicit-IEQ method is the introduction of four auxiliary variables and their associated ordinary differential equations (ODEs). By using these auxiliary variables, the original system is then rewritten into an equivalent form and we design the algorithm for the modified model. The benefit of discretizing the new model is that the auxiliary variables can help to decompose the discretized equations into several linear sub-equations, in which all variables can be solved independently at each time step, thereby greatly reducing the computational cost and improving the efficiency. After combining the decoupling approach with the IEQ approach for linearizing the energy potential, the penalty method of the Navier–Stokes equation, and the operator-splitting method, we get an easy-to-implement numerical algorithm in which all nonlinear terms are discretized using an explicit way, which is why we call this method the explicit-IEQ method. When implementing the scheme at each time step, we first solve the phase-field equation with constant coefficients, then solve the fluid momentum equation, in which only two terms have variable coefficients (associated with the time derivative of the velocity and the viscosity term), and finally solve the pressure Poisson equation with constant coefficients. We also present the rigorous proof of energy stability and solvability and further simulate various numerical examples to numerically demonstrate accuracy and stability.

The rest of the article is organized as follows. In Section 2, a phase-field elastic bending energy model with volume conservation, variable density and viscosity, and hydrodynamics coupling is established for lipid vesicles. Its law of energy dissipation has also been verified. In Section 3, we construct a semi-discrete time marching scheme and describe its implementations in detail. Unconditional energy stability and solvability are also proved rigorously. In Section 4, we perform several accuracy/stability tests and implement various simulations to demonstrate the effectiveness of the scheme. In Section 5, some concluding remarks are given finally.

2. The variable-density/viscosity vesicle model

We first introduce some notations that will be used throughout this article. We assume that the domain $\Omega \in \mathbb{R}^d, d = 2, 3$ is open, rectangular, smooth and bounded. For any two functions $\phi(\boldsymbol{x})$ and $\psi(\boldsymbol{x})$, their L^2-inner product on Ω is denoted by $(\phi, \psi) = \int_\Omega \phi(\boldsymbol{x})\psi(\boldsymbol{x})d\boldsymbol{x}$, and the L^2-norm of $\phi(\boldsymbol{x})$ is denoted by $\|\phi\| = (\phi, \phi)^{\frac{1}{2}}$.

Based on the total free energy given in [7–9], we establish the volume-conserved, flow-coupled phase-field bending energy model with variable density and viscosity as follows.

We define the phase-field variable as $\phi(\boldsymbol{x}) = \tanh\left(\frac{d(\boldsymbol{x})}{\sqrt{2}\epsilon}\right)$ for all $\boldsymbol{x} \in \Omega$, where $d(\boldsymbol{x})$ is the signed distance of the point $\boldsymbol{x}$ from the membrane surface Γ, and ϵ is related to the width of the diffusive interface. After coupling with the fluid flow field, the total free energy of the elastic bending energy model is given as follows [8]:

$$E(\rho, \mathbf{u}, \phi) = \underbrace{\int_\Omega \frac{\rho}{2}|\mathbf{u}|^2 d\boldsymbol{x}}_{\text{part A}} + \lambda\Big(\underbrace{\int_\Omega \frac{\epsilon}{2}(\Delta\phi - f(\phi))^2\, d\boldsymbol{x}}_{\text{part B}} + \underbrace{\frac{1}{2}M(A(\phi) - \beta)^2}_{\text{part C}}\Big), \tag{2.1}$$

where ρ is the density of the fluid mixture, $\mathbf{u}$ is the average velocity field ($\mathbf{u} = (u_1, u_2)$ for 2D, and $\mathbf{u} = (u_1, u_2, u_3)$ for 3D), λ is a normalization constant that characterizes the magnitude of bending energy, $F(\phi) = \frac{1}{4\epsilon^2}(\phi^2 - 1)^2$ is the double-well potential, $f(\phi) = F'(\phi) = \frac{1}{\epsilon^2}\phi(\phi^2 - 1)$, and $A(\phi)$ is the surface area function that is defined as

$$A(\phi) = \epsilon \int_\Omega \Big(\frac{1}{2}|\nabla\phi|^2 + F(\phi)\Big)\, d\boldsymbol{x}. \tag{2.2}$$

The surface area can be given as $\frac{3}{2\sqrt{2}}A(\phi)$. $M \gg 1$ is a positive penalty parameter to enforce the surface area conservation (approximately). β denotes a constant related to the initial surface area. In this paper, $\beta = A(\phi^0)$ with

$\phi^0 = \phi|_{(t=0)}$. In the energy (2.1), part A represents the kinetic energy of the flow field, part B represents the elastic bending energy of lipid vesicles, and part C represents the penalty term to enforce the surface area constraint in an approximate way.

Assuming that the system follows the Allen–Cahn relaxation dynamics (L^2-gradient flow), the governing PDEs read as:

$$\phi_t + \nabla\cdot(\mathbf{u}\phi) + \gamma\left(\mu - \frac{1}{|\Omega|}\int_\Omega \mu d\boldsymbol{x}\right) = 0, \tag{2.3}$$

$$\mu = \epsilon(\Delta - f'(\phi))(\Delta\phi - f(\phi)) + \epsilon M(A(\phi)-\beta)(-\Delta\phi + f(\phi)), \tag{2.4}$$

where $f'(\phi) = \frac{1}{\epsilon^2}(3\phi^2 - 1)$, γ is the relaxation mobility parameter, $\mu = \frac{1}{\lambda}\frac{\delta E}{\delta\phi}$ is the scaled variational derivative or chemical potential, and $\nabla\cdot(\mathbf{u}\phi)$ is the fluid advection term.

Note that the volume of the vesicle is defined as $\int_\Omega \frac{1+\phi}{2} d\boldsymbol{x}$, see [1,7–9], hence, by integrating (2.3) and using the integration by parts and $\nabla\cdot\mathbf{u} = 0$, we derive

$$\frac{d}{dt}\int_\Omega \phi d\boldsymbol{x} = 0. \tag{2.5}$$

This means that the vesicle volume remains the same over time. The volume conservation property is actually achieved by adding the nonlocal term $-\frac{1}{|\Omega|}\int_\Omega \mu d\boldsymbol{x}$ into (2.3). This idea of using such a nonlocal Lagrange multiplier term to conserve the volume of the phase-field variable was first proposed in [25].

We assume that the two fluid components (the fluid inside the vesicle and the ambient fluid) have different density and viscosity, i.e., ρ_1, ν_1 are the density and viscosity of the vesicle, and ρ_2, ν_2 are the density and viscosity of the ambient fluid, respectively. For the case where the density and viscosity of the two fluids are the same, i.e., $\rho_1 = \rho_2 = 1$, $\nu_1 = \nu_2 = \nu$, the Navier–Stokes fluid flow equation reads as:

$$\mathbf{u}_t + (\mathbf{u}\cdot\nabla)\mathbf{u} - \nu\Delta\mathbf{u} + \nabla p + \phi\nabla\mu = 0, \tag{2.6}$$

$$\nabla\cdot\mathbf{u} = 0, \tag{2.7}$$

with $\mathbf{u}|_{\partial\Omega} = \mathbf{0}$ as the boundary condition. The nonlinear term $\phi\nabla\mu$ is a surface tension term, that can be derived from the principle of least action, see [56].

For the case where the density and viscosity of the two fluids are not matched. If the density ratio is low ($\sim O(1)$), one could use the well-known Boussinesq approximation and to simulate the density difference (cf. for instance [56]). When the density ratio is high, the Boussinesq approximation is useless. For this case, we follow the derivation of [53,57] and assume that the density function $\rho(\phi)$ and viscosity $\nu(\phi)$ have the linear relations with the phase-field variable ϕ as:

$$\rho(\phi) = \frac{\rho_1 - \rho_2}{2}\phi + \frac{\rho_1 + \rho_2}{2},\ \nu(\phi) = \frac{\nu_1 - \nu_2}{2}\phi + \frac{\nu_1 + \nu_2}{2}. \tag{2.8}$$

This implies $\rho(\phi) = \rho_1$, $\nu(\phi) = \nu_1$ when $\phi = 1$, and $\rho(\phi) = \rho_2$, $\nu(\phi) = \nu_2$ when $\phi = -1$. Thus, from (2.3) and (2.8), the density function $\rho(\phi)$ satisfies the following equation

$$\rho_t + \nabla\cdot(\mathbf{u}\rho) - J = 0, \tag{2.9}$$

where $J = \gamma\frac{\rho_2-\rho_1}{2}\left(\mu - \frac{1}{|\Omega|}\int_\Omega \mu d\boldsymbol{x}\right)$. By using integration by parts, we derive the following identities

$$\begin{aligned}\frac{d}{dt}\left(\rho, \frac{|\mathbf{u}|^2}{2}\right) &= (\rho\mathbf{u}_t, \mathbf{u}) + (\rho_t, \frac{|\mathbf{u}|^2}{2})\\ &= (\rho\mathbf{u}_t, \mathbf{u}) + (-\nabla\cdot(\mathbf{u}\rho), \frac{|\mathbf{u}|^2}{2}) + (J, \frac{|\mathbf{u}|^2}{2})\\ &= (\rho\mathbf{u}_t, \mathbf{u}) + (\rho\mathbf{u}\cdot\nabla\mathbf{u}, \mathbf{u}) + (\frac{1}{2}J\mathbf{u}, \mathbf{u}).\end{aligned} \tag{2.10}$$

In the above derivations, we use the boundary condition $\mathbf{u}\cdot\mathbf{n}|_{\partial\Omega} = 0$ where $\mathbf{n}$ is the outward-pointing unit normal vector on the domain boundary $\partial\Omega$.

Hence, we obtain the nonlocal volume-conserved hydrodynamically coupled elastic bending energy system with non-matched density and viscosity for lipid vesicles as follows:

$$\phi_t + \nabla\cdot(\mathbf{u}\phi) = -\gamma\left(\mu - \frac{1}{|\Omega|}\int_\Omega \mu d\boldsymbol{x}\right), \tag{2.11}$$

$$\mu = \lambda\epsilon(\Delta - f'(\phi))(\Delta\phi - f(\phi)) + \lambda\epsilon M(A(\phi) - \beta)(-\Delta\phi + f(\phi)), \tag{2.12}$$

$$\rho\mathbf{u}_t + \rho(\mathbf{u}\cdot\nabla)\mathbf{u} + \frac{1}{2}J\mathbf{u} - \nabla\cdot(\nu(\phi)D(\mathbf{u})) + \nabla p + \phi\nabla\mu = 0, \tag{2.13}$$

$$\nabla\cdot\mathbf{u} = 0, \tag{2.14}$$

with $D(\mathbf{u}) = \nabla\mathbf{u} + (\nabla\mathbf{u})^T$.

We consider one of the following two types of boundary conditions:

$$\text{(i) all variables are periodic, or (ii) } \mathbf{u}|_{\partial\Omega} = \mathbf{0},\ \partial_{\mathbf{n}}\phi|_{\partial\Omega} = \partial_{\mathbf{n}}\Delta\phi|_{\partial\Omega} = 0. \tag{2.15}$$

The initial conditions of the system read as

$$\mathbf{u}|_{(t=0)} = \mathbf{u}^0,\ p|_{(t=0)} = p^0,\ \phi|_{(t=0)} = \phi^0. \tag{2.16}$$

The system (2.11)–(2.14) admits the law of energy dissipation which is derived in the following lemma.

Lemma 2.1. *The following energy law holds for the variable density system (2.11)–(2.14):*

$$\frac{d}{dt}E(\rho,\mathbf{u},\phi) = -\gamma\left\|\mu - \frac{1}{|\Omega|}\int_\Omega \mu d\boldsymbol{x}\right\|^2 - \frac{1}{2}\int_\Omega \nu|D(\mathbf{u})|^2 d\boldsymbol{x} \le 0, \tag{2.17}$$

where $E(\rho,\mathbf{u},\phi)$ is given in (2.1).

Proof. By multiplying the inner product of (2.11) with μ in L^2, we derive

$$(\phi_t,\mu) + \gamma\left\|\mu - \frac{1}{|\Omega|}\int_\Omega \mu d\boldsymbol{x}\right\|^2 = \underbrace{\int_\Omega \phi\mathbf{u}\cdot\nabla\mu d\boldsymbol{x}}_{\text{I}}, \tag{2.18}$$

where we use

$$\begin{aligned}
&(\mu - \frac{1}{|\Omega|}\int_\Omega \mu d\boldsymbol{x}, \mu)\\
&= (\mu - \frac{1}{|\Omega|}\int_\Omega \mu d\boldsymbol{x}, \mu - \frac{1}{|\Omega|}\int_\Omega \mu d\boldsymbol{x}) + (\mu - \frac{1}{|\Omega|}\int_\Omega \mu d\boldsymbol{x}, \frac{1}{|\Omega|}\int_\Omega \mu d\boldsymbol{x})\\
&= \left\|\mu - \frac{1}{|\Omega|}\int_\Omega \mu d\boldsymbol{x}\right\|^2,
\end{aligned}$$

since $(\mu - \frac{1}{|\Omega|}\int_\Omega \mu d\boldsymbol{x}, 1) = 0$.

Taking the inner product of (2.4) with $-\phi_t$ in L^2, we get

$$-(\mu,\phi_t) + \lambda\frac{d}{dt}\left(\int_\Omega \frac{\epsilon}{2}(\Delta\phi - f(\phi))^2\, d\boldsymbol{x} + \frac{1}{2}M(A(\phi)-\beta)^2\right) = 0. \tag{2.19}$$

By taking the L^2 inner product of (2.13) with $\mathbf{u}$ and using integration by parts and the divergence-free condition (2.14), we obtain

$$(\rho\mathbf{u}_t,\mathbf{u}) = -\frac{1}{2}\int_\Omega \nu|D(\mathbf{u})|^2 d\boldsymbol{x} - \underbrace{\int_\Omega (\phi\nabla\mu\cdot\mathbf{u}) d\boldsymbol{x}}_{\text{II}} - \underbrace{\int_\Omega \rho(\mathbf{u}\cdot\nabla)\mathbf{u}\cdot\mathbf{u} d\boldsymbol{x}}_{\text{III}} - \frac{1}{2}\int_\Omega J\mathbf{u}\cdot\mathbf{u} d\boldsymbol{x}. \tag{2.20}$$

We multiply (2.9) with $\frac{1}{2}\mathbf{u}$ to derive

$$\frac{1}{2}\rho_t\mathbf{u} + \frac{1}{2}\nabla\cdot(\rho\mathbf{u})\mathbf{u} - \frac{1}{2}J\mathbf{u} = 0. \tag{2.21}$$

Hence, we take the L^2 inner product of (2.21) with $\mathbf{u}$ to derive

$$(\rho_t, \frac{1}{2}|\mathbf{u}|^2) = -\underbrace{\frac{1}{2}\int_\Omega \nabla\cdot(\rho\mathbf{u})\mathbf{u}\cdot\mathbf{u} d\boldsymbol{x}}_{\text{IV}} + \frac{1}{2}\int_\Omega J\mathbf{u}\cdot\mathbf{u} d\boldsymbol{x}. \tag{2.22}$$

By combining (2.18), (2.19), (2.20), and (2.22), we obtain (2.17). □

Remark 2.1. In the process of deducing the PDE energy law (2.17), many nonlinear terms are cancelled, such as

$$\text{I and II}: \quad \int_\Omega \phi \mathbf{u}\cdot\nabla\mu d\boldsymbol{x} - \int_\Omega \phi\nabla\mu\cdot\mathbf{u}d\boldsymbol{x} = 0, \tag{2.23}$$

$$\text{III and IV}: \quad \int_\Omega \rho(\mathbf{u}\cdot\nabla)\mathbf{u}\cdot\mathbf{u}d\boldsymbol{x} + \frac{1}{2}\int_\Omega \nabla\cdot(\rho\mathbf{u})\mathbf{u}\cdot\mathbf{u}d\boldsymbol{x} = 0, \tag{2.24}$$

which are derived by using integration by parts and the boundary condition $\mathbf{u}\cdot\mathbf{n}|_{\partial\Omega} = 0$ (or periodic). This means that despite the complexity of these terms, they do not contribute to the energy diffusivity. Such a kind of cancellation during the process of deriving the energy law can be referred as "zero-energy-contribution" property (see the explicit-auxiliary variable method developed in [38,39,58–61] for the matched density model). It provides some inspiration when we aim to develop decoupling type schemes, see the next section.

Remark 2.2. For completeness, we also present the total free energy of the classical phase-field bending energy model developed in [7–9]. In addition to the penalty potential that enforces conservation of surface area, conservation of volume is also approximated by a similar penalty potential. Namely, the energy reads as

$$\tilde{E}(\mathbf{u},\phi) = E(\mathbf{u},\phi) + \frac{1}{2}\widetilde{M}(V(\phi) - V(\phi^0))^2, \tag{2.25}$$

where $V(\phi) = \int_\Omega \frac{\phi+1}{2}d\boldsymbol{x}$, and $\widetilde{M} \gg 1$ is a penalty parameter to enforce volume conservation (approximately). That is, no matter how big $\widetilde{M}$ is, the vesicle volume is preserved over time in an approximate manner.

