COMMUN. MATH. SCI.
Vol. 20, No. 4, pp. 987–1024

A SECOND ORDER ACCURATE, ENERGY STABLE NUMERICAL SCHEME FOR THE ONE-DIMENSIONAL POROUS MEDIUM EQUATION BY AN ENERGETIC VARIATIONAL APPROACH*

CHENGHUA DUAN†, WENBIN CHEN‡, CHUN LIU§, CHENG WANG¶, AND XINGYE YUE‖

Abstract. The porous medium equation (PME) is a typical nonlinear degenerate parabolic equation. An energetic variational approach (EVA) provides many insights to such a physical model, in which the trajectory equation can be obtained, based on different dissipative energy laws. In this article, we propose and analyze a second order accurate in time numerical scheme for the PME in the EVA approach. A modified Crank-Nicolson temporal discretization is applied, combined with the finite difference over a uniform spatial mesh. Such a numerical scheme is highly nonlinear, and it is proved to be uniquely solvable on an admissible convex set, in which the convexity of the nonlinear implicit terms will play an important role. Subsequently, the energy dissipation property is established, with careful summation by parts formulas applied in the spatial discretization. More importantly, an optimal rate convergence analysis is provided in this work, in which many highly non-standard estimates have to be involved, due to the nonlinear parabolic coefficients. The higher order asymptotic expansion (up to fourth order temporal and spatial accuracy), the rough error estimate (to establish the $W_h^{1,\infty}$ bound for the numerical variable), and the refined error estimate have to be carried out to accomplish such a convergence result. In our knowledge, it will be the first work to combine three theoretical properties for a second order accurate numerical scheme to the PME in the EVA approach: unique solvability, energy stability and optimal rate convergence analysis. A few numerical results are also presented in this article, which demonstrate the robustness of the proposed numerical scheme.

Keywords. Energetic variational approach; porous medium equation; trajectory equation; optimal rate convergence analysis.

AMS subject classifications. 35K35; 35K55; 49J40; 65M06; 65M12.

1. Introduction and background

In this paper, we consider the second order scheme of the porous medium equation (PME):

$$\partial_t f = \Delta_x (f^m),\ x \in \Omega \subset \mathbb{R}^d,\ m > 1,$$

where $f := f(x,t)$ is a non-negative scalar function of space $x \in \mathbb{R}^d$ $(d \geq 1)$ and the time $t \in \mathbb{R}^+$, and m is a constant larger than 1. It has wide application in many physical and biological models, such as an isentropic gas flow through a porous medium, the viscous gravity currents, nonlinear heat transfer and image processing [39], etc.

The basic characteristic of the PME is that it is degenerate at points where $f = 0$. In turn, there are many special features: the finite speed of propagation, the free boundary, and a possible waiting time phenomenon [13, 39]. Many theoretical analyses

*Received: August 13, 2020; Accepted (in revised form): October 15, 2021. Communicated by Jie Shen.

†Department of Mathematics, Shanghai University, Shanghai, 200444, China and Shanghai Center for Mathematical Sciences, Fudan University, Shanghai 200438, China (chduan@shu.edu.cn, chduan@fudan.edu.cn).

‡School of Mathematical Sciences, Fudan University, Shanghai 200438, China (wbchen@fudan.edu.cn).

§Department of Applied Mathematics, Illinois Institute of Technology, Chicago, IL 60616, USA (cliu124@iit.edu).

¶Corresponding author. Department of Mathematics, The University of Massachusetts, North Dartmouth, MA 02747, USA (cwang1@umassd.edu).

‖Department of Mathematics, Soochow University, Suzhou 215006, China (xyyue@suda.edu.cn).

have been derived in the existing literature [1, 26, 32, 36, 37, 39], etc. Meanwhile, various numerical methods have been studied for the PME, such as finite difference approach [17], tracking algorithm method [9], a local discontinuous Galerkin finite element method [45], Variational Particle Scheme (VPS) [44] and an adaptive moving mesh finite element method [31]. Also see other related numerical works [4, 18, 25, 28, 29], etc.

The numerical methods have also been developed for the PME by an Energetic Variational Approach (EnVarA). Instead of solving the original PME in Eulerian coordinate directly, we need two steps: obtaining a trajectory equation by an EnVarA and combining with the mass conservation law. Based on the trajectory of particles, the free boundary and the waiting time can be obtained more accurately. Meanwhile, the numerical solution can naturally keep the physical laws, such as the conservation of mass, energy dissipation and force balance. Moreover, numerical schemes could be constructed based on different dissipative energy laws. These features indicate certain advantages of the Lagrangian method over the traditional Euler method. In [13, 14], two different Lagrange numerical schemes have been derived, based on two different energy dissipation laws. It has also been proved that these two numerical schemes are uniquely solvable on an admissible convex set, and preserve the corresponding discrete energy dissipation laws. Besides a good approximation for the solution without oscillation and the free boundary, the notable advantage is that the waiting time problem could be naturally treated, which has been a well-known difficult issue for all the existing methods.

The aim of the paper is to construct a second order accurate scheme in both time and space. A modified Crank-Nicolson approximation is taken in the temporal discretization, and a finite difference is applied over a uniform spatial mesh. The resulting numerical scheme is highly nonlinear, and its solution is equivalent to a minimization of a discrete functional. In turn, its unique solvability comes from the convexity property associated with the implicit nonlinear parts, combined with the singular nature of the logarithmic terms. Subsequently, an unconditional energy stability is a direct consequence of the convexity analysis of each numerical approximation. More importantly, an optimal rate convergence analysis is provided in this work. In fact, the highly nonlinear nature of the trajectory equation makes the corresponding analysis very challenging. To overcome these subtle difficulties, we make use of a higher order expansion technique to ensure a higher order consistency estimate, which is needed to obtain a discrete $W^{1,\infty}$ bound of the numerical solution. Similar ideas have been reported in earlier literature for incompressible fluid equations [15, 16, 33, 42], non-local gradient flows [19, 21, 27], while the analysis presented in this work turns out to be more complicated, due to the lack of a linear diffusion term in the trajectory equation of the PME and the high order of the numerical scheme. In addition, we have to carry out two-step estimates to recover the nonlinear analysis:

- **Step 1.** A rough estimate for the discrete derivative of numerical solution, namely $(D_h x_h^{n+1})$ at time t_{n+1}, to control the nonlinear term;
- **Step 2.** A refined estimate for the numerical error function to obtain an optimal convergence order.

Different from a standard error estimate, the rough estimate controls the nonlinear term, which is an effective approach to handle the highly nonlinear term. As a result of the rough estimate, the refined error estimate is performed to derive the desired convergence result.

This paper is organized as follows. The trajectory equation of the PME and the numerical scheme are outlined in Section 2 and Section 3, respectively. The proof of unique solvability analysis, unconditional energy stability, and optimal rate convergence analysis are provided in Sections 4, 5 and 6, respectively. In addition, the convergence of Newton's iteration for the nonlinear numerical scheme can be found in Section 7. Finally we present a simple numerical example to demonstrate the convergence rate of the numerical scheme in Section 8. Some concluding remarks are also made in Section 9.

2. Trajectory equation of the PME

In comparison with solving the original PME directly, we split the PME into mass conservation and force balance relationship (trajectory equation in Lagrangian coordinate), which can be regarded as the gradient flow of total energy and contain all the physical information of the system. Based on the trajectory of particles, the free boundary, the finite speed and the waiting time can be computed more effectively. Moreover, the system satisfies some laws of physics, such as the conservation of mass, force balance and the dissipation of energy. In this section, we review the one-dimensional trajectory equation, derived by an Energetic Variational Approach.

Firstly, we briefly introduce the Lagrangian coordinate and the Eulerian coordinate systems.

DEFINITION 2.1. *Suppose that Ω_0^X and $\Omega_t^x \subset \mathbb{R}^m$, $m \in \mathbb{N}^+$, are domains with smooth boundaries, time $t>0$, and $\mathbf{v}$ is a smooth vector field in $\mathbb{R}^m$. The flow map $x(X,t)$: $\Omega_0^X \to \Omega_t^x$ is defined as a solution of*

$$\begin{cases} \dfrac{d}{dt}x(X,t) = \mathbf{v}(x(X,t),t), \quad t>0, \\ x(X,0)=X, \end{cases} \tag{2.1}$$

where $X=(X_1,\ldots,X_m)\in\Omega_0^X$ and $x=(x_1,\ldots,x_m)\in\Omega_t^x$. The coordinate system X is called the Lagrangian coordinate and Ω_0^X is called the reference configuration; the coordinate system x is called the Eulerian coordinate and Ω_t^x is called the deformed configuration [13, 14].

Considering Ω_0^X and Ω_t^x are the same domain described by different coordinate systems, we uniformly denote the domain by Ω in the rest of this paper.

The following initial-boundary problem is formulated:

$$\partial_t f + \partial_x(f\mathbf{v}) = 0, \ x\in\Omega\subset\mathbb{R}, \ t>0, \tag{2.2}$$
$$f\mathbf{v} = -\partial_x(f^m), \ x\in\Omega, \ m>1, \tag{2.3}$$
$$f(x,0) = f_0(x) \ge 0, \ x\in\Omega, \tag{2.4}$$
$$\partial_x f = 0, \ x\in\partial\Omega, \ t>0, \tag{2.5}$$

where f is a non-negative function, Ω is a bounded domain and $\mathbf{v}$ is the velocity. The following lemma is available; the proof has been provided in a recent work [13], here we go over it again.

