

STRUCTURE PRESERVING PRIMAL DUAL METHODS FOR GRADIENT FLOWS WITH NONLINEAR MOBILITY TRANSPORT DISTANCES*

JOSÉ A. CARRILLO[†], LI WANG[‡], AND CHAOZHEN WEI[§]

Abstract. We develop structure preserving schemes for a class of nonlinear mobility continuity equation. When the mobility is a concave function, this equation admits a form of gradient flow with respect to a Wasserstein-like transport metric. Our numerical schemes build upon such formulation and utilize modern large-scale optimization algorithms. There are two distinctive features of our approach compared to previous ones. On the one hand, the essential properties of the solution, including positivity, global bounds, mass conservation, and energy dissipation, are all guaranteed by construction. On the other hand, our approach enjoys sufficient flexibility when applied to a large variety of problems including different free energy functionals, general wetting boundary conditions, and degenerate mobilities. The performance of our methods is demonstrated through a suite of examples.

Key words. structure preserving methods, primal dual methods, Wasserstein-like gradient flows, optimal transport distances, minimizing movements

MSC codes. 35A15, 47J25, 47J35, 49M29, 65K10

DOI. 10.1137/23M1562068

1. Introduction. We consider a nonlinear mobility continuity equation of the form

$$(1.1) \quad \partial_t \rho = -\nabla_x \cdot (M(\rho)v(\rho)),$$

where the velocity $v(\rho)$, in the most general form, may contain

$$(1.2) \quad v(\rho) = -\nabla_x H'(\rho) + \nabla_x \Delta_x \rho(x) - \nabla_x V(x) - \nabla_x (W * \rho).$$

Here $H(\rho)$, $V(x)$, and $W(x)$ are given functions with various meanings depending on the specific context. Typical examples include the following:

- Lubrication model for thin films [10] in which case ρ represents the thickness of the film. The mobility and velocity take the expression

*Received by the editors March 29, 2023; accepted for publication (in revised form) October 23, 2023; published electronically February 5, 2024.
<https://doi.org/10.1137/23M1562068>

Funding: The first author was supported by the ERC Advanced grant 883363 (Nonlocal PDEs for Complex Particle Dynamics (Nonlocal-CPD): Phase Transitions, Patterns and Synchronization) under the European Union's Horizon 2020 research and innovation programme. The first author was also partially supported by EPSRC grants EP/V051121/1 (stability analysis for non-linear partial differential equations across multiscale applications) under the EPSRC lead agency agreement with the NSF, and EP/T022132/1 (spectral element methods for fractional differential equations, with applications in applied analysis and medical imaging). The second author received support from NSF grant DMS-1846854. The third author was supported by the National Natural Science Foundation of China grant 12371392.

[†]Mathematical Institute, University of Oxford, Oxford OX2 6GG, UK (carrillo@maths.ox.ac.uk).

[‡]School of Mathematics, University of Minnesota Twin Cities, Minneapolis, MN 55455 USA (liwang@umn.edu).

[§]Corresponding author. School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu, Sichuan 611731, China (cwei4@uestc.edu.cn).

$$(1.3) \quad M(\rho) = \rho^3, \quad v(\rho) = -(Ca)^{1/2} \nabla_x \rho + \nabla_x \Delta_x \rho$$

in the simplest scenario. Here Ca is the capillary number.

- Cahn–Hilliard equation for phase separation in binary alloys [13]. In this case, ρ is often defined to be the difference of local concentrations of two components in the alloy and therefore is in the range of $[-1, 1]$. The mobility is required to be zero in the pure component, i.e., $\rho = \pm 1$, and strictly positive for $|\rho| < 1$, which naturally leads to the choice

$$(1.4) \quad M(\rho) = 1 - \rho^2.$$

The velocity has the form of

$$(1.5) \quad v(\rho) = -\nabla_x \frac{\delta \mathcal{E}}{\delta \rho}, \quad \mathcal{E}(\rho) := \int_{\Omega} H(\rho(x)) + \frac{\epsilon^2}{2} |\nabla \rho(x)|^2 dx,$$

with $\mathcal{E}(\rho)$ being the Ginzburg–Landau free energy. The specific form of H will be given in section 4.2.

- Chemotaxis with prevention of overcrowding [12] by assuming a saturation of the population density ρ . This then yields the mobility of the form

$$(1.6) \quad M(\rho) = \rho(1 - \rho).$$

The nonlinear mobility always comes with a degeneracy, as explicit from (1.3) when $\rho = 0$, or (1.4) when $\rho = \pm 1$, or (1.6) when $\rho = 0$ or 1 . This degeneracy, although making the development of the well-posedness theory a lot more difficult, has a favorable effect on the global bounds of the solution. In particular, it has been conjectured and proved in certain cases that for (1.1) with (1.3), there is a critical threshold in the power of the mobility such that, when the power is above that threshold, the solution remains positive if it started out positively [8]. This is genuinely not true for the constant mobility case due to the oscillatory feature of the fourth-order heat kernel. Likewise, for the Cahn–Hilliard equation (1.1) and (1.5) with nonlinear mobility (1.4), the solution with initial data $|\rho(0, \cdot)| \leq 1$ has the property that $|\rho(t, \cdot)| \leq 1$ for all later times t [18]. This is again in sharp contrast to the constant mobility case which does not preserve such bounds over time due to the lack of the comparison principle.

The mathematical machinery that produces the above results stems from the seminal papers [9, 18], which established two important Lyapunov functionals. One is the energy functional (1.5) or

$$\mathcal{E}(\rho) := \int_{\Omega} [H(\rho(x)) + V(x)\rho(x)] dx + \frac{1}{2} \int_{\Omega} |\nabla \rho(x)|^2 dx + \frac{1}{2} \int_{\Omega} \rho(x)(W * \rho(x)) dx$$

corresponding to the more general velocity (1.2). This energy shall decay over time and leads to regularity estimates. Another is the entropy-like functional

$$(1.7) \quad \mathcal{U}(\rho) := \int_{\Omega} a(\rho(x)) dx, \quad \text{where } a''(\rho) = \frac{1}{M(\rho)}.$$

When $M(\rho) = \rho$, $a(\rho) = \rho \ln \rho$. If the mobility $M(\rho)$ degenerates strongly at the extreme values of ρ , $\mathcal{U}$ controls ρ close to its extremes and therefore leads to the global bounds.

More recently, with the advent of optimal transport theory, (1.1)–(1.2) can be characterized as a gradient flow with respect to a transport metric [1]. This is particularly true if $M(\rho)$ is concave and satisfies other properties [17, 16]. As a result, the weak solution to (1.1)–(1.2) can be obtained by the minimizing movement scheme [31], which is also commonly termed as the JKO scheme after the seminal paper [27]. When $M(\rho) \equiv 1$, it reduces to the constant metric in Hilbert spaces; when $M(\rho) = \rho$, it is the well-studied Wasserstein-2 metric. This variational viewpoint avoids the cumbersome justification of the propagation of global bounds at the analytical level and will also be the stepping stone of our numerical methods developed in this paper.

Indeed, the complex structure of the equation (1.1)–(1.2), originated from the degeneracy in mobility and high-order derivatives, poses severe challenges in designing reliable numerical solvers that would yield physically relevant solutions. One early attempt is in [6] where a nonnegativity preserving finite element method was proposed. The main idea there is to solve a variational problem with a Lagrangian multiplier to advance the negative solution. Later in [36], the authors showed that, by conducting the discretization following the idea from entropy (defined in (1.7)) dissipation at the continuous level, the so derived finite difference scheme is positivity preserving. More recently, a popular line of research concerns the development of scalar auxiliary variable methods [33], which extend significantly on the idea of convex splitting [19, 20]. This approach, although it has been successfully applied to many examples, is still under development for general variable mobilities [23, 25]. Finally, we point out there is also some literature on accelerated primal-dual methods and generalized computation of JKO schemes [26, 22].

Structure preserving finite volume methods have also been developed for Wasserstein gradient flows of zeroth-order functionals [14, 3], for more general mobilities with saturation [4], and for first-order functionals [5] including Cahn–Hilliard type problems as in the present work. These methods have the advantage of keeping the sharp bounds in case of degenerate mobilities while incorporating convex splitting of the free energy functional to obtain their dissipation property at the fully discrete level. These methods are applicable beyond equations with a gradient flow structure being bound preserving with nonlinear mobilities even for systems [4, 21].

In this paper, we will develop a new approach based on the variational formulation mentioned above. More precisely, we rewrite (1.1) and (1.2) as

$$(1.8) \quad \begin{cases} \partial_t \rho = -\nabla \cdot (M(\rho)v) = \nabla \cdot \left(M(\rho) \nabla \frac{\delta \mathcal{E}}{\delta \rho} \right), \\ \mathcal{E}(\rho) = \int_{\Omega} [H(\rho(x)) + V(x)\rho(x)] dx + \frac{\epsilon^2}{2} \int_{\Omega} |\nabla \rho(x)|^2 dx + \int_{\partial\Omega} f_w(\rho, \beta_w) ds. \end{cases}$$

Here we have omitted the interaction term involving W for simplicity, but the methods to be developed shall directly apply. We also add the surface integral of $f_w(\rho, \beta_w)$ to describe the wall free energy. It is defined piecewisely: on the substrate $\Gamma_w \subset \partial\Omega$, its value depends on the phase-field ρ at the wall and the equilibrium contact angle β_w between the free interface and the substrate, determined by the balance of local surface tensions; it is zero on the nonsubstrate boundaries $\partial\Omega \setminus \Gamma_w$.

Taking the variation of $\mathcal{E}$ with respect to ρ ,

$$\begin{aligned} \frac{d}{ds} \mathcal{E}(\rho + sh) \Big|_{s=0} &= \lim_{s \rightarrow 0} \frac{1}{s} [\mathcal{E}(\rho + sh) - \mathcal{E}(\rho)] \\ &= \int_{\Omega} (H'(\rho) + V - \epsilon^2 \Delta \rho) h dx + \int_{\partial\Omega} (\epsilon^2 \nabla \rho \cdot \nu + f'_w(\rho, \beta_w)) h ds, \end{aligned}$$

where ν is an outward-pointing unit vector normal to the wall and $f'_w(\rho, \beta_w)$ denotes the derivative of $f_w(\rho, \beta_w)$ with respect to ρ . Then the chemical potential, defined as the first variation of $\mathcal{E}$ with respect to ρ , is

$$\frac{\delta \mathcal{E}}{\delta \rho} = H'(\rho) + V - \epsilon^2 \Delta \rho.$$

Further, the boundary conditions for (1.8) are a combination of the equilibrium boundary condition for the wall free energy and the no-flux condition for the chemical potential [29, 2],

$$(1.9) \quad \epsilon^2 \nabla \rho \cdot \nu = -f'_w(\rho, \beta_w), \quad M(\rho) \nabla \frac{\delta \mathcal{E}}{\delta \rho} \cdot \nu = 0.$$

Our approach will then be a numerical realization of the minimizing movement scheme [27, 31]. This is a nontrivial extension to the previous works on Wasserstein gradient flow [15, 30] in the following aspects: (1) we propose a new bound preserving proximal solver for the nonlinear transport metric; (2) a nontrivial boundary condition is integrated to account for the wall effect; (3) a preconditioned version of the original primal dual method is explored to accelerate the convergence; (4) the developed methods have been applied to a number of challenging examples. A related work is in [34], where a mirror descent method is developed for variable metric gradient flows. By building the Hessian information in the mirror variable, it accelerates the convergence in optimization and preserves the solution bounds. Compared to the current paper, the method in [34] is built upon a semi-implicit rather than fully implicit version of the minimizing movement approach, and has only been tested for simple prototype models.

The rest of the paper is organized as follows. In the next section, we provide the semidiscrete variational formulation based on a fluid dynamic version of the new transport metric, followed by a fully discrete scheme in both one and two dimensions. Section 3 is devoted to the computation of the proximal operator and resulting primal dual algorithms. Several numerical tests are conducted in section 4, including various energy functionals and boundary conditions.

