Communications on Applied Mathematics and Computation
https://doi.org/10.1007/s42967-025-00486-2

ORIGINAL PAPER

Structure Preserving Schemes for a Class of Wasserstein Gradient Flows

Shiheng Zhang[1] · **Jie Shen**[2]

Received: 27 March 2024 / Revised: 28 August 2024 / Accepted: 14 October 2024

Abstract
We introduce in this paper two time discretization schemes tailored for a range of Wasserstein gradient flows. These schemes are designed to preserve mass and positivity and to be uniquely solvable. In addition, they also ensure energy dissipation in many typical scenarios. Through extensive numerical experiments, we demonstrate the schemes' robustness, accuracy, and efficiency.

Keywords Wasserstein gradient flow · Positivity preserving · Energy stability · Porous media equation (PME)

Mathematics Subject Classification 65M12 · 35K61 · 35K55 · 65Z05

1 Introduction

Gradient flows play important roles in mathematical models across various scientific and engineering disciplines, especially in materials science and fluid dynamics [1, 2, 8, 12, 15, 21, 30, 46]. A general form of gradient flows can be written as

$$\frac{\partial \rho}{\partial t} = -\mathcal{G}\frac{\delta E}{\delta \rho}, \tag{1}$$

where E is a specific energy functional, and $\mathcal{G}$ is a positive, possibly nonlinear, operator. Typical forms of $\mathcal{G}$ include $\mathcal{G} = I$ [1] or $\mathcal{G} = -\nabla \cdot M\nabla$ where M is a positive mobility function [8]. There are also instances where more intricate metrics are desired. A notable example is the gradient flows over spaces characterized by the Wasserstein metric, leading to the concept of Wasserstein gradient flows, where $\mathcal{G}$ embodies a nonlinear operator such that

✉ Jie Shen
jshen@eitech.edu.cn

Shiheng Zhang
shzhang3@uw.edu

[1] Department of Applied Mathematics, University of Washington, Seattle, WA 98195, USA

[2] School of Mathematical Science, Eastern Institute of Technology, Ningbo 315200, Zhejiang, China

Published online: 08 April 2025

$$\frac{\partial \rho}{\partial t} = \nabla \cdot \left(\rho \nabla \frac{\delta E}{\delta \rho} \right). \tag{2}$$

This framework encapsulates a broad class of nonlinear parabolic equations, including the porous medium equation (PME) [42, 43], the Keller-Segel equation [27, 28, 32], the Fokker-Planck equation [6, 41], and the Poisson-Nernst-Planck (PNP) equation [4, 5], which can all be classified as Wasserstein gradient flows. The surge in interest for Wasserstein gradient flows in recent years is apparent in fields ranging from optimal transport to fluid dynamics and statistical mechanics [11, 33, 44]. Essential properties of the Wasserstein gradient flows include mass conservation, positivity preserving, and energy dissipation.

Designing numerical schemes that can efficiently and accurately capture the essential properties of the Wasserstein gradient flows is still a challenge. Since the Wasserstein gradient flows are a special class of gradient flows, we can consider established techniques for gradient flows. However, energy stable schemes for gradient flows, such as convex splitting [16, 17, 35], IEQ [45, 47], and the scalar auxiliary variable (SAV) [36, 39, 40]), do not inherently guarantee positivity preservation. The celebrated JKO scheme [26] is a widely used approach to solve variational problems associated with Wasserstein gradient flows. This method's variational structure ensures unconditional energy stability and positivity preserving. However, it requires solving a nonlinear minimization problem in the Wasserstein metric. Significant efforts have been made towards efficient computation of the JKO scheme, such as [9, 10, 18, 19, 24, 25, 29, 34]. Despite extensive research on the JKO scheme, efficiently solving it remains challenging due to the inherent complexity of the Wasserstein metric. Some other approaches have been developed recently that can preserve positivity and energy dissipation for Wasserstein gradient flows, but they either are not uniquely solvable [3, 23], or are limited to certain Wasserstein gradient flows [37, 38]. We also refer to more recent work [13, 14] for other efforts in developing structure preserving schemes for Wasserstein gradient flows.

We propose in this paper two approaches to construct time discretization schemes for Wasserstein gradient flows.

(i) In the first approach, we set $E(\rho) = \int_\Omega (H(\rho) + \rho v)\, \mathrm{d}\mathbf{x}$, where $H''(\rho) > 0$ when $\rho > 0$ is assumed, and v is a given potential function. We construct first-order and second-order schemes that can preserve mass and positivity and are uniquely solvable. Furthermore, the first-order scheme is also energy dissipative in some typical scenarios.

(ii) In the second approach, $E(\rho)$ can take a more general form, but can be split into two parts: one part is convex, and the other is bounded from below. This assumption is significantly less restrictive than the assumption required by the convex splitting method since only one part is required to be convex. We construct first-order and second-order schemes by introducing an SAV [40], and show that they preserve mass and positivity, are uniquely solvable, and energy dissipative with respect to a modified energy.

These schemes are nonlinear in nature, but their solutions can all be interpreted as minimizers of strictly convex functionals.

The paper is structured as follows: we introduce in Sect. 2 the two approaches in the semi-discrete form, and establish their essential properties. We present in Sect. 3 a spatial discretization which can inherent essential properties in the space continuous case using the finite-difference method as an example. We carry out a series of numerical experiments in Sect. 4 to validate our schemes' accuracy and efficiency. Finally, we provide some concluding remarks in Sect. 5.

2 Time Discretization

In this section, we present two distinct time discretization approaches, each tailored to specific scenarios of the Wasserstein gradient flows (2) with a smooth functional $E\colon \Omega \times \mathbb{R} \to \mathbb{R}$. We shall assume throughout this paper that the boundary condition for ρ is either periodic or homogeneous Neumann, i.e., $\frac{\partial \rho}{\partial \boldsymbol{n}} = 0$, and use (a, b) to denote the integral $\int_\Omega ab \, \mathrm{d}\boldsymbol{x}$.

2.1 The First Approach (S1)

The Wasserstein gradient flow is as follows:

$$\frac{\partial \rho}{\partial t} = \nabla \cdot \left(\rho \nabla \frac{\delta E}{\delta \rho} \right), \quad \boldsymbol{x} \in \Omega \subset \mathbb{R}^d, t > 0. \tag{3}$$

Under the assumption that $E(\rho) = \int_\Omega (H(\rho) + \rho v) \, \mathrm{d}\boldsymbol{x}$ with $H''(\rho) > 0$ when $\rho > 0$, and using the identity $\nabla \rho = \rho \nabla \log \rho$, (3) can be reformulated as

$$\frac{\partial \rho}{\partial t} = \nabla \cdot \left(\rho \left(H''(\rho) \rho \nabla \log \rho + \nabla v \right) \right). \tag{4}$$

The assumption $H''(\rho) > 0$ when $\rho > 0$ is satisfied in various contexts, including the PME and the Fokker-Planck equation.

A first-order time-discretized scheme can be written as follows:

$$\frac{\rho^{n+1} - \rho^n}{\delta t} = \nabla \cdot \left(\rho^n \left(H'' \left(\rho^n \right) \rho^n \nabla \log \rho^{n+1} + \nabla v \right) \right). \tag{S1}$$

This scheme is a generalization of the schemes introduced in Refs. [20, 38]. It is a nonlinear scheme, but can guarantee mass conservation law, positivity, also the energy dissipation law if $H''(\rho)\rho = 1$ or $\nabla v = 0$, as stated below.

Theorem 1 *Assuming $H''(\rho) > 0$ when $\rho > 0$ and $\rho^n > 0$, the scheme (S1) exhibits the following attributes.*

(i) *Mass conservation*: $\int_\Omega \rho^n \, \mathrm{d}\boldsymbol{x} = \int_\Omega \rho^{n+1} \, \mathrm{d}\boldsymbol{x}$.
(ii) *Positivity preserving*: $\rho^{n+1} > 0$.
(iii) *Unique solvability.*
(iv) *Energy dissipation law*: *if $H(\rho) = \rho(\log \rho - 1) + c\rho$, we have*

$$\begin{aligned} &\int_\Omega \rho^{n+1} (\log \rho^{n+1} - 1 + v) \, \mathrm{d}\boldsymbol{x} - \int_\Omega \rho^n (\log \rho^n - 1 + v) \, \mathrm{d}\boldsymbol{x} \\ \leqslant &-\delta t \int_\Omega \rho^n \left| \nabla (\log \rho^{n+1} + v) \right|^2 \mathrm{d}\boldsymbol{x}; \end{aligned} \tag{5}$$

or if $\nabla v = 0$, we have

$$\begin{aligned} &\int_\Omega \rho^{n+1} (\log \rho^{n+1} - 1) \, \mathrm{d}\boldsymbol{x} - \int_\Omega \rho^n (\log \rho^n - 1) \, \mathrm{d}\boldsymbol{x} \\ \leqslant &-\delta t \int_\Omega \left(\rho^n \right)^2 H''(\rho^n) \left| \nabla \log \rho^{n+1} \right|^2 \mathrm{d}\boldsymbol{x}. \end{aligned} \tag{6}$$

Proof The mass conservation is obtained by integrating (S1) over the domain Ω and utilizing the Neumann boundary conditions or periodic boundary conditions imposed to ρ^{n+1}.

The preservation of positivity for ρ^{n+1} is ensured through the inclusion of $\log \rho^{n+1}$ in our formulation.