LEMMA 2.1. *$f(x,t)$ is a positive solution of (2.2)-(2.5) if and only if $f(x,t)$ satisfies the corresponding energy dissipation law:*

$$\frac{d}{dt}\int_\Omega f\ln f\,dx = -\int_\Omega \frac{f}{mf^{m-1}}|\mathbf{v}|^2 dx. \tag{2.6}$$

Proof. We first prove the energy dissipation law (2.6) if f is the solution of (2.2)-(2.5). Multiplying by $(1+\ln f)$ and integrating on both sides of (2.2), we get

$$\int_\Omega (1+\ln f)\partial_t f\,dx = -\int_\Omega (1+\ln f)\partial_x(f\mathbf{v})dx.$$

Using integration by parts, in combination with (2.3), we have

$$\frac{d}{dt}\int_{\Omega} f\ln f dx = \int_{\Omega}\frac{\partial_x f}{f}(f\mathbf{v})^2 dx = -\int_{\Omega}\frac{f}{mf^{m-1}}|\mathbf{v}|^2 dx \leq 0. \tag{2.7}$$

Subsequently, we are also able to derive (2.3) from the energy dissipation law (2.6) by EnVarA, which include the mass conservation, the least action principle (LAP), the maximum dissipation law (MDL), and force balance.

- **The mass conservation.** We know that the conservation of mass means

$$\int_{E_0^X} f_0(X)dX = \int_{E_t^x} f(x,t)dx = \int_{E_0^X} f(x(X,t),t)\det\frac{\partial x}{\partial X}dX,$$

where $f_0(X)$ is the initial condition. $E_t^x \subset \Omega_t^x$ is the deformed configuration of an arbitrary subdomain $E_0^X \subset \Omega_0^X$, and $\det\frac{\partial x(X,t)}{\partial X}$ is the Jacobian matrix of the map: $X \to x(X,t)$. Thus, in the Lagrangian coordinate, mass conservation leads to

$$f(x(X,t),t) = \frac{f_0(X)}{\det\frac{\partial x(X,t)}{\partial X}}. \tag{2.8}$$

In addition, (2.2) corresponds to a conservation law in the Eulerian coordinate.

- **LAP**. With (2.8) and the total energy

$$E^{total} = \int_{\Omega} f\ln f dx, \tag{2.9}$$

the action functional in Lagrangian coordinate is

$$\mathcal{A}(x) = \int_0^{t^*}(-\mathcal{H})dt = -\int_0^{t^*}\int_{\Omega_0^X} f_0(X)\ln\left(\frac{f_0(X)}{\partial_X x}\right)dX dt,$$

where $t^* > 0$ is a given terminal time. By taking the variational of $\mathcal{A}(x)$ with respect to x, we obtain the conservation force

$$F_{con} = \frac{\delta\mathcal{A}}{\delta x} = -\partial_x f,$$

in Eulerian coordinate, and

$$F_{con} = -\partial_X\left(\frac{f_0(X)}{\partial_X x}\right),$$

in Lagrangian coordinate.

- **MDL.** Consider the entropy production

$$\Delta = \int_{\Omega}\frac{f}{mf^{m-1}}|\mathbf{v}|^2 dx.$$

By taking the variational of $\frac{1}{2}\Delta$ with respect to $\mathbf{v}$, we have the dissipation force

$$F_{dis} = \frac{\delta\frac{1}{2}\Delta}{\delta\mathbf{v}} = \frac{f}{mf^{m-1}}\mathbf{v}$$

in Eulerian coordinate and

$$F_{dis} = \frac{\delta \frac{1}{2}\Delta}{\delta x_t} = \frac{f_0(X)}{m\left(\frac{f_0(X)}{\partial_X x}\right)^{m-1}} \mathbf{x}_t$$

in Lagrangian coordinate.

- **Force balance.** By the force balance law, we obtain the trajectory equation:

$$\frac{f_0(X)}{m\left(\frac{f_0(X)}{\partial_X x}\right)^{m-1}} \partial_t x = -\partial_X \left(\frac{f_0(X)}{\partial_X x}\right), \tag{2.10}$$

in the Lagrangian coordinate, and the Darcy's law in the Eulerian coordinate

$$\frac{f}{mf^{m-1}} \mathbf{u} = -\partial_x f.$$

which is exactly (2.3).

Note that there is an assumption that the value of initial state $f_0(x)$ is non-negative in Ω to make $\int_\Omega f \ln f dx$ well-defined in (2.6).

In turn, the trajectory problem becomes

$$\frac{f_0(X)}{m\left(\frac{f_0(X)}{\partial_X x}\right)^{m-1}} \partial_t x = -\partial_X \left(\frac{f_0(X)}{\partial_X x}\right), \; X \in \Omega, \; t > 0, \tag{2.11}$$

$$x|_{\partial\Omega} = X|_{\partial\Omega}, \; t > 0, \tag{2.12}$$

$$x(X,0) = X, \; X \in \Omega. \tag{2.13}$$

Finally, with a substitution of (2.11) into (2.8), we obtain the solution $f(x,t)$ to (2.2)-(2.5). □

Throughout the rest of this article, the one-dimensional trajectory equation will be considered.

3. The proposed numerical scheme

Consider the trajectory problem for the porous medium equation

$$\frac{f_0(X)}{m\left(\frac{f_0(X)}{\partial_X x}\right)^{m-1}} \partial_t x = -\partial_X \left(\frac{f_0(X)}{\partial_X x}\right), \; X \in \Omega, \; t > 0, \tag{3.1}$$

$$x\,|_{X=0} = 0, \text{ and } x\,|_{X=1} = 1, \, t > 0, \tag{3.2}$$

$$x(X,0) = X, \; X \in \Omega. \tag{3.3}$$

Let X_0 be the left point of Ω and $h = \frac{|\Omega|}{M}$ be the spatial step, $M \in \mathbb{N}^+$. Denote by $X_r = X(r) = X_0 + rh$, where r takes on integer and half integer values. Let $\mathcal{E}_M$ and $\mathcal{C}_M$ be the spaces of functions whose domains are $\{X_i \mid i = 0,...,M\}$ and $\{X_{i-\frac{1}{2}} \mid i = 1,...,M\}$, respectively. In component form, these functions are identified via $l_i = l(X_i)$, $i = 0,...,M$, for $l \in \mathcal{E}_M$, and $\phi_{i-\frac{1}{2}} = \phi(X_{i-\frac{1}{2}})$, $i = 1,...,M$, for $\phi \in \mathcal{C}_M$.

The difference operator $D_h : \mathcal{E}_M \to \mathcal{C}_M$, $d_h : \mathcal{C}_M \to \mathcal{E}_M$, and $\widetilde{D}_h : \mathcal{E}_M \to \mathcal{E}_M$ can be defined as:

$$(D_h l)_{i-\frac{1}{2}} = (l_i - l_{i-1})/h, \; i = 1,...,M, \tag{3.4}$$

$$(d_h\phi)_i = (\phi_{i+\frac{1}{2}} - \phi_{i-\frac{1}{2}})/h, \ i=1,...,M-1, \tag{3.5}$$

$$(\widetilde{D}_h l)_i = \begin{cases} (l_{i+1}-l_{i-1})/2h, & i=1,...,M-1, \\ (4l_{i+1}-l_{i+2}-3l_i)/2h, & i=0, \\ (l_{i-2}-4l_{i-1}+3l_i)/2h, & i=M. \end{cases} \tag{3.6}$$

Let $\mathcal{Q} := \{l \in \mathcal{E}_M \mid l_{i-1} < l_i, \ 1 \le i \le M; \ l_0 = X_0, \ l_M = X_M\}$ be the admissible set, in which the particles are arranged in the order without twisting or exchanging. Its boundary set is $\partial\mathcal{Q} := \{l \in \mathcal{E}_M \mid l_{i-1} \le l_i, \ 1 \le i \le M, \ and \ l_i = l_{i-1}, \ for \ some \ 1 \le i \le M; \ l_0 = X_0, \ l_M = X_M\}$. Then $\bar{\mathcal{Q}} := \mathcal{Q} \cup \partial\mathcal{Q}$ is a closed convex set.

For two grid functions f and g over a uniform numerical mesh h, the discrete ℓ^2 inner product and the associated ℓ^2 norm are defined as

$$\langle f,g \rangle := h\Big(\frac{1}{2}(f_0 g_0 + f_M g_M) + \sum_{i=1}^{M-1} f_i g_i\Big), \quad \|f\|_2 := \sqrt{\langle f,f \rangle}. \tag{3.7}$$

The corresponding ℓ^2 inner product and ℓ^2 norm for the gradient variable could be formulated as

$$\langle D_h f, D_h g \rangle := h \sum_{i=1}^{M} (D_h f)_{i-\frac{1}{2}} (D_h g)_{i-\frac{1}{2}}, \quad \|D_h f\|_2 := \sqrt{\langle D_h f, D_h f \rangle}. \tag{3.8}$$

In addition to the discrete $\|\cdot\|_2$ norm, the discrete maximum norm is defined as $\|f\|_\infty := \max_{0\le i\le M} |f_i|$. An application of inverse inequality gives

$$\|f\|_\infty \le C h^{-\frac{1}{2}} \|f\|_2. \tag{3.9}$$

Notice that this inverse inequality is valid for the spatially discrete function, while it is incorrect for a continuous function.