2. Variational formulation.

2.1. Semidiscrete JKO scheme. Following the dynamic formulation of the JKO scheme [7, 15], we propose the following variational formulation.

PROBLEM 1 (generalized dynamic JKO). *Denote the momentum $m(x, t) = M(\rho)v$. Given $\rho^k(x)$, solve $\rho^{k+1}(x) = \rho(x, 1)$ by*

$$\begin{cases} (\rho(x, t), m(x, t)) \in \arg \inf_{(\rho, m)} \frac{1}{2} d_{\mathcal{W}_m}^2(\rho, \rho^k) + \tau \mathcal{E}(\rho(\cdot, 1)), \\ \text{s.t. } \partial_t \rho + \nabla \cdot m = 0, \quad \rho(x, 0) = \rho^k(x), \quad m \cdot \nu = 0, \end{cases}$$

where

$$d_{\mathcal{W}_m}^2(\rho, \rho^k) = \int_0^1 \int_{\Omega} \phi(\rho, m) dx dt, \quad \phi(\rho, m) = \begin{cases} \frac{|m|^2}{M(\rho)} & \text{if } M(\rho) > 0, \\ 0 & \text{if } (M(\rho), m) = (0, 0), \\ \infty & \text{otherwise.} \end{cases}$$

In this paper, we assume that the nonlinear mobility $M(\rho)$ is a nonnegative concave function within an interval (α, β) and that it could vanish at the boundaries.

For concave mobility $M(\rho)$, Problem 1 is always convex with suitable τ that depends on the convexity property of the energy term $\mathcal{E}$, and it is proved that the solution of the minimization problem exists and weakly converges to the solution of (1.1) as the timestep size $\tau \rightarrow 0$; see, for instance, [16, 31]. Since we only focus on problems where $M(\rho)$ is concave, then the timestep size does not need to be small for the convergence of the optimization steps in simulations in the case when the free energy functional is also convex. The choice of τ is more related to the accuracy of solutions we want to achieve. For more general mobility such that $|m|^2/M(\rho)$ is not convex, the existence of minimizers for such JKO-like schemes and the convergence of the obtained time-discrete sequence to the solution of PDEs is still an open problem.

As with the vanilla JKO formulation, our generalized version shares similar desirable traits such as energy dissipation and mass conservation. Moreover, it also preserves the bounds of the solution automatically.

PROPOSITION 2.1. *The variational formulation has the following properties for any $k \geq 0$:*

- (i) *Energy dissipation:* $\mathcal{E}(\rho^{k+1}) \leq \mathcal{E}(\rho^k)$.
- (ii) *Mass conservation:* $\int_{\Omega} \rho^{k+1} dx = \int_{\Omega} \rho^k dx$.
- (iii) *Bound preservation for nonlinear mobility $M(\rho) = (\rho - \alpha)(\beta - \rho)$:* $\alpha \leq \rho^k \leq \beta$.

Proof. Property (i) is a direct consequence of minimization. Property (ii) is guaranteed by the constraint of continuity equation along with no-flux boundary conditions. Property (iii) comes from the penalization encoded in the definition of $\phi(\rho, m)$. $\square$

Remark 2.2. We would like to emphasize that keeping the mobility term *implicitly* in the definition of the distance is important for maintaining the bound preserving property of the solutions. Only when ρ^{k+1} is put in the mobility $M(\rho)$ in the denominator of the distance functional $\phi(\rho, m)$, can we guarantee that ρ^{k+1} is within the bounded domain such that $M(\rho) > 0$ since the proximal operator of $\phi(\rho, m)$ is bound-preserved (see section 3.3 for detailed discussion). Otherwise, if the mobility is treated *explicitly* such as freezing at the previous time step, and viewing (1.8) as a weighted H^{-1} gradient flow, we will lose the control on confining the solution in a bounded domain as listed in (iii) in the above theorem.

2.2. Fully discrete schemes. We now provide a full discretization to (1). As pointed out in [30], we can remove the artificial time in the dynamic formulation by simply replacing the time derivative with a *one-step* finite difference without violating the first-order accuracy of the JKO scheme, and therefore arrive at the following formulation:

$$\begin{cases} (\rho^{k+1}(x), m^{k+1}(x)) = \arg \inf_{(\rho, m)} \frac{1}{2} d_{\mathcal{W}_m}^2(\rho, \rho^k) + \tau \mathcal{E}(\rho), \\ \text{s.t. } \rho(x) - \rho^k(x) + \nabla \cdot m(x) = 0, \quad m \cdot \nu = 0. \end{cases}$$

In the next two subsections, we will discuss in detail the spatial treatment in the finite difference setting with an effort to conserve mass at the discrete level.

2.2.1. One-dimensional (1D) case. For the 1D problem, we discretize the computational domain $[a, b]$ into N_x cells $C_i = [x_{i-1/2}, x_{i+1/2}]$ with uniform size $\Delta x = (b - a)/N_x$ for $i = 1, \dots, N_x$, and we let $a = x_{1/2}$ and $b = x_{N_x+1/2}$. Then each cell C_i is centered at $x_i = a + (i - 1/2)\Delta x$. We assume the numerical solution $\rho(t, x)$ at each time t is a piecewise constant function with value ρ_i^k in cell C_i at time t^k . In

the following discussion, we may drop the superscript k when it does not cause any confusion.

Then the weighted Wasserstein distance can be approximated by the midpoint rule in space and right endpoint rule in time:

$$(d_{\mathcal{W}_m}^h)^2(\rho^h) = \sum_{i=1}^{N_x} \phi(\rho_i, m_i) \Delta x.$$

The continuity equation is discretized as

$$\rho_i + \frac{1}{2\Delta x}(m_{i+1} - m_{i-1}) = \rho_i^k \quad \text{for } i = 1, \dots, N_x,$$

where m_0 and m_{N_x+1} can be obtained from the no-flux boundary condition. More precisely, using the linear approximation at the boundaries $m_{1/2} = (m_0 + m_1)/2 = 0$ and $m_{N_x+1/2} = (m_{N_x} + m_{N_x+1})/2 = 0$, we have

$$m_0 = -m_1, \quad m_{N_x+1} = -m_{N_x}.$$

Notice that we have used the simplest finite difference approximation of the linear transport equation $\partial_t \rho + \partial_x m = 0$. Any other convergent discretization method for this linear transport equation gives similar results and this does not affect the stability of the simulations as discussed in [15]. Another reason to choose the central finite difference method is its simplicity for the computations of the proximal operators and its easy parallelization at the computational stage. This becomes more apparent below when the proximal operators and their implementation are discussed in section 3.3.

The discretization of the energy functional, denoted as $\mathcal{E}^h(\rho^h)$, reads

$$(2.1) \quad \begin{aligned} \mathcal{E}^h = & \sum_{i=1}^{N_x} \left(H(\rho_i) + V(x_i) \rho_i \right) \Delta x + \frac{\epsilon^2}{4} \left((\nabla \rho)_{\frac{1}{2}}^2 + 2 \sum_{i=1}^{N_x-1} (\nabla \rho)_{i+\frac{1}{2}}^2 + (\nabla \rho)_{N_x+\frac{1}{2}}^2 \right) \Delta x \\ & + f_w(\rho_{\frac{1}{2}}) + f_w(\rho_{N_x+\frac{1}{2}}), \end{aligned}$$

where we have employed the trapezoidal rule for the Dirichlet energy and $(\nabla \rho)_{i+1/2} = (\rho_{i+1} - \rho_i)/\Delta x$ for $i = 0, \dots, N_x$. Here the values on the boundaries $\rho_{1/2} = (\rho_0 + \rho_1)/2$ and $\rho_{N_x+1/2} = (\rho_{N_x} + \rho_{N_x+1})/2$ (or equivalently the values at the ghost points ρ_0 and ρ_{N_x+1}) can be determined via the wetting boundary conditions (see Remark 2.4)

$$(2.2) \quad \epsilon^2 \left(\frac{\rho_1 - \rho_0}{\Delta x} \right) = f'_w(\rho_{1/2}),$$

$$(2.3) \quad \epsilon^2 \left(\frac{\rho_{N_x+1} - \rho_{N_x}}{\Delta x} \right) = -f'_w(\rho_{N_x+1/2}).$$

Then the gradient of $\mathcal{E}^h$ is computed as

$$\frac{\partial \mathcal{E}^h}{\partial \rho_i} = \begin{cases} \left(H'(\rho_1) + V(x_1) \right) \Delta x - \epsilon^2 \left(\frac{\rho_2 - \rho_1}{\Delta x} \right) + f'_w\left(\frac{\rho_0 + \rho_1}{2}\right) & \text{if } i = 1, \\ \left(H'(\rho_{N_x}) + V(x_{N_x}) \right) \Delta x - \epsilon^2 \left(\frac{\rho_{N_x} - \rho_{N_x-1}}{\Delta x} \right) + f'_w\left(\frac{\rho_{N_x} + \rho_{N_x+1}}{2}\right) & \text{if } i = N_x, \\ \left(H'(\rho_i) + V(x_i) \right) \Delta x - \epsilon^2 \left(\frac{\rho_{i+1} - \rho_{i-1}}{\Delta x} \right) & \text{otherwise,} \end{cases}$$

where we have used the linear approximation of ρ_0 and ρ_{N_x} and the wetting boundary conditions (2.2) and (2.3).

In summary, the 1D fully discrete JKO scheme is the following.

PROBLEM 2 (1D discrete generalized dynamic JKO). Given $\{\rho^k\}_{i=1}^{N_x}$, solve $\{\rho^{k+1}\}_{i=1}^{N_x}$ by

$$\begin{cases} (\rho_i^{k+1}, m_i^{k+1}) = \arg \inf_{(\rho^h, m^h)} \sum_{i=1}^{N_x} \frac{1}{2} \phi(\rho_i, m_i) \Delta x + \tau \mathcal{E}^h(\rho^h), \\ \text{s.t. } \rho_i + \frac{1}{2\Delta x} (m_{i+1} - m_{i-1}) = \rho_i^k, \quad m_0 = -m_1, m_{N_x+1} = -m_{N_x}, \end{cases}$$

where $\mathcal{E}^h$ is computed via (2.1).

Remark 2.3 (use of mix boundary conditions). The mix boundary conditions (1.9) are implicitly used in the discrete JKO scheme: the wetting boundary conditions are used in the discretization of free energy functional $\mathcal{E}^h$ and its derivatives $\partial \mathcal{E}^h / \partial \rho_i$; the no-flux boundary conditions are used in the discretization of the constraint of continuity equation.

Remark 2.4 (determination of boundary values using wetting boundary conditions). We can use the wetting boundary conditions (2.2) and (2.3) to determine the values of $\rho_{1/2}$ and $\rho_{N_x+1/2}$ at the boundaries (and hence ρ_0 and ρ_{N_x+1} at the ghost points) in the evaluation of $\mathcal{E}^h$ and $\partial \mathcal{E}^h / \partial \rho_i$. We employ a cubic-polynomial wall energy which can both ensure the vanishing of normal gradient of the phase-field in the bulk region and avoid the formation of the wall layer

$$(2.4) \quad f_w(\rho, \beta_w) = \frac{\epsilon}{\sqrt{2}} \cos \beta_w (\rho^3/3 - \rho).$$

Here we consider the determination of $\rho_{1/2}$ for illustrative purposes. Inserting $\rho_0 = 2\rho_{1/2} - \rho_1$ into (2.2) leads to a quadratic equation for $\rho_{1/2}$:

$$(2.5) \quad \gamma X^2 + \epsilon X - (\epsilon \rho_1 + \gamma) = 0,$$

where $\gamma = \frac{\sqrt{2}\Delta x}{4} \cos \beta_w$. Note that (2.5) has two solutions and $\rho_{1/2}$ is the solution that lies within the range of the phase-field:

$$\rho_{1/2} = \begin{cases} -\frac{\epsilon}{2\gamma} + \sqrt{(\rho_1 + \frac{\epsilon}{2\gamma})^2 + (1 - \rho_1^2)} & \text{if } \cos \beta_w > 0, \\ -\frac{\epsilon}{2\gamma} - \sqrt{(\rho_1 + \frac{\epsilon}{2\gamma})^2 + (1 - \rho_1^2)} & \text{if } \cos \beta_w < 0, \\ \rho_1 & \text{if } \cos \beta_w = 0. \end{cases}$$

Similarly, we can obtain the value of $\rho_{N_x+1/2}$ using (2.3).