For the unique solvability, we define $\mathcal{L}_n$ as a linear operator by $\mathcal{L}_n g = u$ where u is the solution of the following elliptic equation:

$$-\nabla \cdot \left(\left(\rho^n\right)^2 H''(\rho^n) \nabla u \right) = g, \quad \int_\Omega u \, \mathrm{d}\boldsymbol{x} = 0$$

with Neumann boundary conditions or periodic boundary conditions. Hence, $\mathcal{L}_n$ is self-adjoint and semi-positive definite. We then consider a functional

$$\begin{aligned} F[\rho^{n+1}] := & \int_\Omega \rho^{n+1} (\log \rho^{n+1} - 1) \, \mathrm{d}\boldsymbol{x} + \frac{1}{2\delta t} \int_\Omega \left(\rho^{n+1} - \rho^n\right) \mathcal{L}_n \left(\rho^{n+1} - \rho^n\right) \mathrm{d}\boldsymbol{x} \\ & + \int_\Omega \rho^{n+1} \mathcal{L}_n \nabla \cdot \left(\rho^n \nabla v\right) \mathrm{d}\boldsymbol{x}. \end{aligned}$$

Notably, F stands as a strictly convex functional within the set

$$\mathcal{A} := \left\{ \rho^{n+1} \in H^2(\Omega) : \rho^{n+1} > 0 \text{ in } \Omega \right\}.$$

Remarkably, the Euler-Lagrange equation of F under the constraint of mass conservation is

$$\begin{cases} \log \rho^{n+1} + \dfrac{1}{\delta t} \mathcal{L}_n \left(\rho^{n+1} - \rho^n\right) + \mathcal{L}_n \nabla \cdot \left(\rho^n \nabla v\right) = \lambda, \\ \displaystyle\int_\Omega (\rho^{n+1} - \rho^n) \, \mathrm{d}\boldsymbol{x} = 0, \end{cases} \tag{7}$$

where λ is the Lagrange multiplier of the mass conservation. It is equivalent to the equation given in the scheme (S1) with the definition of $\mathcal{L}_n$. Consequently, the unique minimizer of F serves as the unique solution to the scheme (S1).

If $H(\rho) = \rho(\log \rho - 1) + c\rho$, we have $H''(\rho)\rho = 1$. Taking the inner product of (S1) with $\log \rho^{n+1} + v$ and employing integration by parts, we obtain

$$\int_\Omega \left(\rho^{n+1} - \rho^n\right) \left(\log \rho^{n+1} + v\right) \mathrm{d}\boldsymbol{x} = -\delta t \int_\Omega \rho^n \left| \nabla (\log \rho^{n+1} + v) \right|^2 \mathrm{d}\boldsymbol{x}.$$

To facilitate further simplification, we utilize the following inequality which can be derived by the Taylor expansion:

$$(a - b) \log a = (a \log a - a) - (b \log b - b) + \frac{(a-b)^2}{2\xi}, \quad \xi \in [\min\{a, b\}, \max\{a, b\}]. \tag{8}$$

Subsequently, we can infer

$$(\rho^{n+1} \log \rho^{n+1} - \rho^{n+1}) - (\rho^n \log \rho^n - \rho^n) \leqslant \left(\rho^{n+1} - \rho^n\right) \log \rho^{n+1}.$$

Employing this inequality directly leads us to

$$\begin{aligned} & \int_\Omega \left(\rho^{n+1} \left(\log \rho^{n+1} + v - 1\right) - \rho^n \left(\log \rho^n + v - 1\right)\right) \mathrm{d}\boldsymbol{x} \\ \leqslant & \int_\Omega \left(\rho^{n+1} - \rho^n\right) \left(\log \rho^{n+1} + v\right) \mathrm{d}\boldsymbol{x} \\ = & -\delta t \int_\Omega \rho^n \left| \nabla (\log \rho^{n+1} + v) \right|^2 \mathrm{d}\boldsymbol{x} \\ \leqslant & 0, \end{aligned} \tag{9}$$

which is the desired energy dissipation law (5).

Similarly, if $\nabla v = 0$, the energy dissipation law (6) can be obtained by taking the inner product of (S1) with $\log \rho^{n+1}$ and integrating by parts.

We can construct a second-order scheme by combining the second-order BDF with the Adams-Bashforth extrapolation as follows:

$$\frac{3\rho^{n+1} - 4\rho^n + \rho^{n-1}}{2\delta t} = \nabla \cdot (\phi^{*,n+\frac{1}{2}} \nabla \log \rho^{n+1}) + \nabla \cdot (\rho^{*,n+\frac{1}{2}} \nabla v), \tag{10}$$

where $\phi = \rho^2 H''(\rho)$, and for any function ψ,

$$\psi^{*,n+\frac{1}{2}} = \begin{cases} 2\psi^n - \psi^{n-1} & \text{if } \psi^n \geqslant \psi^{n-1}, \\ \frac{1}{2/\psi^n - 1/\psi^{n-1}} & \text{if } \psi^n < \psi^{n-1}, \end{cases} \tag{11}$$

which is a modified Adams-Bashforth extrapolation to preserve the positivity, i.e., $\psi^{*,n+\frac{1}{2}} > 0$ if ψ^n, $\psi^{n-1} > 0$. To find ρ^1, we can use the first-order method (S1).

Theorem 2 *Assume that ρ^1 is obtained from the first-order scheme. The second-order scheme (10) exhibits the following attributes.*

(i) *Mass conservation*: $\int_\Omega \rho^n \, \mathrm{d}\boldsymbol{x} = \int_\Omega \rho^{n+1} \, \mathrm{d}\boldsymbol{x}$.
(ii) *Positivity preserving*: $\rho^{n+1} > 0$.
(iii) *Unique solvability.*

Proof The proof is similar to that of Theorem 1. We only need to modify slightly the definition of the linear operator $\tilde{\mathcal{L}}_n$ as follows: $\tilde{\mathcal{L}}_n g = u$ is defined by the elliptic equation

$$-\nabla \cdot \left(\phi^{*,n+\frac{1}{2}} \nabla u\right) = g, \quad \int_\Omega u \, \mathrm{d}\boldsymbol{x} = 0$$

with the homogeneous Neumann boundary conditions or periodic boundary conditions. We can also define a slightly different convex functional $\tilde{F}$ as follows:

$$\begin{aligned} \tilde{F}[\rho^{n+1}] := & \int_\Omega \rho^{n+1} (\log \rho^{n+1} - 1) \, \mathrm{d}\boldsymbol{x} + \int_\Omega \rho^{n+1} \tilde{\mathcal{L}}_n \nabla \cdot \left(\rho^n \nabla v\right) \, \mathrm{d}\boldsymbol{x} \\ & + \frac{1}{12\delta t} \int_\Omega \left(3\rho^{n+1} - 4\rho^n + \rho^{n-1}\right) \tilde{\mathcal{L}}_n \left(3\rho^{n+1} - 4\rho^n + \rho^{n-1}\right) \, \mathrm{d}\boldsymbol{x}. \end{aligned}$$

Then, it can be shown that the solution of (10) is the unique minimizer of the above convex functional.

Unfortunately, we are unable to show that the above second-order scheme is energy dissipative as in the first-order scheme.

2.2 The Second Approach (S2)

While the scheme (S1) retains many desired properties, its applicability is restricted to limited scenarios. To address the limitation, we propose the scheme (S2). This scheme maintains all the desirable properties and accommodates more general scenarios by splitting the energy and introducing an SAV.

Again, we consider the general Wasserstein gradient flow

$$\frac{\partial \rho}{\partial t} = \nabla \cdot \left(\rho \nabla \frac{\delta E}{\delta \rho}\right) \tag{12}$$

with a smooth functional $E: \Omega \times \mathbb{R} \to \mathbb{R}$ and periodic boundary conditions or Neumann boundary conditions.

Consider the functional E, which can be decomposed into two parts: $E = E_1 + E_2$. Specifically,

$$E_1 = E - \int_\Omega \rho(\log \rho - 1)\,\mathrm{d}\boldsymbol{x} \text{ and } E_2 = \int_\Omega \rho(\log \rho - 1)\,\mathrm{d}\boldsymbol{x}.$$

To employ the SAV method, we assume that E_1 is bounded below, and there exists a positive constant C such that $E_1 + C > 0$. Importantly, this assumption is mild in the context of the typical Wasserstein gradient flow. As an illustrative example, we consider $E(\rho) = \int_\Omega \rho^2 \mathrm{d}x + 1$, which can also be expressed as

$$E(\rho) = \int_\Omega (\rho^2 - \rho(\log \rho - 1))\,\mathrm{d}\boldsymbol{x} + 1 + \int_\Omega \rho(\log \rho - 1)\,\mathrm{d}\boldsymbol{x}.$$

With $E_1 = \int_\Omega (\rho^2 - \rho(\log \rho - 1))\,\mathrm{d}\boldsymbol{x} + 1$, it is evident from the condition $\rho^2 - \rho(\log \rho - 1) > 0$ for $\rho > 0$ that $E_1 > 1$.

Building on this foundation, we can introduce a scalar variable $r = \sqrt{E_1 + C}$, then (12) can be expressed as follows:

$$\frac{\partial \rho}{\partial t} = \nabla \cdot \left(\rho \nabla \left(\frac{r}{\sqrt{E_1 + C}} \frac{\delta E_1}{\delta \rho} + \frac{\delta E_2}{\delta \rho} \right) \right). \tag{13}$$

Noting that $\frac{\delta E_2}{\delta \rho} = \log \rho$, we can deal with r and $\frac{\delta E_2}{\delta \rho}$ implicitly, and obtain

$$\frac{\rho^{n+1} - \rho^n}{\delta t} = \nabla \cdot \left(\rho^n \nabla \left(\frac{r^{n+1}}{\sqrt{E_1^n + C}} \frac{\delta E_1^n}{\delta \rho} + \log \rho^{n+1} \right) \right), \tag{S2-1}$$

$$\frac{r^{n+1} - r^n}{\delta t} = \frac{1}{2\sqrt{E_1^n + C}} \left(\frac{\delta E_1^n}{\delta \rho}, \frac{\rho^{n+1} - \rho^n}{\delta t} \right). \tag{S2-2}$$

Theorem 3 *The scheme (S2) exhibits the following attributes.*

(i) *Mass conservation*: $\int_\Omega \rho^n \,\mathrm{d}\boldsymbol{x} = \int_\Omega \rho^{n+1} \,\mathrm{d}\boldsymbol{x}$.
(ii) *Positivity preserving*: $\rho^{n+1} > 0$ *if* $\rho^n > 0$.
(iii) *Unique solvability.*
(iv) *Energy dissipation law*:

$$\tilde{E}^{n+1} - \tilde{E}^n \leqslant -\delta t \int_\Omega \rho^n \left| \nabla \left(\frac{r^{n+1}}{\sqrt{E_1^n + C}} \frac{\delta E_1^n}{\delta \rho} + \log \rho^{n+1} \right) \right|^2 \mathrm{d}\boldsymbol{x}, \tag{14}$$

where the discrete energy $\tilde{E}^n = \int_\Omega \rho^n \,(\log \rho^n - 1)\,\mathrm{d}\boldsymbol{x} + (r^n)^2$.