We propose the second order numerical scheme as follows, based on a modified Crank-Nicolson approach. Given the positive initial state $f_0(X) \in \mathcal{E}_M$ and the particle position $x^n, x^{n-1} \in \mathcal{Q}$, find $x^{n+1} = (x_0^{n+1}, ..., x_M^{n+1}) \in \mathcal{Q}$ such that

$$\begin{aligned} &\frac{f_0(X_i)}{m\big(\frac{f_0(X)}{S_h(x^n,x^{n-1})}\big)_i^{m-1}} \cdot \frac{x_i^{n+1}-x_i^n}{\Delta t} \\ &= -d_h\Big[\Big(f_0(X)\frac{\ln(D_h x^{n+1}) - \ln(D_h x^n)}{D_h x^{n+1} - D_h x^n}\Big) \\ &\quad -A_0 \Delta t D_h(x^{n+1}-x^n) + \Delta t^2 \big(\frac{1}{D_h x^{n+1}} - \frac{1}{D_h x^n}\big)\Big]_i, \\ &\text{with} \quad S_h(x^n,x^{n-1}) = \max(\widetilde{D}_h(\frac{3}{2}x^n - \frac{1}{2}x^{n-1}), \Delta t^2), \end{aligned} \tag{3.10}$$

for $1 \le i \le M-1$, and we take $x_0^{n+1} = 0$ and $x_M^{n+1} = 1$, $n = 0, \cdots, N-1$. Notice that A_0 is a second order artificial regularization parameter.

To solve the nonlinear Equation (3.10), we use Newton's iteration.

Newton's iteration. Set $x^{n+1,0} = x^n$. For $k = 0,1,2,\cdots$, update $x^{n+1,k+1} = x^{n+1,k} + \delta_x$, which is the solution of

$$\frac{f_0(X_i)}{m\big(\frac{f_0(X)}{S_h(x^n,x^{n-1})}\big)_i^{m-1}} \cdot \frac{\delta_{x_i}}{\Delta t} - d_h\Big[\Big(f_0(X)\cdot W^{n+1,k} + A_0\Delta t + \frac{\Delta t^2}{(D_h x^{n+1,k})^2}\Big) D_h \delta_x\Big]_i$$

$$= -\frac{f_0(X_i)}{m\big(\frac{f_0(X)}{S_h(x^n,x^{n-1})}\big)_i^{m-1}}\frac{x_i^{n+1,k}-x_i^n}{\Delta t} - d_h\Big[f_0(X)R^{n+1,k}$$

$$-A_0\Delta t D_h(x^{n+1,k}-x^n)+\Delta t^2\big(\frac{1}{D_h x^{n+1,k}}-\frac{1}{D_h x^n}\big)\Big]_i, \quad 1\le i\le M-1, \tag{3.11}$$

with $\delta_{x_0}=\delta_{x_M}=0$,

where, for $i=1,\cdots,M$,

$$W_{i-\frac12}^{n+1,k} = \left[\frac{(1-\frac{D_h x^n}{D_h x^{n+1,k}})+\ln(\frac{D_h x^n}{D_h x^{n+1,k}})}{(D_h x^{n+1,k}-D_h x^n)^2}\right]_{i-\frac12},$$

and

$$R_{i-\frac12}^{n+1,k} = \Big[\frac{\ln(D_h x^{n+1,k})-\ln(D_h x^n)}{D_h x^{n+1,k}-D_h x^n}\Big]_{i-\frac12}.$$

Let $w_i^{n+1,k}=|D_h x_{i-\frac12}^{n+1,k}-D_h x_{i-\frac12}^n|$, $1\le i\le M$. Then we have

$$\lim_{w_i^{n+1,k}\to 0} W_{i-\frac12}^{n+1,k} = -\frac{1}{2(D_h x^{n+1,k})^2},$$

and

$$\lim_{w_i^{n+1,k}\to 0} R_{i-\frac12}^{n+1,k} = \frac{1}{D_h x^{n+1,k}}.$$

Then we obtain the numerical solution $f(x^n,t^n):=f_i^n$ by

$$f_i^n = \frac{f_0(X)}{\widetilde{D}_h x_i^n}, \quad 0\le i\le M, \tag{3.12}$$

which is the discrete scheme of (2.8).

REMARK 3.1. In the numerical scheme (3.10), the numerical discretization to the nonlinear chemical potential term, namely $\frac{\ln(D_h x^{n+1})-\ln(D_h x^n)}{D_h x^{n+1}-D_h x^n}$, turns out to be a second order approximation to $\frac{1}{\partial_X x}$ at the middle time instant $t^{n+1/2}$, due to the fact that $\ln y$ is the primitive function of $\frac{1}{y}$ (in terms of $y=\partial_X x>0$). Because of this subtle fact, such a second order numerical discretization is labeled as a modified Crank-Nicolson approximation. In particular, this approximation makes its inner product with $D_h(x^{n+1}-x^n)$ exactly the difference of the nonlinear energy functionals between two consecutive time steps, as will be observed in the later sections. Also see the related works [6, 7, 8, 10, 11, 12, 20, 22, 23, 24, 35] for various gradient flows, in particular for the modified Crank-Nicolson approximation to the nonlinear energy potential, which takes a form of $\frac{F(\phi^{n+1})-F(\phi^n)}{\phi^{n+1}-\phi^n}$ (with ϕ being the phase variable). In comparison with the standard Crank-Nicolson formula, such a modified Crank-Nicolson approximation greatly facilitates the energy stability analysis, while the standard formula would face serious difficulty in the theoretical justification of this property.

REMARK 3.2. The term $S_h(x^n,x^{n-1})$, given by $S_h(x^n,x^{n-1})=\max(\widetilde{D}_h(\frac32 x^n-\frac12 x^{n-1}),\Delta t^2)$ in (3.10), stands for a second order approximation to $\partial_X x$ at the middle time instant $t^{n+1/2}$ in the nonlinear mobility part on the left-hand side. The mobility function has to be explicitly updated in the numerical approximation, to ensure the convexity of the temporal differentiation term and the unique solvability of the numerical scheme. Therefore, we have to use a second order explicit extrapolation formula in such a numerical approximation, which in turn gives the extrapolation weight coefficients as $\frac{3}{2}$ and $-\frac{1}{2}$ at time steps t^n, t^{n-1}, respectively. In addition, a maximum value of Δt^2 is taken, to ensure the point-wise positivity of such a mobility function. On the other hand, if the exact PDE solution preserves a separation property, i.e, $\partial_X x \geq \epsilon_0$ for a fixed $\epsilon_0 > 0$, we see that $\frac{3}{2}\widetilde{D}_h x^n - \frac{1}{2}\widetilde{D}_h x^{n-1}$ will always be greater than another fixed constant, if the numerical solution is a sufficiently accurate approximation to the exact PME solution. In other words, if the exact PME solution preserves a phase separation property and the numerical solution is sufficiently accurate, the value of $S_h(x^n, x^{n-1})$ will always be given by $\frac{3}{2}\widetilde{D}_h x^n - \frac{1}{2}\widetilde{D}_h x^{n-1}$, a second order approximation to $\partial_X x$ at the middle time instant $t^{n+1/2}$.

REMARK 3.3. A similar energy dissipative numerical method has been reported in [18], based on Eulerian coordinate. The authors replace the original problem with a perturbed problem by a positivity-preserving perturbation term. In comparison with this method in Eulerian coordinate, the trajectory equation in the Lagrangian coordinate can capture the trajectory of particles along the direction of the maximum energy dissipation. Therefore, the numerical scheme in this paper can naturally satisfy the original discrete energy dissipation law, preserve the positivity of the solution and the force balance law. In particular, the degenerate feature, such as the free boundary and the waiting time, can be solved more effectively. In addition, another advantage of the Lagrange method could be observed from the fact that, numerical schemes can be established [13] based on different dissipation laws.

In addition, we consider the trajectory equation and numerical scheme for the free boundaries. If the initial data is given with a compact support in Ω, the left and right interfaces can be defined as

$$\xi_1^t := \inf\{x \in \Omega : u(x,t) > 0, t \geq 0\}, \tag{3.13}$$

$$\xi_2^t := \sup\{x \in \Omega : u(x,t) > 0, t \geq 0\}. \tag{3.14}$$

We denote $\Gamma^t := [\xi_1^t, \xi_2^t] \subset \Omega$. In this case, we shall solve the initial-boundary value problem (2.11) with the boundary condition:

$$\partial_t x = -\frac{m}{m-1}\frac{\partial_X[f_0(X)^{m-1}]}{(\partial_X x)^m}, \; X = \xi_1^t, \xi_2^t, \tag{3.15}$$

and the initial condition

$$x(X,0) = X, \; X \in \Gamma^0. \tag{3.16}$$

The Equation (3.15) is derived from the trajectory Equation (2.11) by the fact that the initial data $f_0(X) = 0$, $X \in \partial\Omega$.

Based on the Crank-Nicolson temporal discretization, the full numerical scheme becomes

$$\frac{x^{n+1} - x^n}{\Delta t} = -\frac{m}{m-1}\frac{D_h[f_0(X)^{m-1}]}{\left(\frac{\widetilde{D}_h x_i^{n+1} + \widetilde{D}_h x_i^n}{2}\right)^m}, \; i = 1, M. \tag{3.17}$$

Again, the Newton's iteration is applied in the numerical implementation.