2.2.2. Two-dimensional (2D) case. Consider the computational domain $\Omega = [a, b] \times [c, d]$, where the substrate boundary is $\Gamma_w = [a, b] \times \{y = c\}$ and the nonsubstrate boundary is $\partial\Omega \setminus \Gamma_w = \{x = a \text{ or } x = b\} \times [c, d] \cup [a, b] \times \{y = d\}$. We divide the domain Ω into $N_x \times N_y$ cells $C_{i,j} = [x_{i-1/2}, x_{i+1/2}] \times [y_{j-1/2}, y_{j+1/2}]$ with uniform size $\Delta x \Delta y = \left(\frac{b-a}{N_x}\right) \left(\frac{d-c}{N_y}\right)$ for $i = 1, \dots, N_x$ and $j = 1, \dots, N_y$. Then $a = x_{1/2}$, $b = x_{N_x+1/2}$, $c = y_{1/2}$, $d = y_{N_y+1/2}$, and the center of the cell $C_{i,j}$ is (x_i, y_i) with $x_i = a + (i-1/2)\Delta x$ and $y_j = c + (j-1/2)\Delta y$.

The discrete energy is obtained by applying the mid-point rule for the integral of $H(\rho)$, $V(x)\rho$, and $f_w(\rho)$ and the trapezoidal rule for the integral of $\|\nabla\rho\|$:

$$\begin{aligned}
 \mathcal{E}^h = & \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} \left(H(\rho_{i,j}) + V(x_{i,j})\rho_{i,j} \right) \Delta x \Delta y \\
 & + \frac{\epsilon^2}{2} \left(\sum_{j=1}^{N_y} \sum_{i=1}^{N_x-1} (\nabla_x \rho)_{i+\frac{1}{2},j}^2 + \sum_{i=1}^{N_x} \sum_{j=1}^{N_y-1} (\nabla_y \rho)_{i,j+\frac{1}{2}}^2 \right) \Delta x \Delta y \\
 & + \frac{\epsilon^2}{4} \sum_{i=1}^{N_x} (\nabla_y \rho)_{i,\frac{1}{2}}^2 \Delta x \Delta y + \sum_{i=1}^{N_x} f_w(\rho_{i,\frac{1}{2}}) \Delta x,
 \end{aligned}
 \tag{2.6}$$

where we have used the nonsubstrate wetting boundary condition $\nabla\rho \cdot \hat{n} = 0$ on $\partial\Omega \setminus \Gamma_w$:

$$(\nabla_x \rho)_{\frac{1}{2},j} = (\nabla_x \rho)_{N_x+\frac{1}{2},j} = (\nabla_y \rho)_{i,N_y+\frac{1}{2}} = 0.$$

The gradient of ρ is approximated by second-order finite difference:

$$(\nabla_x \rho)_{i+\frac{1}{2},j} = \frac{\rho_{i+1,j} - \rho_{i,j}}{\Delta x}, \quad (\nabla_y \rho)_{i,j+\frac{1}{2}} = \frac{\rho_{i,j+1} - \rho_{i,j}}{\Delta y}.$$

Then the gradient of $\mathcal{E}^h$ is

$$\frac{\partial \mathcal{E}^h}{\partial \rho_{i,j}} = \left(H'(\rho_{i,j}) + V(x_{i,j})\rho_{i,j} - \epsilon^2 L_{i,j} \right) \Delta x \Delta y + W_{i,j} \Delta x,
 \tag{2.7}$$

where $L_{i,j} = (L^x)_{i,j} + (L^y)_{i,j}$ is

$$\begin{aligned}
 (L^x)_{i,j} &= \begin{cases} \frac{\rho_{2,j} - \rho_{1,j}}{\Delta x^2} & \text{if } i = 1, \\ \frac{\rho_{i-1,j} - \rho_{i,j}}{\Delta x^2} & \text{if } i = N_x, \\ \frac{\rho_{i+1,j} - 2\rho_{i,j} + \rho_{i-1,j}}{\Delta x^2} & \text{otherwise,} \end{cases} \\
 (L^y)_{i,j} &= \begin{cases} \frac{\rho_{i,2} - \rho_{i,1}}{\Delta y^2} & \text{if } j = 1, \\ \frac{\rho_{i,j-1} - \rho_{i,j}}{\Delta y^2} & \text{if } j = N_y, \\ \frac{\rho_{i,j+1} - 2\rho_{i,j} + \rho_{i,j-1}}{\Delta y^2} & \text{otherwise,} \end{cases}
 \end{aligned}$$

and the wall-energy part $W_{i,j}$ is

$$W_{i,j} = \begin{cases} f'_w(\rho_{i,1/2}) & \text{if } j = 1, \\ 0 & \text{otherwise.} \end{cases}$$

Note that we have used the approximation $\rho_{i,1/2} = (\rho_{i,0} + \rho_{i,1})/2$ and the wetting boundary condition on solid substrate Γ_w in the derivation of $\nabla \mathcal{E}^h(\rho^h)$:

$$\epsilon^2 \left(\frac{\rho_{i,1} - \rho_{i,0}}{\Delta y} \right) = f'_w \left(\frac{\rho_{i,0} + \rho_{i,1}}{2} \right).$$

Again, we obtain the values $\rho_{i,1/2}$ (and hence $\rho_{i,0}$) involved in (2.6) and (2.7) by the above boundary condition according to Remark 2.4. Then we have the following discrete JKO scheme.

PROBLEM 3 (2D discrete generalized dynamic JKO). Given $\{\rho^k\}_{i,j}$, solve $\{\rho^{k+1}\}_{i,j}$ by

$$\begin{cases} (\rho_{i,j}^{k+1}, \mathbf{m}_{i,j}^{k+1}) = \arg \inf_{(\rho^h, \mathbf{m}^h)} \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} \frac{1}{2} \phi(\rho_{i,j}, \mathbf{m}_{i,j}) \Delta x \Delta y + \tau \mathcal{E}^h(\rho^h), \\ \text{s.t. } \rho_{i,j} + \frac{1}{2\Delta x} (m_{i+1,j}^x - m_{i-1,j}^x) + \frac{1}{2\Delta y} (m_{i,j+1}^y - m_{i,j-1}^y) = \rho_{i,j}^k, \\ m_{0,j}^x = -m_{1,j}^x, m_{N_x+1,j}^x = -m_{N_x,j}^x, m_{i,0}^y = -m_{i,1}^y, m_{i,N_y+1}^y = -m_{i,N_y}^y. \end{cases}$$

3. Primal-dual algorithm. Upon discretization, the discrete generalized dynamic JKO scheme amounts to solving an optimization problem subject to a linear constraint:

$$\min_u \Phi(u) + E(u), \quad \text{s.t. } Au = b,$$

where we have rewritten the constraint of the discretized continuity equation in the form $Au = b$ and we define

$$\begin{aligned} u &= (\rho^h, (m^x)^h, (m^y)^h) = ((\rho_{i,j})_{1 \leq i \leq N_x, 1 \leq j \leq N_y}, (m_{i,j}^x)_{1 \leq i \leq N_x, 1 \leq j \leq N_y}, (m_{i,j}^y)_{1 \leq i \leq N_x, 1 \leq j \leq N_y}), \\ \Phi(u) &= \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} \frac{1}{2\tau} \phi(\rho_{i,j}, \mathbf{m}_{i,j}) \Delta x \Delta y, \\ E(u) &= \mathcal{E}^h(\rho^h). \end{aligned}$$

This minimization problem can be reformulated as an unconstrained optimization problem:

$$\min_u \Phi(u) + E(u) + i_\delta(Au), \quad i_\delta(y) = \begin{cases} 0 & \text{if } \|y - b\|_2 \leq \delta, \\ \infty & \text{otherwise.} \end{cases}$$

Here we relax the equality of the linear constraint at the fully discrete level to an inequality by a small parameter δ , given that even an exact solution of the continuity equation at the continuum level will only satisfy the discrete linear constraint up to an error term depending on the order the finite difference operators.

3.1. Primal-dual method for three operators. We can apply the primal dual splitting scheme for three operators in [35] to solve this minimization problem:

$$(3.1) \quad \begin{cases} \varphi^{(l+1)} = \text{Prox}_{\sigma i_\delta^*}(\varphi^{(l)} + \sigma A \bar{u}^{(l)}), \\ u^{(l+1)} = \text{Prox}_{\lambda \Phi}(u^{(l)} - \lambda \nabla E(u^{(l)}) - \lambda A^t \varphi^{(l+1)}), \\ \bar{u}^{(l+1)} = 2u^{(l+1)} - u^{(l)} + \lambda \nabla E(u^{(l)}) - \lambda \nabla E(u^{(l+1)}), \end{cases}$$

where we require $\sigma\lambda < 1/\lambda_{\max}(AA^t)$ for the convergence. The primal-dual method for one step of the discrete dynamic JKO scheme is shown in Algorithm 1 (PDJKO), where we choose the initial guesses as follows (here we take 2D case for illustrative purposes):

$$u^0 = (\rho^0, \mathbf{0}_{N_x \times N_y}, \mathbf{0}_{N_x \times N_y}), \quad \varphi^0 = \mathbf{0}_{N_x \times N_y}.$$

Algorithm 1. Primal-dual method for one step of dynamic JKO.

Input: $u^0, \varphi^0, \text{Iter}_{\max}, \lambda, \sigma, \tau > 0$
Output: u^*, φ^*

```

1 Let  $\bar{u}^0 = u^0$  and  $l = 0$ ;
2 while  $l < \text{Iter}_{\max}$  do
3   repeat
4      $\varphi^{(l+1)} = \text{Prox}_{\sigma i_\delta^*}(\varphi^{(l)} + \sigma A \bar{u}^{(l)}),$ 
5      $u^{(l+1)} = \text{Prox}_{\lambda \Phi}(u^{(l)} - \lambda \nabla E(u^{(l)}) - \lambda A^t \varphi^{(l+1)}),$ 
6      $\bar{u}^{(l+1)} = 2u^{(l+1)} - u^{(l)} + \lambda \nabla E(u^{(l)}) - \lambda \nabla E(u^{(l+1)}),$ 
7   until stopping criteria are achieved;
8    $u^* = u^{(l+1)}$ 
9    $\varphi^* = \varphi^{(l+1)}$ 
10 end

```

We update the variables until achieving the stopping criteria that consist of the constraint and the convergence monitors:

$$\|Au^{(l+1)} - b\|_2 = \left| \rho_{i,j} - \rho_{i,j}^0 + \frac{m_{i+1,j}^x - m_{i-1,j}^x}{2\Delta x} + \frac{m_{i,j+1}^y - m_{i,j-1}^y}{2\Delta y} \right|^2 \Delta x \Delta y \leq \delta,$$

$$\max \left\{ \frac{\|u^{(l+1)} - u^{(l)}\|}{\|u^{(l+1)}\|}, \frac{\|\varphi^{(l+1)} - \varphi^{(l)}\|}{\|\varphi^{(l+1)}\|} \right\} \leq \text{TOL},$$

$$\max \left\{ \frac{|E(u^{(l+1)}) - E(u^{(l)})|}{|E(u^{(l+1)})|}, \frac{|\Phi(u^{(l+1)}) - \Phi(u^{(l)})|}{|\Phi(u^{(l+1)})|} \right\} \leq \text{TOL}.$$

The success of this algorithm depends on the ease of computing the two proximal operators, which in general is not trivial. Fortunately, we can compute $\text{Prox}_{\lambda \Phi}$ easily by performing a Newton iteration method with a strategy for choosing initial guesses that guarantees the convergence (which is discussed in section 3.3), and we also have an explicit formula for $\text{Prox}_{\sigma i_\delta^*}$. By Moreau's identity, we can write $\text{Prox}_{\sigma i_\delta^*}$ in terms of projections onto balls of radius σ centered at b :

$$\text{Prox}_{\sigma i_\delta^*}(y) = y - \sigma \text{Proj}_{B_\delta}(y/\sigma), \quad \text{Proj}_{B_\delta}(y) = \begin{cases} y, & \|y - b\|_2 \leq \delta, \\ \delta \frac{y - b}{\|y - b\|_2} + b & \text{otherwise.} \end{cases}$$

3.2. Accelerated primal-dual method by preconditioning. For the phase-separation simulation for the 2D Cahn–Hilliard equation (Figure 7), the PDJKO method (Algorithm 1) may converge slowly (see Figure 9). Inspired by the recent work on the acceleration of the original primal-dual method for two operators by preconditioning [32], we propose the preconditioned primal-dual method for one step of the dynamic JKO scheme (PrePDJKO; see Algorithm 2) that converges much faster:

$$\begin{cases} \varphi^{(l+1)} = \text{Prox}_{i_\delta^*}^{M_2}(\varphi^{(l)} + M_2^{-1} A \bar{u}^{(l)}), \\ u^{(l+1)} = \text{Prox}_{\Phi}^{M_1}(u^{(l)} - M_1^{-1} \nabla E(u^{(l)}) - M_1^{-1} A^t \varphi^{(l+1)}), \\ \bar{u}^{(l+1)} = 2u^{(l+1)} - u^{(l)} + M_1^{-1} \nabla E(u^{(l)}) - M_1^{-1} \nabla E(u^{(l+1)}), \end{cases}$$

Algorithm 2. Preconditioned primal-dual method for one step of dynamic JKO.