3. Numerical scheme

In this section, we try to develop a time-marching algorithm for the variable-density/viscosity vesicle model (2.11)–(2.14) with the focus on achieving a decoupling type scheme. We also expect that the scheme can be linear, provably solvable, and unconditionally energy stable. The detailed process is as follows.

3.1. Equivalent form of system

We notice that one of the main challenges of numerical discretizations is on the nonlinear term μ in (2.12). Using $f'(\phi) = \frac{1}{\epsilon^2}(3\phi^2 - 1)$, we rewrite the energy potential part A in (2.1) as follows,

$$\begin{aligned}\epsilon\int_\Omega \Big(\frac{1}{2}(\Delta\phi - f(\phi))\Big)^2 d\boldsymbol{x} &= \epsilon\int_\Omega \underbrace{\Big(\frac{1}{2}|\Delta\phi|^2 - \frac{1}{\epsilon^2}|\nabla\phi|^2 + \frac{e}{2}\phi^2\Big)}_{\text{linear}} d\boldsymbol{x} \\ &+ \epsilon\int_\Omega \underbrace{\Big(\frac{3}{\epsilon^2}\phi^2|\nabla\phi|^2 + \frac{1}{2}f(\phi)^2 - \frac{e}{2}\phi^2\Big)}_{\text{nonlinear}} d\boldsymbol{x},\end{aligned} \tag{3.1}$$

where e is a positive constant. Remarkably, we add a zero term $\frac{e}{2}\phi^2 - \frac{e}{2}\phi^2$ in the total free energy. Since $f(\phi)^2$ is a sixth-order polynomial with the positive leading term, it can bound the negative term $-\frac{e}{2}\phi^2$ from below for any e.

First, we define an auxiliary variable $U(\boldsymbol{x},t)$ (local) to "quadratize" the nonlinear part of (3.1) as

$$U(\boldsymbol{x},t) = \sqrt{\frac{3}{\epsilon^2}\phi^2|\nabla\phi|^2 + \frac{1}{2}f(\phi)^2 - \frac{e}{2}\phi^2 + B_1}, \tag{3.2}$$

where the constant $B_1 > 0$ is added to ensure the radicand further positive. Since the nonlinear part in (3.1) is bounded from below for any e, the existence of the constant B_1 is obvious.

For the surface area potential (part C in (2.1)), we further define another auxiliary variable $V(t)$ (nonlocal) to quadratize it as

$$V(t) = \sqrt{\frac{1}{2\epsilon}M(A(\phi) - \beta)^2 + B_2}, \tag{3.3}$$

where $B_2 > 0$ is added to make the radicand further positive.

The usage of auxiliary variables U and V are the so-called IEQ/SAV method which is an efficient method to linearize the nonlinear terms induced by the energy potentials, see [37,62–65]. Using the two variables U and V, we reformulate (2.12) to the following form:

$$\mu = \lambda\epsilon\left(\Delta^2\phi + e\phi + \frac{2}{\epsilon^2}\Delta\phi\right) + \lambda\epsilon H U + \lambda\epsilon V K, \tag{3.4}$$

$$U_t = \frac{1}{2}H\phi_t, \tag{3.5}$$

$$V_t = \frac{1}{2}(K, \phi_t), \tag{3.6}$$

where the two functions $H(\phi)$ and $K(\phi)$ are defined as:

$$\begin{cases} H(\phi) = \dfrac{\frac{6}{\epsilon^2}(\phi|\nabla\phi|^2 - \nabla\cdot(\phi^2|\nabla\phi|)) + f(\phi)f'(\phi) - e\phi}{\sqrt{\frac{3}{\epsilon^2}\phi^2|\nabla\phi|^2 + \frac{1}{2}f(\phi)^2 - \frac{e}{2}\phi^2 + B_1}}, \\ K(\phi) = \dfrac{M(A(\phi)-\beta)(-\Delta\phi + f(\phi))}{\sqrt{\frac{1}{2\epsilon}M(A(\phi)-\beta)^2 + B_2}}. \end{cases} \tag{3.7}$$

Here, Eq. (2.12) is reformulated by using the new variables U, V, and the two ODEs ((3.5) and (3.6)) are obtained by taking the time derivative of U and V.

Second, we introduce a variable $Q(t)$ (nonlocal) and an ODE system related to it, that reads as:

$$\begin{cases} Q_t = (\nabla\cdot(\mathbf{u}\phi), \mu) + (\phi\nabla\mu, \mathbf{u}) + \lambda\epsilon(UH, \phi_t) - \lambda\epsilon(H\phi_t, U) \\ \qquad\qquad + \lambda\epsilon V(K, \phi_t) - \lambda\epsilon(K, \phi_t)V, \\ Q|_{(t=0)} = 1, \\ \mathbf{u}|_{\partial\Omega} = \mathbf{0}, \text{ or all variables are periodic.} \end{cases} \tag{3.8}$$

By utilizing the "*zero-energy-contribution*" property (2.23), it is straightforward to see that the ODE (3.8) is the same as the trivial ODE ($Q_t = 0$, $Q|_{(t=0)} = 1$), which has $Q(t) = 1$ as the exact solution.

Third, we introduce another nonlocal variable $R(t)$ (nonlocal) and an ODE system related to it, that reads as:

$$\begin{cases} R_t = (\rho(\mathbf{u}\cdot\nabla)\mathbf{u}, \mathbf{u}) + \dfrac{1}{2}(\nabla\cdot(\rho\mathbf{u})\mathbf{u}, \mathbf{u}), \\ R|_{(t=0)} = 1, \\ \mathbf{u}|_{\partial\Omega} = \mathbf{0}, \text{ or all variables are periodic.} \end{cases} \tag{3.9}$$

Similarly, by using (2.24), one can see that (3.9) is equivalent to a trivial ODE ($R_t = 0$, $R|_{(t=0)} = 1$), which has $R(t) = 1$ as the exact solution.

Finally, using the new variables U, V, Q and R, we rewrite the PDE system (2.11)–(2.14) as:

$$\phi_t + \underbrace{Q\nabla\cdot(\mathbf{u}\phi)}_{\text{Q-reform}} + \gamma\left(\mu - \frac{1}{|\Omega|}\int_\Omega \mu d\boldsymbol{x}\right) = 0, \tag{3.10}$$

$$\mu = \lambda\epsilon(\Delta^2\phi + e\phi + \frac{2}{\epsilon^2}\Delta\phi) + \underbrace{\lambda\epsilon QHU}_{\text{Q-reform}} + \underbrace{\lambda\epsilon QVK}_{\text{Q-reform}}, \tag{3.11}$$

$$U_t = \frac{1}{2}\underbrace{QH\phi_t}_{\text{Q-reform}}, \tag{3.12}$$

$$V_t = \frac{1}{2}\underbrace{Q(K, \phi_t)}_{\text{Q-reform}}, \tag{3.13}$$

$$\rho\mathbf{u}_t + \underbrace{R\rho(\mathbf{u}\cdot\nabla)\mathbf{u}}_{\text{R-reform}} + \frac{1}{2}J\mathbf{u} + \overbrace{\frac{1}{2}\rho_t\mathbf{u} + \underbrace{\frac{1}{2}R\nabla\cdot(\rho\mathbf{u})\mathbf{u}}_{\text{R-reform}} - \frac{1}{2}J\mathbf{u}}^{cf.\ (2.21)} \tag{3.14}$$

$$-\nabla\cdot(\nu D(\mathbf{u}))+\nabla p+\underbrace{Q\phi\nabla\mu}_{\text{Q-reform}}=0,$$

$$\nabla\cdot\mathbf{u}=0, \tag{3.15}$$

$$Q_t=(\nabla\cdot(\mathbf{u}\phi),\mu)+(\phi\nabla\mu,\mathbf{u})+\lambda\epsilon(UH,\phi_t)-\lambda\epsilon(H\phi_t,U) \tag{3.16}$$
$$+\lambda\epsilon V(K,\phi_t)-\lambda\epsilon(K,\phi_t)V,$$

$$R_t=(\rho(\mathbf{u}\cdot\nabla)\mathbf{u},\mathbf{u})+\frac{1}{2}(\nabla\cdot(\rho\mathbf{u})\mathbf{u},\mathbf{u}), \tag{3.17}$$

with the boundary conditions that read as

$$\text{(i) all variables are periodic, or (ii) } \mathbf{u}|_{\partial\Omega}=\mathbf{0},\ \partial_{\mathbf{n}}\phi|_{\partial\Omega}=\partial_{\mathbf{n}}\Delta\phi|_{\partial\Omega}=0, \tag{3.18}$$

and initial conditions that read as

$$\mathbf{u}|_{(t=0)}=\mathbf{u}^0,\ p|_{(t=0)}=p^0,\ \phi|_{(t=0)}=\phi^0,\ Q|_{(t=0)}=1,\ R|_{(t=0)}=1,\ U|_{(t=0)}=U(\phi^0). \tag{3.19}$$

Remark 3.1. Using the combinations of the new variables and their associated ODEs, some modifications are applied to the original system (2.11)–(2.14) and a new system (3.10)–(3.17) is formed. The terms $\nabla\cdot(\mathbf{u}\phi)$ and $\phi\nabla\mu$ are multiplied by Q, and the term $\rho(\mathbf{u}\cdot\nabla)\mathbf{u}$ is multiplied by R. This kind of modification of multiplying Q or R does not bring up any changes to the system since $Q(t)=1$ and $R(t)=1$. Second, in the fluid momentum equation, we add $\frac{1}{2}\rho_t\mathbf{u}+\frac{1}{2}R\nabla\cdot(\rho\mathbf{u})\mathbf{u}+\frac{1}{2}J\mathbf{u}$, which is actually zero using the fact of $R(t)=1$ and (2.21). In addition, in (3.14), the two terms $(\frac{1}{2}J\mathbf{u}-\frac{1}{2}J\mathbf{u})$ will be offset, so they will no longer appear in the algorithm given in next section. Finally, we multiply the nonlocal variable Q to the nonlinear terms (the under braced terms by "Q-reform") in (3.11)–(3.12)–(3.13) due to $Q(t)=1$.

Remark 3.2. When developing numerical algorithms, if the algorithm design of the original system is very challenging, a commonly used strategy is to use some techniques to rewrite it to an equivalent but relatively simple system for discretization. This strategy is used in this article for the aim of developing a scheme with a fully-decoupled structure and unconditional energy stability. This strategy had also been adopted in algorithm designs of many other models. For example, the advective term in the Navier–Stokes equation with variable density is reformulated for stability reasons (cf. [42,57]); the penalty method reformulates the divergence-free condition of the velocity field of the Navier–Stokes equation to avoid solving the pressure Poisson equation with variable coefficients (cf. [43,45,53]); the energy potentials in gradient flow models had been reformulated in the IEQ or SAV methods (cf. [36,50,62,63,66,67]) for the reasons of obtaining linear schemes, etc. Based on this basic logic, we reformulate the original PDE system to obtain an equivalent system, which is relatively easier to obtain the desired property of numerical schemes, i.e., linear, fully-decoupled, and unconditionally energy stable.

Using a similar process obtaining the energy law (2.17), the new system (3.10)–(3.17) can be also proven to hold the similar energy dissipation law. Since the energy stability proof process of the discrete algorithm follows the same line, the detailed process are given in the following lemma to make it clear.

Lemma 3.1. *The PDE system (3.10)–(3.17) holds a law of energy dissipation:*

$$\frac{d}{dt}E_{tot}(\rho,\phi,\mathbf{u},U,V,Q,R)=-\gamma\left\|\mu-\frac{1}{|\Omega|}\int_\Omega\mu d\boldsymbol{x}\right\|^2-\frac{1}{2}\int_\Omega\nu|D(\mathbf{u})|^2d\boldsymbol{x}\le 0, \tag{3.20}$$

where E_{tot} reads as

$$E_{tot}(\rho,\phi,\mathbf{u},U,V,Q,R)=\int_\Omega\frac{1}{2}\rho|\mathbf{u}|^2d\boldsymbol{x}+\lambda\epsilon\Big(\frac{1}{2}\|\Delta\phi\|^2+\frac{e}{2}\|\phi\|^2-\frac{1}{\epsilon^2}\|\nabla\phi\|^2 \tag{3.21}$$
$$+\|U\|^2+|V|^2\Big)+\frac{|Q|^2}{2}+\frac{|R|^2}{2}-B_1|\Omega|-B_2-1.$$

Meanwhile, E_{tot} is bounded from below when the parameter e is sufficiently large.

Proof. We multiply the inner product of (3.14) with $\mathbf{u}$ i L^2 space, and using integration by parts and the divergence-free condition (3.15), we obtain

$$\begin{aligned}\frac{d}{dt}\int_\Omega \frac{1}{2}\rho|\mathbf{u}|^2 d\boldsymbol{x} = & -\frac{1}{2}\int_\Omega \nu|D(\mathbf{u})|^2 d\boldsymbol{x} - \underbrace{Q(\phi\nabla\mu,\mathbf{u})}_{\mathrm{I}_1} \\ & - \underbrace{R(\rho(\mathbf{u}\cdot\nabla)\mathbf{u},\mathbf{u})}_{\mathrm{II}_1} - \underbrace{\frac{1}{2}R(\nabla\cdot(\rho\mathbf{u})\mathbf{u},\mathbf{u})}_{\mathrm{III}_1}.\end{aligned} \tag{3.22}$$

We multiply the inner product of (3.10) by μ in L^2 to derive

$$(\phi_t,\mu) = -\gamma\left\|\mu - \frac{1}{|\Omega|}\int_\Omega \mu d\boldsymbol{x}\right\|^2 - \underbrace{Q(\nabla\cdot(\mathbf{u}\phi),\mu)}_{\mathrm{IV}_1}. \tag{3.23}$$

Taking the inner product of (3.11) with $-\phi_t$ in L^2, we get

$$-(\mu,\phi_t) = -\lambda\epsilon\frac{d}{dt}\Big(\frac{1}{2}\|\Delta\phi\|^2 + \frac{e}{2}\|\phi\|^2 - \frac{1}{\epsilon^2}\|\nabla\phi\|^2\Big) \underbrace{-\lambda\epsilon Q(UH,\phi_t)}_{\mathrm{V}_1} \underbrace{-\lambda\epsilon QV(K,\phi_t)}_{\mathrm{VI}_1}. \tag{3.24}$$

By taking the L^2 inner product of (3.12) with $2\lambda\epsilon U$, we obtain

$$\frac{d}{dt}\Big(\lambda\epsilon\|U\|^2\Big) = \underbrace{\lambda\epsilon Q(H\phi_t,U)}_{\mathrm{VII}_1}. \tag{3.25}$$

By multiplying (3.13) with $2\lambda\epsilon V$, we obtain

$$\frac{d}{dt}\Big(\lambda\epsilon|V|^2\Big) = \underbrace{\lambda\epsilon Q(K,\phi_t)V}_{\mathrm{VIII}_1}. \tag{3.26}$$

By multiplying (3.16) with Q, we obtain

$$\begin{aligned}\frac{d}{dt}\Big(\frac{|Q|^2}{2}\Big) = & \underbrace{Q(\nabla\cdot(\mathbf{u}\phi),\mu)}_{\mathrm{IV}_2} + \underbrace{Q(\phi\nabla\mu,\mathbf{u})}_{\mathrm{I}_2} \\ & + \underbrace{\lambda\epsilon Q(UH,\phi_t)}_{\mathrm{V}_2} - \underbrace{\lambda\epsilon Q(H\phi_t,U)}_{\mathrm{VII}_2} + \underbrace{\lambda\epsilon QV(K,\phi_t)}_{\mathrm{VI}_2} \underbrace{-\lambda\epsilon Q(K,\phi_t)V}_{\mathrm{VIII}_2}.\end{aligned} \tag{3.27}$$

By multiplying (3.17) with R, we obtain

$$\frac{d}{dt}\Big(\frac{|R|^2}{2}\Big) = \underbrace{R(\rho(\mathbf{u}\cdot\nabla)\mathbf{u},\mathbf{u})}_{\mathrm{II}_2} + \underbrace{\frac{1}{2}R(\nabla\cdot(\rho\mathbf{u})\mathbf{u},\mathbf{u})}_{\mathrm{III}_2}. \tag{3.28}$$

By combining all obtained Eqs. (3.23)–(3.28), and noting that the two terms marked with the same Roman numerals cancel each other out, we obtain the energy law (3.20).