4. Unique solvability analysis

THEOREM 4.1. *Given any x^n, x^{n-1}, with $D_h x^n > 0$, $D_h x^{n-1} > 0$ at a point-wise level. In more details, we denote $Q^{(k),1} = \min_{i=1,\dots,M}(D_h x^k)_{i-1/2}$ and $Q^{(k),2} = \max_{i=1,\dots,M}(D_h x^k)_{i-1/2}$, $k = n, n-1$, so that the following inequality is valid:*

$$0 < Q^{(k),1} \le D_h x^k \le Q^{(k),2}, \quad k = n, n-1. \tag{4.1}$$

The proposed numerical scheme (3.10) is uniquely solvable, with $D_h x^{n+1} > 0$ at a point-wise level.

Proof. With an introduction $\hat{x} = x^{n+1} - X$, it is clear that (3.10) could be rewritten as

$$\frac{f_0(X_i)}{m\big(\frac{f_0(X)}{S_h(x^n,x^{n-1})}\big)_i^{m-1}} \cdot \frac{X_i + \hat{x}_i - x_i^n}{\Delta t} = -d_h\Big[\Big(f_0(X)\frac{\ln(1+D_h\hat{x}) - \ln(D_h x^n)}{1 + D_h\hat{x} - D_h x^n}\Big) - A_0 \Delta t D_h(\hat{x} - x^n) + \Delta t^2\big(\frac{1}{1+D_h\hat{x}} - \frac{1}{D_h x^n}\big)\Big]_i. \tag{4.2}$$

Because of the fact that $\hat{x}_0 = \hat{x}_M = 0$, we see that the solution of (4.2) is equivalent to a minimization of the following discrete functional:

$$F(\hat{x}) = \sum_{j=1}^{4} F_j(\hat{x}), \quad \text{with } F_1(\hat{x}) = \frac{1}{2\Delta t}\Big\langle \frac{f_0(X)}{m\big(\frac{f_0(X)}{S_h(x^n,x^{n-1})}\big)^{m-1}}, (X + \hat{x} - x^n)^2\Big\rangle, \tag{4.3}$$

$$F_2(\hat{x}) = \Big\langle f_0(X) G(D_h\hat{x}, D_h x^n), \mathbf{1}\Big\rangle, \tag{4.4}$$

$$F_3(\hat{x}) = A_0 \Delta t\Big(\frac{1}{2}\|D_h\hat{x}\|_2^2 - \langle D_h\hat{x}, D_h x^n\rangle\Big), \tag{4.5}$$

$$F_4(\hat{x}) = \Delta t^2\Big(-\langle \ln(1 + D_h\hat{x}), \mathbf{1}\rangle + \langle D_h\hat{x}, \frac{1}{D_h x^n}\rangle\Big), \tag{4.6}$$

in which $G(x, x_0)$ is given by the primitive function of $-\frac{\ln(1+x) - \ln x_0}{1 + x - x_0}$, for a fixed x_0:

$$G(x, x_0) = \int_x^0 \frac{\ln(1+t) - \ln x_0}{1 + t - x_0}\, dt, \quad \text{for } x \ge -1. \tag{4.7}$$

The convexity of F_1, F_3 and F_4 (in terms of $\hat{x}$) is obvious. For the functional F_2, we have the following observation, for $x > -1$:

$$G''(x, x_0) = \Big(-\frac{\ln(1+x) - \ln x_0}{1 + x - x_0}\Big)'_x = \frac{-\frac{1}{1+x}(1 + x - x_0) + (\ln(1+x) - \ln x_0)}{(1 + x - x_0)^2} \ge 0, \tag{4.8}$$

in which the convexity of $-\ln(1+x)$ has been used. This fact implies the convexity of F_2. Therefore, we conclude that F is convex in terms of $\hat{x}$, provided that $D_h\hat{x} > -1$ at a point-wise level. Furthermore, F is strictly convex, because of the strict convexity of F_1.

In the next step, we wish to prove that there exists a minimizer of F at an interior point of $\mathcal{Q}$. To this end, consider the following closed domain: for a given $\delta > 0$,

$$\mathcal{Q}_\delta := \Big\{X + \hat{x} \in \mathcal{Q} \,:\, 1 + (D_h\hat{x})_{i+1/2} \ge \delta, \forall 0 \le i \le M-1\Big\} \subset \mathcal{Q}. \tag{4.9}$$

Since $\mathcal{Q}_\delta$ is a compact and convex set in $\mathbb{R}^{M-1}$, there exists a (not necessarily unique) minimizer of F over $\mathcal{Q}_\delta$. The key point of our positivity analysis is that such a minimizer could not occur on the boundary of $\mathcal{Q}_\delta$, if δ is small enough.

Assume a minimizer of F over $\mathcal{Q}_\delta$, denoted by $\hat{x}^\star$, occurs at a boundary point. There is at least one grid point such that $1+(D_h\hat{x}^*)_{i_0+1/2}=\delta$. Next we estimate the value of $F(\hat{x}^*)$. For the F_1 part, the following bound is available, for any $X+\hat{x}\in\mathcal{Q}$:

$$
\begin{aligned}
0\le F_1(\hat{x}) &= \frac{1}{2\Delta t}\Big\langle \frac{f_0(X)}{m\big(\frac{f_0(X)}{S_h(x^n,x^{n-1})}\big)^{m-1}},(X+\hat{x}-x^n)^2\Big\rangle \\
&\le \frac{1}{2m\Delta t}\Big(\tilde{C}_1^{|m-2|}\big(\frac{3}{2}Q^{(n),2}+\frac{1}{2}Q^{(n-1),2}\big)^{m-1}\Big) := A^{(1)}, \qquad (4.10)\\
\text{with}\quad \tilde{C}_1 &= \max\Big(\max_\Omega f_0(X),\frac{1}{\min_\Omega f_0(X)}\Big),
\end{aligned}
$$

in which the assumption (4.1) has been recalled, and we have made use of the following fact:

$$
0\le X+\hat{x}\le 1,\ \text{so that}\ -1\le X+\hat{x}-x^n\le 1,\quad \text{at a point-wise level.} \qquad (4.11)
$$

For the F_2 part, we observe that $G(x,x_0)\ge 0$ for $-1\le x\le 0$, and

$$
G(x,x_0)=-\int_0^x \frac{\ln(1+t)-\ln x_0}{1+t-x_0}\,dt \ge -\int_0^x \frac{1}{1+t}\,dt=-\ln(1+x),\quad \text{for } x\ge 0, \qquad (4.12)
$$

in which the convexity of $\ln(1+t)$ has been applied. Meanwhile, by the fact that $X+\hat{x}\in\mathcal{Q}$, we have the following observation:

$$
0<1+(D_h\hat{x})_{i+1/2}\le \frac{1}{h},\ \forall 0\le i\le M-1,\quad \text{since } 0\le x_i\le 1,\ 0\le x_{i+1}\le 1. \qquad (4.13)
$$

In turn, its substitution into (4.12) implies that

$$
G(D_h\hat{x},D_hx^n)\ge -\ln\frac{1}{h}=\ln h,\quad \text{at any grid point.} \qquad (4.14)
$$

As a consequence, we obtain a lower bound for F_2:

$$
F_2(\hat{x}^*)=\Big\langle f_0(X)G(D_h\hat{x}^*,D_hx^n),\mathbf{1}\Big\rangle \ge \|f_0(X)\|_\infty\cdot\ln h, \qquad (4.15)
$$

The derivation for a lower bound of F_3 is straightforward:

$$
F_3(\hat{x}^*)\ge -A_0\Delta t\langle D_h\hat{x},D_hx^n\rangle \ge -\frac{A_0\Delta t}{h^2}, \qquad (4.16)
$$

in which the inequality (4.13) has been applied. For the functional F_4, we see that the second part has the following lower bound:

$$
\Delta t^2\langle D_h\hat{x},\frac{1}{D_hx^n}\rangle \ge -\Delta t^2\cdot\frac{1}{h}\cdot\frac{1}{Q^{(n),1}}=-\frac{\Delta t^2}{Q^{(n),1}h}, \qquad (4.17)
$$

in which the inequality (4.13) and the assumption (4.1) have been used. For the first part of F_4, we recall that $1+(D_h\hat{x}^*)_{i_0+1/2}=\delta$, and the following estimate is available:

$$
\begin{aligned}
&-\langle \ln(1+D_h\hat{x}^*,\mathbf{1}\rangle = -h\Big(\ln(1+(D_h\hat{x}^*)_{i_0+1/2})+\sum_{i\ne i_0}\ln(1+(D_h\hat{x}^*)_{i+1/2})\Big)\\
&\ge -h\Big(\ln\delta+(M-1)\ln\frac{1}{h}\Big)=h\Big(\ln\frac{1}{\delta}+(M-1)\ln h\Big)\ge h\ln\frac{1}{\delta}+\ln h, \qquad (4.18)
\end{aligned}
$$

in which the inequality (4.13) has been applied in the second step, and we have used the fact that $h\cdot M = 1$ in the last step. In turn, we get a lower bound for $F_4(\hat{x}^*)$:

$$F_4(\hat{x}^*) = \Delta t^2\Big(-\langle \ln(1+D_h\hat{x},\mathbf{1}\rangle + \langle D_h\hat{x}, \frac{1}{D_hx^n}\rangle\Big) \geq \Delta t^2 h \ln\frac{1}{\delta} - \frac{\Delta t^2}{Q^{(n),1}h} + \Delta t^2 \ln h. \tag{4.19}$$