Proof Integrating (S2-1) over the domain Ω and applying Neumann or periodic boundary conditions to both $\frac{\delta E_1}{\delta \rho}$ and $\log \rho$ enables us to establish the mass conservation law.

Furthermore, the inclusion of $\log \rho^{n+1}$ in our formulation guarantees the preservation of positivity for ρ^{n+1}.

Next, we show that (S2) is uniquely solvable. We first plug (S2-1) into (S2-2), and obtain

$$\frac{\rho^{n+1} - \rho^n}{\delta t} = \frac{r^n + \frac{1}{2\sqrt{E_1^n + C}} \left(\frac{\delta E_1^n}{\delta \rho}, \rho^{n+1} - \rho^n \right)}{\sqrt{E_1^n + C}} \nabla \cdot \left(\rho^n \nabla \frac{\delta E_1^n}{\delta \rho} \right) + \nabla \cdot \left(\rho^n \nabla \log \rho^{n+1} \right). \tag{15}$$

We then define $\mathcal{L}_n$ by $\mathcal{L}_n g = u$ where u is the solution of the following elliptic equation:

$$-\nabla \cdot \left(\rho^n \nabla u\right) = g, \quad \int_\Omega u \, \mathrm{d}\boldsymbol{x} = 0$$

with Neumann boundary conditions or periodic boundary conditions. Hence, $\mathcal{L}_n$ is self-adjoint and semi-positive definite. Similarly, we define a functional

$$\begin{aligned} F[\rho^{n+1}] := &\left(\rho^{n+1}(\log \rho^{n+1} - 1), 1\right) + \frac{1}{2\delta t}\left(\rho^{n+1} - \rho^n, \mathcal{L}_n\left(\rho^{n+1} - \rho^n\right)\right) \\ &+ \left(\frac{r^n}{\sqrt{E_1^n + C}} \frac{\delta E_1^n}{\delta \rho}, \rho^{n+1}\right) + \frac{1}{4(E_1^n + C)}\left(\frac{\delta E_1^n}{\delta \rho}, \rho^{n+1} - \rho^n\right)^2. \end{aligned} \tag{16}$$

The Euler-Lagrange equation of the above functional under mass conservation is

$$\begin{cases} \dfrac{1}{\delta t}\mathcal{L}_n(\rho^{n+1} - \rho^n) + \dfrac{r^n}{\sqrt{E_1^n + C}} \dfrac{\delta E_1^n}{\delta \rho} + \dfrac{1}{2E_1^n}\left(\dfrac{\delta E_1^n}{\delta \rho}, \rho^{n+1} - \rho^n\right) \dfrac{\delta E_1^n}{\delta \rho} + \log \rho^{n+1} = \lambda, \\ \displaystyle\int_\Omega (\rho^{n+1} - \rho^n) \, \mathrm{d}\boldsymbol{x} = 0, \end{cases} \tag{17}$$

where λ is the Lagrange multiplier to enforce the mass conservation. It is easy to see from the definition of $\mathcal{L}_n$ that the above is equivalent to (15).

On the other hand, F is clearly a strictly convex functional on

$$\mathcal{A} = \left\{\rho^{n+1} \in H^2(\Omega) \colon \rho^{n+1} > 0\right\}.$$

Hence, F admits a unique minimizer, i.e., (17) is uniquely solvable, which implies that (S2) is uniquely solvable. Furthermore, the minimizer ρ^{n+1} is positive since the derivative of the term $\rho^{n+1}(\log \rho^{n+1} - 1)$ tends to $-\infty$ at zero.

To show the energy dissipation, we take the inner product of (S2-1) with $\frac{r^{n+1}}{\sqrt{E_1^n + C}} \frac{\delta E_1^n}{\delta \rho} + \log \rho^{n+1}$, we obtain

$$\begin{aligned} &\left(\frac{r^{n+1}}{\sqrt{E_1^n + C}} \frac{\delta E_1^n}{\delta \rho} + \log \rho^{n+1}, \frac{\rho^{n+1} - \rho^n}{\delta t}\right) \\ &= -\left(\rho^n \nabla\left(\frac{r^{n+1}}{\sqrt{E_1^n + C}} \frac{\delta E_1^n}{\delta \rho} + \log \rho^{n+1}\right), \nabla\left(\frac{r^{n+1}}{\sqrt{E_1^n + C}} \frac{\delta E_1^n}{\delta \rho} + \log \rho^{n+1}\right)\right). \end{aligned}$$

On the other hand, taking the inner product of (S2-2) with $2r^{n+1}$, and using a Taylor expansion, we can rewrite the first term in the above to

$$\begin{aligned} &\left(\frac{r^{n+1}}{\sqrt{E_1^n + C}} \frac{\delta E_1^n}{\delta \rho} + \log \rho^{n+1}, \frac{\rho^{n+1} - \rho^n}{\delta t}\right) \\ &= \frac{2(r^{n+1} - r^n) r^{n+1}}{\delta t} + \left(\log \rho^{n+1}, \frac{\rho^{n+1} - \rho^n}{\delta t}\right) \\ &= \frac{1}{\delta t}\left(\left(r^{n+1}\right)^2 - \left(r^n\right)^2 + \left(r^{n+1} - r^n\right)^2\right) \\ &\quad + \frac{1}{\delta t}\left(\left(\log \rho^{n+1} - 1\right)\rho^{n+1} - \left(\log \rho^n - 1\right)\rho^n + \frac{\left(\rho^{n+1} - \rho^n\right)^2}{2\xi}\right). \end{aligned} \tag{18}$$

Combining the above two identities, we have

$$\begin{aligned}
&\tilde{E}^{n+1} - \tilde{E}^n \\
&\leqslant -\delta t \left(\rho^n \nabla \left(\frac{r^{n+1}}{\sqrt{E_1^n + C}} \frac{\delta E_1^n}{\delta \rho} + \log \rho^{n+1} \right), \nabla \left(\frac{r^{n+1}}{\sqrt{E_1^n + C}} \frac{\delta E_1^n}{\delta \rho} + \log \rho^{n+1} \right) \right) \\
&\leqslant 0,
\end{aligned}$$

where $\tilde{E}^n = \int_\Omega \rho^n (\log \rho^n - 1)\, \mathrm{d}\boldsymbol{x} + (r^n)^2$.

We can also construct a second-order scheme with the BDF and the Adams-Bashforth extrapolation as follows:

$$\frac{3\rho^{n+1} - 4\rho^n + \rho^{n-1}}{2\delta t} = \nabla \cdot \left(\rho^{*,n+\frac{1}{2}} \nabla \left(\frac{r^{n+1}}{\sqrt{E_1^{*,n+\frac{1}{2}} + C}} \frac{\delta E_1^{*,n+\frac{1}{2}}}{\delta \rho} + \log \rho^{n+1} \right) \right), \tag{19}$$

$$\frac{3r^{n+1} - 4r^n + r^{n-1}}{2\delta t} = \frac{1}{2\sqrt{E_1^{*,n+\frac{1}{2}} + C}} \left(\frac{\delta E_1^{*,n+\frac{1}{2}}}{\delta \rho}, \frac{3\rho^{n+1} - 4\rho^n + \rho^{n-1}}{2\delta t} \right), \tag{20}$$

where $E_1^{*,n+\frac{1}{2}} = E_1(\rho^{*,n+\frac{1}{2}})$, $\frac{\delta E_1^{*,n+\frac{1}{2}}}{\delta \rho} = \frac{\delta E_1}{\delta \rho}(\rho^{*,n+\frac{1}{2}})$, and $\rho^{*,n+\frac{1}{2}}$ is defined as (11).

Theorem 4 *Assume that ρ^1 is obtained from the first-order scheme. The above second-order scheme (19)–(20) exhibits the following attributes.*

(i) *Mass conservation*: $\int_\Omega \rho^n \,\mathrm{d}\boldsymbol{x} = \int_\Omega \rho^{n+1} \,\mathrm{d}\boldsymbol{x}$.
(ii) *Positivity preserving*: $\rho^{n+1} > 0$.
(iii) *Unique solvability.*

Proof The proof is similar to that of Theorem 3. We can define a slightly different linear operator $\tilde{\mathcal{L}}_n$ as follows: $\tilde{\mathcal{L}}_n g = u$ is defined by the elliptic equation

$$-\nabla \cdot \left(\rho^{*,n+\frac{1}{2}} \nabla u \right) = g, \quad \int_\Omega u \,\mathrm{d}\boldsymbol{x} = 0$$

with the homogeneous Neumann boundary conditions or periodic boundary conditions. We can also define a slightly different convex functional $\tilde{F}$ as follows:

$$\begin{aligned}
F[\rho^{n+1}] := &\left(\rho^{n+1}(\log \rho^{n+1} - 1), 1\right) + \left(\frac{4r^n - r^{n-1}}{3\sqrt{E_1^{*,n+\frac{1}{2}} + C}} \frac{\delta E_1^{*,n+\frac{1}{2}}}{\delta \rho}, \rho^{n+1} \right) \\
&+ \frac{1}{12\delta t} \left(3\rho^{n+1} - 4\rho^n + \rho^{n-1}, \tilde{\mathcal{L}}_n \left(3\rho^{n+1} - 4\rho^n + \rho^{n-1}\right)\right) \\
&+ \frac{1}{36(E_1^{*,n+\frac{1}{2}} + C)} \left(\frac{\delta E_1^{*,n+\frac{1}{2}}}{\delta \rho}, 3\rho^{n+1} - 4\rho^n + \rho^{n-1} \right)^2.
\end{aligned}$$

Then it can be shown that the solution of (19)–(20) is the unique minimizer of the above convex functional.

For the same reason as the second-order scheme of the first approach (S1), we are unable to show that the above second-order scheme is energy dissipative.