We further show that the reformulated total free energy E_{tot} is bounded from below. By using the boundary condition (2.15) and the Cauchy–Schwarz inequality, we derive

$$\|\nabla\phi\|^2 = (\nabla\phi,\nabla\phi) = -(\Delta\phi,\phi) \le \frac{\eta}{2}\|\Delta\phi\|^2 + \frac{1}{2\eta}\|\phi\|^2, \forall\eta > 0. \tag{3.29}$$

Hence, we deduce

$$\frac{1}{2}\|\Delta\phi\|^2 + \frac{e}{2}\|\phi\|^2 - \frac{1}{\epsilon^2}\|\nabla\phi\|^2 \ge (\frac{1}{2} - \frac{\eta}{2\epsilon^2})\|\Delta\phi\|^2 + (\frac{e}{2} - \frac{1}{2\eta\epsilon^2})\|\phi\|^2. \tag{3.30}$$

We note that the parameter e is adjustable since in the definition of $U(\boldsymbol{x},t)$ in (3.2), $f(\phi)^2$ is the sixth-order polynomial and thus it can always bound the term $-\frac{e}{2}\phi^2$ from below for any e. Therefore, as long as we adjust the magnitude of η and e, we can always get $\frac{1}{2}\|\Delta\phi\|^2+\frac{e}{2}\|\phi\|^2-\frac{1}{\epsilon^2}\|\nabla\phi\|^2\geq 0$. For instance, we let $\eta=\frac{\epsilon^2}{2}, e=\frac{4}{\epsilon^4}$, then

$$\frac{1}{2}\|\Delta\phi\|^2+\frac{e}{2}\|\phi\|^2-\frac{1}{\epsilon^2}\|\nabla\phi\|^2\geq\frac{1}{4}\|\Delta\phi\|^2+\frac{1}{\epsilon^4}\|\phi\|^2\geq 0. \tag{3.31}$$

Therefore, E_{tot} is shown to be bounded from below when the parameter e is large enough. □

Remark 3.3. It is worth noting that when $e\sim O(\frac{1}{\epsilon^4})$, its magnitude is the same as the nonlinear term $f(\phi)f'(\phi)$ in (2.4) or in $H(\phi)$. This means that in (3.11), the linear term $e\phi$ acts as a stabilizer to balance the nonlinear term $f(\phi)f'(\phi)$ that is embedded in QHU.

3.2. Numerical scheme

Now, in this subsection, we construct the time-marching algorithm for (3.10)–(3.17) as follows. We let $\delta t>0$ be a time step size, $t^n=n\delta t$ for $0\leq n\leq N$ with $T=N\delta t$, ψ^n be the numerical approximation to the analytic function $\psi(\cdot,t)|_{t=t^n}$. Given $\mathbf{u}^n$, p^n, μ^n, ϕ^n, ϕ^n, Q^n, R^n, U^n, V^n, we update $\mathbf{u}^{n+1}$, p^{n+1}, μ^{n+1}, ϕ^{n+1}, ρ^{n+1}, Q^{n+1}, R^{n+1}, U^{n+1}, V^{n+1} by the following:

Step 1:

$$\frac{\phi^{n+1}-\phi^n}{\delta t}+Q^{n+1}\nabla\cdot(\phi^n\mathbf{u}^n)=-\gamma\Big(\mu^{n+1}-\frac{1}{|\Omega|}\int_\Omega\mu^{n+1}d\boldsymbol{x}\Big), \tag{3.32}$$

$$\mu^{n+1}=\lambda\epsilon\Big(\Delta^2\phi^{n+1}+e\phi^{n+1}+\frac{2}{\epsilon^2}\Delta\phi^{n+1}\Big)+\lambda\epsilon Q^{n+1}H^nU^n+\lambda\epsilon Q^{n+1}V^nK^n, \tag{3.33}$$

$$\frac{U^{n+1}-U^n}{\delta t}=\frac{1}{2}Q^{n+1}H^n\phi_t^n, \tag{3.34}$$

$$\frac{V^{n+1}-V^n}{\delta t}=\frac{1}{2}Q^{n+1}(K^n,\phi_t^n), \tag{3.35}$$

$$\rho^n\frac{\tilde{\mathbf{u}}^{n+1}-\mathbf{u}^n}{\delta t}+Q^{n+1}\phi^n\nabla\mu^n=0, \tag{3.36}$$

$$\begin{aligned}\frac{Q^{n+1}-Q^n}{\delta t}&=(\nabla\cdot(\mathbf{u}^n\phi^n),\mu^{n+1})+(\phi^n\nabla\mu^n,\tilde{\mathbf{u}}^{n+1})\\&\quad+\lambda\epsilon(U^nH^n,\frac{\phi^{n+1}-\phi^n}{\delta t})-\lambda\epsilon(H^n\phi_t^n,U^{n+1})\\&\quad+\lambda\epsilon V^n(K^n,\frac{\phi^{n+1}-\phi^n}{\delta t})-\lambda\epsilon(K^n,\phi_t^n)V^{n+1}.\end{aligned} \tag{3.37}$$

Step 2:

$$\begin{aligned}&\rho^n\frac{\mathbf{u}^{n+1}-\tilde{\mathbf{u}}^{n+1}}{\delta t}+\frac{1}{2}\frac{\rho^{n+1}-\rho^n}{\delta t}\mathbf{u}^{n+1}-\nabla\cdot(\nu^{n+1}D(\mathbf{u}^{n+1}))+\nabla(2p^n-p^{n-1})\\&\quad+R^{n+1}\rho^n(\mathbf{u}^n\cdot\nabla)\mathbf{u}^n+\frac{1}{2}R^{n+1}\nabla\cdot(\rho^n\mathbf{u}^n)\mathbf{u}^n=0,\end{aligned} \tag{3.38}$$

$$\frac{R^{n+1}-R^n}{\delta t}=(\rho^n(\mathbf{u}^n\cdot\nabla)\mathbf{u}^n,\mathbf{u}^{n+1})+\frac{1}{2}(\nabla\cdot(\rho^n\mathbf{u}^n)\mathbf{u}^n,\mathbf{u}^{n+1}). \tag{3.39}$$

Step 3:

$$\Delta(p^{n+1}-p^n)=\frac{\chi}{\delta t}\nabla\cdot\mathbf{u}^{n+1}. \tag{3.40}$$

The notations used in the scheme read as

$$\begin{cases} H^n = H(\phi^n),\ K^n = K(\phi^n),\ \phi_t^n = \dfrac{\phi^n - \phi^{n-1}}{\delta t}, \\ \chi = \dfrac{1}{2}\min(\rho_1, \rho_2), \\ \hat{\phi} = \begin{cases} \phi, & |\phi| \le 1, \\ \text{sign}(\phi), & |\phi| \rangle 1, \end{cases} \\ \rho^{n+1} = \dfrac{\rho_1 - \rho_2}{2}\hat{\phi}^{n+1} + \dfrac{\rho_1 + \rho_2}{2},\ \nu^{n+1} = \dfrac{\nu_1 - \nu_2}{2}\hat{\phi}^{n+1} + \dfrac{\nu_1 + \nu_2}{2}. \end{cases} \tag{3.41}$$

The boundary conditions are either periodic for all variables or

$$\mathbf{u}^{n+1}|_{\partial\Omega} = \mathbf{0},\ \partial_{\mathbf{n}}\phi^{n+1}|_{\partial\Omega} = \partial_{\mathbf{n}}\Delta\phi^{n+1}|_{\partial\Omega} = \partial_{\mathbf{n}}p^{n+1}|_{\partial\Omega} = 0. \tag{3.42}$$

We give some detailed explanations and comments on the scheme (3.32)–(3.40) in the following remarks.

Remark 3.4. In step 1, we note that the term $Q^{n+1}\phi^n\nabla\mu^n$ (surface tension term) is actually extracted from the fluid equation, which is the application of the first-order operator splitting method. The advantage of using this method is to decouple the computation of the phase-field variable ϕ^{n+1} from the momentum equation. We recall that in [52–55], a similar splitting method had been applied as well. But there exists some essential difference between the method therein and our developed scheme of this article on the discretization approach for coupled nonlinear terms. By taking the surface tension term $\phi\nabla\mu$ as an example, it was discretized as $\phi^n\nabla\mu^{n+1}$ in [52–55], namely, by using the combination of implicit and explicit approaches. In our scheme, we discretize it as $Q^{n+1}\phi^n\nabla\mu^n$ where ϕ and $\nabla\mu$ are both discretized explicitly. From the final structure, although the scheme given in [52–55] can also generate a decoupling scheme, due to the implicit processing of μ^{n+1}, we have to solve a variable-coefficient phase-field equation at each time step, resulting in costly computations in practice.

Remark 3.5. Due to the discretization method used in (3.37) and (3.39), it is easy to see that Q^{n+1} and R^{n+1} are no longer equal to 1, because the discrete form of nonlinear terms contained in (3.37) and (3.39) are no longer equal to zero. For the benefit of the reader, we explain here why we allow this to happen. As we all know, the numerical solution is only an approximation of the exact solution of the original PDE system with a certain accuracy. Therefore, in the actual calculations, Q^{n+1} and R^{n+1} will be the numerical approximations of $Q(t)|_{(t=t^{n+1})}$, $R(t)|_{(t=t^{n+1})}$. This is similar to the fact of that $\mathbf{u}^{n+1}$ is the approximation of $\mathbf{u}(t)|_{(t=t^{n+1})}$, because ψ^{n+1} is numerical approximation to $\psi|_{(t=t^{n+1})}$ for any ψ. Namely, Q^{n+1} and R^{n+1} are the approximate solutions to $Q(t)|_{(t=t^{n+1})}$ and $R(t)|_{(t=t^{n+1})}$ in the first-order temporal accuracy, as shown in the accuracy test in Section 4, cf. Fig. 4.1 (a).

Remark 3.6. The penalty method is a widely-used approach to solve the variable-density Navier–Stokes equations, see [43,53]. Its advantage is to avoid solving the variable-coefficient $(1/\rho)$ pressure Poisson equation. Note that in our scheme, only one constant-coefficient pressure Poisson equation needs to be solved, which is very efficient in practice.

Remark 3.7. The scheme is linear. It seems that the two terms $\rho^{n+1}\frac{\mathbf{u}^{n+1}-\mathbf{u}^n}{\delta t}$ and $\nabla\cdot(\nu^{n+1}D(\mathbf{u}^{n+1}))$ in step 2 contains two implicit terms, but in fact, ϕ^{n+1} is already obtained in step 1, i.e., ρ^{n+1} and ν^{n+1} are known terms in step 2. Moreover, from the definition of the cut-off function $\hat{\phi}$ given in (3.41), we have

$$\rho^{n+1} \ge \min(\rho_1, \rho_2), \quad \nu^{n+1} \ge \min(\nu_1, \nu_2). \tag{3.43}$$

3.3. Decoupled implementation

Now, we give the process of how to implement the scheme (3.32)–(3.40). Since step 3 only includes a Poisson type Eq. (3.40) with constant coefficients, therefore we only discuss the implementation of step 1 and 2 here. The schemes in (3.32) and (3.39) contain many nonlocal terms, so it is usually time-consuming to solve them directly. Moreover, the four ODEs of U, V, Q, R need to be solved, which will also cause a lot of time to calculate. We

design the decoupling process as follows by using the nonlocal properties of auxiliary variables Q, R to realize the fully-decoupled structure. For example, since Q^{n+1} is a nonlocal scalar, any variable ψ can always be split into the form of

$$\psi = \psi_1 + Q^{n+1}\psi_2. \tag{3.44}$$

Using this linear combination form, we split all variables into multiple variables and then merge them back.

3.3.1. Implementation of step 1

First, we use the nonlocal variable Q^{n+1} to split ϕ^{n+1}, μ^{n+1}, $\tilde{\mathbf{u}}^{n+1}$, U^{n+1}, and V^{n+1} into a linear combination form that reads as

$$\begin{cases} \phi^{n+1} = \phi_1^{n+1} + Q^{n+1}\phi_2^{n+1},\ \mu^{n+1} = \mu_1^{n+1} + Q^{n+1}\mu_2^{n+1}, \\ \tilde{\mathbf{u}}^{n+1} = \tilde{\mathbf{u}}_1^{n+1} + Q^{n+1}\tilde{\mathbf{u}}_2^{n+1},\ U^{n+1} = U_1^{n+1} + Q^{n+1}U_2^{n+1}, \\ V^{n+1} = V_1^{n+1} + Q^{n+1}V_2^{n+1}. \end{cases} \tag{3.45}$$

Then the scheme (3.32), (3.33), and (3.36) can be rewritten as

$$\begin{cases} \dfrac{\phi_1^{n+1} + Q^{n+1}\phi_2^{n+1} - \phi^n}{\delta t} + Q^{n+1}\nabla\cdot(\mathbf{u}^n\phi^n) \\ \qquad + \gamma\left(\mu_1^{n+1} + Q^{n+1}\mu_2^{n+1} - \dfrac{1}{|\Omega|}\displaystyle\int_\Omega (\mu_1^{n+1} + Q^{n+1}\mu_2^{n+1})d\boldsymbol{x}\right) = 0, \\ \mu_1^{n+1} + Q^{n+1}\mu_2^{n+1} = \lambda\epsilon\left(\Delta^2(\phi_1^{n+1} + Q^{n+1}\phi_2^{n+1}) + e(\phi_1^{n+1} + Q^{n+1}\phi_2^{n+1})\right. \\ \qquad \left. + \dfrac{2}{\epsilon^2}\Delta(\phi_1^{n+1} + Q^{n+1}\phi_2^{n+1})\right) + \lambda\epsilon Q^{n+1}H^nU^n + \lambda\epsilon Q^{n+1}V^nK^n, \\ \rho^n\dfrac{\tilde{\mathbf{u}}_1^{n+1} + Q^{n+1}\tilde{\mathbf{u}}_2^{n+1} - \mathbf{u}^n}{\delta t} + Q^{n+1}\phi^n\nabla\mu^n = 0. \end{cases} \tag{3.46}$$

According to Q^{n+1}, the three equations in the system (3.46) can be split into three sub-systems as

$$\begin{cases} \dfrac{1}{\delta t}\phi_1^{n+1} + \gamma\left(\mu_1^{n+1} - \dfrac{1}{|\Omega|}\displaystyle\int_\Omega \mu_1^{n+1}d\boldsymbol{x}\right) = \dfrac{\phi^n}{\delta t}, \\ \mu_1^{n+1} = \lambda\epsilon(\Delta^2\phi_1^{n+1} + e\phi_1^{n+1} + \dfrac{2}{\epsilon^2}\Delta\phi_1^{n+1}), \end{cases} \tag{3.47}$$

and

$$\begin{cases} \dfrac{1}{\delta t}\phi_2^{n+1} + \gamma\left(\mu_2^{n+1} - \dfrac{1}{|\Omega|}\displaystyle\int_\Omega \mu_2^{n+1}d\boldsymbol{x}\right) = -\nabla\cdot(\mathbf{u}^n\phi^n), \\ \mu_2^{n+1} = \lambda\epsilon(\Delta^2\phi_2^{n+1} + e\phi_2^{n+1} + \dfrac{2}{\epsilon^2}\Delta\phi_2^{n+1}) + \lambda\epsilon H^nU^n + \lambda\epsilon V^nK^n, \end{cases} \tag{3.48}$$

$$\begin{cases} \tilde{\mathbf{u}}_1^{n+1} = \mathbf{u}^n, \\ \rho^n\dfrac{\tilde{\mathbf{u}}_2^{n+1}}{\delta t} + \phi^n\nabla\mu^n = 0. \end{cases} \tag{3.49}$$

By taking the L^2 inner product of the first equation in (3.47) and (3.48) with 1, we immediately get

$$\int_\Omega \phi_1^{n+1}d\boldsymbol{x} = \int_\Omega \phi^n d\boldsymbol{x} = \cdots = \int_\Omega \phi^0 d\boldsymbol{x}, \qquad \int_\Omega \phi_2^{n+1}d\boldsymbol{x} = 0. \tag{3.50}$$

The boundary conditions of the ϕ_1^{n+1} and ϕ_2^{n+1} are either periodic or

$$\partial_{\mathbf{n}}\phi_1^{n+1}|_{\partial\Omega} = \partial_{\mathbf{n}}\Delta\phi_1^{n+1}|_{\partial\Omega} = 0,\ \partial_{\mathbf{n}}\phi_2^{n+1}|_{\partial\Omega} = \partial_{\mathbf{n}}\Delta\phi_2^{n+1}|_{\partial\Omega} = 0. \tag{3.51}$$

In fact, the two equations in (3.47) and (3.48) can be combined by eliminating μ_1^{n+1} and μ_2^{n+1} together so that one does not need to solve a coupled system. More precisely, using (3.50), we find that (3.47) and (3.48) become

$$\begin{cases} \dfrac{1}{\delta t}\phi_1^{n+1} + \lambda\epsilon\gamma(\Delta^2\phi_1^{n+1} + e\phi_1^{n+1} + \dfrac{2}{\epsilon^2}\Delta\phi_1^{n+1}) = G_1^n, \\ \dfrac{1}{\delta t}\phi_2^{n+1} + \lambda\epsilon\gamma(\Delta^2\phi_2^{n+1} + e\phi_2^{n+1} + \dfrac{2}{\epsilon^2}\Delta\phi_2^{n+1}) = G_2^n, \end{cases} \tag{3.52}$$

where

$$\begin{cases} G_1^n = \dfrac{\phi^n}{\delta t} + \dfrac{\lambda\gamma\epsilon e}{|\Omega|}\displaystyle\int_\Omega \phi^0 d\boldsymbol{x}, \\ G_2^n = -\nabla\cdot(\mathbf{u}^n\phi^n) - \lambda\epsilon\gamma\Big(H^nU^n - \dfrac{1}{|\Omega|}\displaystyle\int_\Omega H^nU^n d\boldsymbol{x}\Big) - \lambda\epsilon\gamma V^n\Big(K^n - \dfrac{1}{|\Omega|}\displaystyle\int_\Omega K^n d\boldsymbol{x}\Big). \end{cases} \tag{3.53}$$

We can solve the two linear and constant-coefficient biharmonic equations in (3.52) directly to obtain ϕ_1^{n+1} and ϕ_2^{n+1}.