Input: $u^0, \varphi^0, \text{Iter}_{\max}, \lambda, \sigma, \tau > 0$
Output: u^*, φ^*

```

1 Let  $\bar{u}^0 = u^0$  and  $l = 0$ ;
2 while  $l < \text{Iter}_{\max}$  do
3   repeat
4      $\varphi^{(l+1)} = \text{Prox}_{i_\delta^*}^{M_2}(\varphi^{(l)} + M_2^{-1}A\bar{u}^{(l)}),$ 
5      $u^{(l+1)} = \text{Prox}_{\Phi}^{M_1}(u^{(l)} - M_1^{-1}\nabla E(u^{(l)}) - M_1^{-1}A^t\varphi^{(l+1)}),$ 
6      $\bar{u}^{(l+1)} = 2u^{(l+1)} - u^{(l)} + M_1^{-1}\nabla E(u^{(l)}) - M_1^{-1}\nabla E(u^{(l+1)}),$ 
7     where  $M_1 = \frac{1}{\lambda}I$  and  $M_2 = \lambda AA^t$ ,
8   until stopping criteria are achieved;
9    $u^* = u^{(l+1)}$ 
10   $\varphi^* = \varphi^{(l+1)}$ 
11 end

```

where the extended proximal operator is defined as

$$\text{Prox}_f^M(y) = \arg \min_x \frac{1}{2}\|x - y\|_M^2 + f(x), \text{ where } \|x\|_M^2 = x^t M x.$$

Here we use $M_1 = \frac{1}{\lambda}I$ and $M_2 = \lambda AA^t$ with λ being a tuning parameter to achieve acceleration. Then $\text{Prox}_{\Phi}^{M_1} = \text{Prox}_{\lambda\Phi}$ as given in section 3.3. Moreover, one can show the Moreau-like identity for the extended proximal operator

$$\text{Prox}_{f^*}^{M_2}(y) = y - M_2^{-1}\text{Prox}_f^{M_2^{-1}}(M_2 y),$$

which provides the expression of $\text{Prox}_{i_\delta^*}^{M_2}(y)$ in terms of projections onto balls of radius δ centered at b :

$$\text{Prox}_{i_\delta^*}^{M_2}(y) = y - M_2^{-1}\text{Proj}_{B_\delta}(M_2 y), \quad \text{Proj}_{B_\delta}(y) = \begin{cases} y, & \|y - b\|_2 \leq \delta, \\ \delta \frac{y - b}{\|y - b\|_2} + b & \text{otherwise.} \end{cases}$$

3.3. Computing $\text{Prox}_{\lambda\Phi}(u)$. The efficiency of the method relies largely on the computation of the proximal operator $\text{Prox}_{\lambda\Phi}$. For linear mobility $M(\rho) = \rho$, we obtained an explicit formula for the proximal operator; see [15]. This is, however, not true for the nonlinear mobility case with $M(\rho) = (\rho - \alpha)(\beta - \rho)$. Nevertheless, Newton's iteration provides a viable surrogate for computing the proximal. In addition, we can prove that with appropriate choice of initial guess, the Newton iteration converges to a solution that lies within the desired range and thus makes the whole solver bound preserving.

Since $\Phi(u) = \sum_{i,j} \frac{1}{2\tau} \phi(\rho_{i,j}, m_{i,j})$ is separable, its proximal operator is componentwise, i.e., $\text{Prox}_{\lambda\Phi}(u) = (\text{Prox}_{\frac{\lambda}{2\tau}\phi}(\rho_{i,j}, m_{i,j}))_{1 \leq i \leq N_x}^{1 \leq j \leq N_y}$. We compute the proximal operator $\text{Prox}_{\frac{\lambda}{2\tau}\phi}(\rho, m)$ by using Newton iteration. For $M(\rho) = (\rho - \alpha)(\beta - \rho)$, we can choose appropriate initial values to guarantee the convergence of Newton iteration, regardless of the time step τ . Furthermore, we can guarantee that the convergent solution satisfies the desired bounds, $\rho \in [\alpha, \beta]$.

The proximal operator of $\phi(\rho, m)$ with a general λ is given by

$$(3.2) \quad \text{Prox}_{\frac{\lambda}{2}\phi}(\rho, m) = \arg \min_{\tilde{\rho}, \tilde{m}} \frac{1}{2} |\tilde{\rho} - \rho|^2 + \frac{1}{2} \|\tilde{m} - m\|^2 + \frac{\lambda}{2} \phi(\tilde{\rho}, \tilde{m}),$$

where

$$\phi(\rho, m) = \begin{cases} \frac{\|m\|^2}{M(\rho)} & \text{if } M(\rho) > 0, \\ 0 & \text{if } (M(\rho), m) = (0, 0), \\ \infty & \text{otherwise.} \end{cases}$$

The definition of $\phi(\rho, m)$ guarantees the preserving of the bounds of $\alpha \leq \rho^* \leq \beta$ for $(\rho^*, m^*) = \text{Prox}_{\frac{\lambda}{2}\phi}(\rho, m)$.

Let us first restrict our consideration to $\alpha < \tilde{\rho} < \beta$, for which we consider the minimization

$$\min F(\tilde{\rho}, \tilde{m}) = \frac{1}{2} |\tilde{\rho} - \rho|^2 + \frac{1}{2} \|\tilde{m} - m\|^2 + \frac{\lambda \|\tilde{m}\|^2}{2M(\tilde{\rho})}.$$

The optimal conditions for minimization yield

$$\begin{cases} \frac{\partial F}{\partial \tilde{\rho}} = \tilde{\rho} - \rho - \lambda M'(\tilde{\rho}) \frac{\|\tilde{m}\|^2}{2M^2(\tilde{\rho})} = 0, \\ \frac{\partial F}{\partial \tilde{m}} = \tilde{m} - m + \lambda \frac{\tilde{m}}{M(\tilde{\rho})} = 0, \end{cases}$$

which reduces to

$$f(\tilde{\rho}) = \tilde{\rho} - \rho - \lambda M'(\tilde{\rho}) \frac{\|m\|^2}{2(\lambda + M(\tilde{\rho}))^2} = 0.$$

We use the Newton iteration method to find the root $\rho^* \in (\alpha, \beta)$ of $f(\tilde{\rho})$. Depending on the monotonicity and concavity of $f(\tilde{\rho})$, we can choose the appropriate initial values to guarantee the convergence of the Newton iteration. Taking derivatives of $f(\tilde{\rho})$ gives

$$\begin{cases} f'(\tilde{\rho}) = 1 + (M'(\tilde{\rho}))^2 \frac{\lambda \|m\|^2}{(\lambda + M(\tilde{\rho}))^3} - M''(\tilde{\rho}) \frac{\lambda \|m\|^2}{2(\lambda + M(\tilde{\rho}))^2}, \\ f''(\tilde{\rho}) = M'(\tilde{\rho}) \frac{3\lambda \|m\|^2}{(\lambda + M(\tilde{\rho}))^3} \left(M''(\tilde{\rho}) - \frac{(M'(\tilde{\rho}))^2}{\lambda + M(\tilde{\rho})} \right) - M'''(\tilde{\rho}) \frac{\lambda \|m\|^2}{2(\lambda + M(\tilde{\rho}))^2}. \end{cases}$$

For $M(\tilde{\rho}) = (\tilde{\rho} - \alpha)(\beta - \tilde{\rho})$ and $\lambda > 0$, we can show that

(1) $f'(\tilde{\rho}) > 0$ for $\alpha < \tilde{\rho} < \beta$;

(2) $f''(\tilde{\rho}) < 0$ for $\alpha < \tilde{\rho} < (\alpha + \beta)/2$ and $f''(\tilde{\rho}) > 0$ for $(\alpha + \beta)/2 < \tilde{\rho} < \beta$.

Given that $f(\tilde{\rho})$ is monotonically increasing in (α, β) , if there exists a subinterval $(a, b) \subset (\alpha, \beta)$ such that $f(a) < 0$ and $f(b) > 0$, we have $\rho^* \in (a, b)$. Then we can choose the initial guess $\tilde{\rho}_0 = a$ ($\tilde{\rho}_0 = b$) if $f''(\tilde{\rho}) < 0$ ($f''(\tilde{\rho}) > 0$) to guarantee the convergence of the Newton iteration. By extending the domain of $f(\tilde{\rho})$ to where it has meaning, we can evaluate the following values:

$$\begin{aligned} f(\alpha) &= \alpha - (\beta - \alpha) \frac{\|m\|^2}{2\lambda} - \rho, \\ f\left(\frac{\alpha + \beta}{2}\right) &= \frac{\alpha + \beta}{2} - \rho, \\ f(\rho) &= \frac{\lambda \|m\|^2}{(\lambda + M(\rho))^2} \left(\rho - \frac{\alpha + \beta}{2} \right), \\ f(\beta) &= \beta + (\beta - \alpha) \frac{\|m\|^2}{2\lambda} - \rho. \end{aligned}$$

Notice that the input of the proximal operator ρ can be outside of $[\alpha, \beta]$ (see (3.1)). Depending on the value of ρ , we have the following strategy for choosing the initial guess $\tilde{\rho}_0$ for the Newton iteration that converges to desired solution $\rho^* \in (\alpha, \beta)$:

- Case 1:** When $\alpha \leq \rho < (\alpha + \beta)/2$, we have $f(\rho) < 0$ and $f((\alpha + \beta)/2) > 0$, and hence $\rho^* \in (\rho, (\alpha + \beta)/2) \subset (\alpha, \beta)$. Since $f''(\tilde{\rho}) < 0$ on $\rho < \tilde{\rho} < (\alpha + \beta)/2$, we set the initial guess $\tilde{\rho}_0 = \rho$.
- Case 2:** When $(\alpha + \beta)/2 < \rho \leq \beta$, we have $f((\alpha + \beta)/2) < 0$, $f(\rho) > 0$, and hence $\rho^* \in ((\alpha + \beta)/2, \rho) \subset (\alpha, \beta)$. Since $f''(\tilde{\rho}) > 0$ on $(\alpha + \beta)/2 < \tilde{\rho} < \rho$, we set the initial guess $\tilde{\rho}_0 = \rho$.
- Case 3:** When $\rho = (\alpha + \beta)/2$, $f((\alpha + \beta)/2) = 0$, then the optimal solution is $\rho^* = (\alpha + \beta)/2$.
- Case 4:** When $\alpha - (\beta - \alpha) \frac{\|m\|^2}{2\lambda} < \rho < \alpha$, $f(\alpha) < 0$ and $f((\alpha + \beta)/2) > 0$; hence $\rho^* \in (\alpha, (\alpha + \beta)/2) \subset (\alpha, \beta)$. Since $f''(\tilde{\rho}) < 0$ on $\alpha < \tilde{\rho} < (\alpha + \beta)/2$, we set the initial guess $\tilde{\rho}_0 = \alpha$.
- Case 5:** When $\beta < \rho < \beta + (\beta - \alpha) \frac{\|m\|^2}{2\lambda}$, $f((\alpha + \beta)/2) < 0$ and $f(\beta) > 0$; hence $\rho^* \in ((\alpha + \beta)/2, \beta) \subset (\alpha, \beta)$. Since $f''(\tilde{\rho}) > 0$ on $(\alpha + \beta)/2 < \tilde{\rho} < \beta$, we set the initial guess $\tilde{\rho}_0 = \beta$.