2.3 Application to Onsager Gradient Flows

The Onsager gradient flow, as presented in Ref. [11], is as follows:

$$\frac{\partial \rho}{\partial t} = \nabla \cdot \left(V_1(\rho) \nabla \frac{\delta E(\rho)}{\delta \rho} \right) - V_2(\rho) \frac{\delta E(\rho)}{\delta \rho}, \tag{21}$$

where $E(\rho)$ is an energy functional, and $V_1, V_2 \colon \mathbb{R} \to \mathbb{R}_+$ are positive mobility functions. By decomposing the energy functional $E(\rho)$ into two components: $E(\rho) = E_1(\rho) + E_2(\rho)$, where E_1 is bounded from below with a constant, E_2 is convex with respect to ρ, and introducing a scalar variable, r, defined as $r = \sqrt{E_1 + C}$, where C is a constant such that $E_1 + C > 0$, (21) can be written in the following manner:

$$\frac{\partial \rho}{\partial t} = \nabla \cdot \left(V_1 \nabla \left(\frac{r}{\sqrt{E_1 + C}} \frac{\delta E_1}{\delta \rho} + \frac{\delta E_2}{\delta \rho} \right) \right) - V_2 \left(\frac{r}{\sqrt{E_1 + C}} \frac{\delta E_1}{\delta \rho} + \frac{\delta E_2}{\delta \rho} \right).$$

Then, following the construction strategy of (S1) and denoting $V_1^n = V_1(\rho^n)$ and $V_2^n = V_2(\rho^n)$, we construct the following scheme:

$$\begin{aligned} \frac{\rho^{n+1} - \rho^n}{\delta t} = & \nabla \cdot \left(V_1^n \nabla \left(\frac{r^{n+1}}{\sqrt{E_1^n + C}} \frac{\delta E_1^n}{\delta \rho} + \frac{\delta E_2^{n+1}}{\delta \rho} \right) \right) \\ & + V_2^n \left(\frac{r^{n+1}}{\sqrt{E_1^n + C}} \frac{\delta E_1^n}{\delta \rho} + \frac{\delta E_2^{n+1}}{\delta \rho} \right), \end{aligned} \tag{22}$$

$$\frac{r^{n+1} - r^n}{\delta t} = \frac{1}{2\sqrt{E_1^n + C}} \left(\frac{\delta E_1^n}{\delta \rho}, \frac{\rho^{n+1} - \rho^n}{\delta t} \right). \tag{23}$$

Taking the inner products of (22) with $\frac{r^{n+1}}{\sqrt{E_1^n + C}} \frac{\delta E_1^n}{\delta \rho} + \frac{\delta E_2^{n+1}}{\delta \rho}$, of (23) with $2r^{n+1}$, we obtain

$$\begin{aligned} & \left(r^{n+1}\right)^2 - \left(r^n\right)^2 + \left(r^{n+1} - r^n\right)^2 + E_2^{n+1} - E_2^n \\ \leqslant & -\delta t \left(V_1^n \nabla \left(\frac{r^{n+1}}{\sqrt{E_1^n + C}} \frac{\delta E_1^n}{\delta \rho} + \frac{\delta E_2^{n+1}}{\delta \rho} \right), \nabla \left(\frac{r^{n+1}}{\sqrt{E_1^n + C}} \frac{\delta E_1^n}{\delta \rho} + \frac{\delta E_2^{n+1}}{\delta \rho} \right) \right) \\ & - \delta t \left(V_2^n \left(\frac{r^{n+1}}{\sqrt{E_1^n + C}} \frac{\delta E_1^n}{\delta \rho} + \frac{\delta E_2^{n+1}}{\delta \rho} \right), \left(\frac{r^{n+1}}{\sqrt{E_1^n + C}} \frac{\delta E_1^n}{\delta \rho} + \frac{\delta E_2^{n+1}}{\delta \rho} \right) \right) \leqslant 0. \end{aligned}$$

Through appropriate choice of the form of E_2, we can impose specific attributes for the scheme (22)–(23). As an illustrative example, by selecting $E_2(\rho) = \int_\Omega \rho(\log \rho - 1)\, \mathrm{d}\boldsymbol{x}$, we can ensure that ρ remains positive.

Theorem 5 *With $E_2(\rho) = \int_\Omega \rho(\log \rho - 1)\, \mathrm{d}\boldsymbol{x}$, the scheme (22)–(23) exhibits the following attributes.*

(i) *Positivity preserving*: $\rho^{n+1} > 0$.
(ii) *Unique solvability.*
(iii) *Energy dissipation law*:

$$\tilde{E}^{n+1} - \tilde{E}^n \leqslant 0, \tag{24}$$

where the discrete energy is defined as $\tilde{E}^n = \int_\Omega \rho^n (\log \rho^n - 1)\, \mathrm{d}\boldsymbol{x} + (r^n)^2$.

Proof The proof is almost the same as Theorem 3 excluding the mass conservation.

3 Finite-Difference Discretization in Space

In this section, we turn our attention to the construction of finite-difference schemes that can maintain the properties of the time discretizations discussed in the previous section. Since integration by parts is essential in mass conservation and energy dissipation in the spatial continuous case, a critical aspect in full discretization is to make sure that the implementation of boundary conditions satisfies suitable summation by parts formulae. This task is relatively straightforward in rectangular domains. Since the construction methodologies for (S1) are essentially the same as those presented in [38]. Our subsequent discussion primarily focuses on the 2D implementation of (S2).

Let the domain $[0, L]^2$ be discretized into M^2 points, designated as $x_{j,k} = \left(\left(j - \frac{1}{2}\right)\delta x, \left(k - \frac{1}{2}\right)\delta x\right)$ for $j, k = 1, \cdots, M$, with $\delta x = L/M$. Then, a fully discrete version of the scheme (S2-1)–(S2-2) is

$$\begin{aligned}\frac{\rho_{j,k}^{n+1} - \rho_{j,k}^{n}}{\delta t} = \frac{1}{\delta x^2}\Bigg[&\frac{\rho_{j+1,k}^{n} + \rho_{j,k}^{n}}{2}\left(\xi^{n+1}\varphi_{j+1,k}^{n} + \mu_{j+1,k}^{n+1} - \xi^{n+1}\varphi_{j,k}^{n} - \mu_{j,k}^{n+1}\right)\\ &- \frac{\rho_{j,k}^{n} + \rho_{j-1,k}^{n}}{2}\left(\xi^{n+1}\varphi_{j,k}^{n} + \mu_{j,k}^{n+1} - \xi^{n+1}\varphi_{j-1,k}^{n} - \mu_{j-1,k}^{n+1}\right),\\ &+ \frac{\rho_{j,k+1}^{n} + \rho_{j,k}^{n}}{2}\left(\xi^{n+1}\varphi_{j,k+1}^{n} + \mu_{j,k+1}^{n+1} - \xi^{n+1}\varphi_{j,k}^{n} - \mu_{j,k}^{n+1}\right)\\ &- \frac{\rho_{j,k}^{n} + \rho_{j,k-1}^{n}}{2}\left(\xi^{n+1}\varphi_{j,k}^{n} + \mu_{j,k}^{n+1} - \xi^{n+1}\varphi_{j,k-1}^{n} - \mu_{j,k-1}^{n+1}\right)\Bigg],\end{aligned} \tag{25}$$

$$r^{n+1} - r^{n} = \frac{\delta x^2}{2\sqrt{E_1^n + C}}\sum_{j,k=1}^{M}\varphi_{j,k}^{n}\left(\rho_{j,k}^{n+1} - \rho_{j,k}^{n}\right), \tag{26}$$

where $\xi^{n+1} = \frac{r^{n+1}}{\sqrt{E_1^n + C}}$, $\varphi_{j,k}^{n} = (\frac{\delta E_1}{\delta \rho})_{j,k}^{n}$, and $\mu_{j,k}^{n} = (\log \rho)_{j,k}^{n}$ for $1 \leqslant j, k \leqslant M$. To illustrate the handling of boundary conditions, we consider, as an example, the Neumann boundary conditions. For achieving summation by parts, it is necessary to impose boundary terms as follows:

$$\begin{cases}\dfrac{\varphi_{0,k}^{n+1} - \varphi_{1,k}^{n+1}}{\delta x} = 0, & \dfrac{\varphi_{M+1,k}^{n+1} - \varphi_{M,k}^{n+1}}{\delta x} = 0,\\ \dfrac{\mu_{0,k}^{n+1} - \mu_{1,k}^{n+1}}{\delta x} = 0, & \dfrac{\mu_{M+1,k}^{n+1} - \mu_{M,k}^{n+1}}{\delta x} = 0.\end{cases} \tag{27}$$

By representing $\tilde{\rho}^n$ as a vector composed of the elements $\rho_{j,k}^{n}$, $\tilde{\varphi}^n$ as a vector consisting of $\varphi_{j,k}^{n}$, and $\tilde{\mu}^n$ as a vector formed from $\mu_{j,k}^{n+1}$ for $1 \leqslant j, k \leqslant M$ arranged in the lexicographical order, we can rewrite the above into a matrix form as follows:

$$\frac{\delta x^2}{\delta t}\left(\tilde{\rho}^{n+1} - \tilde{\rho}^{n}\right) = -A^n\left(\frac{r^n}{\sqrt{E_1^n + C}}\tilde{\varphi}^n + \frac{\delta x^2}{2(E_1^n + C)}(\tilde{\varphi}^n)^{\mathrm{T}}\left(\tilde{\rho}^{n+1} - \tilde{\rho}^{n}\right)\tilde{\varphi}^n + \tilde{\mu}^{n+1}\right). \tag{28}$$

Here, A^n is a sparse matrix with nonzero elements adjacent to its diagonal. Two indices (j, k) and (j', k') are considered adjacent if $\left|j - j'\right| + \left|k - k'\right| = 1$. The diagonal entry of A^n is

$$2\rho_{j,k}^{n} + \frac{1}{2}\left(\rho_{j+1,k}^{n} + \rho_{j-1,k}^{n} + \rho_{j,k+1}^{n} + \rho_{j,k-1}^{n}\right). \tag{29}$$

A^n is symmetric and diagonally dominant, with positive diagonal entries, ensuring it is positive semi-definite with non-negative eigenvalues. The eigen-decomposition of A^n is $A^n = W^{\mathrm{T}}\Lambda W$ where $\Lambda = \mathrm{diag}\left(0, \mu_2, \cdots, \mu_{M^2}\right)$ and $\mu_j > 0$ for $j = 2, \cdots, M^2$. The pseudo-inverse of A^n, denoted $(A^n)^*$, is defined as $(A^n)^* = W^{\mathrm{T}}\,\mathrm{diag}\left(0, \mu_2^{-1}, \cdots, \mu_{M^2}^{-1}\right) W$. For the term $(\tilde{\varphi}^n)^{\mathrm{T}}\left(\tilde{\rho}^{n+1} - \tilde{\rho}^n\right)\tilde{\varphi}^n$, we have

$$(\tilde{\varphi}^n)^{\mathrm{T}}\left(\tilde{\rho}^{n+1} - \tilde{\rho}^n\right)\tilde{\varphi}^n = \tilde{\varphi}^n(\tilde{\varphi}^n)^{\mathrm{T}}\left(\tilde{\rho}^{n+1} - \tilde{\rho}^n\right).$$

Defining $\Psi^n = \tilde{\varphi}^n(\tilde{\varphi}^n)^{\mathrm{T}}$, a positive semi-definite matrix, and multiplying (28) by the pseudo-inverse $(A^n)^*$, we obtain

$$\frac{\delta x^2}{\delta t}(A^n)^*\left(\tilde{\rho}^{n+1} - \tilde{\rho}^n\right) + a_1^n\tilde{\varphi}^n + a_2^n\Psi^n\left(\tilde{\rho}^{n+1} - \tilde{\rho}^n\right) + \tilde{\mu}^{n+1} = 0, \tag{30}$$

where $a_1^n = \frac{r^n}{\sqrt{E_1^n + C}}$ and $a_2^n = \frac{\delta x^2}{2(E_1^n + C)}$.