Second, for the system (3.34)–(3.35), we use (3.45) to replace U^{n+1}, V^{n+1} and decompose the obtained equations according to Q^{n+1} into the following sub-systems:

$$\begin{cases} U_1^{n+1} = U^n, \\ \dfrac{1}{\delta t}U_2^{n+1} = \dfrac{1}{2}H^n\phi_t^n, \end{cases} \tag{3.54}$$

$$\begin{cases} V_1^{n+1} = V^n, \\ \dfrac{1}{\delta t}V_2^{n+1} = \dfrac{1}{2}(K^n, \phi_t^n). \end{cases} \tag{3.55}$$

These equations are algebraic and so they are very straight forward to be solved.

Third, Q^{n+1} is computed from (3.37). With the splitting form given in (3.45) for the variables μ^{n+1}, $\tilde{\mathbf{u}}^{n+1}$, U^{n+1}, V^{n+1}, ϕ^{n+1}, one can rewrite (3.37) as the following form:

$$(\frac{1}{\delta t} - \theta_2)Q^{n+1} = \frac{1}{\delta t}Q^n + \theta_1, \tag{3.56}$$

where

$$\begin{cases} \theta_1 = (\nabla\cdot(\mathbf{u}^n\phi^n), \mu_1^{n+1}) + (\phi^n\nabla\mu^n, \tilde{\mathbf{u}}_1^{n+1}) + \lambda\epsilon(U^nH^n, \dfrac{\phi_1^{n+1}-\phi^n}{\delta t}) \\ \qquad - \lambda\epsilon(H^n\phi_t^n, U_1^{n+1}) + \lambda\epsilon V^n(K^n, \dfrac{\phi_1^{n+1}-\phi^n}{\delta t}) - \lambda\epsilon(K^n, \phi_t^n)V_1^{n+1}, \\ \theta_2 = (\nabla\cdot(\mathbf{u}^n\phi^n), \mu_2^{n+1}) + (\phi^n\nabla\mu^n, \tilde{\mathbf{u}}_2^{n+1}) + \lambda\epsilon(U^nH^n, \dfrac{\phi_2^{n+1}}{\delta t}) \\ \qquad - \lambda\epsilon(H^n\phi_t^n, U_2^{n+1}) + \lambda\epsilon V^n(K^n, \dfrac{\phi_2^{n+1}}{\delta t}) - \lambda\epsilon(K^n, \phi_t^n)V_2^{n+1}. \end{cases} \tag{3.57}$$

We need to verify that (3.56) is solvable by showing $\frac{1}{\delta t} - \theta_2 \neq 0$. By taking the L^2-inner product of the first equation in (3.48) with μ_2^{n+1}, of the second equation of (3.48) with $-\frac{1}{\delta t}\phi_2^{n+1}$, combining the obtained two equations and using Lemma 3.1, we get

$$\begin{aligned} &-(\nabla\cdot(\mathbf{u}^n\phi^n), \mu_2^{n+1}) - \lambda\epsilon(H^nU^n, \frac{\phi_2^{n+1}}{\delta t}) - \lambda\epsilon V^n(K^n, \frac{\phi_2^{n+1}}{\delta t}) \\ &= \gamma\Big\|\mu_2^{n+1} - \frac{1}{|\Omega|}\int_\Omega \mu_2^{n+1}d\boldsymbol{x}\Big\|^2 + \lambda\epsilon\frac{1}{\delta t}\Big(\|\Delta\phi_2^{n+1}\|^2 + e\|\phi_2^{n+1}\|^2 - \frac{1}{\epsilon^2}\|\nabla\phi_2^{n+1}\|^2\Big) \geq 0. \end{aligned} \tag{3.58}$$

By taking the L^2 inner product of the second equation in (3.49) with $\tilde{\mathbf{u}}_2^{n+1}$ and using (3.43), we deduce

$$-(\phi^n\nabla\mu^n, \tilde{\mathbf{u}}_2^{n+1}) = (\rho^n\frac{\tilde{\mathbf{u}}_2^{n+1}}{\delta t}, \tilde{\mathbf{u}}_2^{n+1}) \geq 0. \tag{3.59}$$

By taking the L^2 inner product of the second equations of (3.54) with $2\lambda\epsilon U_2^{n+1}$ and multiplying the second equation of (3.55) with $2\lambda\epsilon V_2^{n+1}$, we derive

$$\begin{cases} \lambda\epsilon(H^n\phi_t^n, U_2^{n+1}) = 2\lambda\epsilon\dfrac{1}{\delta t}\|U_2^{n+1}\|^2 \geq 0, \\ \lambda\epsilon(K^n, \phi_t^n)V_2^{n+1} = 2\lambda\epsilon\dfrac{1}{\delta t}|V_2^{n+1}|^2 \geq 0. \end{cases} \tag{3.60}$$

From (3.58), (3.59), and (3.60), we get $-\theta_2 \geq 0$ that ensures (3.56) is solvable. After obtaining Q^{n+1} from (3.56), we can get ϕ^{n+1}, μ^{n+1}, $\tilde{\mathbf{u}}^{n+1}$, U^{n+1}, V^{n+1} from (3.45), that completes the step 1.

3.3.2. Implementation of step 2

We further give a detailed implementation of step 2, i.e., (3.38)–(3.39).

First, we use the nonlocal variable R^{n+1} to rewrite $\mathbf{u}^{n+1}$ as a linear combination form that reads as

$$\mathbf{u}^{n+1} = \mathbf{u}_1^{n+1} + R^{n+1}\mathbf{u}_2^{n+1}. \tag{3.61}$$

Using (3.61), the scheme (3.38) can be split into the following two equations according to R^{n+1},

$$\begin{cases} \dfrac{1}{2\delta t}(\rho^{n+1}+\rho^n)\mathbf{u}_1^{n+1} - \nabla\cdot(\nu^{n+1}D(\mathbf{u}_1^{n+1})) = \dfrac{\rho^n}{\delta t}\tilde{\mathbf{u}}^{n+1} - \nabla(2p^n - p^{n-1}), \\ \dfrac{1}{2\delta t}(\rho^{n+1}+\rho^n)\mathbf{u}_2^{n+1} - \nabla\cdot(\nu^{n+1}D(\mathbf{u}_2^{n+1})) = -\rho^n(\mathbf{u}^n\cdot\nabla)\mathbf{u}^n - \dfrac{1}{2}\nabla\cdot(\rho^n\mathbf{u}^n)\mathbf{u}^n. \end{cases} \tag{3.62}$$

The boundary conditions of these two equations are either periodic or

$$\mathbf{u}_1^{n+1}|_{\partial\Omega} = \mathbf{u}_2^{n+1}|_{\partial\Omega} = \mathbf{0}. \tag{3.63}$$

The two elliptic equations in (3.62) are always solvable since the variable coefficients $\rho^{n+1}+\rho^n$ and ν^{n+1} are non-negative.

Second, the nonlocal variable R^{n+1} is computed from (3.39). With the splitting form given in (3.61) for the variable $\mathbf{u}^{n+1}$, one can rewrite (3.39) as the following form:

$$(\frac{1}{\delta t} - \eta_2)R^{n+1} = \frac{1}{\delta t}R^n + \eta_1, \tag{3.64}$$

where

$$\begin{cases} \eta_1 = (\rho^n(\mathbf{u}^n\cdot\nabla)\mathbf{u}^n, \mathbf{u}_1^{n+1}) + \dfrac{1}{2}(\nabla\cdot(\rho^n\mathbf{u}^n)\mathbf{u}^n, \mathbf{u}_1^{n+1}), \\ \eta_2 = (\rho^n(\mathbf{u}^n\cdot\nabla)\mathbf{u}^n, \mathbf{u}_2^{n+1}) + \dfrac{1}{2}(\nabla\cdot(\rho^n\mathbf{u}^n)\mathbf{u}^n, \mathbf{u}_2^{n+1}). \end{cases}$$

We need to verify that (3.64) is solvable by showing $\frac{1}{\delta t} - \eta_2 \neq 0$. By taking the L^2 inner product of the second equation in (3.62) with $\mathbf{u}_2^{n+1}$, and using integration by parts and (3.43), we derive

$$-\eta_2 = \frac{1}{2\delta t}\int_\Omega (\rho^{n+1}+\rho^n)|\mathbf{u}_2^{n+1}|^2 d\boldsymbol{x} + \frac{1}{2}\int_\Omega \nu^{n+1}|D(\mathbf{u}_2^{n+1})|^2 d\boldsymbol{x} \geq 0, \tag{3.65}$$

that implies the solvability of (3.64). Once R^{n+1} is obtained, $\mathbf{u}^{n+1}$ is updated using (3.61).

To summarize, the total computing cost at each time step includes solving two biharmonic equations (with constant coefficients) in (3.52), two elliptic equations (with variable coefficients) in (3.62), and one Poisson type equation (with constant coefficients) in (3.40). The linearity and decoupling nature of these equations means that the designed scheme is very effective in actual implementation.

Remark 3.8. For the hydrodynamics coupled phase-field model, we recall that in [52–55], a linear scheme with fully-decoupled nature had been developed using the explicit-stabilization method (abbreviated as Ex-Stab for short). For comparison, we briefly introduce its key idea of discretization.

First, the Ex-Stab method split the surface tension term from the momentum equation and then discretize the phase-field equation as

$$\begin{cases} \dfrac{\phi^{n+1}-\phi^n}{\delta t} + \nabla\cdot(\tilde{\mathbf{u}}^{n+1}\phi^n) = -\gamma(\mu^{n+1} - \dfrac{1}{|\Omega|}\displaystyle\int_\Omega \mu^{n+1}d\boldsymbol{x}), \\ \rho^n\dfrac{\tilde{\mathbf{u}}^{n+1}-\mathbf{u}^n}{\delta t} + \underbrace{\phi^n\nabla\mu^{n+1}}_{\text{explicit}-\text{implicit}} = 0 \text{ (:Operator-Splitting method).} \end{cases} \tag{3.66}$$

Here, the application of the operator splitting method helps to isolate the surface tension term from the momentum equation, which is the same as our method. But there exists a substantial difference between the Ex-Stab scheme

and our method. The Ex-Stab scheme in (3.66) discretizes the advection and surface tension term using the implicit–explicit combination, while our scheme discretizes them explicitly, as $Q^{n+1}\nabla\cdot(\mathbf{u}^n\phi^n)$ and $Q^{n+1}\phi^n\nabla\mu^n$.

The decoupling nature of the Ex-Stab scheme can be achieved after combining the two equations of (3.66) together, that reads as

$$\frac{\phi^{n+1}-\phi^n}{\delta t}+\nabla\cdot((\mathbf{u}^n-\frac{\delta t}{\rho^n}\phi^n\nabla\mu^{n+1})\phi^n)=-\gamma(\mu^{n+1}-\frac{1}{|\Omega|}\int_\Omega\mu^{n+1}d\boldsymbol{x}). \tag{3.67}$$

Therefore, on the one hand, (3.67) clearly tells us the decoupling structure of the phase field variable and the velocity field. On the other hand, the scheme (3.67) also shows one shortcoming, that is, the phase-field equation has become variable-coefficient (for the term $\nabla\mu^{n+1}$). So in actual calculations, the calculation cost is relatively high.

Second, in the Ex-stab method given in [52–55], the fluid momentum equation is discretized as

$$\begin{aligned}\rho^n\frac{\mathbf{u}^{n+1}-\tilde{\mathbf{u}}^{n+1}}{\delta t}+\frac{1}{2}\frac{\rho^{n+1}-\rho^n}{\delta t}\mathbf{u}^{n+1}-\nabla\cdot(\nu^{n+1}D(\mathbf{u}^{n+1}))+\nabla(2p^n-p^{n-1})\\ \underbrace{+\rho^n(\mathbf{u}^n\cdot\nabla)\mathbf{u}^{n+1}+\frac{1}{2}\nabla\cdot(\rho^n\mathbf{u}^n)\mathbf{u}^{n+1}}_{\text{explicit}-\text{implicit}}=0.\end{aligned} \tag{3.68}$$

Here, the advection terms are discretized using the explicit–implicit method, which also leads to many variable coefficients (for $\mathbf{u}^{n+1}$).

Therefore, although the numerical scheme constructed using the Ex-Stab method can also be full decoupled while maintaining the energy stability unconditionally, its design leads to a large number of variable coefficients, so it will lead to costly computations. To illustrate this, in Fig. 4.2 (b), we compare the average number of iterations per time step required by the Ex-Stab method and our developed scheme (3.32)–(3.40). The comparison clearly show that our developed scheme has much higher effectiveness.

3.3.3. Unconditional energy stability

In this subsection, we show that the scheme (3.32)–(3.40) is unconditionally energy stable. We will use the following identity repeatedly:

$$2(a-b)a=|a|^2-|b|^2+|a-b|^2. \tag{3.69}$$

Theorem 3.1. *The scheme (3.32)–(3.40) holds the following discrete energy dissipation law:*

$$E^{n+1}\le E^n-\frac{1}{2}\delta t\|\sqrt{\nu^{n+1}}D(\mathbf{u}^{n+1})\|^2-\gamma\delta t\left\|\mu^{n+1}-\frac{1}{|\Omega|}\int_\Omega\mu^{n+1}d\boldsymbol{x}\right\|^2, \tag{3.70}$$

where

$$\begin{aligned}E^{n+1}=&\frac{1}{2}\|\sigma^{n+1}\mathbf{u}^{n+1}\|^2+\frac{\delta t^2}{2\chi}\|\nabla p^{n+1}\|^2\\ &+\lambda\epsilon\Big(\frac{1}{2}\|\Delta\phi^{n+1}\|^2+\frac{e}{2}\|\phi^{n+1}\|^2-\frac{1}{\epsilon^2}\|\nabla\phi^{n+1}\|^2||+\|U^{n+1}\|^2+|V^{n+1}|^2\Big)\\ &+\frac{1}{2}|Q^{n+1}|^2+\frac{1}{2}|R^{n+1}|^2-B_1|\Omega|-B_2-1,\end{aligned} \tag{3.71}$$

and $\sigma^k=\sqrt{\rho^k}$ for any k. Moreover, E^{n+1} is bounded from below when the parameter e is sufficiently large as described in Lemma 3.1.

Proof. We take the inner product of (3.38) with $2\delta t\mathbf{u}^{n+1}$ in the L^2-space and use (3.69) to obtain

$$\begin{aligned}&\|\sigma^{n+1}\mathbf{u}^{n+1}\|^2-\|\sigma^n\tilde{\mathbf{u}}^{n+1}\|^2+\|\sigma^n(\mathbf{u}^{n+1}-\tilde{\mathbf{u}}^{n+1})\|^2+\delta t\|\sqrt{\nu^{n+1}}D(\mathbf{u}^{n+1})\|^2\\ &\quad+2\delta t(p^{n+1}-2p^n+p^{n-1},\nabla\cdot\mathbf{u}^{n+1})-2\delta t(p^{n+1},\nabla\cdot\mathbf{u}^{n+1})\\ &\quad=-2\delta tR^{n+1}(\rho^n(\mathbf{u}^n\cdot\nabla)\mathbf{u}^n,\mathbf{u}^{n+1})+\frac{1}{2}(\nabla\cdot(\rho^n\mathbf{u}^n)\mathbf{u}^n,\mathbf{u}^{n+1}).\end{aligned} \tag{3.72}$$

We take the L^2-inner product of (3.36) with $2\delta t\tilde{\mathbf{u}}^{n+1}$ and use (3.69) to obtain

$$\|\sigma^n\tilde{\mathbf{u}}^{n+1}\|^2 - \|\sigma^n\mathbf{u}^n\|^2 + \|\sigma^n(\tilde{\mathbf{u}}^{n+1}-\mathbf{u}^n)\|^2 = -2\delta t\, Q^{n+1}(\phi^n\nabla\mu^n, \tilde{\mathbf{u}}^{n+1}). \tag{3.73}$$

We combine (3.72) and (3.73) to obtain

$$\begin{aligned}\|\sigma^{n+1}\mathbf{u}^{n+1}\|^2 \; & -\|\sigma^n\mathbf{u}^n\|^2 + \|\sigma^n(\mathbf{u}^{n+1}-\tilde{\mathbf{u}}^{n+1})\|^2 + \|\sigma^n(\tilde{\mathbf{u}}^{n+1}-\mathbf{u}^n)\|^2\\ &+\delta t\|\sqrt{\nu^{n+1}}D(\mathbf{u}^{n+1})\|^2 + 2\delta t(p^{n+1}-2p^n+p^{n-1}, \nabla\cdot\mathbf{u}^{n+1}) - 2\delta t(p^{n+1}, \nabla\cdot\mathbf{u}^{n+1})\\ &= -2\delta t\, R^{n+1}(\rho^n(\mathbf{u}^n\cdot\nabla)\mathbf{u}^n, \mathbf{u}^{n+1}) + \frac{1}{2}(\nabla\cdot(\rho^n\mathbf{u}^n)\mathbf{u}^n, \mathbf{u}^{n+1})\\ &\quad -2\delta t\, Q^{n+1}(\phi^n\nabla\mu^n, \tilde{\mathbf{u}}^{n+1}).\end{aligned} \tag{3.74}$$