Therefore, a combination of (4.10), (4.15), (4.16) and (4.19) yields a lower bound for $F(\hat{x}^*)$:

$$F(\hat{x}^*) \geq \Delta t^2 h \ln\frac{1}{\delta} - A_{\Delta t,h}, \quad \text{with } A_{\Delta t,h} = \frac{A_0\Delta t}{h^2} + \frac{\Delta t^2}{Q^{(n),1}h} - (\Delta t^2 + \|f_0(X)\|_\infty)\ln h. \tag{4.20}$$

Meanwhile, we observe that, by taking $\hat{x}^0 = \mathbf{0}$, so that $X + \hat{x}^0 \in \mathcal{Q}_\delta$, the following estimates are available:

$$0 \leq F_1(\hat{x}^0) \leq A^{(1)}, \text{ (by (4.10))}, \quad F_2(\hat{x}^0) = 0,\ F_3(\hat{x}^0) = 0,\ F_4(\hat{x}^0) = 0, \tag{4.21}$$

so that

$$0 \leq F(\hat{x}^0) \leq A^{(1)}. \tag{4.22}$$

We also notice that both $A_{\Delta t,h}$ and $A^{(1)}$ are independent of δ. Consequently, by taking $\delta > 0$ sufficiently small so that

$$\Delta t^2 h \ln\frac{1}{\delta} - A_{\Delta t,h} > A^{(1)}, \quad \text{i.e. } 0 < \delta < \exp\Big(-\frac{A_{\Delta t,h} + A^{(1)}}{\Delta t^2 h}\Big). \tag{4.23}$$

This yields a contradiction that F takes a global minimum at $\hat{x}^\star$ over $\mathcal{Q}_\delta$, because $F(\hat{x}^\star) > F(\hat{x}^0)$. As a result, the global minimum of F over $\mathcal{Q}_\delta$ could only possibly occur at an interior point, with δ satisfying (4.23). We conclude that there must be a solution $\hat{x} \in (\mathcal{Q}_\delta)^{\mathrm{o}}$, the interior region of $\mathcal{Q}_\delta$, so that for all $\psi \in \mathcal{C}_{\mathrm{per}}$,

$$0 = d_s F(\hat{x} + s\psi)|_{s=0}, \tag{4.24}$$

which is equivalent to the numerical solution of (4.2), provided that δ satisfies (4.23). The existence of a numerical solution of (3.10), with "positive" gradient, is established. In addition, since F is a strictly convex function over $\mathcal{Q}$, the uniqueness analysis for this numerical solution is straightforward. ∎

5. Unconditional energy stability

With the positivity-preserving and unique solvability properties for the numerical scheme (3.10) established, we now prove energy stability. For any grid function x with $D_hx > 0$ at a point-wise level, the following discrete energy functional is introduced:

$$E_h(x) := \langle f_0(X)\ln(D_hx), \mathbf{1}\rangle. \tag{5.1}$$

In fact, such a discrete energy functional is a second order numerical approximation to the continuous version of the free energy, namely $\int_\Omega f_0(X)\ln(\partial_X x)\,dX$, associated with the trajectory Equation (2.10) of the PME. Notice that this energy is only involved with $\partial_X x$ (and D_hx in the finite difference approximation), not with x itself.

THEOREM 5.1. *The proposed numerical scheme (3.10) is unconditionally energy stable: $E_h(x^{n+1}) \leq E_h(x^n)$.*

Proof. Taking a discrete inner product with (3.10) by $x^{n+1}-x^n$, making use of the summation by parts formula (because of the boundary condition $(x^{n+1}-x^n)_0=(x^{n+1}-x^n)_M=0$), we get

$$\begin{aligned}&\frac{1}{\Delta t}\Big\langle \frac{(S_h(x^n,x^{n-1}))^{m-1}}{m(f_0(X))^{m-2}},(x^{n+1}-x^n)^2\Big\rangle\\ &\quad-\Big\langle f_0(X)\frac{\ln(D_hx^{n+1})-\ln(D_hx^n)}{D_hx^{n+1}-D_hx^n},D_hx^{n+1}-D_hx^n\Big\rangle\\ &\quad+A_0\Delta t\|D_h(x^{n+1}-x^n)\|_2^2-\Delta t^2\Big\langle\frac{1}{D_hx^{n+1}}-\frac{1}{D_hx^n},D_hx^{n+1}-D_hx^n\Big\rangle=0. \qquad (5.2)\end{aligned}$$

The first term on the left-hand side turns out to be non-negative, since $S_h(x^n,x^{n-1})>0$, $f_0(X)>0$ at a point-wise level:

$$\Big\langle \frac{(S(x^n,x^{n-1}))^{m-1}}{m(f_0(X))^{m-2}},(x^{n+1}-x^n)^2\Big\rangle\geq 0. \qquad (5.3)$$

The second term exactly gives the difference between the discrete energy values at time steps t^{n+1} and t^n:

$$\begin{aligned}&-\Big\langle f_0(X)\frac{\ln(D_hx^{n+1})-\ln(D_hx^n)}{D_hx^{n+1}-D_hx^n},D_hx^{n+1}-D_hx^n\Big\rangle\\ &=-\Big\langle f_0(X)\ln(D_hx^{n+1}),\mathbf{1}\Big\rangle+\Big\langle f_0(X)\ln(D_hx^n),\mathbf{1}\Big\rangle=E_h(x^{n+1})-E_h(x^n). \qquad (5.4)\end{aligned}$$

The third term is clearly non-negative, and the last term turns out to be non-negative as well:

$$-\Big\langle\frac{1}{D_hx^{n+1}}-\frac{1}{D_hx^n},D_hx^{n+1}-D_hx^n\Big\rangle=\Big\langle\frac{1}{D_hx^{n+1}D_hx^n},(D_hx^{n+1}-D_hx^n)^2\Big\rangle\geq 0, \qquad (5.5)$$

in which we have made use of the unique solvability result, $D_hx^{n+1}>0$, $D_hx^n>0$, at the point-wise level, as given by Theorem 4.1. As a consequence, a substitution of (5.3)-(5.5) into (5.2) reveals an unconditional energy stability of the numerical scheme:

$$E_h(x^{n+1})-E_h(x^n)\leq -A_0\Delta t\|D_h(x^{n+1}-x^n)\|_2^2\leq 0, \text{so that } E_h(x^{n+1})\leq E_h(x^n). \qquad (5.6)$$

This completes the proof of Theorem 5.1. □

6. Optimal rate convergence analysis

Now we proceed into the convergence analysis. Let x_e be the exact solution for the PME Equation (3.1)-(3.3). With sufficiently regular initial data, we could assume that the exact solution has regularity of class $\mathcal{R}$:

$$\begin{aligned}&x_e\in\mathcal{R}:=H^6(0,T;C(\Omega))\cap H^4\left(0,T;C^2(\Omega)\right)\cap L^\infty\left(0,T;C^6\Omega)\right), \quad \text{with}\\ &\|f\|_{H^k(0,T;C^m(\Omega))}:=\Big\|\|f(\cdot,t)\|_{C^m(\Omega)}\Big\|_{H^k(0,T)}, \ \|f\|_{L^\infty(0,T;C^m(\Omega))}:=\Big\|\|f(\cdot,t)\|_{C^m(\Omega)}\Big\|_{L^\infty(0,T)},\\ &\|g\|_{C^m(\Omega)}:=\sum_{\ell=0}^m\|\partial_x^\ell g\|_{C(\Omega)}, \quad \|\partial_x^\ell g\|_{C(\Omega)}:=\max_\Omega|\partial_x^\ell g|. \qquad (6.1)\end{aligned}$$

In addition, we assume that the following separation property is valid for the exact solution, in terms of its gradient:

$$\partial_X x_e\geq\epsilon_0, \quad \text{for } \epsilon_0>0, \qquad (6.2)$$

at a point-wise level. The following theorem is the convergence result of the proposed scheme.

THEOREM 6.1. *Given initial data $x_e(\cdot,t=0)\in C^6(\Omega)$, suppose the exact solution for the PME Equation (3.1)-(3.3) is of regularity class $\mathcal{R}$. Define the numerical error function as $e_j^n=(x_e)_j^n-x_j^n$, at a point-wise level. Then, provided Δt and h are sufficiently small, and under the linear refinement requirement $C_1h\le\Delta t\le C_2h$, we have*

$$\|e^n\|_2+\Big(\Delta t\sum_{m=0}^{n-1}\|\frac{1}{2}D_h(e^m+e^{m+1})\|_2^2\Big)^{1/2}\le C(\Delta t^2+h^2), \tag{6.3}$$

for all positive integers n, such that $t_n=n\Delta t\le T$, where $C>0$ is independent of n, Δt, and h.