For the above scheme, we have

Theorem 6 *The finite-difference scheme (25)–(26) enjoys the following properties.*

(i) *Mass conservation*

$$\delta x^2 \sum_{j,k=1}^{M} \rho_{j,k}^{n+1} = \delta x^2 \sum_{j,k=1}^{M} \rho_{j,k}^{n},\ 1 \leqslant i \leqslant N.$$

(ii) *Positivity preserving*: *if* $\rho_{j,k}^n > 0$ *for* (j,k), *we have* $\rho_{j,k}^{n+1} > 0$ *for all* (j,k).

(iii) *Unique solvability*: *the scheme (25) possesses a unique solution* $\rho_{j,k}^{n+1}$.

(iv) *Energy dissipation law*:

$$\begin{aligned}
&\tilde{E}^{n+1} - \tilde{E}^n \\
\leqslant& -\frac{\delta t}{\delta x^2}\sum_{1\leqslant j\leqslant M-1, 1\leqslant k\leqslant M}\frac{\rho_{j+1,k}^n + \rho_{j,k}^n}{2}\left(\xi^{n+1}\varphi_{j+1,k}^n + \mu_{j+1,k}^{n+1} - \xi^{n+1}\varphi_{j,k}^n - \mu_{j,k}^{n+1}\right)^2 \\
&+ \sum_{1\leqslant j\leqslant M, 1\leqslant k\leqslant M-1}\frac{\rho_{j,k+1}^n + \rho_{j,k}^n}{2}\left(\xi^{n+1}\varphi_{j,k+1}^n + \mu_{j,k+1}^{n+1} - \xi^{n+1}\varphi_{j,k}^n - \mu_{j,k}^{n+1}\right)^2,
\end{aligned}$$

where

$$\tilde{E}^n = \sum_{j,k=1}^{M}\rho_{j,k}^n\left(\log\rho_{j,k}^n - 1\right) + (r^n)^2.$$

Proof The mass conservation is obtained by taking the sum over $1 \leqslant j, k \leqslant M$ on (25) and using the boundary conditions of $(\varphi)_{j,k}^n$ and $(\mu)_{j,k}^{n+1}$ in (27).

To prove the unique solvability and positivity preservation, we define

$$\begin{aligned}
F[\tilde{\rho}^{n+1}] = &\frac{\delta x^2}{2\delta t}\left(\tilde{\rho}^{n+1} - \tilde{\rho}^n\right)^{\mathrm{T}}(A^n)^*\left(\tilde{\rho}^{n+1} - \tilde{\rho}^n\right) + a_1^n\left(\tilde{\varphi}^n\right)^{\mathrm{T}}\tilde{\rho}^{n+1} \\
&+ \frac{a_2^n}{2}\left(\tilde{\rho}^{n+1} - \tilde{\rho}^n\right)^{\mathrm{T}}\Psi^n\left(\tilde{\rho}^{n+1} - \tilde{\rho}^n\right) + \left(\tilde{\rho}^{n+1}\right)^{\mathrm{T}}\left(\log\tilde{\rho}^{n+1} - 1\right).
\end{aligned} \tag{31}$$

It is evident that $F[\tilde{\rho}^{n+1}]$ is strictly convex with respect to $\tilde{\rho}^{n+1}$, and (30) represents its Euler-Lagrange equation. Consequently, the solution to (30) is the unique minimizer of $F[\tilde{\rho}^{n+1}]$ with $\tilde{\rho}^{n+1} > 0$. As any element of $\tilde{\rho}^{n+1}$ approaches zero, the gradient of $F[\tilde{\rho}^{n+1}]$ tends

towards negative infinity, implying that the function's value decreases when incrementing elements near 0 in $\tilde{\rho}^{n+1}$. Therefore, a minimizer with any zero element in $\tilde{\rho}^{n+1}$ is not attainable.

The energy dissipation law can be obtained by multiplying (25) with $\xi^{n+1}\varphi^n_{j,k}+\mu^{n+1}_{j,k}$ and multiplying (26) with $2r^{n+1}$. The summation by parts is used with the employed boundary conditions of $\varphi^n_{j,k}$ and $\mu^{n+1}_{j,k}$.

4 Numerical Experiments

In this section, we present various numerical experiments to validate the theoretical results discussed previously. It is important to note that our schemes require solving a nonlinear system at every time step. To solve these nonlinear equations, we employ the damped Newton's iteration method [22]. In scenarios where the density must remain positive, we enforce the positivity by replacing ρ with $\max(\rho, \epsilon)$, where $0 < \epsilon \ll 1$ is a minimal threshold. We set $\epsilon = 10^{-6}$ in the following experiments.

4.1 Accuracy Test

First, we test the accuracy in time of the schemes (S1) and (S2) by solving a heat equation with Neumann boundary conditions:

$$\begin{cases} \dfrac{\partial \rho}{\partial t} = \dfrac{1}{50}\rho_{xx} = \dfrac{1}{50}\partial_x(\rho\partial_x \log(\rho)), \quad x \in [0,1], t > 0, \\ \rho_x(0,t) = \rho_x(1,t) = 0, \end{cases} \tag{32}$$

which can be expressed as a Wasserstein gradient flow with $E(\rho) = \int_\Omega \frac{1}{50}\rho(\log\rho - 1)\,\mathrm{d}\boldsymbol{x}$. The exact solution is taken as $\rho(x,t) = \mathrm{e}^{-\pi^2 t/50}\cos(\pi x) + 1.1$. To assess the accuracy of our model, we employ both the L_∞ error and the L_2 error. These are defined by

$$e^N_\infty := \max_i \left|\rho^N_i - \rho\left(x_i, T\right)\right|,$$

$$e^N_2 := \left(\delta x \sum_{i=1}^{I-1}\left(\rho^N_i - \rho\left(x_i, T\right)\right)^2\right)^{\frac{1}{2}},$$

where ρ^N_i is the value of ρ^N at x_i.

To determine the convergent rate of different schemes, we utilize a fine mesh with a finite-difference method, specifically setting $N = 50\,000$. The time steps chosen are $\delta t = 0.1, 0.05, 0.025, 0.012\,5$, with a final time of $T = 1$. For the scheme (S2), we decompose the energy as

$$E(\rho) = \int_\Omega \frac{1}{50}\rho(\log\rho - 1)\,\mathrm{d}\boldsymbol{x} = \int_\Omega \frac{1}{100}\rho(\log\rho - 1)\,\mathrm{d}\boldsymbol{x} + \int_\Omega \frac{1}{100}\rho(\log\rho - 1)\,\mathrm{d}\boldsymbol{x}.$$

It is observed that the schemes both (S1) and (S2) consistently demonstrate first-order convergent rates in time (Tables 1 and 2).

Table 1 Heat equation, the first-order convergent rate of (S1) in time

$\delta x = 1/50\,000$	e_∞^N	e_2^N	Order of e_∞^N	Order of e_2^N
$\delta t = 0.1$	8.154 0E−03	3.047 3E−03	–	–
$\delta t = 0.05$	4.110 1E−03	1.545 6E−03	0.988 3	0.979 4
$\delta t = 0.025$	2.057 8E−03	7.774 6E−04	0.998 1	0.991 3
$\delta t = 0.012\,5$	1.024 4E−03	3.889 0E−04	1.006 3	0.999 4

Table 2 Heat equation, the first-order convergent rate of (S2) in time

$\delta x = 1/50\,000$	e_∞^N	e_2^N	Order of e_∞^N	Order of e_2^N
$\delta t = 0.1$	3.279 8E−03	1.306 0E−03	–	–
$\delta t = 0.05$	1.649 7E−03	6.581 5E−04	0.979 4	0.979 4
$\delta t = 0.025$	8.224 1E−04	3.292 6E−04	0.991 3	0.991 3
$\delta t = 0.012\,5$	4.055 6E−04	1.635 2E−04	0.999 4	0.999 4

4.2 Barenblatt Solution

The Barenblatt solution, a fundamental benchmark for the PME, is widely used to test the accuracy and efficiency of numerical methods. The PME, denoted by $\frac{\partial \rho}{\partial t} = \Delta \rho^m$, integrates energy as $E = \int_\Omega \frac{1}{m-1} \rho^m(\boldsymbol{x})\ \mathrm{d}\boldsymbol{x}$ and takes the explicit form:

$$B_{m,d}(\boldsymbol{x}, t) = (t+1)^{-\alpha} \left[1 - \frac{\alpha(m-1)}{2md} \frac{\|\boldsymbol{x}\|^2}{(t+1)^{2\alpha/d}} \right]_+^{1/(m-1)},$$

where $(s)_+ = \max(s, 0)$ and $\alpha = \frac{d}{d(m-1)+2}$. In this context, $d = 2$ represents the spatial dimensions of the problem. To illustrate the precision and energy dissipation efficiency of our proposed numerical schemes, we evaluate the solution from $t_0 = 0$ to $T = 1$, with a time step $\delta t = 0.001$ and spatial step $\delta x = 0.25$, over the domain $\Omega = (-10, 10)^2$. We utilize schemes (S1) and (S2) for $m = 3$. For the scheme (S1), the energy is defined as

$$E_{\mathrm{S1}} = \int_\Omega \rho(\boldsymbol{x}) \left(\log \rho(\boldsymbol{x}) - 1 \right) \mathrm{d}\boldsymbol{x}.$$