We take the L^2 inner product of (3.40) with $\frac{2\delta t^2}{\chi}(p^{n+1}-2p^n+p^{n-1})$ and use (3.69) to obtain

$$\begin{aligned}2\delta t(p^{n+1}-2p^n+p^{n-1}, \nabla\cdot\mathbf{u}^{n+1}) = \; & -\frac{\delta t^2}{\chi}\Big(\|\nabla p^{n+1}-\nabla p^n\|^2 - \|\nabla p^n-\nabla p^{n-1}\|^2\\ & +\|\nabla p^{n+1}-2\nabla p^n+\nabla p^{n-1}\|^2\Big).\end{aligned} \tag{3.75}$$

We take the L^2 inner product of (3.40) with $-\frac{2\delta t^2}{\chi}p^{n+1}$ and use (3.69) to obtain

$$-2\delta t(p^{n+1}, \nabla\cdot\mathbf{u}^{n+1}) = \frac{\delta t^2}{\chi}\Big(\|\nabla p^{n+1}\|^2 - \|\nabla p^n\|^2 + \|\nabla p^{n+1}-\nabla p^n\|^2\Big). \tag{3.76}$$

By combining (3.75) and (3.76), we derive

$$\begin{aligned}&2\delta t(p^{n+1}-2p^n+p^{n-1}, \nabla\cdot\mathbf{u}^{n+1}) - 2\delta t(p^{n+1}, \nabla\cdot\mathbf{u}^{n+1})\\ &\quad = \frac{\delta t^2}{\chi}\Big(\|\nabla p^{n+1}\|^2 - \|\nabla p^n\|^2\Big) + \frac{\delta t^2}{\chi}\|\nabla p^n-\nabla p^{n-1}\|^2\\ &\qquad -\frac{\delta t^2}{\chi}\|\nabla p^{n+1}-2\nabla p^n+\nabla p^{n-1}\|^2.\end{aligned} \tag{3.77}$$

We subtract (3.40) at $(n+1)$-step and n-step to obtain

$$\Delta(p^{n+1}-2p^n+p^{n-1}) = \frac{\chi}{\delta t}\nabla\cdot(\mathbf{u}^{n+1}-\mathbf{u}^n). \tag{3.78}$$

We multiply the L^2 inner product of (3.78) with $p^{n+1}-2p^n+p^{n-1}$ to derive

$$\begin{aligned}\|\nabla(p^{n+1}-2p^n+p^{n-1})\|^2 \; &= -\frac{\chi}{\delta t}(\mathbf{u}^{n+1}-\mathbf{u}^n, \nabla(p^{n+1}-2p^n+p^{n-1}))\\ &\le \frac{\chi^2}{2\delta t^2}\|\mathbf{u}^{n+1}-\mathbf{u}^n\|^2 + \frac{1}{2}\|\nabla(p^{n+1}-2p^n+p^{n-1})\|^2,\end{aligned} \tag{3.79}$$

where we use integration by parts and Cauchy–Schwarz inequality. Hence (3.79) implies

$$\|\nabla(p^{n+1}-2p^n+p^{n-1})\|^2 \le \frac{\chi^2}{\delta t^2}\|\mathbf{u}^{n+1}-\mathbf{u}^n\|^2. \tag{3.80}$$

By multiplying (3.80) with $\frac{\delta t^2}{\chi}$ and applying $\chi = \frac{1}{2}\min(\rho_1, \rho_2) \le \frac{1}{2}\rho^n$, we derive

$$\frac{\delta t^2}{\chi}\|\nabla(p^{n+1}-2p^n+p^{n-1})\|^2 \le \chi\|\mathbf{u}^{n+1}-\mathbf{u}^n\|^2 \le \frac{1}{2}\|\sigma^n(\mathbf{u}^{n+1}-\mathbf{u}^n)\|^2. \tag{3.81}$$

Using the inequality $\frac{1}{2}(a+b)^2 \le a^2+b^2$, we obtain

$$\frac{1}{2}\|\sigma^n(\mathbf{u}^{n+1}-\mathbf{u}^n)\|^2 \le \|\sigma^n(\mathbf{u}^{n+1}-\tilde{\mathbf{u}}^{n+1})\|^2 + \|\sigma^n(\tilde{\mathbf{u}}^{n+1}-\mathbf{u}^n)\|^2. \tag{3.82}$$

Therefore, by combining (3.77), (3.81) and (3.82), we obtain

$$\begin{aligned}
&2\delta t(p^{n+1}-2p^n+p^{n-1},\nabla\cdot\mathbf{u}^{n+1})-2\delta t(p^{n+1},\nabla\cdot\mathbf{u}^{n+1})\\
&\quad\geq\frac{\delta t^2}{\chi}\Big(\|\nabla p^{n+1}\|^2-\|\nabla p^n\|^2\Big)+\frac{\delta t^2}{\chi}\|\nabla p^n-\nabla p^{n-1}\|^2\\
&\quad\quad-\|\sigma^n(\mathbf{u}^{n+1}-\tilde{\mathbf{u}}^{n+1})\|^2-\|\sigma^n(\tilde{\mathbf{u}}^{n+1}-\mathbf{u}^n)\|^2.
\end{aligned}\tag{3.83}$$

Hence, from (3.74) and (3.83), we derive

$$\begin{aligned}
\|\sigma^{n+1}\mathbf{u}^{n+1}\|^2\ &-\|\sigma^n\mathbf{u}^n\|^2+\delta t\|\sqrt{\nu^{n+1}}D(\mathbf{u}^{n+1})\|^2\\
&+\frac{\delta t^2}{\chi}\Big(\|\nabla p^{n+1}\|^2-\|\nabla p^n\|^2\Big)+\frac{\delta t^2}{\chi}\|\nabla p^n-\nabla p^{n-1}\|^2\\
&\leq\underbrace{-2\delta t\,R^{n+1}(\rho^n(\mathbf{u}^n\cdot\nabla)\mathbf{u}^n,\mathbf{u}^{n+1})}_{\mathrm{I}_1}\underbrace{-\delta t\,R^{n+1}(\nabla\cdot(\rho^n\mathbf{u}^n)\mathbf{u}^n,\mathbf{u}^{n+1})}_{\mathrm{II}_1}\\
&\quad\underbrace{-2\delta t\,Q^{n+1}(\phi^n\nabla\mu^n,\tilde{\mathbf{u}}^{n+1})}_{\mathrm{III}_1}.
\end{aligned}\tag{3.84}$$

We take the L^2 inner product of (3.32) with $2\delta t\mu^{n+1}$ and use integration by parts to get

$$\begin{aligned}
2(\phi^{n+1}-\phi^n,\mu^{n+1})\ &+2\gamma\delta t\left\|\mu^{n+1}-\frac{1}{|\Omega|}\int_\Omega\mu^{n+1}d\boldsymbol{x}\right\|^2\\
&=\underbrace{-2\delta t\,Q^{n+1}(\nabla\cdot(\mathbf{u}^n\phi^n),\mu^{n+1})}_{\mathrm{IV}_1}.
\end{aligned}\tag{3.85}$$

By taking the L^2 inner product of (3.33) with $-2(\phi^{n+1}-\phi^n)$, and using integration by parts and (3.69), we get

$$\begin{aligned}
-2(\phi^{n+1}-\phi^n,\mu^{n+1})=\ &-\lambda\epsilon(\|\Delta\phi^{n+1}\|^2-\|\Delta\phi^n\|^2+\|\Delta(\phi^{n+1}-\phi^n)\|^2)\\
&-\lambda\epsilon(e\|\phi^{n+1}\|^2-e\|\phi^n\|^2+e\|\phi^{n+1}-\phi^n\|^2)\\
&-\lambda\epsilon(-\frac{2}{\epsilon^2}\|\nabla\phi^{n+1}\|^2+\frac{2}{\epsilon^2}\|\nabla\phi^n\|^2-\frac{2}{\epsilon^2}\|\nabla(\phi^{n+1}-\phi^n)\|^2)\\
&\underbrace{-2\lambda\epsilon\,Q^{n+1}(H^nU^n,\phi^{n+1}-\phi^n)}_{\mathrm{V}_1}\underbrace{-2\lambda\epsilon\,Q^{n+1}V^n(K^n,\phi^{n+1}-\phi^n)}_{\mathrm{VI}_1}.
\end{aligned}\tag{3.86}$$

By taking the L^2 inner product of (3.34) with $4\lambda\epsilon\delta tU^{n+1}$ and using (3.69), we get

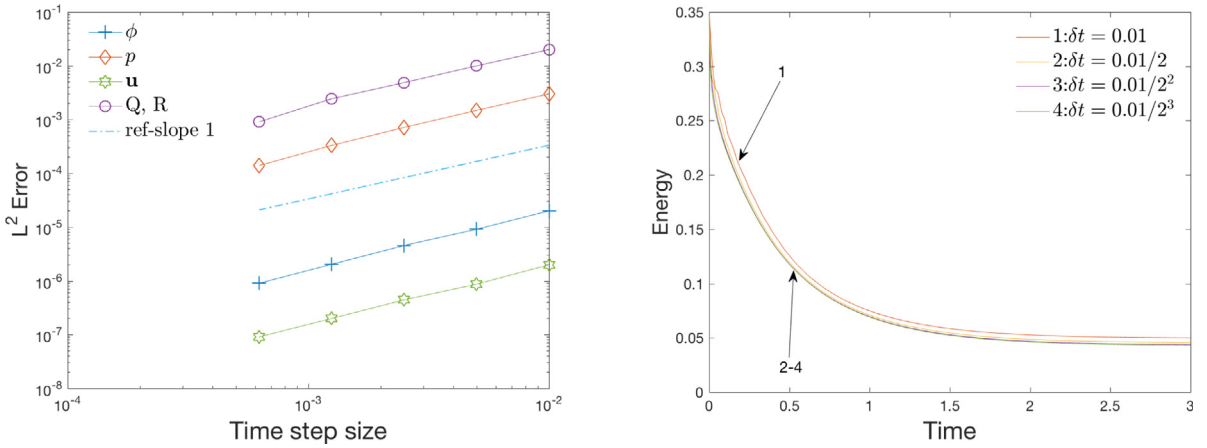
$$2\lambda\epsilon(\|U^{n+1}\|^2-\|U^n\|^2+\|U^{n+1}-U^n\|^2)=\underbrace{2\delta t\lambda\epsilon\,Q^{n+1}(H^n\phi_t^n,U^{n+1})}_{\mathrm{VII}_1}.\tag{3.87}$$

By multiplying (3.35) with $4\lambda\epsilon\delta tV^{n+1}$ and using (3.69), we get

$$2\lambda\epsilon(|V^{n+1}|^2-|V^n|^2+|V^{n+1}-V^n|^2)=\underbrace{2\delta t\lambda\epsilon\,Q^{n+1}(K^n,\phi_t^n)V^{n+1}}_{\mathrm{VIII}_1}.\tag{3.88}$$

We multiply (3.37) with $2\delta tQ^{n+1}$ and use (3.69) to get

$$\begin{aligned}
|Q^{n+1}|^2\ &-|Q^n|^2+|Q^{n+1}-Q^n|^2\\
&=\underbrace{2\delta t\,Q^{n+1}(\nabla\cdot(\mathbf{u}^n\phi^n),\mu^{n+1})}_{\mathrm{IV}_2}+\underbrace{2\delta t\,Q^{n+1}(\phi^n\nabla\mu^n,\tilde{\mathbf{u}}^{n+1})}_{\mathrm{III}_2}\\
&\quad+\underbrace{2\lambda\epsilon\,Q^{n+1}(U^nH^n,\phi^{n+1}-\phi^n)}_{\mathrm{V}_2}\underbrace{-2\delta t\lambda\epsilon\,Q^{n+1}(H^n\phi_t^n,U^{n+1})}_{\mathrm{VII}_2}\\
&\quad+\underbrace{2\lambda\epsilon V^nQ^{n+1}(K^n,\phi^{n+1}-\phi^n)}_{\mathrm{VI}_2}\underbrace{-2\delta t\lambda\epsilon\,Q^{n+1}(K^n,\phi_t^n)V^{n+1}}_{\mathrm{VIII}_2}.
\end{aligned}\tag{3.89}$$



(a) Accuracy test. (b) Stability test.

Fig. 4.1. (a) The numerical errors of $\phi, p, \mathbf{u}$ in L^2 and the nonlocal variables Q, R that are calculated by using the scheme EIEQ with different time steps; (b) the total free energy (3.71) that is evolved over time and computed by using the scheme EIEQ with different time steps.

We multiply (3.39) with $2\delta t R^{n+1}$ and use (3.69) to get

$$\begin{aligned} &|R^{n+1}|^2 - |R^n|^2 + |R^{n+1} - R^n|^2 \\ &= \underbrace{2\delta t R^{n+1}(\rho^n(\mathbf{u}^n \cdot \nabla)\mathbf{u}^n, \mathbf{u}^{n+1})}_{\mathrm{I}_2} + \underbrace{\delta t R^{n+1}(\nabla \cdot (\rho^n \mathbf{u}^n)\mathbf{u}^n, \mathbf{u}^{n+1})}_{\mathrm{II}_2}. \end{aligned} \tag{3.90}$$

By combining (3.84)–(3.90), we derive

$$\begin{aligned} \|\sigma^{n+1}\mathbf{u}^{n+1}\|^2 &- \|\sigma^n \mathbf{u}^n\|^2 + \frac{\delta t^2}{\chi}\Big(\|\nabla p^{n+1}\|^2 - \|\nabla p^n\|^2\Big) + \frac{\delta t^2}{\chi}\|\nabla p^n - \nabla p^{n-1}\|^2 \\ &+ \lambda\epsilon(\|\Delta\phi^{n+1}\|^2 - \|\Delta\phi^n\|^2 + \|\Delta(\phi^{n+1} - \phi^n)\|^2) \\ &+ \lambda\epsilon(e\|\phi^{n+1}\|^2 - e\|\phi^n\|^2 + e\|\phi^{n+1} - \phi^n\|^2) \\ &+ \lambda\epsilon(-\frac{2}{\epsilon^2}\|\nabla\phi^{n+1}\|^2 + \frac{2}{\epsilon^2}\|\nabla\phi^n\|^2 - \frac{2}{\epsilon^2}\|\nabla(\phi^{n+1} - \phi^n)\|^2) \\ &+ 2\lambda\epsilon(\|U^{n+1}\|^2 - \|U^n\|^2 + \|U^{n+1} - U^n\|^2) + 2\lambda\epsilon(|V^{n+1}|^2 - |V^n|^2 + |V^{n+1} - V^n|^2) \\ &+ (|Q^{n+1}|^2 - |Q^n|^2 + |Q^{n+1} - Q^n|^2) + (|R^{n+1}|^2 - |R^n|^2 + |R^{n+1} - R^n|^2) \\ &\le -2\gamma\delta t \left\| \mu^{n+1} - \frac{1}{|\Omega|}\int_\Omega \mu^{n+1} d\boldsymbol{x} \right\|^2 - \delta t \|\sqrt{\nu^{n+1}} D(\mathbf{u}^{n+1})\|^2. \end{aligned} \tag{3.91}$$

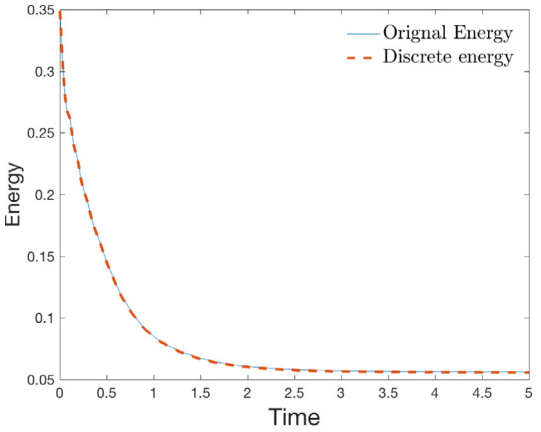
Using (3.31), we deduce that when e is sufficiently large, for any k, the following inequality holds:

$$\|\Delta(\phi^{k+1} - \phi^k)\|^2 + e\|\phi^{k+1} - \phi^k\|^2 - \frac{2}{\epsilon^2}\|\nabla\phi^{k+1} - \nabla\phi^k\|^2 \ge 0. \tag{3.92}$$

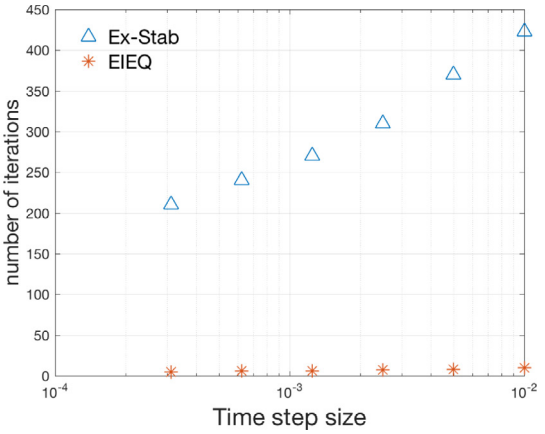
Therefore, after we divide 2 to each side of (3.91) and remove some unnecessary positive terms, we obtain (3.70). In addition, it is easy to prove that E^{n+1} is bounded from below by following a process similar to (3.31). □

4. Numerical simulations

In this section, several numerical tests are first carried out to verify the accuracy and energy stability of the proposed explicit-IEQ scheme (3.32)–(3.40) (referred to as EIEQ for short). Then, by performing numerous 2D and 3D numerical simulations of the vesicle sedimentation dynamics under gravity, we demonstrate the effectiveness of the variable-density model and the proposed scheme EIEQ.