When $\rho \leq \alpha - (\beta - \alpha) \frac{\|m\|^2}{2\lambda}$ ($\rho \geq \beta + (\beta - \alpha) \frac{\|m\|^2}{2\lambda}$), we have $f(\tilde{\rho}) > f(\alpha) \geq 0$ ($f(\tilde{\rho}) < f(\beta) \leq 0$) on (α, β) , and hence there exists no root for $f(\tilde{\rho})$ within (α, β) . Then the solution ρ^* to the proximal operator (3.2) must be obtained at the endpoints of (α, β) , as follows:

- Case 6:** When $\rho \leq \alpha - (\beta - \alpha) \frac{\|m\|^2}{2\lambda}$, we have $(\rho^*, m^*) = (\alpha, 0)$.
- Case 7:** When $\rho \geq \beta + (\beta - \alpha) \frac{\|m\|^2}{2\lambda}$, we have $(\rho^*, m^*) = (\beta, 0)$.

3.4. Computational cost. The computational complexity of the proposed algorithms primarily hinges on the number of iterations needed for convergence within the one-step JKO scheme, which unfortunately lacks a rigorous theoretical estimate. Nonetheless, assuming that N_{iter} represents the iteration count needed for convergence, we can readily deduce that the computational complexity of the primal-dual method for the one-step JKO scheme (Algorithm 1) is on the order of $O(N_x \times N_y)$. This holds true, given that the proximal operator $\text{Prox}_{\sigma_i^*}$ entails solely sparse matrix-vector multiplication and vector l^2 -norm calculations, while the proximal operator $\text{Prox}_{\lambda\Phi}$ involves only elementwise computations. Note that the iteration required for the convergence of the Newton method for calculating $\text{Prox}_{\lambda\Phi}$ is not an issue here since we have provided a provable convergent method for choosing the initial guess.

Our approach holds significant potential for large-scale computations demanding strict structure preserving properties, particularly in the context of three-dimensional problems. We've leveraged the preconditioned primal-dual method for optimization, which necessitates significantly fewer iterations compared to the primal-dual method. Additionally, the application of parallel algorithms for computing the proximal operator of the distance functional further enhances efficiency. One bottleneck in our current method (Algorithm 2) arises from the matrix inversion required when computing the proximal operator of the indicator function. However, it's worth noting that there exist various efficient algorithms for matrix inversion, such as FFT-based methods (see [26]). This capability augments the promise of our approach for addressing large-scale computational problems. This direction will be explored elsewhere.

4. Numerical results. This section is devoted to showcasing the flexibility and applicability of our proposed approach to several challenging problems.

4.1. 1D saturation experiment. In the first example we investigate the saturation effect due to the nonlinear mobility. Consider the equation

$$(4.1) \quad \rho_t = \nabla \cdot \left(\rho(\alpha - \rho) \nabla \left(D \ln(\rho) + \frac{C}{2} |x|^2 \right) \right),$$

whose corresponding energy is

$$\mathcal{E} = D\rho(\ln \rho - 1) + \frac{C}{2} |x|^2 \rho,$$

with nonlinear degenerate mobility $M(\rho) = \rho(\alpha - \rho)$. The steady state of this problem depends on the conserved mass of the solution $M = \|\rho_0\|_{L^1}$ in the sense that

$$(4.2) \quad \rho_\infty(x) = \begin{cases} M \sqrt{\frac{C}{2\pi D}} \exp\left(-\frac{C}{2D} x^2\right) & \text{if } M \leq M_c, \\ \alpha \exp\left(-\frac{C}{2D} (x^2 - l^2)^+\right) & \text{if } M > M_c, \end{cases}$$

where $M_c = \alpha \sqrt{\frac{2\pi D}{C}}$ is a critical mass, $(s)^+ = \max\{s, 0\}$, and l is a positive constant to make sure that $\|\rho_\infty\|_{L^1} = M$. Clearly, (4.2) indicates an upper bound on ρ_∞ , $\rho \leq \alpha$. More particularly, when M is beyond the critical value M_c , ρ_∞ has two segments: constant α when $|x| \leq l$ and an exponential decay function when $|x| \geq l$. This is the saturation effect.

Numerically, we solve (4.1) over the domain $\Omega = [-4, 4]$ with parameters $\alpha = 1$, $C = 1$, $D = 1$, and start from a uniform initial density $\rho_0 \in [-4, 4]$ with the supercritical mass $M = 3.32$. We plot the evolution of ρ for $t \in [0, 15]$ computed by the generalized dynamics JKO scheme (see Problem 2) for various Δx , as shown in Figure 1. We observe oscillations as ρ approaches 1, which can be reduced by finer mesh (see the comparison between the results for $\Delta x = 0.04$ and $\Delta x = 0.01$).

Moreover, we can reduce the oscillation near $\rho = 1$ by computing the evolution by the generalized Schrödinger bridge (GSB) scheme (see SBP scheme in Remark 4.1), which is equivalent to the Fisher information regularization [30]. We implement the GSB scheme with adaptive regularization coefficient $\eta = 1/(1 - \|\rho\|_\infty)$ starting from $\eta_0 = 80$, shown in Figure 2.

1D Saturation experiment by JKO scheme

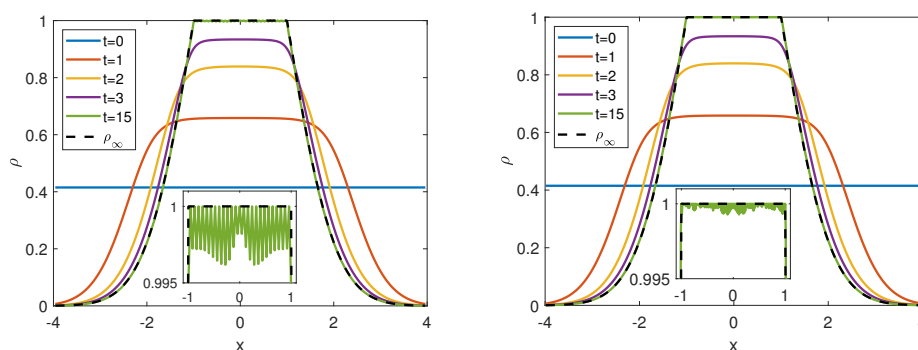


FIG. 1. Evolution of solutions to the 1D saturation equation with mobility $M(\rho) = \rho(1 - \rho)$ for $t \in [0, 15]$. Left: $\Delta x = 0.04$, $\tau = 0.01$; right: $\Delta x = 0.01$, $\tau = 0.01$. The insets are the zoom-in figures for the oscillation when ρ is close to 1.

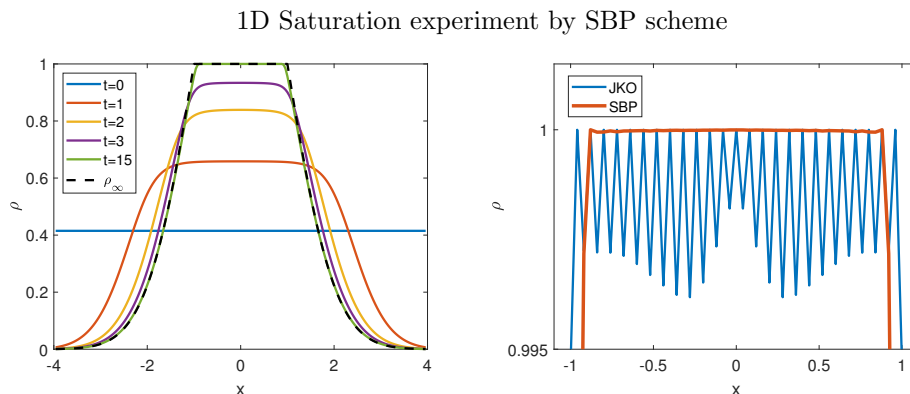


FIG. 2. Left: Evolution of solutions for 1D saturation experiment by the generalized Schrödinger bridge scheme with $\Delta x = 0.04$, $\tau = 0.01$, $\eta_0 = 80$; Right: Comparison with generalized dynamic JKO scheme.

Remark 4.1 (generalized Schrödinger bridge problem). To avoid oscillation as $\rho \rightarrow \alpha$ and $\rho \rightarrow \beta$ that may appear (for example, in the 1D saturation experiment), we propose the following scheme inspired by the Schrödinger bridge problem:

Defining $\mathcal{H}(\rho) = \int_{\Omega} \frac{1}{\beta - \alpha} ((\rho - \alpha) \ln(\rho - \alpha) + (\beta - \rho) \ln(\beta - \rho)) dx$, we solve $\rho^{k+1}(x) = \rho(x, 1)$ by

$$\begin{cases} (\rho(x, t), m(x, t)) = \arg \inf_{(\rho, m)} \int_0^1 \int_{\Omega} \frac{\|m\|^2}{2M(\rho)} dx dt + \tau \left(\mathcal{E}(\rho(\cdot, 1)) - \eta^{-1} \mathcal{H}(\rho(\cdot, 1)) \right), \\ \text{s.t. } \partial_t \rho + \nabla \cdot m = \tau \eta^{-1} \Delta \rho, \quad \rho(x, 0) = \rho^k(x), \quad (m - \tau \eta^{-1} \nabla \rho) \cdot \nu = 0. \end{cases}$$

The auxiliary entropy $\mathcal{H}$ keeps ρ away from α and β . The generalized Schrödinger bridge problem (SBP) is equivalent to the JKO scheme with Fisher information regularization (FIR), and it does not violate the first-order accuracy of the JKO scheme. As $\eta \rightarrow \infty$, the SBP (or FIR) recovers the JKO scheme.

4.2. 1D Cahn–Hilliard equation. In the second example, we consider the Cahn–Hilliard equation

$$\rho_t = \nabla \cdot \left(M(\rho) \nabla (H'(\rho) - \epsilon^2 \Delta \rho) \right),$$

with nonlinear mobility $M(\rho) = (1 + \rho)(1 - \rho)$. This is the model used to study the phase separation in binary alloys, where ρ is the difference of the mass density of the two components of the alloy. The corresponding energy is

$$\mathcal{E}(\rho) = \int_{\Omega} \left(H(\rho) + \frac{\epsilon^2}{2} |\nabla \rho|^2 \right) dx,$$

where H is either the Ginzburg–Landau double-well potential

$$(4.3) \quad H_{GL}(\rho) = \frac{1}{4} (\rho^2 - 1)^2,$$

or the logarithmic potential $H_{log}(\rho)$,

$$(4.4) \quad H_{log}(\rho) = \frac{\theta}{2} \left[(1 + \rho) \ln \left(\frac{1 + \rho}{2} \right) + (1 - \rho) \ln \left(\frac{1 - \rho}{2} \right) \right] + \frac{\theta_c}{2} (1 - \rho^2).$$

The Dirichlet energy $\int_{\Omega} |\nabla \rho|^2 dx$ is to penalize large gradients with strength ϵ^2 .