For the scheme (S2), the original energy is defined as

$$E_{\mathrm{S2}} = \int_\Omega \frac{1}{m-1} \rho^m(\boldsymbol{x})\ \mathrm{d}\boldsymbol{x},$$

and E_{S2} is decomposed into $E_{\mathrm{S2}} = E_{1,\mathrm{S2}} + \int_\Omega \rho(\boldsymbol{x}) \left(\log \rho(\boldsymbol{x}) - 1 \right)\ \mathrm{d}\boldsymbol{x}$. The modified energy for (S2) is then defined as

$$\tilde{E}_{\mathrm{S2}} = r^2 + \int_\Omega \rho(\boldsymbol{x}) \left(\log \rho(\boldsymbol{x}) - 1 \right)\ \mathrm{d}\boldsymbol{x}.$$

It is demonstrated that the quantity $E_{1,\mathrm{S2}}$, defined by

$$E_{1,\mathrm{S2}} = \int_\Omega \left(\frac{\rho^m(\boldsymbol{x})}{m-1} - \rho(\boldsymbol{x}) \left(\log \rho(\boldsymbol{x}) - 1 \right) \right) \mathrm{d}\boldsymbol{x},$$

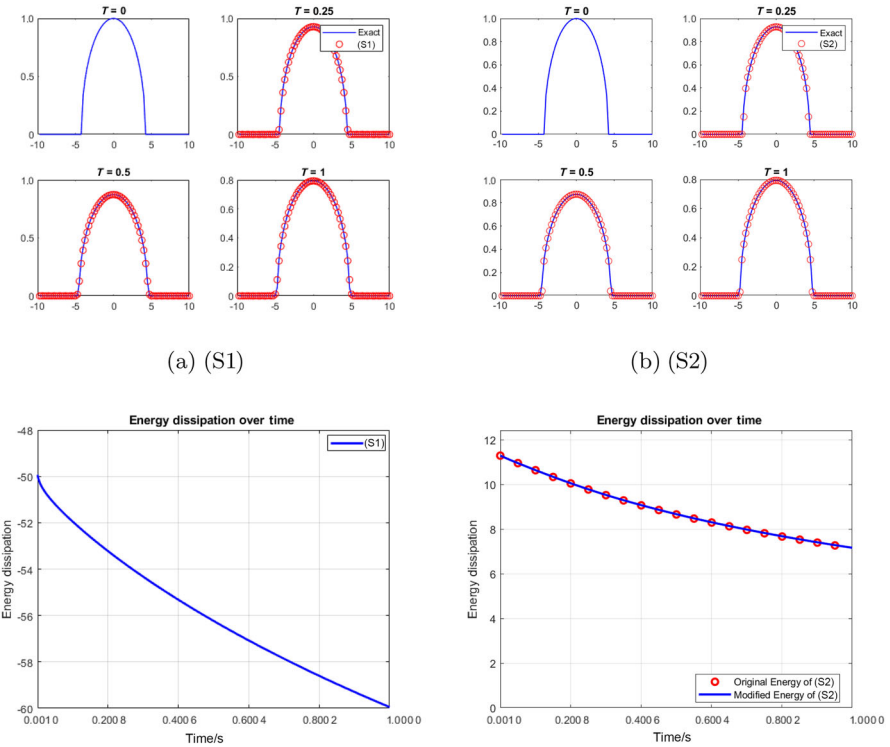


Fig. 1 Visualizations of the PME evolution with (S1) and (S2) at a cross-section $y = 0$ and energy dissipation comparison for $m = 3$

is bounded from below when $m \geqslant 2$. In this context, setting $C = 0$ for (S2) is sufficient to ensure $E_{1,S2}$ remains positive.

Figures 1a and b display the solution's evolution at the cross-section $y = 0$ for (S1) and (S2), respectively. These figures confirm that our schemes can closely approximate the exact solution, demonstrating their accuracy. Figures 1c and d illustrate the energy dissipation for (S1) and (S2), respectively, with Fig. 1d also comparing the original and modified energy, which are observed to be consistent.

To evaluate the computational efficiency of our schemes, we compared the average number of Newton iterations required every 50 steps (Fig. 2). Both (S1) and (S2) achieve convergence with several Newton iterations, where only the initial few steps require slightly more iterations. Thanks to its simpler structure, the scheme (S1) generally needs fewer iterations, making it the preferred option when feasible.

4.3 Fokker-Planck Equation

To test the proposed schemes in the context of potential-influenced equations, we turn our attention to the linear Fokker-Planck equation presented as

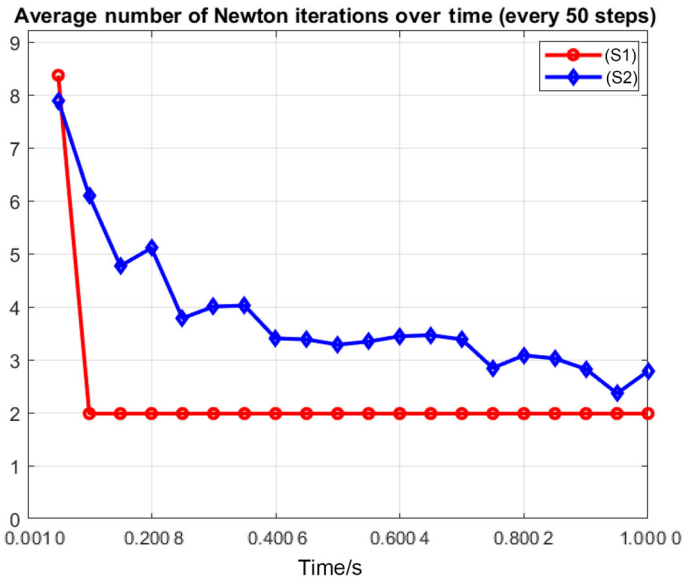


Fig. 2 Comparative analysis of the average number of Newton iterations required for schemes (S1) and (S2) every 50 computational steps over the simulation time from $t_0 = 0$ to $T = 1$, demonstrating the relative computational efficiency of each scheme

$$\frac{\partial \rho}{\partial t} = \Delta \rho + \nabla \cdot (\rho \nabla V) = \nabla \cdot (\rho \nabla (\log \rho + V)),$$

which is a Wasserstein gradient flow with the energy functional $E(\rho)$ as

$$E(\rho) = \int_{\Omega} (\rho(\log(\rho) - 1) + \rho V) \, \mathrm{d}\boldsymbol{x}.$$

In this example, we take $V(x, y) = \frac{x^2+y^2}{2}$ which will guide the system towards a unique and globally stable equilibrium, irrespective of the initial conditions. The equilibrium state, also known as the heat kernel, is analytically defined as

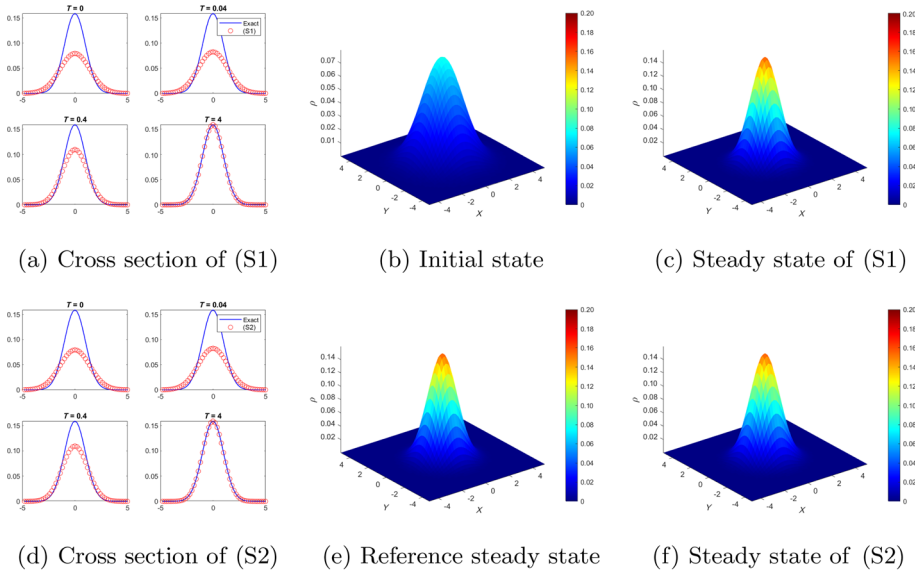
$$\rho(x, y, t) = \frac{1}{4\,\pi\, t} \exp\left(-\frac{x^2 + y^2}{4t}\right), \tag{33}$$

evaluated at $t = \frac{1}{2}$, i.e., $\rho_\infty = \rho(x, y, 1/2)$.

For our numerical experiments, we adopt an initial state $\rho(x, y, 1)$ to approximate the above-mentioned steady state, $\rho(x, y, \infty)$. Both schemes, (S1) and (S2), are employed for this purpose, and the resultant solutions at the final time $T = 4$ are presented in Fig. 3. For the scheme (S2), the decomposed energy functional $E_{2,\mathrm{S2}}$ is specifically taken as $E_{2,\mathrm{S2}}(\rho) = \int_\Omega \rho(\log \rho - 1) \, \mathrm{d}\boldsymbol{x}$. We also set C as 10 to keep $E_{1,\mathrm{S2}} + C$ positive, where $E_{1,\mathrm{S2}} = E - E_{2,\mathrm{S2}}$. As illustrated in Fig. 4, (S1) and (S2) exhibit comparable averages in the number of Newton iterations throughout most of the simulation.