(a) Energy evolution of (2.1) and (3.71) using $\delta t = 0.01$.



(b) Number of iterations.

Fig. 4.2. (a) The free energy in the original form (2.1) and the energy (3.71) in the discrete modified form, which are calculated using $\delta t = 0.01$; and (b) comparisons of the EIEQ and Ex-Stab schemes in computational efficiency.

In all the numerical examples below, the computational domain is set to a 2D or 3D rectangular region. We can use periodic boundary conditions or the boundary conditions given in (3.42). We adopt the Legendre–Galerkin method to discretize directions with the boundary conditions given in (3.42). The stable inf–sup pair (P_N, P_{N-2}), where N is the degree of the Legendre polynomial, is adopted for the velocity field $\mathbf{u}$ and pressure p, respectively; and P_N is adopted for discretizations of the phase-field variable ϕ and the chemical potential μ. We adopt the Fourier-Spectral method to discretize directions satisfying the periodic boundary conditions.

4.1. Accuracy and stability test

In this subsection, we carry out several accuracy and stability tests to verify the accuracy and stability of the scheme EIEQ (3.32)–(3.40). We set the 2D computational domain as $\Omega = [0, 2\pi]^2$. The system is assumed to satisfy the periodic boundary conditions and so space is discretized by the Fourier-Spectral method.

We set the initial condition of $\phi^0, \mathbf{u}^0, p^0$ as follows:

$$\phi^0(\boldsymbol{x}) = 1 + \sum_{i=1}^{2} \tanh\left(\frac{r_i - |\boldsymbol{x} - \boldsymbol{x}_i|}{\sqrt{2}\epsilon}\right), \mathbf{u}^0(\boldsymbol{x}) = \mathbf{0}, p^0 = 0, \tag{4.1}$$

where the profile of ϕ^0 is set as two circular kissing vesicles, $(r_1, r_2) = (0.28\pi, 0.28\pi)$, $\boldsymbol{x}_1 = (0.71\pi, \pi), \boldsymbol{x}_2 = (1.29\pi, \pi)$. The model parameters are set as

$$\begin{cases} \rho_1 = 100, \rho_2 = 1, \nu_1 = 100, \nu_2 = 1, \lambda = 0.01, \\ \epsilon = 0.08, \gamma = 0.1, e = \dfrac{4}{\epsilon^4}, M = 1\text{e}5, B_1 = 100, B_2 = 100. \end{cases} \tag{4.2}$$

We use 256 Fourier modes to discretize each direction of the space. Thus, if compared with the error caused by time discretization, the error caused by spatial discretization is relatively very small and can be ignored. Moreover, since the exact solutions equipped with the above initial conditions are unknown, for simplicity, we consider the numerical solution calculated with a tiny time step $\delta t = 1\text{e}{-7}$ as the roughly exact solution. The time step refinement method is adopted to verify the numerical accuracy caused by time discretization.

We perform an accuracy test to investigate the convergence order of EIEQ. In Fig. 4.1 (a), the L^2 errors of ϕ, $\mathbf{u}$, p and the errors of Q, R ($Q(t) = R(t) = 1$ is the exact solution) at $t = 0.5$ are obtained by using different time step sizes. It can be seen that all variables provide a very good first-order temporal convergence order. In Fig. 4.1 (b), to verify the energy stability, we plot the total free energy (3.71) that evolves over time, which are calculated by using four different time steps. All obtained energy curves exhibit monotonic decaying trends, which illustrates the unconditional energy stability that the EIEQ scheme has always maintained.

In Fig. 4.2 (a), we compare the temporal evolution curve of the total free energy (2.1) (in the original form) and (3.71) (in the modified discrete form) using $\delta t = 0.01$. We can see that the two energy curves are very consistent. To further illustrate the effectiveness of the scheme, in Fig. 4.2 (b), the average number of iterations per time step needed by EIEQ and that needed by the scheme developed in [53,55,68] (see Remark 3.8, denoted by Ex-Stab for short) are compared. It can be seen that the efficiency of EIEQ is much higher than that of Ex-Stab. We take the result obtained using $\delta t = 0.01$ as an example, where at each time step, the scheme Ex-Stab requires around 450 iterations, and only around 10 iterations are required when using the scheme EIEQ. The reason is that in the scheme Ex-Stab, the linear system for ϕ^{n+1} and $\mathbf{u}^{n+1}$, i.e., (3.67) and (3.68), have many variable coefficients.

4.2. 2D vesicle cell sedimentation

The vesicle model has been extensively used in simulating the dynamical motion of red blood cells [24,34]. When the vesicle cell sinks due to the density difference of the inner and outer fluids, the so-called "reorientation" or "rotation" phenomenon caused by the gravity force happens during the sedimentation process. Due to the asymmetry of the vesicle shape, it has been observed that the trajectory of the vesicle will be significantly different depending on the initial angle between the main axis and the gravity direction. In this subsection, we use the developed variable-density vesicle model and our constructed scheme EIEQ to investigate the dynamical sedimentation process of a single vesicle driven by the gravity force.

We first perform 2D simulations in this subsection, and impose the gravity field to the fluid equation as follows:

$$\rho(\mathbf{u}_t + (\mathbf{u}\cdot\nabla)\mathbf{u}) + \frac{1}{2}J\mathbf{u} - \nabla\cdot(\nu(\phi)D(\mathbf{u})) + \nabla p + \phi\nabla\mu = \rho\mathbf{g}, \tag{4.3}$$

where $\mathbf{g} = (0, g_0)$ and g_0 is the gravity constant. The variables/parameters are rescaled as:

$$\hat{t} = \frac{t}{t_0}, \hat{\boldsymbol{x}} = \frac{\boldsymbol{x}}{d_0}, \hat{\mathbf{u}} = \frac{\mathbf{u}}{u_0}, \hat{\nu} = \frac{\nu}{\sqrt{d_0^3 g_0}}, \hat{\gamma} = \gamma, \hat{\lambda} = \lambda, \hat{p} = p, \hat{\rho} = \rho, \hat{\phi} = \phi, \tag{4.4}$$

where $t_0 = \sqrt{d/g_0}$, $u_0 = \sqrt{dg_0}$, $d_0 = d$. Note that if we omit $\hat{}$, the dimensionless governing system (2.11)–(2.12)–(4.3)–(2.14) keeps the same form.

We set the rectangular computed region as $(x, y) \in \Omega = [0, 3\pi]\times[0, 4\pi]$. We assume both the x and y-directions satisfy the non-periodic boundary conditions specified in (3.42), and adopt Legendre polynomial up to the degree of 256 for each direction.

The initial profile of the vesicle cell is set to the capsule shape (consistent with [24]). To obtain such a profile, we set the pre-initial condition of $\hat{\phi}^0$ to be 2D ellipse

$$\tilde{\phi}^0(x, y) = \tanh\Big(\frac{1 - \tilde{x}^2/4.5 - \tilde{y}^2}{\sqrt{2}\epsilon}\Big), \tag{4.5}$$

where

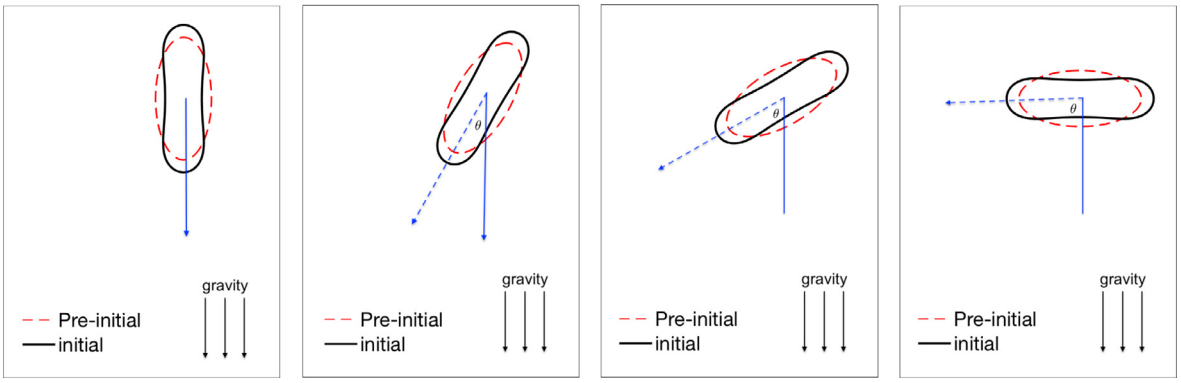
$$\tilde{x} = (x - x_0)\sin\theta + (y - y_0)\cos\theta, \tilde{y} = (x - x_0)\cos\theta + (y - y_0)\sin\theta. \tag{4.6}$$

Here the angle θ can be adjusted to obtain different oblique shapes of the ellipse, see four different dashed ellipses plotted in Fig. 4.3.

We set the model parameters as

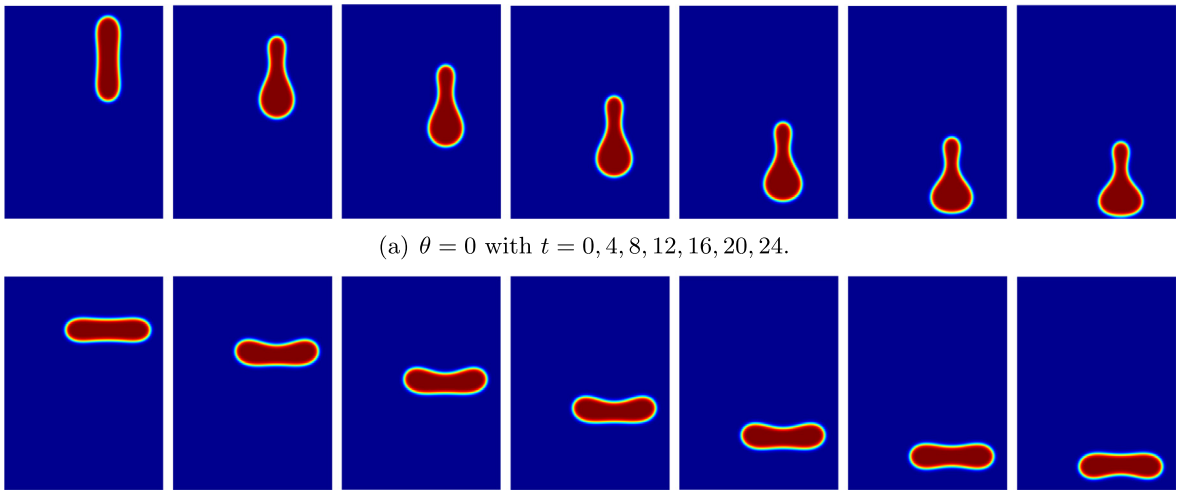
$$\lambda = 0.01, \epsilon = 0.14, \gamma = 0.01, e = \frac{4}{\epsilon^4}, M = 1e5, B_1 = 100, B_2 = 100. \tag{4.7}$$

and implement the no-flow simulation (set $\mathbf{u} = \mathbf{0}$, $p = 0$) to long enough time. The obtained steady-state solution of ϕ is used as the initial condition ϕ^0 to simulate the sedimentation process. In Fig. 4.3, we select four different angles of $\theta = \pi/2, \pi/3, \pi/6, 0$ and plot the contour of the pre-initial condition $\tilde{\phi}^0$ using the dashed lines. The real initial condition of ϕ^0 used to simulate the sediment process are plotted by the solid lines in Fig. 4.3, where we can see that these steady-state contours are all capsule-like.



(a) $\theta = 0$. (b) $\theta = \pi/6$ (c) $\theta = \pi/3$ (d) $\theta = \pi/2$.

Fig. 4.3. The pre-initial (dashed lines) and initial (solid lines) profiles of four different oblique ellipses where (a) $\theta = 0$, (b) $\theta = \pi/6$, (c) $\theta = \pi/3$, (a) $\theta = \pi/2$.



(a) $\theta = 0$ with $t = 0, 4, 8, 12, 16, 20, 24$.

(b) $\theta = \pi/2$ with $t = 0, 10, 20, 30, 40, 50, 60$.

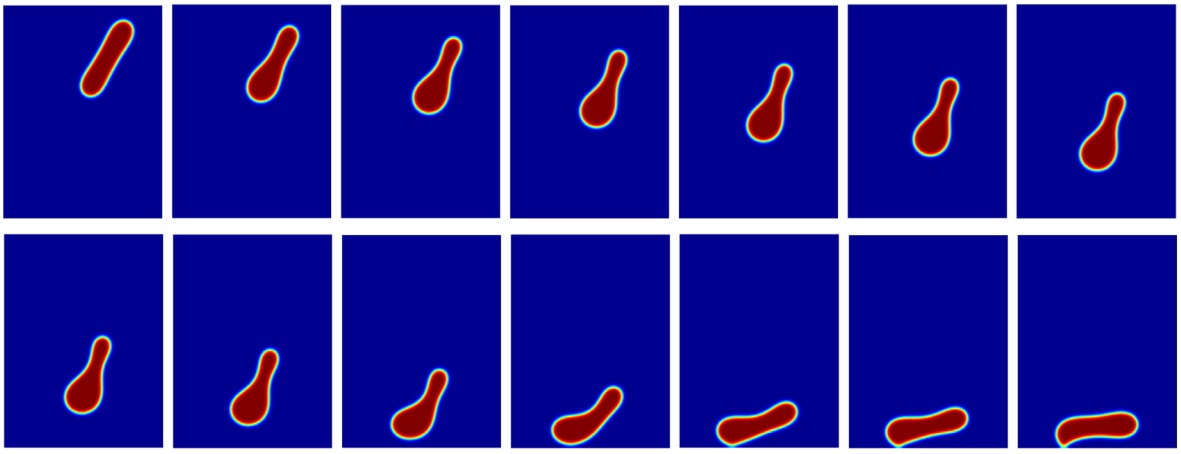
Fig. 4.4. Dynamics of a 2D vesicle sink to the bottom wall under the action of gravity at different times where (a) $\theta = 0$ and (b) $\theta = \pi/2$.

Using the obtained initial conditions of $\mathbf{u}^0 = \mathbf{0}$, $p^0 = 0$, ϕ^0 (the solid contour in Fig. 4.3), and setting other model parameters as

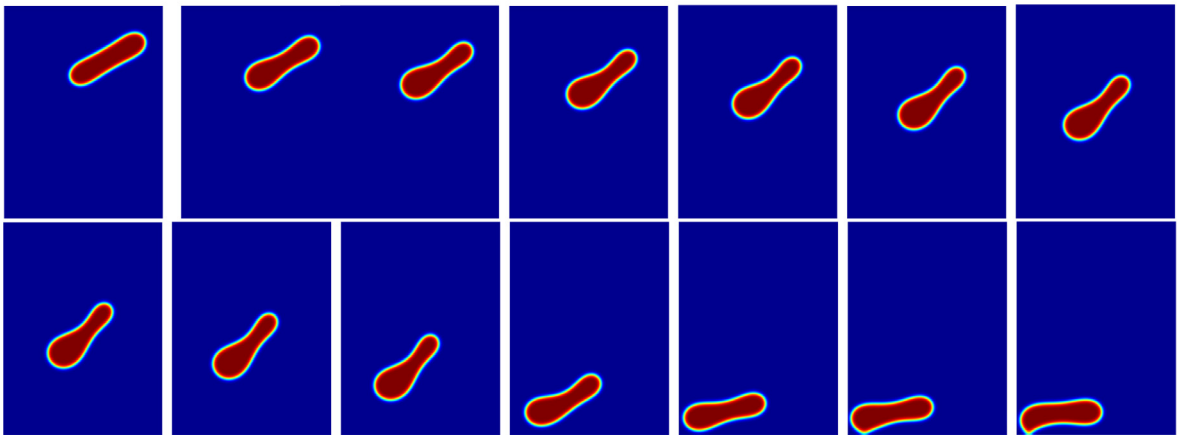
$$\begin{cases} g_0 = 1, \rho_1 = 40, \rho_2 = 1, \nu_1 = 100, \nu_2 = 1, \lambda = 10, \epsilon = 0.14, \gamma = 0.01, \\ e = \dfrac{4}{\epsilon^4}, M = 1\text{e}5, B_1 = 100, B_2 = 100, \end{cases} \tag{4.8}$$

we start with the two coaxial cases, where the long axis of the capsule-like vesicle cell is parallel or perpendicular to the direction of gravity. Snapshots of the phase-field variable ϕ at various times are shown in Fig. 4.4(a) and (b). The common point of these two cases is that the main axis of the vesicle always maintains the same angle with the gravity direction when the vesicle is falling. The difference between these two cases is, when $\theta = 0$, as shown in Fig. 4.4(a), the lower part of the vesicle is enlarged, and the upper part becomes thinner and shorter. When $\theta = 0$, as shown in Fig. 4.4(b), the shape of the vesicle remains unchanged when it falls. Note that no "reorientation" behaviors occur in either case.