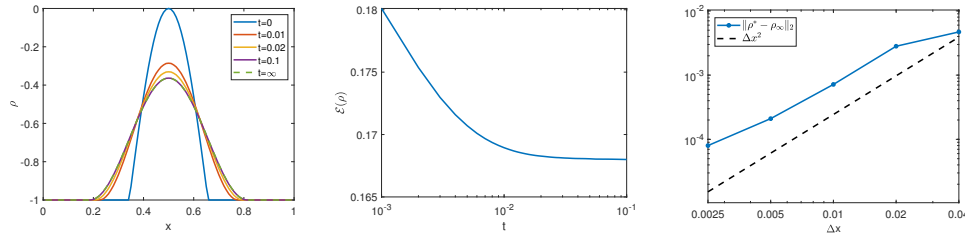


FIG. 3. Evolution for the 1D Cahn-Hilliard equation with logarithmic potential $H_{log} = (1 - \rho^2)/2$. Left: Evolution of $\rho(x, t)$. $\Delta x = 0.01$, $\tau = 0.001$. Center: Free energy decay. Right: Convergence to exact steady solution for various Δx .

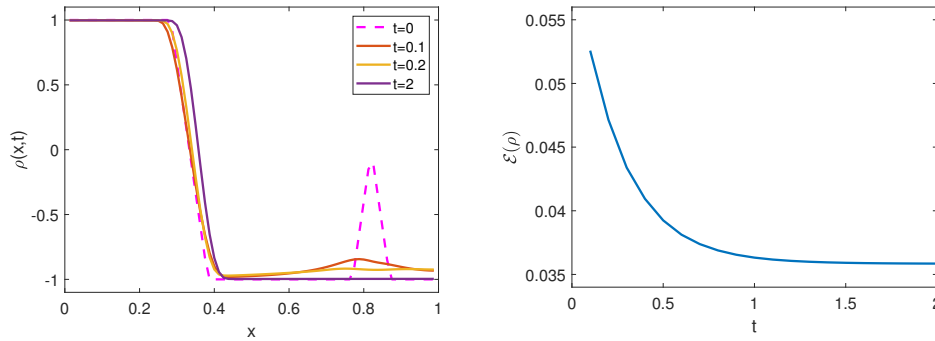


FIG. 4. Evolution for the 1D Cahn-Hilliard equation with logarithmic potential, where the parameters are $\theta = 0.3$, $\theta_c = 1$, and $\epsilon = \sqrt{10^{-3}}$. Left: Evolution of $\rho(x, t)$. $\Delta x = 0.0125$, $\tau = 0.1$. Right: Free energy decay.

The first test aims to verify the order of accuracy to our scheme (Problem 2) using the analytical form of the steady state with a carefully chosen initial condition [6]. Take the logarithmic potential H_{log} with $\theta = 0$ and $\theta_c = 1$ and the Dirichlet energy with $\epsilon = 0.1$, and set the initial condition to be

$$(4.5) \quad \rho_0(x) = \begin{cases} \cos\left(\frac{x-1/2}{\epsilon}\right) - 1 & \text{if } \left|x - \frac{1}{2}\right| \leq \frac{\pi\epsilon}{2}, \\ -1 & \text{otherwise.} \end{cases}$$

The corresponding steady state is given by

$$(4.6) \quad \rho_\infty(x) = \begin{cases} \frac{1}{\pi} \left[1 + \cos\left(\frac{x-1/2}{\epsilon}\right) \right] - 1 & \text{if } \left|x - \frac{1}{2}\right| \leq \frac{\pi\epsilon}{2}, \\ -1 & \text{otherwise.} \end{cases}$$

The simulation results and the convergence with respect to the spacial discretization Δx are shown in Figure 3. In particular, we observe a second-order convergence in space for our fully discrete scheme by plotting the error between our numerical solution $\rho^* = \rho(x, 1)$ at $t = 1$ and the analytical solution for steady state ρ_∞ in the l_2 -norm:

$$\|\rho^* - \rho_\infty\|_2 = \sqrt{\sum_i |\rho^*(x_i) - \rho_\infty(x_i)|^2 \Delta x}.$$

As a second test, we consider the logarithmic potential (4.4) with $\theta = 0.3$ and $\theta_c = 1$ and the Dirichlet energy with $\epsilon = \sqrt{10^{-3}}$. With the initial condition given by

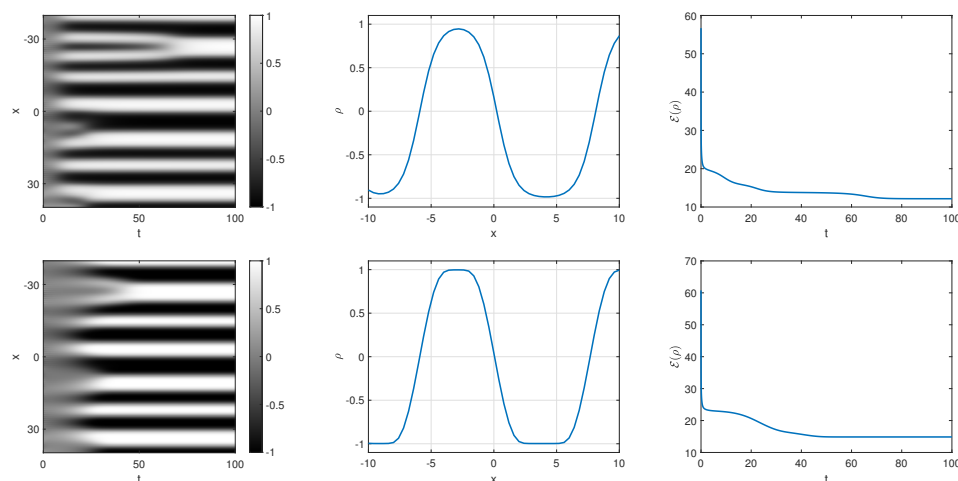


FIG. 5. Evolution for the 1D Cahn-Hilliard equation with logarithmic potential ($\theta = 0.3$, $\theta_c = 1$) (top row) and double-well potential (bottom row) and $\epsilon = 1$. Left: Evolution of $\rho(x, t)$. $\Delta x = 0.4$, $\tau = 0.01$. Center: Zoom-in phase-field solution at $t = 100$. Right: Free energy decay.

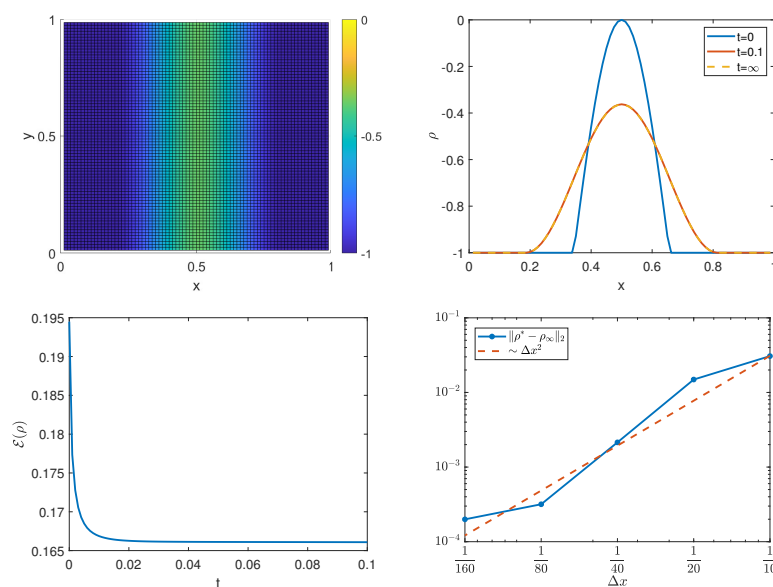


FIG. 6. Evolution for the 2D Cahn-Hilliard equation with logarithmic potential ($\theta = 0$, $\theta_c = 1$, $\epsilon = 0.1$). Top left: Evolution of $\rho(x, t)$. $\Delta x = 0.0125$, $\tau = 0.01$. Top right: Sideview of phase-field solution. Bottom left: Free energy decay. Bottom right: Convergence to exact steady solution for various Δx .

$$\rho_0(x) = \begin{cases} 1 & \text{if } 0 \leq x \leq \frac{1}{3} - \frac{1}{20}, \\ 20\left(\frac{1}{3} - x\right) & \text{if } \left|x - \frac{1}{3}\right| \leq \frac{1}{20}, \\ -20\left|x - \frac{41}{50}\right| & \text{if } \left|x - \frac{41}{50}\right| \leq \frac{1}{20}, \\ -1 & \text{otherwise,} \end{cases}$$

the evolution of the solution $\rho(x, t)$ and the free energy $\mathcal{E}(t)$ are displayed in Figure 4.

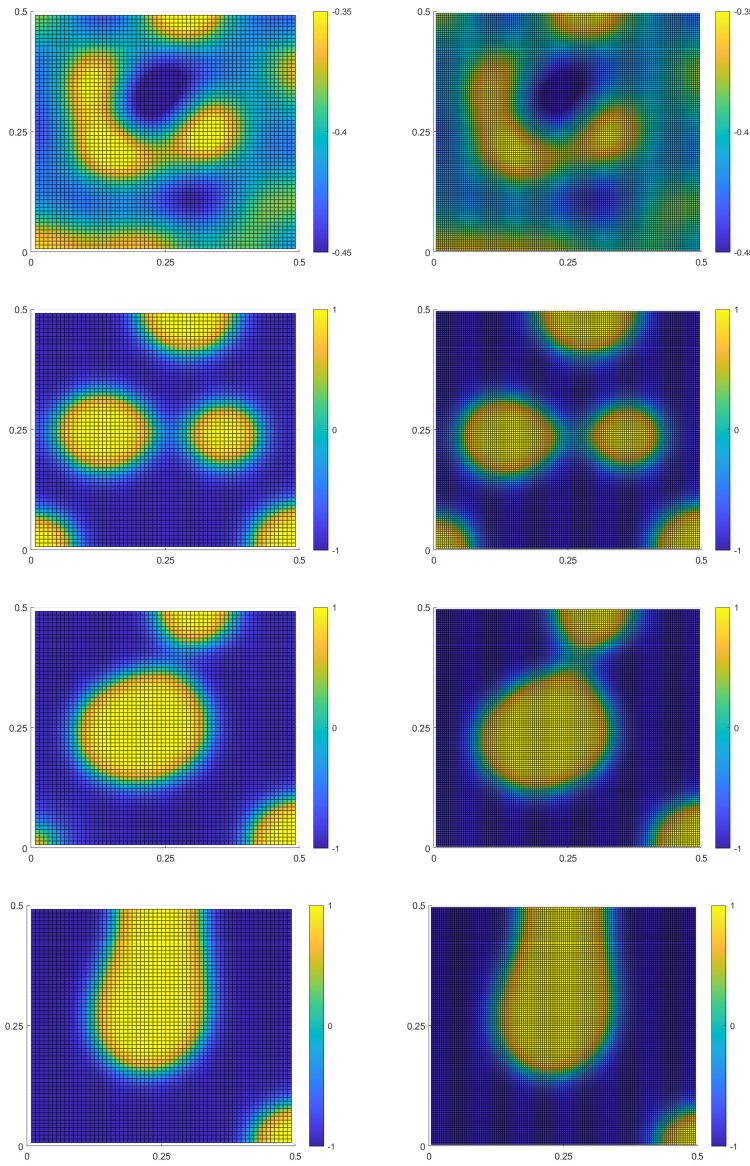


FIG. 7. Phase separation of the initially randomized phase-field for the 2D Cahn-Hilliard equation with double-well potential and $\epsilon = 0.018$. Left column: Computed by $\Delta x = \Delta y = 1/64$. Right column: Computed by $\Delta x = \Delta y = 1/128$. From top to bottom: Solution at $t = 0.01, 0.2, 0.4, 1$, where $\tau = 0.001$.