6.1. Higher order consistency analysis of (3.10): asymptotic expansion of the numerical solution. By consistency, the exact solution x_e solves the discrete Equation (3.10) with second order accuracy in both time and space. Meanwhile, it is observed that this leading local truncation error will not be enough to recover an a-priori $W_h^{1,\infty}$ bound for the temporal derivative of the numerical solution, which is needed in the nonlinear error estimate. To remedy this, we use a higher order consistency analysis, via a perturbation argument, to recover such a bound in later analysis. In more details, we need to construct supplementary fields, $x_{h,1}$, $x_{\Delta t,1}$, $x_{\Delta t,2}$, and W, satisfying

$$W=x_e+h^2x_{h,1}+\Delta t^2x_{\Delta t,1}+\Delta t^3x_{\Delta t,2}, \tag{6.4}$$

so that a higher $O(\Delta t^4+h^4)$ consistency is satisfied with the given numerical scheme (3.10). The constructed fields $x_{h,1}$, $x_{\Delta t,1}$, $x_{\Delta t,2}$, which will be found using a perturbation expansion, will depend solely on the exact solution x_e.

In other words, we introduce a higher order approximate expansion of the exact solution, since a leading order consistency estimate gives a second order accuracy in both time and space, which is not able to control the discrete $W_h^{1,\infty}$ norm of the numerical solution. Instead of substituting the exact solution into the numerical scheme, a careful construction of an approximate profile is performed by adding $O(\Delta t^2)$, $O(\Delta t^3)$ and $O(h^2)$ correction terms to the exact solution to satisfy an $O(\Delta t^4+h^4)$ truncation error. In turn, we estimate the numerical error function between the constructed profile and the numerical solution, instead of a direct comparison between the numerical solution and exact solution. Such a higher order consistency enables us to derive a higher order convergence estimate in the $\|\cdot\|_2$ norm, which in turn leads to a desired $W_h^{1,\infty}$ bound of the numerical solution, via an application of inverse inequality. This approach has been reported for a wide class of nonlinear PDEs; see the related works for the incompressible fluid equation [15, 16, 33, 34, 40, 41, 42], various gradient equations [3, 19, 21, 27], the porous medium equation based on the energetic variational approach [14], the Poisson-Nernst-Planck system [28], the nonlinear wave equation [43], etc.

The following truncation error analysis for the temporal discretization can be obtained by using a straightforward Taylor expansion

$$\begin{aligned}&\frac{(S_e(x_e^n,x_e^{n-1}))^{m-1}}{m(f_0(X))^{m-2}}\cdot\frac{x_e^{n+1}-x_e^n}{\Delta t}\\&=-\partial_X\Big[\Big(f_0(X)\frac{\ln(\partial_Xx_e^{n+1})-\ln(\partial_Xx_e^n)}{\partial_X(x_e^{n+1}-x_e^n)}\Big)-A_0\Delta t\partial_X(x_e^{n+1}-x_e^n)+\Delta t^2(\frac{1}{\partial_Xx_e^{n+1}}-\frac{1}{\partial_Xx_e^n})\Big]\\&\quad+\Delta t^2(\boldsymbol{g}_1^{(0)})^{n+1/2}+\Delta t^3(\boldsymbol{g}_1^{(1)})^{n+1/2}+O(\Delta t^4),\end{aligned} \tag{6.5}$$

$$\text{with}\quad S_e(x_e^n,x_e^{n-1})=\partial_X(\frac{3}{2}x_e^n-\frac{1}{2}x_e^{n-1}).$$

Here the functions $\boldsymbol{g}_1^{(j)}$ $(j=0,\ 1)$ are smooth enough in the sense that their derivatives, both the temporal and spatial ones, are bounded in the $\|\cdot\|_{L^\infty}$ norm. In more details, $\boldsymbol{g}_1^{(j)}$ turns out to be only dependent on the higher order derivatives (both spatial and temporal) of the exact solution x_e, as indicated by the Taylor expansion. In fact, both $\boldsymbol{g}_1^{(0)}$ and $\boldsymbol{g}_1^{(j)}$ are space-time functions, i.e., they are time and space dependent, due to the involved higher order spatial and temporal derivatives of the exact solution. Meanwhile, for a fixed time instant $t^{n+1/2}$, $(\boldsymbol{g}_1^{(0)})^{n+1/2}$ and $(\boldsymbol{g}_1^{(1)})^{n+1/2}$ become spatial functions.

The temporal correction function $x_{\Delta t,1}$ is given by solving the following equation:

$$\begin{aligned}&\frac{(\partial_X x_e)^{m-1}}{m(f_0(X))^{m-2}}\partial_t x_{\Delta t,1}+\frac{(m-1)(\partial_X x_e)^{m-2}\partial_X x_{\Delta t,1}}{m(f_0(X))^{m-2}}\partial_t x_e\\&=-\partial_X\Big(-f_0(X)\frac{1}{(\partial_X x_e)^2}\partial_X x_{\Delta t,1}\Big)-(\boldsymbol{g}_1^{(0)}),\qquad (6.6)\\&x_{\Delta t,1}(0)=x_{\Delta t,1}(1)=0,\quad x_{\Delta t,1}(t=0)=0.\qquad (6.7)\end{aligned}$$

In fact, (6.6) is a linear parabolic PDE, with sufficiently regular coefficient functions. The existence and uniqueness of its solution could be derived by making use of a standard Galerkin procedure and Sobolev estimates, following the classical techniques for time-dependent parabolic equation [38]. Such a solution depends solely on the exact profile x_e and is regular enough. In addition, the derivatives of $x_{\Delta t,1}$ of various orders are bounded. Of course, an application of the semi-implicit discretization to (6.6)-(6.7) implies that

$$\begin{aligned}&\frac{(S_e(x_e^n,x_e^{n-1}))^{m-1}}{m(f_0(X))^{m-2}}\cdot\frac{x_{\Delta t,1}^{n+1}-x_{\Delta t,1}^n}{\Delta t}+\frac{(m-1)(S_e(x_e^n,x_e^{n-1}))^{m-2}S_e(x_{\Delta t,1}^n,x_{\Delta t,1}^{n-1})}{m(f_0(X))^{m-2}}\cdot\frac{x_e^{n+1}-x_e^n}{\Delta t}\\&=-\partial_X\Big(-f_0(X)\frac{1}{(\partial_X(\frac12 x_e^{n+1}+x_e^n))^2}\cdot\frac12\partial_X(x_{\Delta t,1}^{n+1}+x_{\Delta t,1}^n)\Big)-(\boldsymbol{g}_1^{(0)})^{n+1/2}+O(\Delta t^2),\qquad (6.8)\\&\text{with}\quad S_e(x_{\Delta t,1}^n,x_{\Delta t,1}^{n-1})=\partial_X(\frac32 x_{\Delta t,1}^n-\frac12 x_{\Delta t,1}^{n-1}).\qquad (6.9)\end{aligned}$$

Therefore, a combination of (6.5) and (6.8) leads to the third order temporal truncation error for $W_1=x_e+\Delta t^2 x_{\Delta t,1}$:

$$\begin{aligned}&\frac{(S_e(W_1^n,W_1^{n-1}))^{m-1}}{m(f_0(X))^{m-2}}\cdot\frac{W_1^{n+1}-W_1^n}{\Delta t}\\&=-\partial_X\Big[\Big(f_0(X)\frac{\ln(\partial_X W_1^{n+1})-\ln(\partial_X W_1^n)}{\partial_X(W_1^{n+1}-W_1^n)}\Big)\\&\quad-A_0\Delta t\partial_X(W_1^{n+1}-W_1^n)+\Delta t^2(\frac{1}{\partial_X W_1^{n+1}}-\frac{1}{\partial_X W_1^n})\Big]\\&\quad+\Delta t^3(\boldsymbol{g}_1^{(1)})^{n+1/2}+O(\Delta t^4),\qquad (6.10)\\&\quad\text{with}\quad S_e(W_1^n,W_1^{n-1})=\partial_X(\frac32 W_1^n-\frac12 W_1^{n-1}).\end{aligned}$$

In the derivation of (6.10), the following linearized expansions have been utilized:

$$\frac{W_1^{n+1}-W_1^n}{\Delta t}=\frac{x_e^{n+1}-x_e^n}{\Delta t}+O(\Delta t^2),\qquad (6.11)$$

$$\begin{aligned}&(S_e(x_e^n,x_e^{n-1}))^{m-1}+(m-1)(S_e(x_e^n,x_e^{n-1}))^{m-2}S_e(x_{\Delta t,1}^n,x_{\Delta t,1}^{n-1})\cdot\Delta t^2 x_{\Delta t,1}\\&=S_e(W_1^n,W_1^{n-1})+O(\Delta t^4),\qquad (6.12)\end{aligned}$$

$$\begin{aligned}&\frac{\ln(\partial_X x_e^{n+1})-\ln(\partial_X x_e^n)}{\partial_X(x_e^{n+1}-x_e^n)}-\Delta t^2\frac{1}{(\partial_X(\frac12 x_e^{n+1}+x_e^n))^2}\cdot\frac12\partial_X(x_{\Delta t,1}^{n+1}+x_{\Delta t,1}^n)\\&=\frac{\ln(\partial_X W_1^{n+1})-\ln(\partial_X W_1^n)}{\partial_X(W_1^{n+1}-W_1^n)}+O(\Delta t^4).\qquad (6.13)\end{aligned}$$

Similarly, the temporal correction function $x_{\Delta t,2}$ is given by solving the following equation:

$$\frac{(\partial_X x_e)^{m-1}}{m(f_0(X))^{m-2}}\partial_t x_{\Delta t,2}+\frac{(m-1)(\partial_X x_e)^{m-2}\partial_X x_{\Delta t,2}}{m(f_0(X))^{m-2}}\partial_t x_e$$

$$= -\partial_X \Big(-f_0(X) \frac{1}{(\partial_X x_e)^2} \partial_X x_{\Delta t,2} \Big) - \boldsymbol{g}_1^{(1)}, \tag{6.14}$$

$$x_{\Delta t,2}(0) = x_{\Delta t,2}(1) = 0, \quad x_{\Delta t,2}(t=0) = 0. \tag{6.15}$$