We continue to simulate other two tilted cases, that is, the long axis of the capsule-like vesicle has a certain angle with the direction of gravity. Snapshots of the phase-field variable ϕ at various times are shown in Fig. 4.5(a)



(a) $\theta = \pi/6$, snapshots are taken at $t = 0, 2, 4, 6, 8, 10, 12, 16, 20, 24, 28, 30, 34$.



(b) $\theta = \pi/3$, snapshots are taken at $t = 0, 2, 4, 6, 8, 10, 12, 14, 16, 20, 28, 34, 36, 40$.

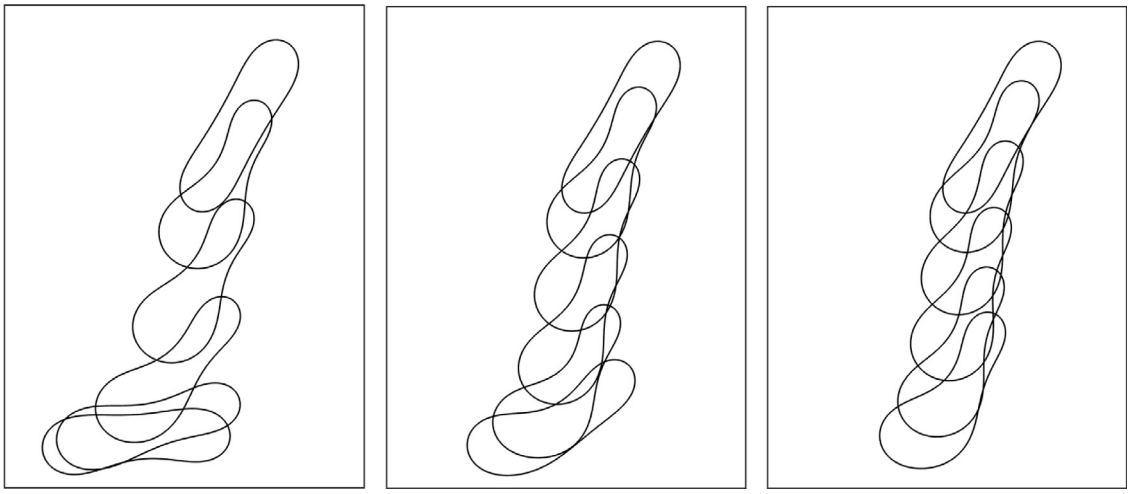
Fig. 4.5. Dynamics of a 2D vesicle sink to the bottom wall under the action of gravity at different times where (a) $\theta = \pi/6$ and (b) $\theta = \pi/3$.

and (b). The common point of these two cases is that the vesicle reorients the direction of the long axis when they approach near the bottom wall, and the long axis direction is almost perpendicular to the direction of gravity, which is the so-called "reorientation" phenomenon, see also [24]. The difference between these two cases is that, the vesicle with $\theta = \pi/3$, shown in Fig. 4.5(b), drifts a greater distance horizontally before touching the bottom strength than the vesicle with $\theta = \pi/6$, shown in Fig. 4.5(a), see also [24].

Finally, using ϕ^0 of $\theta = \pi/6$ as the initial condition (Fig. 4.3(b)), we only adjust the viscosity ν_1 and ν_2 while keeping other parameters unchanged. In Fig. 4.6, the contours of the level set $\{\phi = 0\}$ are plotted at different times to give the trajectory of sinking vesicles for three cases of $\nu_1 : \nu_2 = 50 : 1$, $\nu_1 : \nu_2 = 80 : 1$, and $\nu_1 : \nu_2 = 100 : 1$. It can be seen that when the viscosity is smaller, the horizontal drifted distance of the vesicle is larger. These results are qualitatively consistent with the numerical simulations given by using a different vesicle cell model in [24].

4.3. 3D vesicle cell sedimentation

In this subsection, we perform 3D simulations to investigate the capsule-like vesicle cell sedimentation process. The computational domain is set as $(x, y, z) \in \Omega = [0, 3\pi] \times [0, 4\pi] \times [0, 3\pi]$, where the x and z-directions are assumed to satisfy periodic boundary conditions, and the y-direction is assumed to be non-periodic boundary



(a) $\nu_1 : \nu_2 = 50 : 1$. (b) $\nu_1 : \nu_2 = 80 : 1$. (c) $\nu_1 : \nu_2 = 100 : 1$.

Fig. 4.6. Three simulations of the vesicle sedimentation process using different viscosity ratio, where (a) $\nu_1 : \nu_2 = 50 : 1$, (b) $\nu_1 : \nu_2 = 80 : 1$, and (c) $\nu_1 : \nu_2 = 100 : 1$. In each subfigure, the contour of level set $\{\phi = 0\}$ at various times are shown.

conditions (3.42). The model parameters are set as

$$\begin{cases} g_0 = 1, \rho_1 = 40, \rho_2 = 1, \nu_1 = 100, \nu_2 = 1, \lambda = 40, \epsilon = 0.14, \gamma = 0.01, \\ e = \dfrac{4}{\epsilon^4}, M = 1\text{e}5, B_1 = 100, B_2 = 100. \end{cases} \tag{4.9}$$

The initial conditions of the capsule-like vesicle is still obtained by using the similar process as the 2D simulations. We set the pre-initial condition of $\tilde{\phi}^0$ to be a 3D ellipsoid, that reads as

$$\tilde{\phi}^0(x, y, z) = \tanh\Big(\frac{1 - \tilde{x}^2 - \tilde{y}^2 - \tilde{z}^2/5.5}{\sqrt{2}\epsilon}\Big), \tag{4.10}$$

where

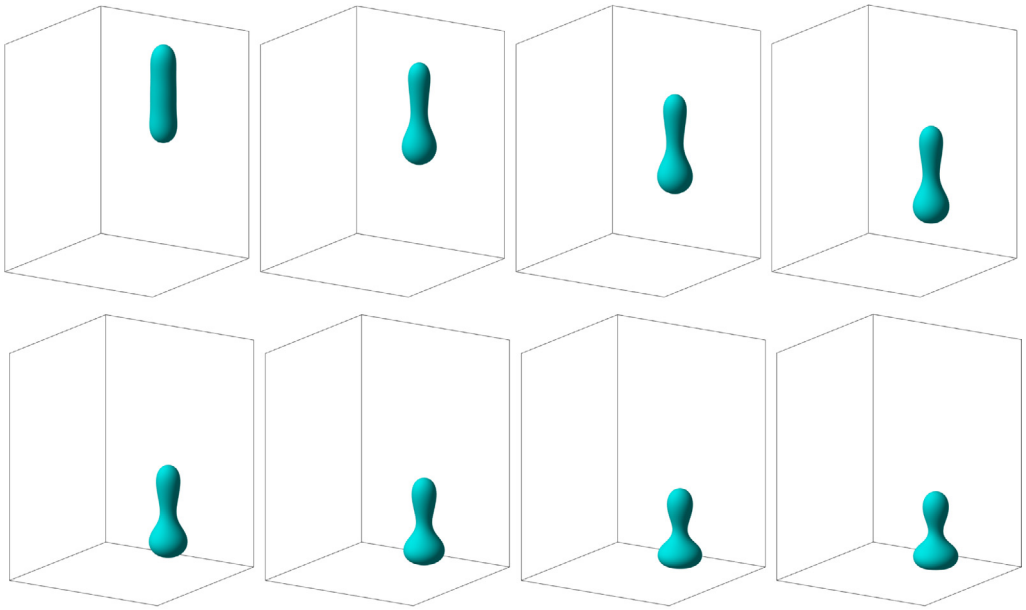
$$(\tilde{x}, \tilde{y}, \tilde{z}) = \tilde{\boldsymbol{x}} = R_z(\theta_z)R_y(\theta_y)R_x(\theta_x)(\boldsymbol{x} - \boldsymbol{x}_0), \tag{4.11}$$

$$R_z(\theta_z) = \begin{pmatrix} \cos\theta_z & -\sin\theta_z & 0 \\ \sin\theta_z & \cos\theta_z & 0 \\ 0 & 0 & 1 \end{pmatrix}, R_y(\theta_y) = \begin{pmatrix} \cos\theta_y & 0 & \sin\theta_y \\ 0 & 1 & 0 \\ -\sin\theta_y & 0 & \cos\theta_y \end{pmatrix},$$
$$R_x(\theta_x) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_x & -\sin\theta_x \\ 0 & \sin\theta_x & \cos\theta_x. \end{pmatrix}$$

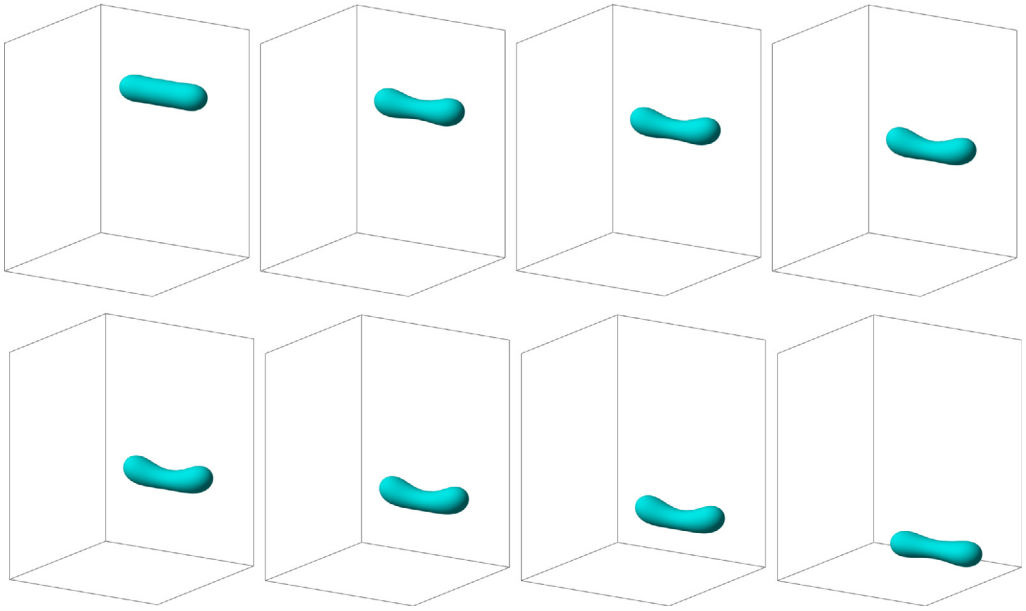
Here R_x, R_y, R_z are three extrinsic rotation matrices, $\theta_x, \theta_y, \theta_z$ are Euler angles, $\boldsymbol{x} - \boldsymbol{x}_0 = (x - x_0, y - y_0, z - z_0)^T$, $x_0 = 1.95\pi$, $y_0 = 2\pi$, $z_0 = 1.95\pi$.

We discretize x and z-directions using 256 Fourier modes, and discretize the y-direction using the Legendre polynomials up to the degree of 512. We set the pre-initial conditions of $\tilde{\phi}^0$ with four different set of rotational angles of $(\theta_x, \theta_y, \theta_z)_1 = (0, \pi/2, 0)$, $(\theta_x, \theta_y, \theta_z)_2 = (-\pi/3, \pi/6, -\pi/4)$, $(\theta_x, \theta_y, \theta_z)_3 = (-5\pi/12, \pi/4, -\pi/12)$, $(\theta_x, \theta_y, \theta_z)_4 = (-\pi/2, \pi/2, 0)$, and implement the no-flow simulation (set $\mathbf{u} = \mathbf{0}$, $p = 0$) to long enough time with other model parameters as (4.7). The obtained steady-state solutions which are shown in the first subfigure in Fig. 4.7 (a)–(b) and Fig. 4.8 (a)–(b), will be used as the initial condition ϕ^0 to simulate the process of 3D vesicle sedimentation.

Using the obtained four different initial conditions ϕ^0 and setting $\mathbf{u}^0 = \mathbf{0}$, $p^0 = 0$, we investigate the sediment process in 3D. First, for the two coaxial cases with $(\theta_x, \theta_y, \theta_z)_1$ and $(\theta_x, \theta_y, \theta_z)_4$, we plot the isosurface of $\{\phi = 0\}$



(a) $(\theta_x, \theta_y, \theta_z)_1$ with $t = 0, 6, 12, 18, 24, 27, 30, 31$.

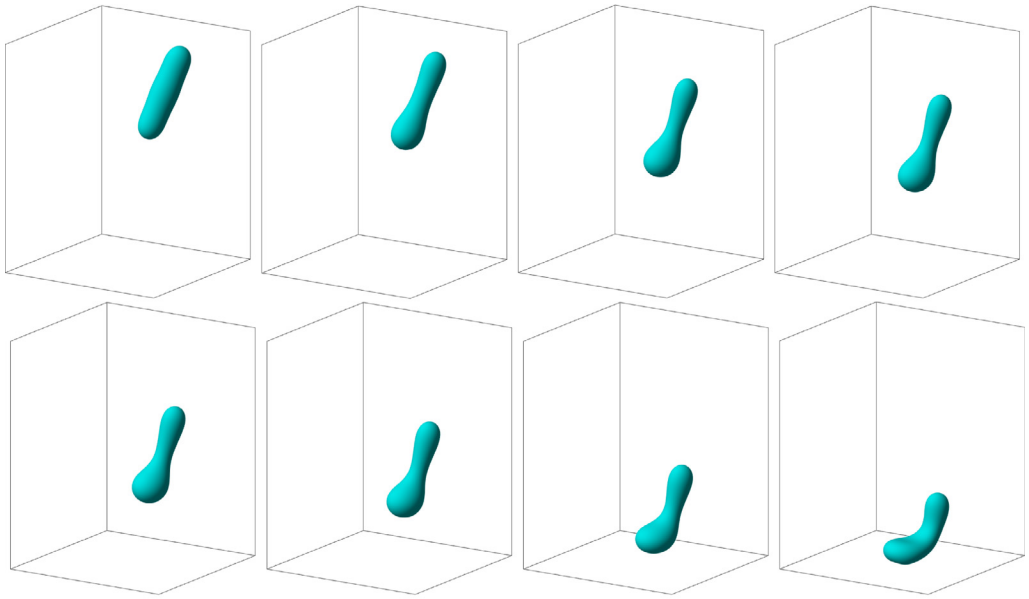


(b) $(\theta_x, \theta_y, \theta_z)_4$ with $t = 0, 6, 12, 18, 24, 30, 36, 50$.

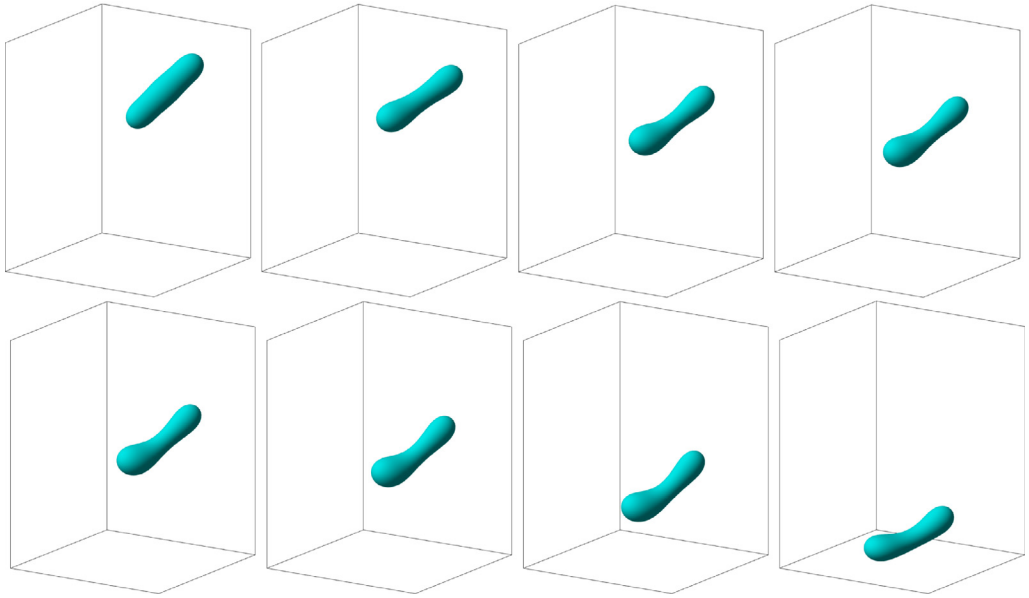
Fig. 4.7. Dynamics of a 3D vesicle sink to the bottom wall under the action of gravity at different times, where the initial conditions with Euler angles are set as (a) $(\theta_x, \theta_y, \theta_z)_1$ and (b) $(\theta_x, \theta_y, \theta_z)_4$.

at various times in Fig. 4.7(a) and (b). Similar to 2D simulations, the dynamical motion of the vesicle roughly maintains its initial shape. For the case of $(\theta_x, \theta_y, \theta_z)_1$, the lower part of the capsule-like vesicle is enlarged slightly during the falling process.