In the third test, we examine the phase separation with emerging clusters at $\rho = \pm 1$ during temporal evolution, and its dependence on the choice of potential. The initial phase-field is taken as a randomized field such the local value of $\rho(x, t = 0)$ that follows uniform distribution in $[-0.5, 0.5]$. The spatial domain is $[-40, 40]$ and is discretized uniformly with $\Delta x = 0.4$. The results, with both logarithmic potential ($\theta = 0.3$ and $\theta_c = 1$) and double-well potential and $\epsilon = 1$, are collected in Figure 5. In

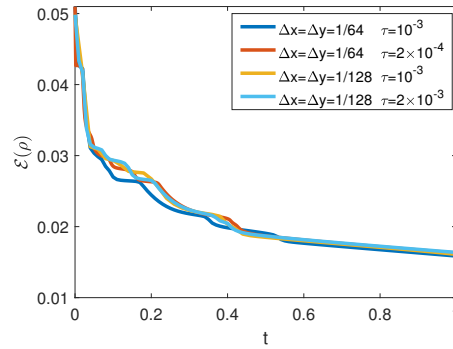


FIG. 8. Free energy decay in time for the 2D Cahn-Hilliard equation with double-well potential and $\epsilon = 0.018$. We compare the temporal evolution of free energy by simulations with different Δx , Δy , and τ .

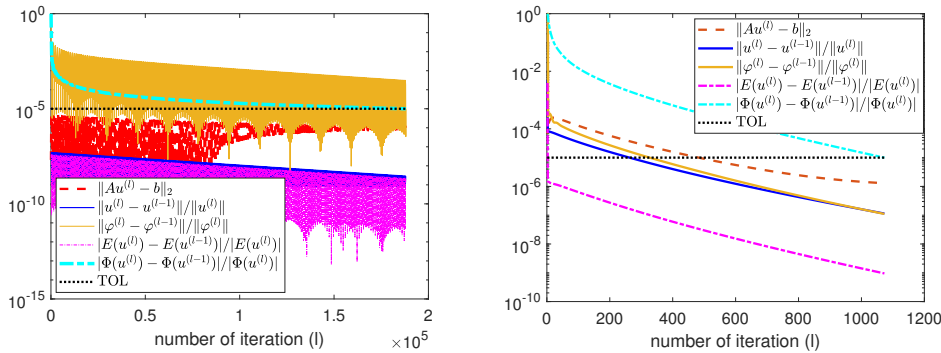


FIG. 9. Comparison of convergence rate of two algorithms PDJKO and PrePDJKO for computing one JKO step of simulations in Figure 7. We take $\Delta x = \Delta y = 1/64$, $\lambda = 0.001$ (and $\sigma = 0.99/(\lambda \lambda_{\max}(AA^T))$ for PDJKO, and $\lambda = 50$ for PrePDJKO, where the values of λ are chosen for the “optimal” convergence rate of each algorithm.

both cases, an initial phase separation is observed followed by a coarsening process with merging phases. The middle column represent the zoom-in plot of the phase-field solution at $t = 100$, where it is shown that with logarithmic potential, a plateau forms at the local maximum and minimum of ρ and is connected by a sharper transition than the double-well potential.

4.3. 2D Cahn-Hilliard equation. As with the 1D Cahn-Hilliard equation case, we first test the order of convergence by applying the logarithmic potential $H_{\log} = (1 - \rho^2)/2$ with $\theta = 0$ and $\theta_c = 1$ and the Dirichlet energy with $\epsilon = 0.1$. The initial conditions are set as in (4.5) and the corresponding steady state is given in (4.6). The evolution and convergence results are shown in Figure 6. Again, we observe a second-order convergence in space for our fully discrete scheme.

We also simulate the phase separation with the double-well potential and $\epsilon = 0.018$. The initial phase-field is taken as a randomized field such that the local value of $\rho(x, t = 0) = -0.4 + r$, where r follows uniform distribution in $[-0.1, 0.1]$. We compute the evolution of phase-field solutions in a domain of $[0, 1] \times [0, 1]$ with 64 and 128 cells. The temporal evolution of the phase-field solutions is almost identical for different mesh choices, as shown in Figure 7. The free energy evolution for different

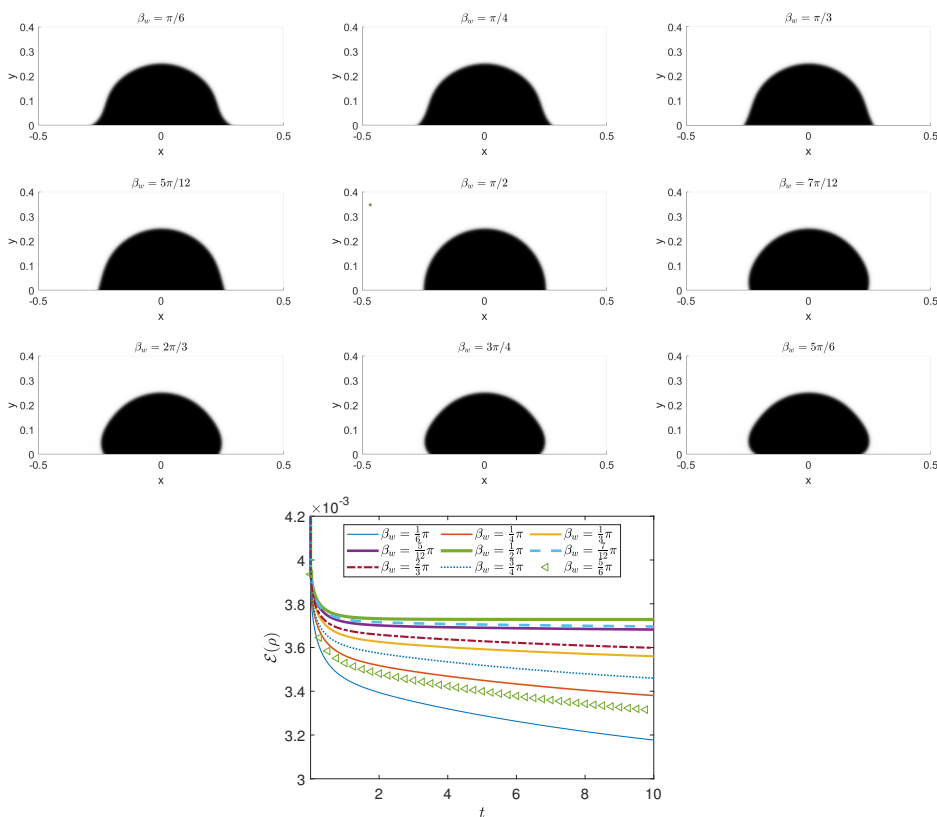


FIG. 10. Phase-fields of sessile droplets at $t = 0.1$ with different contact angles and the evolution of their total free energy: $\beta_w = \pi/6, \pi/4, \pi/3, 5\pi/12, \pi/2, 7\pi/12, 2\pi/3, 4\pi/5, 5\pi/6$. We take $\epsilon = 0.012$, $\tau = 0.01$, and 256×256 cells for simulations.

mesh choices and time steps is shown in Figure 8 to confirm that our simulations are indeed convergent to the real solutions.

For the simulation with random initial condition, the original PDJKO (Algorithm 1) converges slowly for each JKO time step at the early stage due to the randomness. Instead, we implement the PrePDJKO (Algorithm 2) for faster convergence. To compare the convergence for two primal-dual algorithms, we plot the convergence monitors versus iteration number for one time step until they achieve the same stopping criteria with $\text{TOL} = 10^{-5}$ in Figure 9. We observe a much faster convergence rate for PrePDJKO reaching the stopping criteria with around 1100 iterations, while PDJKO needs more than 180000 iterations. Note that the choice of the parameter λ is really problem-dependent. Although we do not have a specific methodology for its selection, we observe from Figure 9 that we might choose small λ and hence large σ for accelerating the convergence for the dual variable for primal-dual method; in contrast, we might choose fairly large λ for fast convergence of the primal variable for the preconditioned primal-dual method since we have done preconditioning on the proximal operator for updating the dual variable.

4.4. Wetting phenomenon of droplets. Now we consider the 2D Cahn–Hilliard equation with double-well potential (4.3) and the wall free energy (2.4). We

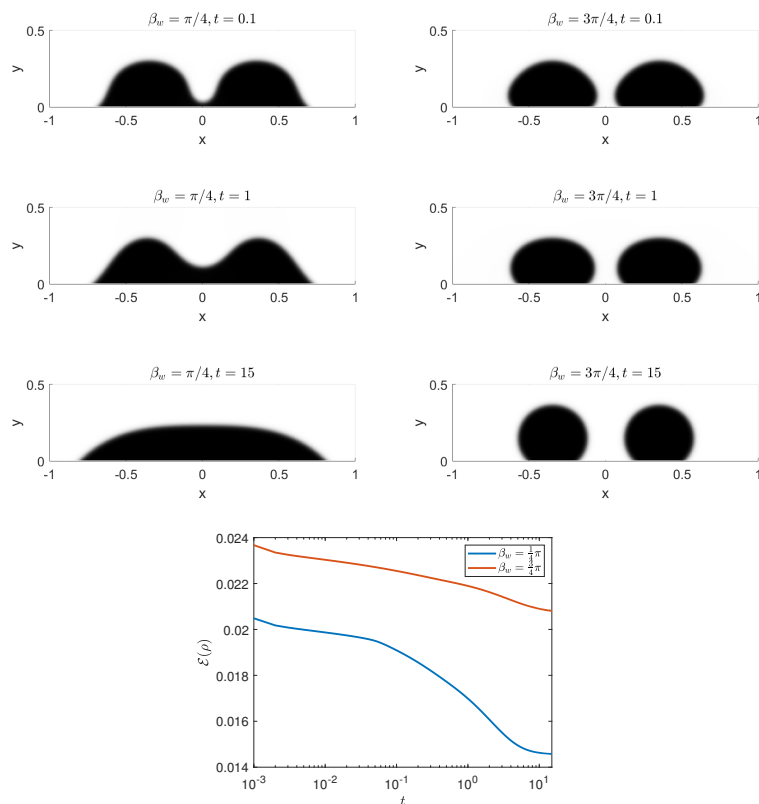


FIG. 11. Temporal evolution of two droplets with contact angles $\beta_w = \pi/4$ (left) and $\beta_w = 3\pi/4$ (right) and their free energy. We take $\epsilon = 0.005$, $\tau = 0.005$, and 256×64 cells for simulations.

simulate the equilibrium phase-fields of sessile droplets on a flat substrate with different contact angles: $\beta_w = \pi/6, \pi/4, \pi/3, 5\pi/12, \pi/2, 7\pi/12, 2\pi/3, 4\pi/5, 5\pi/6$. The equilibrium phase-fields at $t = 0.1$ and the evolution of their energy are shown in Figure 10. The smoothed initial phase is given by a sharp-interface phase convolution with a mollifier

$$\rho_0(x, y) = \tilde{\rho}_0(x, y) * W(x, y) - 1,$$

$$\tilde{\rho}_0(x, y) = \begin{cases} 2 & \text{if } x^2 + y^2 < 0.25^2, \\ 0 & \text{if } x^2 + y^2 > 0.25^2, \end{cases} \quad W(x, y) = \frac{1}{4\pi\epsilon^2} e^{-\frac{x^2+y^2}{4\epsilon^2}}.$$

Then we simulate the dynamics of two droplets with two choices of contact angle: $\beta_w = \pi/4, 3\pi/4$. The temporal evolution of the droplets and their energies are shown in Figure 11. We observe that the two droplets merge and form a single phase on a hydrophilic substrate ($\beta_w = \pi/4$), while they remain separated with some distance on a hydrophobic substrate ($\beta_w = 3\pi/4$). Consequently, we observe two stages of energy decay during the evolution for $\beta_w = \pi/4$, where the first mild decay corresponds to two droplets adjusting the contact angle and the second dramatic decay corresponds to the coalescing of the two droplets. The initial phase is given by the convolution with the mollifier