4.4 PME with Various *m* and a Slow Drift

The PME with a slow drift has earned notable attention in optimal transfers, particularly when accompanied by a high value of m [24]. A general energy functional for such scenarios



(a) Cross section of (S1) (b) Initial state (c) Steady state of (S1)

(d) Cross section of (S2) (e) Reference steady state (f) Steady state of (S2)

Fig. 3 Solution profiles for the Fokker-Planck equation. The simulation begins with the initial condition $\rho(x, y, 1)$ as described in (33). The reference steady state is given by ρ_∞. Additionally, cross-sectional views of the solution at $y = 0$ are presented in **a**, **d**

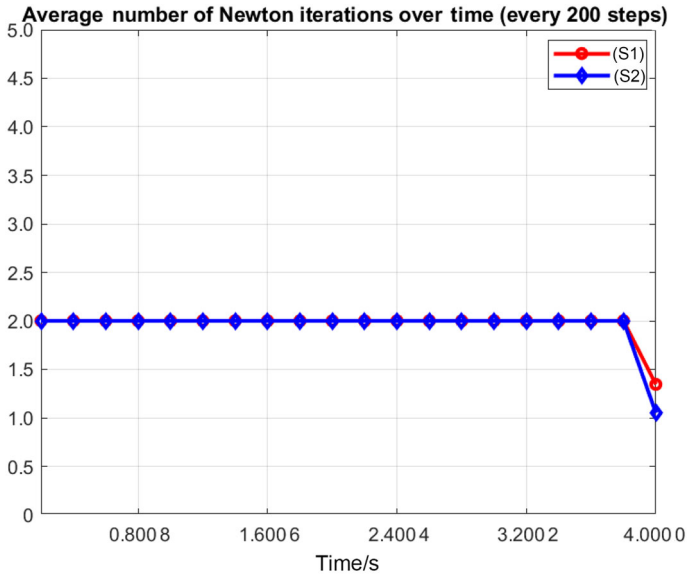


Fig. 4 The average number of Newton iterations for schemes (S1) and (S2), measured every 200 computational steps from $t_0 = 0$ to $T = 4$

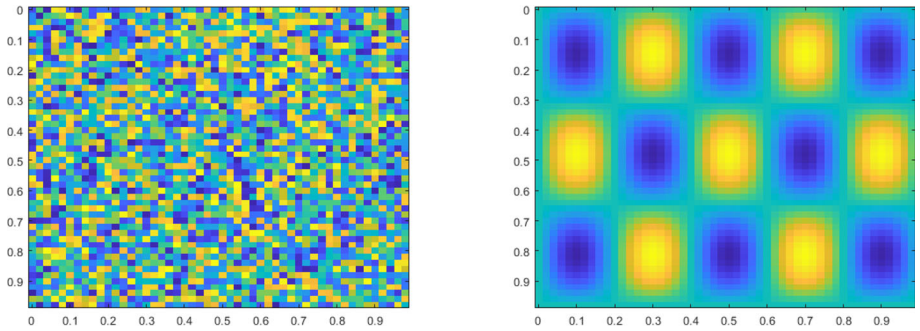


Fig. 5 On the left, the initial state is set to a random configuration using the seed(1). On the right, we depict the potential function given by $V(x, y) = 1 - \sin(5\,\pi\,x)\sin(3\,\pi\,y)$

can be expressed as

$$E(\rho) = \int_{\Omega} \left(\frac{1}{m-1} \rho^m(\boldsymbol{x}) + \rho(\boldsymbol{x}) V(\boldsymbol{x}) \right) \mathrm{d}\boldsymbol{x}, \tag{34}$$

where $V(\boldsymbol{x})$ is a given potential function. For the scope of this study, we take $V(x, y) = 1 - \sin(5\,\pi\,x)\sin(3\,\pi\,y)$. Our simulations are conducted on a domain of $[-1, 1]^2$, using a grid of 50×50 spatial points. The initial conditions are randomly set within the $[-1, 1]^2$ range. The graphical representations of the initial density and the potential function are available in Fig. 5.

Our scheme (S1) does not enjoy the energy dissipation law when $m \geqslant 2$, whereas the scheme (S2) is energy dissipative with respect to a modified energy. Hence, we test our scheme (S2) across different values of m: $m = 2, 4, 6, 20, 50, 100$ to understand how the results change as m increases. For the scheme (S2), the decomposed energy functional E_2 is specifically taken as $E_2(\rho) = \int_{\Omega} \rho(\log \rho - 1)\,\mathrm{d}\boldsymbol{x}$. We also set C as 5 to keep $E_1 + C$ positive, where $E_1 = E - E_2$. For all values of m, we run simulations from $T = 0$ to $T = 0.04$ to observe the evolution over time and the approach to the steady state. The results in Fig. 6 indicate that regions with lower potential attract higher density and that the steady state with larger m appears to be significantly more dispersed than the steady state with smaller m. Figure 7 shows the evolution of the modified energy and original energy for the scheme (S2). Although the modified energy decays slightly faster than the original energy with larger m, both energies consistently exhibit dissipative behavior throughout the evolution.

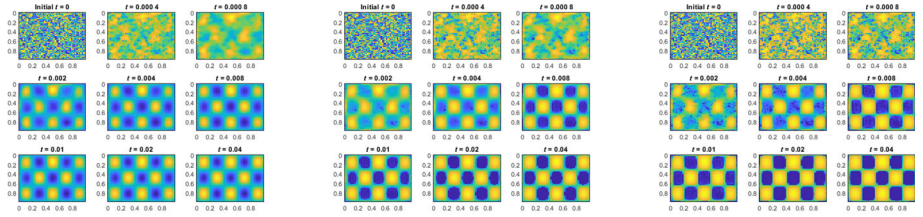
4.5 Fisher-KPP Equation

As a special case of Onsager gradient flows (21), we consider the Fisher-KPP equation [7, 31] characterized by potentials $V_1(\rho) = \alpha\rho$ and $V_2(\rho) = \frac{\rho(\rho-1)}{2\log(\rho)}$. The energy associated with this system is given by

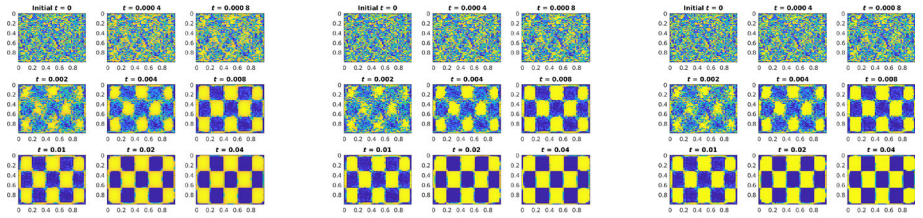
$$E(\rho) := \int_{\Omega} 2\rho(\log(\rho) - 1)\,\mathrm{d}\boldsymbol{x} + C.$$

Then, the equation can be presented as follows:

$$\frac{\partial \rho}{\partial t} = \nabla \cdot (2\alpha\rho \nabla \log \rho) + \rho(1 - \rho).$$

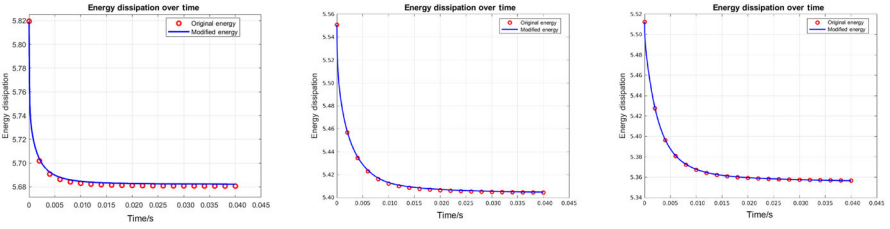


(a) $m = 2, T = 0.04$ (b) $m = 4, T = 0.04$ (c) $m = 6, T = 0.04$

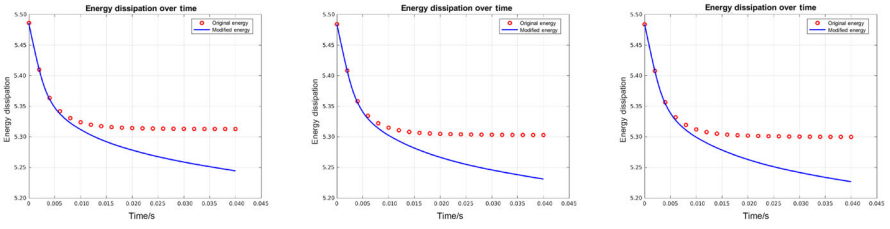


(d) $m = 20, T = 0.04$ (e) $m = 50, T = 0.04$ (f) $m = 100, T = 0.04$

Fig. 6 Evolution of the PME with a slow drift for various values of m. The time evolution is observed until $T = 0.04$ for all values of m. The results show that regions with lower potential attract higher density, with the steady state appearing more dispersed for larger m values compared to smaller m values



(a) $m = 2, T = 0.04$ (b) $m = 4, T = 0.04$ (c) $m = 6, T = 0.04$



(d) $m = 20, T = 0.04$ (e) $m = 50, T = 0.04$ (f) $m = 100, T = 0.04$

Fig. 7 Energy dissipation of the PME with a slow drift for various values of m

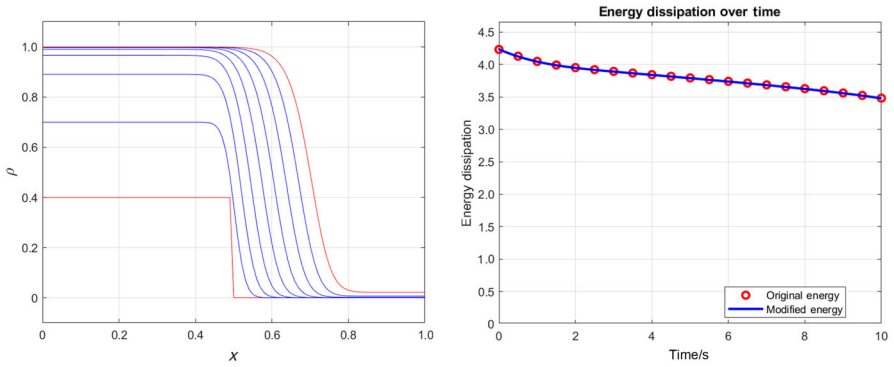


Fig. 8 Evolutionary behavior of the Fisher-KPP equation. The left figure delineates the density evolution from $t_0 = 0$ to $T = 10$, emphasizing the initial and terminal states via the red line. The right figure compares the dissipation of the modified energy and original energy

We fix the domain $\Omega = [0, 1]$, and set the parameters $\alpha = 10^{-4}$, $C = 5$. The initial condition for the density ρ is defined as $\rho(x, 0) = 0.4$ for $0 \leqslant x < 1/2$ and $\rho(x, 0) = 0$ otherwise. For the numerical simulation of the system, we employ the scheme (S2) with $E_1 = \int_\Omega \rho(\log(\rho) - 1)\,\mathrm{d}\boldsymbol{x} + C$ and $E_2 = \int_\Omega \rho(\log(\rho) - 1)\,\mathrm{d}\boldsymbol{x}$ combined with a finite-difference discretization. In this example, we use $N = 100$ spatial grid points and a time step size of $\delta t = 10^{-4}$.