Similarly, (6.14) is a linear parabolic PDE, with sufficiently regular coefficient functions, and its unique solution depends solely on the exact profile x_e and is smooth enough, with derivatives of various orders staying bounded. In turn, an application of the semi-implicit discretization to (6.14)-(6.15) implies that

$$\frac{(S_e(x_e^n, x_e^{n-1}))^{m-1}}{m(f_0(X))^{m-2}} \cdot \frac{x_{\Delta t,2}^{n+1} - x_{\Delta t,2}^n}{\Delta t} + \frac{(m-1)(S_e(x_e^n, x_e^{n-1}))^{m-2} S_e(x_{\Delta t,2}^n, x_{\Delta t,2}^{n-1})}{m(f_0(X))^{m-2}} \cdot \frac{x_e^{n+1} - x_e^n}{\Delta t}$$
$$= -\partial_X \Big(-f_0(X) \frac{1}{(\partial_X (\frac{1}{2} x_e^{n+1} + x_e^n))^2} \cdot \frac{1}{2} \partial_X (x_{\Delta t,2}^{n+1} + x_{\Delta t,2}^n) \Big) - (\boldsymbol{g}_1^{(1)})^{n+1/2} + O(\Delta t^2), \tag{6.16}$$

$$\text{with} \quad S_e(x_{\Delta t,2}^n, x_{\Delta t,2}^{n-1}) = \partial_X (\frac{3}{2} x_{\Delta t,2}^n - \frac{1}{2} x_{\Delta t,2}^{n-1}). \tag{6.17}$$

Subsequently, a combination of (6.10) and (6.16) yields the fourth order temporal truncation error for $W_2 = W_1 + \Delta t^3 x_{\Delta t,2} = x_e + \Delta t^2 x_{\Delta t,1} + \Delta t^2 x_{\Delta t,2}$:

$$\begin{aligned} &\frac{(S_e(W_2^n, W_2^{n-1}))^{m-1}}{m(f_0(X))^{m-2}} \cdot \frac{W_2^{n+1} - W_2^n}{\Delta t} \\ = &-\partial_X \Big[\Big(f_0(X) \frac{\ln(\partial_X W_2^{n+1}) - \ln(\partial_X W_2^n)}{\partial_X (W_2^{n+1} - W_2^n)} \Big) \\ &- A_0 \Delta t \partial_X (W_2^{n+1} - W_2^n) + \Delta t^2 (\frac{1}{\partial_X W_2^{n+1}} - \frac{1}{\partial_X W_2^n}) \Big] + O(\Delta t^4), \end{aligned} \tag{6.18}$$

$$\text{with} \quad S_e(W_2^n, W_2^{n-1}) = \partial_X (\frac{3}{2} W_2^n - \frac{1}{2} W_2^{n-1}),$$

in which the linearized expansions have been extensively applied.

Next, we construct the spatial correction term $x_{h,1}$ to upgrade the spatial accuracy order. The following truncation error analysis for the spatial discretization can be obtained by using a straightforward Taylor expansion for the constructed profile W_2, and exact solution x_e:

$$\begin{aligned} &\frac{(S_h(W_2^n, W_2^{n-1}))^{m-1}}{m(f_0(X))^{m-2}} \cdot \frac{W_2^{n+1} - W_2^n}{\Delta t} \\ = &- d_h \Big[\Big(f_0(X) \frac{\ln(D_h W_2^{n+1}) - \ln(D_h W_2^n)}{D_h (W_2^{n+1} - W_2^n)} \Big) \\ &- A_0 \Delta t D_h (W_2^{n+1} - W_2^n) + \Delta t^2 (\frac{1}{D_h W_2^{n+1}} - \frac{1}{D_h W_2^n}) \Big] \\ &+ h^2 (\boldsymbol{h}^{(0)})^{n+1/2} + O(h^4) + O(\Delta t^4), \\ &\text{with} \quad S_h(W_2^n, W_2^{n-1}) = \widetilde{D}_h (\frac{3}{2} W_2^n - \frac{1}{2} W_2^{n-1}). \end{aligned} \tag{6.19}$$

Similarly, the function $\boldsymbol{h}^{(0)}$ is smooth enough in the sense that its derivatives are bounded in the $\|\cdot\|_{L^\infty}$ norm, and it is only dependent on the higher order derivatives (both spatial and temporal) of the exact solution W_2, henceforth only dependent on the exact solution x_e. It is a space-time function, while for a fixed time instant $t^{n+1/2}$, $(\boldsymbol{h}^{(0)})^{n+1/2}$ becomes a spatial function. We also notice that there is no $O(h^3)$ truncation error term, due to the fact that the centered difference used in the spatial discretization gives local truncation errors with only even order terms, $O(h^2)$, $O(h^4)$, etc. Subsequently, the spatial correction function $x_{h,1}$ is given by solving the following linear PDE:

$$\frac{(\partial_X x_e)^{m-1}}{m(f_0(X))^{m-2}} \partial_t x_{h,1} + \frac{(m-1)(\partial_X x_e)^{m-2} \partial_X x_{h,1}}{m(f_0(X))^{m-2}} \partial_t x_e$$

$$= -\partial_X \Big(-f_0(X) \frac{1}{(\partial_X x_e)^2} \partial_X x_{h,1} \Big) - \boldsymbol{h}^{(0)}, \tag{6.20}$$

$$x_{h,1}(0) = x_{h,1}(1) = 0, \quad x_{h,1}(t=0) = 0, \tag{6.21}$$

and the solution depends only on the exact solution x_e, with the divided differences of various orders staying bounded. In turn, an application of a full discretization to (6.20) implies that

$$\begin{aligned} &\frac{(S_h(x_e^n, x_e^{n-1}))^{m-1}}{m(f_0(X))^{m-2}} \cdot \frac{x_{h,1}^{n+1} - x_{h,1}^n}{\Delta t} + \frac{(m-1)(S_h(x_e^n, x_e^{n-1}))^{m-2} S_h(x_{h,1}^n, x_{h,1}^{n-1})}{m(f_0(X))^{m-2}} \cdot \frac{x_e^{n+1} - x_e^n}{\Delta t} \\ &= -d_h \Big(-f_0(X) \frac{1}{(D_h(\frac{1}{2} x_e^{n+1} + x_e^n))^2} \cdot \frac{1}{2} D_h(x_{h,1}^{n+1} + x_{h,1}^n) \Big) - (\boldsymbol{h}^{(0)})^{n+1/2} + O(h^2), \end{aligned} \tag{6.22}$$

$$\text{with} \quad S_h(x_{h,1}^n, x_{h,1}^{n-1}) = \widetilde{D}_h(\frac{3}{2} x_{h,1}^n - \frac{1}{2} x_{h,1}^{n-1}). \tag{6.23}$$

Finally, a combination of (6.18) and (6.22) yields the fourth order temporal truncation error for $W = W_2 + h^2 x_{h,1} = x_e + h^2 x_{h,1} + \Delta t^2 x_{\Delta t,1} + \Delta t^2 x_{\Delta t,2}$ (as given by (6.4)):

$$\begin{aligned} &\frac{(S_h(W^n, W^{n-1}))^{m-1}}{m(f_0(X))^{m-2}} \cdot \frac{W^{n+1} - W^n}{\Delta t} \\ &= -d_h \Big[\Big(f_0(X) \frac{\ln(D_h W^{n+1}) - \ln(D_h W^n)}{D_h(W^{n+1} - W^n)} \Big) \\ &\quad - A_0 \Delta t D_h(W^{n+1} - W^n) + \Delta t^2 (\frac{1}{D_h W^{n+1}} - \frac{1}{D_h W^n}) \Big] \\ &\quad + \tau^n, \quad \text{with } \|\tau^n\|_2 \le C(\Delta t^4 + h^4), \end{aligned} \tag{6.24}$$

$$\text{and } S_h(W^n, W^{n-1}) = \widetilde{D}_h(\frac{3}{2} W^n - \frac{1}{2} W^{n-1}).$$

Again, the linearized expansions have been extensively applied.

REMARK 6.1. Since the temporal and spatial correction functions, namely $x_{\Delta t,1}$, $x_{\Delta t,2}$ and $x_{h,1}$, are bounded, we recall the separation property (6.2) for the exact solution, and obtain a similar property for the constructed profile W:

$$D_h W \ge \epsilon_0^*, \quad \text{for } \epsilon_0^* = \tfrac{\epsilon_0}{2} > 0. \tag{6.25}$$

Such a uniform bound will be used in the convergence analysis.

For the the constructed profile W, we also assume its discrete $W^{2,\infty}$ bound, as well as the $W^{1,\infty}$ bound for its temporal derivatives:

$$\|D_h W\|_\infty + \|D_h^2 W\|_\infty \le C^*, \quad \|D_t W^n\|_\infty + \|D_h D_t W^n\|_\infty + \big\|D_h(D_t^2 W^n)\big\|_\infty \le C^*, \tag{6.26}$$

$$\text{with} \quad D_t W^n := \frac{W^{n+1} - W^n}{\Delta t}, \ D_t^2 W^n = \frac{W^{n+1} - 2W^n + W^{n-1}}{\Delta t^2},$$

which comes from the regularity of the exact solution x_e and the correction functions.