For the other two tilted cases of $(\theta_x, \theta_y, \theta_z)_2$ and $(\theta_x, \theta_y, \theta_z)_3$, shown in Fig. 4.8(a) and (b), the vesicle flips the direction of the rotation when they fall to the bottom wall, i.e., the "reorientation" phenomenon occurs. To compare their drifted distance, in Fig. 4.9, we plot the profiles of isosurfaces $\{\phi = 0\}$ before the vesicle touches the bottom



(a) $(\theta_x, \theta_y, \theta_z)_2$ with $t = 0, 3, 9, 12, 15, 18, 27, 34$.



(b) $(\theta_x, \theta_y, \theta_z)_3$ with $t = 0, 3, 9, 12, 15, 18, 27, 42$.

Fig. 4.8. Dynamics of a 3D vesicle sink to the bottom wall under the action of gravity at different times, where the initial conditions with Euler angles are set as (a) $(\theta_x, \theta_y, \theta_z)_2$ and (b) $(\theta_x, \theta_y, \theta_z)_3$.

wall together. We find that the drift distance of the case of $(\theta_x, \theta_y, \theta_z)_3$ is slightly more than that of $(\theta_x, \theta_y, \theta_z)_2$, which are also consistent with the 2D simulations and the numerical results given in [24].

5. Concluding remarks

In this paper, for the hydrodynamics-coupled phase-field variable elastic bending energy model for lipid vesicles, we have established the volume-conserved, variable-density and viscosity version, and constructed an efficient numerical algorithm for it. Many desired properties are embedded in the algorithm, including linearity, decoupling

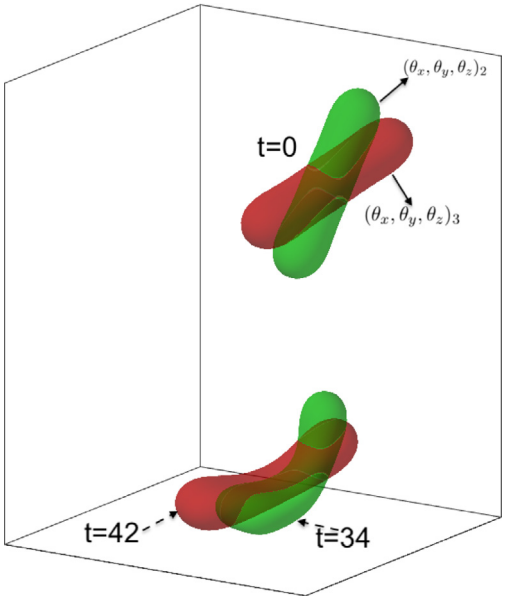


Fig. 4.9. The comparisons of drifted distance for the two tilted angle cases $(\theta_x, \theta_y, \theta_z)_2$ (color green) and $(\theta_x, \theta_y, \theta_z)_3$ (color red), where the isosurfaces of ϕ^0 and the snapshots when the vesicle cell touches the bottom wall. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

nature, and unconditional energy stability. The algorithm is built upon on the effective union of several algorithms using a new decoupling technology, including the projection method for the incompressible Navier–Stokes equation, and the explicit IEQ method for the nonlinear terms. The realization of unconditional energy stability and full decoupling nature is based on the introduction of some auxiliary variables and the design of their ODEs with specific forms. The system is reformulated based on the "zero-energy-contribution" feature satisfied by the coupled nonlinear terms. The decoupling structure of the scheme enables independent calculation of all variables and at each time step we only need to calculate several linear equations independently. The detailed implementation, solvability, as well as the rigorous proof of unconditional energy stability are also given. Numerous numerical examples are simulated, including the benchmark sedimentation process of a vesicle cell under gravity to illustrate the effectiveness of the proposed scheme.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Xiaofeng Yang reports financial support was provided by National Science Foundation. Xiaofeng Yang reports a relationship with University of South Carolina that includes: employment.

Data availability

No data was used for the research described in the article.

Acknowledgments

X. Yang' work was partially supported by the National Science Foudnation with the grant number DMS-2012490, USA.

References

[1] S. Aland, S. Egerer, J. Lowengrub, A. Voigt, Diffuse interface models of locally inextensible vesicles in a viscous fluid, J. Comput. Phys. 277 (2014) 32–47.

[2] J. Lowengrub, A. Ratz, A. Voigt, Phase-field modeling of the dynamics of multicomponent vesicles: Spinodal decomposition, coarsening, budding, and fission, Phys. Rev. E. 79 (2009) 031926.
[3] X. Wang, Q. Du, Modelling and simulations of multi-component lipid membranes and open membranes via diffusive interface approaches, J. Math. Biol. 56 (2008) 347–371.
[4] C. Funkhouser, F. Solis, K. Thorton, Coupled composition-deformation phase-field method for multicomponent lipid membranes, Phys. Rev. E 76 (2007) 011912.
[5] D. Siegel, M. Kozlov, The Gaussian curvature elastic modulus of N-monomethylated dioleoylphosphatidylethanolamine: Relevance to membrane fusion and lipid phase behavior, Biophys. J. 87 (2004) 366–374.
[6] J. Zhang, S. Das, Q. Du, A phase field model for vesicle-substrate adhesion, J. Comput. Phys. 228 (2009) 7837–7849.
[7] Q. Du, M. Li, C. Liu, Analysis of a phase field navier-stokes vesicle-fluid interaction model, Discrete Contin. Dyn. Syst. Ser. B 8 (3) (2007) 539–556.
[8] Q. Du, C. Liu, X. Wang, A phase field approach in the numerical study of the elastic bending energy for vesicle membranes, J. Comput. Phys. 198 (2004) 450–468.
[9] Q. Du, C. Liu, X. Wang, Simulating the deformation of vesicle membranes under elastic bending energy in three dimensions, J. Comput. Phys. 212 (2005) 757–777.
[10] S.K. Veerapaneni, A. Rahimian, G. Biros, D. Zorin, A fast algorithm for simulating vesicle flows in three dimensions, J. Comput. Phys. 230 (14) (2011) 5610–5634.
[11] T. Biben, K. Kassner, C. Misbah, Phase-field approach to three-dimensional vesicle dynamics, Phys. Rev. E 72 (2005) 041921.
[12] C. Pozrikidis, Numerical simulation of the flow-induced deformation of red blood cells, Ann. Biomed. Eng. 31 (2003) 1194–1205.
[13] Y.L. Han, H. Lin, M.M. Ding, R. Li, T.F. Shi, Flow-induced translocation of vesicles through a narrow pore, Soft Matter 15 (2019) 3307–3314.
[14] M. Mikucki, Y. Zhou, Fast simulation of lipid vesicle deformation using spherical harmonic approximation, Commun. Comput. Phys. 15 (2017) 40–64.
[15] B. Kaoui, J. Harting, Two-dimensional lattice boltzmann simulations of vesicles with viscosity contrast, Rheol. Acta 55 (2016) 465–475.
[16] S.K. Veerapaneni, D. Gueyffier, G. Biros, D. Zorin, A numerical method for simulating the dynamics of 3d axisymmetric vesicles suspended in viscous flows, J. Comput. Phys. 228 (2009) 7233–7249.
[17] M. Kraus, W. Wintz, U. Seifert, R. Lipowsky, Fluid vesicles in shear flow, Phys. Rev. Lett. 77 (1996) 3685–3688.
[18] Z.-H. Huang, A. Manouk, V. Annie, Sedimentation of vesicles: From pear-like shapes to microtether extrusion, New J. Phys. 77 (2011) 035026.
[19] N. Kern, B. Fourcade, Vesicles in linearly forced motion, Europhys. Lett. 46 (1999) 262–267.
[20] I.V. Pivkin, G. Karniadakis, Accurate coarse-grained modeling of red blood cells, Phy. Rev. Lett. 101 (2008) 118105.
[21] J. Deschamps, V. Kantsler, V. Steinberg, Phase diagram of single vesicle dynamical states in shear flow, Phys. Rev. Lett. 102 (2009) 118105.
[22] V. Kantsler, V. Steinberg, Orientation and dynamics of a vesicle in tank-treading motion in shear flow, Phys. Rev. Lett. 95 (2005) 258101.
[23] I.R. Suarez, C. Leidy, G. Tellez, G. Gay, Slow sedimentation and deformability of charged lipid vesicles, PLoS One 8 (07) (2013) 1–12.
[24] D. Matsunaga, Y. Imai, C. Wagner, T. Ishikawa, Reorientation of a single red blood cell during sedimentation, J. Fluid Mech. 806 (2016) 102–128.
[25] J. Rubinstein, P. Sternberg, Nonlocal reaction–diffusion equations and nucleation, IMA J. Appl. Math. 48 (1992) 249–264.
[26] Z. Chai, D. Sun, H. Wang, B. Shi, A comparative study of local and nonlocal Allen-Cahn equations with mass conservation, Int. J. Heat Mass Transfer 122 (2018) 631–642.
[27] S. Aiharaa, T. Takakib, N. Takadac, Multi-phase-field modeling using a conservative Allen–Cahn equation for multiphase flow, Comput. & Fluids 178 (2019) 141–151.
[28] J. Kim, S. Lee, Y. Choi, A conservative Allen–Cahn equation with a space–time dependent lagrange multiplier, Internat. J. Engrg. Sci. 84 (2014) 11–17.
[29] M. Brassel, E. Bretin, A modified phase field approximation for mean curvature flow with conservation of the volume, Math. Methods Appl. Sci. 34 (10) (2011) 1157–1180.
[30] D. Jeong, J. Kim, Conservative Allen–Cahn–Navier–Stokes system for incompressible two-phase fluid flows, Comput. & Fluids 156 (2017) 239–246.
[31] P.-H. Chiu, Y.-T. Lin, A conservative phase field method for solving incompressible two-phase flows, J. Comput. Phys. 230 (1) (2011) 185–204.
[32] X. Wang, L. Ju, Q. Du, Efficient and stable exponential time differencing Runge–Kutta methods for phase field elastic bending energy models, J. Comput. Phys. 316 (2016) 21–38.
[33] F. Guillen-Gonzalez, G. Tierra, Unconditionally energy stable numerical schemes for phase-field vesicle membrane model, J. Comput. Phys. 354 (2018) 67–85.
[34] W. Marth, S. Aland, A. Voigt, Margination of white blood cells: A computational approach by a hydrodynamic phase field model, J. Fluid Mech. 79 (2016) 389–406.
[35] R. Chen, G. Ji, X. Yang, H. Zhang, Decoupled energy stable schemes for phase-field vesicle membrane model, J. Comput. Phys. 302 (2015) 509–523.
[36] X. Yang, L. Ju, Efficient linear schemes with unconditionally energy stability for the phase field elastic bending energy model, Comput. Methods Appl. Mech. Engrg. 315 (2017) 691–712.

[37] Q. Cheng, J. Shen, Multiple scalar auxiliary variable (MSAV) approach and its application to the phase-field vesicle membrane model, SIAM J. Sci. Comput. 40 (2018) A3982–A4006.
[38] X. Yang, Numerical approximations of the Navier–Stokes equation coupled with volume-conserved multi-phase-field vesicles system: Fully-decoupled, linear, unconditionally energy stable and second-order time-accurate numerical scheme, Comput. Methods Appl. Mech. Engrg. 375 (2021) 113600.
[39] X. Yang, A novel fully-decoupled, second-order time-accurate, unconditionally energy stable scheme for a flow-coupled volume-conserved phase-field elastic bending energy model, J. Comput. Phys. 432 (2021) 110015.
[40] R. Nochetto, J.-H. Pyo, The gauge-uzawa finite element method part I: The Navier-Stokes equations, SIAM J. Numer. Anal. 43 (2005) 1043–1068.
[41] J.L. Guermond, P. Minev, J. Shen, An overview of projection methods for incompressible flows, Comput. Methods Appl. Mech. Engrg. 195 (2006) 6011–6045.
[42] J.L. Guermond, L. Quartapelle, A projection FEM for variable density incompressible flows, J. Comput. Phys. 165 (1) (2000) 167–188.
[43] J.-L. Guermond, A. Salgado, A splitting method for incompressible flows with variable density based on a pressure Poisson equation, J. Comput. Phys. 228 (2009) 2834–2846.
[44] J.-L. Guermond, A. Salgado, Error analysis of a fractional time-stepping technique for incompressible flows with variable density, SIAM J. Numer. Anal. 49 (3) (2011) 917–944.
[45] J. Shen, X. Yang, A phase-field model and its numerical approximation for two-phase incompressible flows with different densities and viscosities, SIAM J. Sci. Comput. 32 (2010) 1159–1179.
[46] J. Pyo, J. Shen, Gauge-Uzawa methods for incompressible flows with variable density, J. Comput. Phys. 221 (2007) 181–197.
[47] X. Feng, Fully discrete finite element approximations of the Navier–Stokes-Cahn-Hilliard diffuse interface model for two-phase fluid flows, M2AN Math. Model. Numer. Anal. 44 (2006) 1049–1072.
[48] A. Diegel, C. Wang, X. Wang, S. Wise, Convergence analysis and error estimates for a second order accurate finite element method for the Cahn-Hilliard-Navier–Stokes system, Numer. Math. 137 (2017) 495–534.
[49] D. Han, X. Wang, A second order in time, uniquely solvable, unconditionally stable numerical scheme for Cahn–Hilliard–Navier–Stokes equation, J. Comput. Phys. 290 (2015) 139–156.
[50] D. Han, A. Brylev, X. Yang, Z. Tan, Numerical analysis of second order, fully discrete energy stable schemes for phase field models of two phase incompressible flows, J. Sci. Comput. 70 (2017) 965–989.
[51] D. Han, X. Wang, A second order in time, uniquely solvable, unconditionally stable numerical scheme for the Cahn–Hilliard–Darcy system, J. Sci. Comput. 14 (2018) 1210–1233.
[52] S. Minjeaud, An unconditionally stable uncoupled scheme for a triphasic Cahn–Hilliard/Navier–Stokes model, Numer. Methods Partial Differential Equations 29 (2) (2013) 584–618.
[53] J. Shen, X. Yang, Decoupled, energy stable schemes for phase-field models of two-phase incompressible flows, SIAM J. Numer. Anal. 53 (1) (2015) 279–296.
[54] J. Shen, X. Yang, Decoupled energy stable schemes for phase filed models of two phase complex fluids, SIAM J. Sci. Comput. 36 (2014) B122–B145.
[55] C. Liu, J. Shen, X. Yang, Decoupled energy stable schemes for a phase-field model of two-phase incompressible flows with variable density, J. Sci. Comput. 62 (2015) 601–622.
[56] C. Liu, J. Shen, A phase field model for the mixture of two incompressible fluids and its approximation by a Fourier-spectral method, Physica D 179 (3–4) (2003) 211–228.
[57] H. Abels, H. Garcke, G. Grün, Thermodynamically consistent, frame indifferent diffuse interface models for incompressible two-phase flows with different densities, Math. Models Methods Appl. Sci. 22 (2012) 1150013.
[58] X. Yang, X. He, A fully-discrete decoupled finite element method for the conserved Allen-Cahn type phase-field model of three-phase fluid flow system, Comput. Methods Appl. Mech. Engrg. 389 (2022) 114376.
[59] G.-D. Zhang, X. He, X. Yang, A fully decoupled linearized finite element method with second-order temporal accuracy and unconditional energy stability for incompressible MHD equations, J. Comput. Phys. 448 (2022) 110752.
[60] X. Yang, On a novel fully-decoupled, second-order accurate energy stable numerical scheme for a binary fluid-surfactant phase-field model, SIAM J. Sci. Comput. 43 (2021) B479–B507.
[61] C. Chen, X. Yang, Fully-discrete finite element numerical scheme with decoupling structure and energy stability for the Cahn-Hilliard phase-field model of two-phase incompressible flow system with variable density and viscosity, ESAIM Math. Model. Numer. Anal. 55 (2021) 2323–2347.
[62] J. Shen, J. Xu, J. Yang, The scalar auxiliary variable (SAV) approach for gradient flows, J. Comput. Phys. 353 (2018) 407–416.
[63] C. Chen, X. Yang, Fast, provably unconditionally energy stable, and second-order accurate algorithms for the anisotropic Cahn-Hilliard model, Comput. Methods Appl. Mech. Engrg. 351 (2019) 35–59.
[64] J. Zhang, X. Yang, Decoupled, non-iterative, and unconditionally energy stable large time stepping method for the three-phase Cahn-Hilliard phase-field model, J. Comput. Phys. 404 (2020) 109115.
[65] J. Zhang, X. Yang, Unconditionally energy stable large time stepping method for the L2-gradient flow based ternary phase-field model with precise nonlocal volume conservation, Comput. Methods Appl. Mech. Engrg. 361 (2020) 112743.
[66] J. Shen, X. Yang, The IEQ and SAV approaches and their extensions for a class of highly nonlinear gradient flow systems, Contemp. Math. 754 (2020) 217–245.
[67] J. Shen, J. Xu, J. Yang, A new class of efficient and robust energy stable schemes for gradient flows, SIAM Rev. 61 (2019) 474–506.
[68] F. Boyer, S. Minjeaud, Numerical schemes for a three component Cahn-Hilliard model, ESAIM Math. Model. Numer. Anal. 45 (4) (2011) 697–738.