Figure 8 provides a visual representation of the system's behavior. The left figure illustrates the evolution of the system for time intervals ranging from $t_0 = 0$ to $T = 10$, with the initial and final states being demarcated by the red line. The right figure contrasts the modified energy with the original energy, showcasing the energy dissipation as time progresses.

5 Concluding Remarks

In this paper, we introduced two novel numerical schemes specifically tailored for Wasserstein gradient flows. The first scheme is a generalization of the schemes proposed in Refs. [20, 38], while the second scheme is based on energy splitting along with a scalar auxiliary variable to ensure energy dissipation. We demonstrated that both schemes ensure mass conservation, positivity preserving, unique solvability, and the first scheme is energy dissipative in some special cases while the second scheme is energy dissipative with a modified energy.

These schemes were designed to address the challenges associated with Wasserstein gradient flows, particularly in preserving positivity and energy dissipation. Each scheme was rigorously tested through a series of numerical experiments, affirming their theoretical precision and computational efficiency. The results confirmed that our schemes not only align with theoretical predictions but also demonstrate significant computational improvements.

In summary, the schemes proposed in this work are both robust and practically efficient for solving a class of Wasserstein gradient flows, paving the way for further exploration in diverse scientific and engineering fields.

Funding This work was partially supported by the NSFC (Grant No. 12371409).

Declarations

Conflict of Interest This is no conflict of interest.

References

1. Allen, S.M., Cahn, J.W.: A microscopic theory for antiphase boundary motion and its application to antiphase domain coarsening. Acta Metall. **27**, 1085–1095 (1979)
2. Anderson, D.M., McFadden, G.B., Wheeler, A.A.: Diffuse-interface methods in fluid mechanics. Annu. Rev. Fluid Mech. **30**, 139–165 (1998)
3. Bailo, R., Carrillo, J.A., Hu, J.: Fully discrete positivity-preserving and energy-dissipating schemes for aggregation-diffusion equations with a gradient-flow structure. Commun. Math. Sci. **18**, 1259–1303 (2020)
4. Bazant, M.Z., Thornton, K., Ajdari, A.: Diffuse-charge dynamics in electrochemical systems. Phys. Rev. E **70**, 021506 (2004)
5. Biler, P., Hebisch, W., Nadzieja, T.: The Debye system: existence and large time behavior of solutions. Nonlinear Anal. Theory Methods Appl. **23**, 1189–1209 (1994)
6. Bolley, F., Canizo, J.A., Carrillo, J.A.: Stochastic mean-field limit: non-Lipschitz forces and swarming. Math. Models Methods Appl. Sci. **21**, 2179–2210 (2011)
7. Bonizzoni, F., Braukhoff, M., Jüngel, A., Perugia, I.: A structure-preserving discontinuous Galerkin scheme for the Fisher-KPP equation. Numer. Math. **146**, 119–157 (2020)
8. Cahn, J.W., Hilliard, J.E.: Free energy of a nonuniform system. I. Interfacial free energy. J. Chem. Phys. **28**, 258–267 (1958)
9. Carrillo, J.A., Craig, K., Wang, L., Wei, C.: Primal dual methods for Wasserstein gradient flows. Found. Comput. Math. **2022**, 1–55 (2022)
10. Carrillo, J.A., Wang, L., Xu, W., Yan, M.: Variational asymptotic preserving scheme for the Vlasov-Poisson-Fokker-Planck system. Multiscale Model. Simul. **19**, 478–505 (2021)
11. Doi, M.: Onsager's variational principle in soft matter. J. Phys.: Condens. Matter **23**, 284118 (2011)
12. Doi, M., Edwards, S.F.: The Theory of Polymer Dynamics, vol. 73. Oxford University Press, Oxford (1988)
13. Duan, C., Chen, W., Liu, C., Wang, C., Zhou, S.: Convergence analysis of structure-preserving numerical methods for nonlinear Fokker-Planck equations with nonlocal interactions. Math. Methods Appl. Sci. **45**, 3764–3781 (2022)
14. Duan, C., Chen, W., Liu, C., Yue, X., Zhou, S.: Structure-preserving numerical methods for nonlinear Fokker-Planck equations with nonlocal interactions by an energetic variational approach. SIAM J. Sci. Comput. **43**, B82–B107 (2021)
15. Elder, K., Katakowski, M., Haataja, M., Grant, M.: Modeling elasticity in crystal growth. Phys. Rev. Lett. **88**, 245701 (2002)
16. Elliott, C.M., Stuart, A.: The global dynamics of discrete semilinear parabolic equations. SIAM J. Numer. Anal. **30**, 1622–1663 (1993)
17. Eyre, D.J.: Unconditionally gradient stable time marching the Cahn-Hilliard equation. MRS Online Proc. Lib. (OPL) **529**, 39 (1998)
18. Fu, G., Liu, S., Osher, S., Li, W.: High order computation of optimal transport, mean field planning, and mean field games. arXiv:2302.02308 (2023)
19. Fu, G., Osher, S., Li, W.: High order spatial discretization for variational time implicit schemes: Wasserstein gradient flows and reaction-diffusion systems. arXiv:2303.08950 (2023)
20. Gu, Y., Shen, J.: Bound preserving and energy dissipative schemes for porous medium equation. J. Comput. Phys. **410**, 109378 (2020)
21. Gurtin, M.E., Polignone, D., Vinals, J.: Two-phase binary fluids and immiscible fluids described by an order parameter. Math. Models Methods Appl. Sci. **6**, 815–831 (1996)
22. Harker, P., Pang, J.-S.: A damped-Newton method for the linear complementarity problem. Numer. Algorithms **26**(01), 265–284 (1990)
23. Hu, J., Zhang, X.: Positivity-preserving and energy-dissipative finite difference schemes for the Fokker-Planck and Keller-Segel equations. IMA J. Numer. Anal. **43**, 1450–1484 (2023)
24. Jacobs, M., Lee, W., Léger, F.: The back-and-forth method for Wasserstein gradient flows. ESAIM Control Optim. Calc. Var. **27**, 28 (2021)
25. Jacobs, M., Léger, F.: A fast approach to optimal transport: the back-and-forth method. Numer. Math. **146**, 513–544 (2020)

26. Jordan, R., Kinderlehrer, D., Otto, F.: The variational formulation of the Fokker-Planck equation. SIAM J. Math. Anal. **29**, 1–17 (1998)
27. Keller, E.F., Segel, L.A.: Initiation of slime mold aggregation viewed as an instability. J. Theor. Biol. **26**, 399–415 (1970)
28. Keller, E.F., Segel, L.A.: Model for chemotaxis. J. Theor. Biol. **30**, 225–234 (1971)
29. Leclerc, H., Mérigot, Q., Santambrogio, F., Stra, F.: Lagrangian discretization of crowd motion and linear diffusion. SIAM J. Numer. Anal. **58**, 2093–2118 (2020)
30. Leslie, F.M.: Theory of flow phenomena in liquid crystals. In: Advances in Liquid Crystals, vol. 4, pp. 1–81. Elsevier, London (1979)
31. Li, W., Lee, W., Osher, S.: Computational mean-field information dynamics associated with reaction-diffusion equations. J. Comput. Phys. **466**, 111409 (2022)
32. Patlak, C.S.: Random walk with persistence and external bias. Bull. Math. Biophys. **15**, 311–338 (1953)
33. Peletier, M.A.: Variational modelling: energies, gradient flows, and large deviations. arXiv:1402.1990 (2014)
34. Peyré, G.: Entropic approximation of Wasserstein gradient flows. SIAM J. Imaging Sci. **8**, 2323–2351 (2015)
35. Shen, J., Wang, C., Wang, X., Wise, S.M.: Second-order convex splitting schemes for gradient flows with Ehrlich-Schwoebel type energy: application to thin film epitaxy. SIAM J. Numer. Anal. **50**, 105–125 (2012)
36. Shen, J., Xu, J.: Convergence and error analysis for the scalar auxiliary variable (SAV) schemes to gradient flows. SIAM J. Numer. Anal. **56**, 2895–2912 (2018)
37. Shen, J., Xu, J.: Unconditionally bound preserving and energy dissipative schemes for a class of Keller-Segel equations. SIAM J. Numer. Anal. **58**, 1674–1695 (2020)
38. Shen, J., Xu, J.: Unconditionally positivity preserving and energy dissipative schemes for Poisson-Nernst-Planck equations. Numer. Math. **148**, 671–697 (2021)
39. Shen, J., Xu, J., Yang, J.: The scalar auxiliary variable (SAV) approach for gradient flows. J. Comput. Phys. **353**, 407–416 (2018)
40. Shen, J., Xu, J., Yang, J.: A new class of efficient and robust energy stable schemes for gradient flows. SIAM Rev. **61**, 474–506 (2019)
41. Sznitman, A.-S.: Topics in propagation of chaos. Lect. Notes Math. **1991**, 165–251 (1991)
42. Vázquez, J.L.: An introduction to the mathematical theory of the porous medium equation. In: Shape Optimization and Free Boundaries, pp. 347–389. Springer, London (1992)
43. Vázquez, J.L.: The Porous Medium Equation: Mathematical Theory. Oxford University Press, Oxford (2007)
44. Villani, C.: Topics in Optimal Transportation, vol. 58. American Mathematical Soc., London (2021)
45. Yang, X.: Linear, first and second-order, unconditionally energy stable numerical schemes for the phase field model of homopolymer blends. J. Comput. Phys. **327**, 294–316 (2016)
46. Yue, P., Feng, J.J., Liu, C., Shen, J.: A diffuse-interface method for simulating two-phase flows of complex fluids. J. Fluid Mech. **515**, 293–317 (2004)
47. Zhao, J., Wang, Q., Yang, X.: Numerical approximations for a phase field dendritic crystal growth model based on the invariant energy quadratization approach. Int. J. Numer. Methods Eng. **110**, 279–300 (2017)