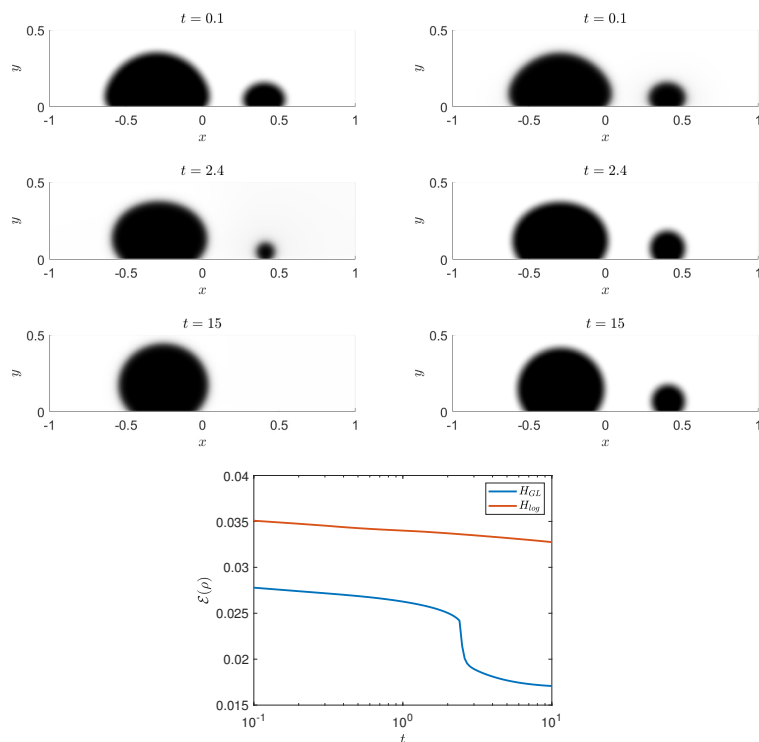


FIG. 12. Temporal evolution of two droplets in different sizes by the Ginzburg–Landau double-well potential H_{GL} (left) and logarithmic potential H_{log} (right) and their free energy (bottom). We take $\epsilon = 0.02$, $\beta_w = 3\pi/4$, $\tau = 0.1$, and 256×96 cells.

$$\rho_0(x, y) = \tilde{\rho}_0(x, y) * W(x, y) - 1,$$

$$\tilde{\rho}_0(x, y) = \begin{cases} 2 & \text{if } (x + 0.35)^2 + y^2 < 0.3^2 \text{ or } (x - 0.35)^2 + y^2 < 0.3^2, \\ 0 & \text{elsewhere.} \end{cases}$$

In the end, we consider the wetting phenomenon of two droplets with different sizes. We investigate the different dynamics of the droplets induced by the Ginzburg–Landau double-well potential H_{GL} (4.3) and the logarithmic potential H_{log} (4.4) ($\theta = 0.3, \theta_c = 1$) with the nonlinear degenerate mobility $M(\rho) = (1 - \rho^2)$. It was shown formally that the Cahn–Hilliard equation with the pair of H_{log} and $M(\rho)$ converges to the sharp limit motion of surface diffusion flow [18], while the pair of H_{GL} and $M(\rho)$ leads to the motion driven by both surface diffusion and additional bulk diffusion [28, 24, 11]. The simulation results in Figure 12 show that the small droplet is gradually absorbed by the large droplet due to the additional bulk diffusion induced by H_{GL} , and a dramatic energy decay occurs corresponding to the disappearance of the small droplet. However, the two droplets remain distant and the small droplet does not disappear driven by only surface diffusion with the pair of H_{log} and $M(\rho)$.

Acknowledgment. The authors would like to thank the anonymous reviewers for their useful observations and suggestions, which greatly improved this work.

REFERENCES

- [1] L. AMBROSIO, N. GIGLI, AND G. SAVARÉ, *Gradient Flows in Metric Spaces and in the Space of Probability Measures*, Birkhäuser Basel, 2005, <https://doi.org/10.1007/978-3-7643-8722-8>.
- [2] B. AYMARD, U. VAES, M. PRADAS, AND S. KALLIADASIS, *A linear, second-order, energy stable, fully adaptive finite element method for phase-field modelling of wetting phenomena*, J. Comput. Phys. X, 2 (2019), 100010, <https://doi.org/10.1016/j.jcp.x.2019.100010>.
- [3] R. BAILO, J. A. CARRILLO, AND J. HU, *Fully discrete positivity-preserving and energy-dissipating schemes for aggregation-diffusion equations with a gradient-flow structure*, Commun. Math. Sci., 18 (2020), pp. 1259–1303, <https://doi.org/10.4310/cms.2020.v18.n5.a5>.
- [4] R. BAILO, J. A. CARRILLO, AND J. HU, *Bound-preserving finite-volume schemes for systems of continuity equations with saturation*, SIAM J. Appl. Math., 83 (2023), pp. 1315–1339, <https://doi.org/10.1137/22M1488703>.
- [5] R. BAILO, J. A. CARRILLO, S. KALLIADASIS, AND S. P. PEREZ, *Unconditional Bound-Preserving and Energy-Dissipating Finite-Volume Schemes for the Cahn-Hilliard Equation*, preprint, arXiv:2105.05351, 2021.
- [6] J. W. BARRETT, J. F. BLOWEY, AND H. GARCKE, *Finite element approximation of the Cahn-Hilliard equation with degenerate mobility*, SIAM J. Numer. Anal., 37 (1999), pp. 286–318, <https://doi.org/10.1137/S0036142997331669>.
- [7] J.-D. BENAMOU AND Y. BRENIER, *A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem*, Numer. Math., 84 (2000), pp. 375–393.
- [8] E. BERETTA, M. BERTSCH, AND R. DAL PASSO, *Nonnegative solutions of a fourth-order nonlinear degenerate parabolic equation*, Arch. Ration. Mech. Anal., 129 (1995), pp. 175–200.
- [9] F. BERNIS AND A. FRIEDMAN, *Higher order nonlinear degenerate parabolic equations*, J. Differential Equations, 83 (1990), pp. 179–206.
- [10] A. L. BERTOZZI, *The mathematics of moving contact lines in thin liquid films*, Notices Amer. Math. Soc., 45 (1998), pp. 689–697.
- [11] E. BRETIN, S. MASNOU, A. SENGERS, AND G. TERII, *Approximation of surface diffusion flow: A second-order variational Cahn-Hilliard model with degenerate mobilities*, Math. Models Methods Appl. Sci., 32 (2022), pp. 1–37, <https://doi.org/10.1142/S0218202522500178>.
- [12] M. BURGER, M. DI FRANCESCO, AND Y. DOLAK-STRUSS, *The Keller-Segel model for chemotaxis with prevention of overcrowding: Linear vs. nonlinear diffusion*, SIAM J. Math. Anal., 38 (2006), pp. 1288–1315, <https://doi.org/10.1137/050637923>.
- [13] J. W. CAHN, *On spinodal decomposition*, Acta Metall., 9 (1961), pp. 795–801.
- [14] J. A. CARRILLO, A. CHERTOCK, AND Y. HUANG, *A finite-volume method for nonlinear nonlocal equations with a gradient flow structure*, Commun. Comput. Phys., 17 (2015), pp. 233–258.
- [15] J. A. CARRILLO, K. CRAIG, L. WANG, AND C. WEI, *Primal dual methods for Wasserstein gradient flows*, Found. Comput. Math., 22 (2022), pp. 389–443.
- [16] J. A. CARRILLO, S. LISINI, G. SAVARÉ, AND D. SLEPČEV, *Nonlinear mobility continuity equations and generalized displacement convexity*, J. Funct. Anal., 258 (2010), pp. 1273–1309.
- [17] J. DOLBEAULT, B. NAZARET, AND G. SAVARÉ, *A new class of transport distances between measures*, Calc. Var. Partial Differential Equations, 34 (2009), pp. 193–231.
- [18] C. M. ELLIOTT AND H. GARCKE, *On the Cahn-Hilliard equation with degenerate mobility*, SIAM J. Math. Anal., 27 (1996), pp. 404–423, <https://doi.org/10.1137/S0036141094267662>.
- [19] C. M. ELLIOTT AND A. STUART, *The global dynamics of discrete semilinear parabolic equations*, SIAM J. Numer. Anal., 30 (1993), pp. 1622–1663, <https://doi.org/10.1137/0730084>.
- [20] D. J. EYRE, *Unconditionally gradient stable time marching the Cahn-Hilliard equation*, in Computational and Mathematical Models of Microstructural Evolution (San Francisco, CA, 1998), Mater. Res. Soc. Sympos. Proc. 529, Materials Research Society, Warrendale, PA, 1998.
- [21] C. FALCÓ, R. E. BAKER, AND J. A. CARRILLO, *A local continuum model of cell-cell adhesion*, SIAM J. Appl. Math., (2023), pp. S17–S42, <https://doi.org/10.1137/22M1506079>.
- [22] G. FU, S. OSHER, AND W. LI, *High order spatial discretization for variational time implicit schemes: Wasserstein gradient flows and reaction-diffusion systems*, J. Comput. Phys., 491 (2023), 112375, <https://doi.org/10.1016/j.jcp.2023.112375>.
- [23] F. HUANG, J. SHEN, AND K. WU, *Bound/positivity preserving and unconditionally stable schemes for a class of fourth order nonlinear equations*, J. Comput. Phys., 460 (2022), 111177, <https://doi.org/10.1016/j.jcp.2022.111177>.
- [24] Q.-A. HUANG, W. JIANG, AND J. YANG, *An efficient and unconditionally energy stable scheme for simulating solid-state dewetting of thin films with isotropic surface energy*, Commun. Comput. Phys., 26 (2019), pp. 1444–1470, <https://doi.org/10.4208/cicp.2019.js60.07>.

- [25] Q.-A. HUANG, W. JIANG, J. Z. YANG, AND C. YUAN, *A Structure-Preserving, Upwind-SAV Scheme for the Degenerate Cahn–Hilliard Equation with Applications to Simulating Surface Diffusion*, preprint, arXiv:2210.16017, 2023.
- [26] M. JACOBS, F. LÉGER, W. LI, AND S. OSHER, *Solving large-scale optimization problems with a convergence rate independent of grid size*, SIAM J. Numer. Anal., 57 (2019), pp. 1100–1123, <https://doi.org/10.1137/18M118640X>.
- [27] R. JORDAN, D. KINDERLEHRER, AND F. OTTO, *The variational formulation of the Fokker–Planck equation*, SIAM J. Math. Anal., 29 (1998), pp. 1–17, <https://doi.org/10.1137/S0036141096303359>.
- [28] A. LEE, A. MÜNCH, AND E. SÜLI, *Degenerate mobilities in phase field models are insufficient to capture surface diffusion*, Appl. Phys. Lett., 107 (2015), 081603, <https://doi.org/10.1063/1.4929696>.
- [29] H. G. LEE AND J. KIM, *Accurate contact angle boundary conditions for the Cahn–Hilliard equations*, Comput. & Fluids, 44 (2011), pp. 178–186, <https://doi.org/10.1016/j.compfluid.2010.12.031>.
- [30] W. LI, J. LU, AND L. WANG, *Fisher information regularization schemes for Wasserstein gradient flows*, J. Comput. Phys., 416 (2020), 109449.
- [31] S. LISINI, D. MATTHES, AND G. SAVARÉ, *Cahn–Hilliard and thin film equations with nonlinear mobility as gradient flows in weighted-Wasserstein metrics*, J. Differential Equations, 253 (2012), pp. 814–850.
- [32] Y. LIU, Y. XU, AND W. YIN, *Acceleration of primal–dual methods by preconditioning and simple subproblem procedures*, J. Sci. Comput., 86 (2021), 21.
- [33] J. SHEN, J. XU, AND J. YANG, *A new class of efficient and robust energy stable schemes for gradient flows*, SIAM Rev., 61 (2019), pp. 474–506, <https://doi.org/10.1137/17M1150153>.
- [34] L. WANG AND M. YAN, *Hessian informed mirror descent*, J. Sci. Comput., 92 (2022), 90.
- [35] M. YAN, *A new primal–dual algorithm for minimizing the sum of three functions with a linear operator*, J. Sci. Comput., 76 (2018), pp. 1698–1717.
- [36] L. ZHORNITSKAYA AND A. L. BERTOZZI, *Positivity-preserving numerical schemes for lubrication-type equations*, SIAM J. Numer. Anal., 37 (2000), pp. 523–555, <https://doi.org/10.1137/S0036142998335698>.