REMARK 6.2. The aim for such a higher order asymptotic expansion and truncation error estimate is to justify an a-priori $W_h^{1,\infty}$ bound of the numerical solution, which is needed to obtain the phase separation property, similarly formulated as (6.25) for the constructed approximate solution. In addition, a discrete $W_h^{1,\infty}$ bound for the temporal derivatives of the numerical solution is also needed in the nonlinear analysis, which turns out to be the key reason to derive a fourth order consistency estimate for the constructed solution.

6.2. A preliminary rough error estimate. Instead of a direct analysis for the error function defined as $e^m = x_e^m - x^m$, we introduce an alternate numerical error function:

$$\tilde{x}^m := W^m - x^m. \tag{6.27}$$

The advantage of such a numerical error function is associated with its higher order accuracy, which comes from the higher order consistency estimate (6.24). Moreover, the following notations are introduced, for the convenience of the analysis presented later.

$$x^{n+1/2} = \frac{1}{2}(x^{n+1} + x^n), \ W^{n+1/2} = \frac{1}{2}(W^{n+1} + W^n), \tag{6.28}$$

$$\breve{x}^{n+1/2} = \frac{3}{2}x^n - \frac{1}{2}x^{n-1}, \ \breve{W}^{n+1/2} = \frac{3}{2}W^n - \frac{1}{2}W^{n-1}, \tag{6.29}$$

$$\tilde{x}^{n+1/2} = W^{n+1/2} - x^{n+1/2} = \frac{1}{2}(\tilde{x}^{n+1} + \tilde{x}^n), \ \breve{\tilde{x}}^{n+1/2} = \breve{W}^{n+1/2} - \breve{x}^{n+1/2} = \frac{3}{2}\tilde{x}^n - \frac{1}{2}\tilde{x}^{n-1}. \tag{6.30}$$

In turn, subtracting the numerical scheme (3.10) from the consistency estimate (6.24) yields

$$\begin{aligned}
&\frac{(S_h(x^n,x^{n-1}))^{m-1}}{m(f_0(X))^{m-2}} \cdot \frac{\tilde{x}^{n+1} - \tilde{x}^n}{\Delta t} + \frac{(S_h(W^n,W^{n-1}))^{m-1} - (S_h(x^n,x^{n-1}))^{m-1}}{m(f_0(X))^{m-2}} \cdot \frac{W^{n+1} - W^n}{\Delta t} \\
&= -d_h\Big[f_0(X)\Big(\frac{\ln(D_hW^{n+1}) - \ln(D_hW^n)}{D_h(W^{n+1} - W^n)} - \frac{\ln(D_hx^{n+1}) - \ln(D_hx^n)}{D_h(x^{n+1} - x^n)}\Big) \\
&\quad -A_0\Delta t D_h(\tilde{x}^{n+1} - \tilde{x}^n) - \Delta t^2\Big(\frac{D_h\tilde{x}^{n+1}}{D_hW^{n+1}D_hx^{n+1}} - \frac{D_h\tilde{x}^n}{D_hW^nD_hx^n}\Big)\Big] \\
&\quad +\tau^n, \quad \text{with } \|\tau^n\|_2 \le C(\Delta t^4 + h^4).
\end{aligned} \tag{6.31}$$

To proceed with the nonlinear analysis, we make the following a-priori assumption at the previous time steps, for $k = n, n-1, n-2$:

$$\|\tilde{x}^k\|_2 \le \mathcal{C}(\Delta t^4 + h^4), \quad \text{with } \mathcal{C} \text{ uniform for a fixed final time } T. \tag{6.32}$$

Such an a-priori assumption will be recovered by the optimal rate convergence analysis at the next time step, as will be demonstrated later. With this assumption, the following bounds for the numerical solution are obtained, with the help of inverse inequality:

$$\|D_h\tilde{x}^k\|_2 \le \frac{C\|\tilde{x}^k\|_2}{h} \le C\mathcal{C}(\Delta t^3 + h^3), \tag{6.33}$$

$$\|D_h\tilde{x}^k\|_\infty \le \frac{C\|\tilde{x}^k\|_2}{h^{\frac{3}{2}}} \le C\mathcal{C}(\Delta t^{\frac{5}{2}} + h^{\frac{5}{2}}) \le \frac{\epsilon_0^*}{2}, \tag{6.34}$$

$$\text{so that } \frac{\epsilon_0^*}{2} \le D_hx^k = D_hW^k - D_h\tilde{x}^k \le C^* + \frac{\epsilon_0^*}{2} := \tilde{C}^*, \tag{6.35}$$

for $k = n, n-1, n-2$, in which the lower and upper bounds (6.25), (6.26), for the constructed profile W, have been used. Notice that the separation property (6.2) is valid for the exact PDE solution, with a fixed constant $\epsilon_0 > 0$, while the constructed solution W preserves a similar separation property (6.25), with an alternate constant $\epsilon_0^* > 0$. Without loss of generality, such a constant could be taken as $\epsilon_0^* = \frac{\epsilon_0}{2} > 0$. As a result, the bound (6.34) is always available, provided that Δt and h are sufficiently small. In addition, the following observation is made, motived by the preliminary estimate (6.34):

$$\begin{aligned}
\|D_hx^k - D_hx^{k-1}\|_\infty &\le \|D_hW^k - D_hW^{k-1}\|_\infty + \|D_h\tilde{x}^k - D_h\tilde{x}^{k-1}\|_\infty \\
&\le C^*\Delta t + C\mathcal{C}(\Delta t^{\frac{5}{2}} + h^{\frac{5}{2}}) \le (C^* + 1)\Delta t,
\end{aligned} \tag{6.36}$$

for $k=n,n-1$, in which the regularity requirement (6.26) for the constructed solution has been applied again. In turn, we conclude that

$$\widetilde{D}_h(\frac{3}{2}x^k-\frac{1}{2}x^{k-1})=\widetilde{D}_h x^k+\frac{1}{2}\widetilde{D}_h(x^k-x^{k-1})\geq\frac{\epsilon_0^*}{2}-\frac{\epsilon_0^*}{4}=\frac{\epsilon_0^*}{4}\geq\Delta t^2, \tag{6.37}$$

$$\text{so that } S_h(x^k,x^{k-1})=\widetilde{D}_h\breve{x}^{n+1/2}=\widetilde{D}_h(\frac{3}{2}x^k-\frac{1}{2}x^{k-1}),\text{ for } k=n,n-1. \tag{6.38}$$

The following preliminary estimates are needed in the nonlinear error analysis for the two terms on the left-hand side of (6.31).

Lemma 6.1. *For the constructed profile W satisfying (6.26), and the numerical error function with a discrete $W^{1,\infty}$ bound given by (6.34), for $k=n,n-1,n-2$, we have*

$$\begin{aligned}&\|(S_h(x^n,x^{n-1}))^{m-1}-(S_h(x^{n-1},x^{n-2}))^{m-1}\|_\infty\\ \leq&\tilde{C}_1\Delta t,\end{aligned} \tag{6.39}$$

$$\begin{aligned}&(S_h(W^n,W^{n-1}))^{m-1}-(S_h(x^n,x^{n-1}))^{m-1}\\ =&\mathcal{N}^{n+1/2}\widetilde{D}_h\breve{\tilde{x}}^{n+1/2},\quad \textit{with } \|\mathcal{N}^{n+1/2}\|_\infty\leq\tilde{C}_2,\end{aligned} \tag{6.40}$$

$$\begin{aligned}&\mathcal{N}^{n+1/2}:=(m-1)(\xi^{(2)})^{m-2},\\ &\textit{for some } \xi^{(2)} \textit{ between } \widetilde{D}_h\left(\tfrac{3}{2}W^n-\tfrac{1}{2}W^{n-1}\right) \textit{ and } \widetilde{D}_h\left(\tfrac{3}{2}x^n-\tfrac{1}{2}x^{n-1}\right),\\ &\textit{and } \|D_h\mathcal{N}^{n+1/2}\|_\infty\leq\tilde{C}_3,\\ &\textit{if we define } (\mathcal{N}^{n+1/2})_0=(\mathcal{N}^{n+1/2})_1,\,(\mathcal{N}^{n+1/2})_M=(\mathcal{N}^{n+1/2})_{M-1},\end{aligned} \tag{6.41}$$

in which $\tilde{C}_1$, $\tilde{C}_2$ and $\tilde{C}_3$ are only dependent on the exact solution, and independent of Δt and h.

Proof. Based on the representation identity (6.38), we apply the intermediate value theorem and see that

$$\begin{aligned}&(S_h(x^n,x^{n-1}))^{m-1}-(S_h(x^{n-1},x^{n-2}))^{m-1}\\ =&(m-1)(\xi^{(1)})^{m-2}\widetilde{D}_h\left(\frac{3}{2}(x^n-x^{n-1})-\frac{1}{2}(x^{n-1}-x^{n-2})\right),\end{aligned} \tag{6.42}$$

with $\xi^{(1)}$ between $\widetilde{D}_h\left(\frac{3}{2}x^n-\frac{1}{2}x^{n-1}\right)$ and $\widetilde{D}_h\left(\frac{3}{2}x^{n-1}-\frac{1}{2}x^{n-2}\right)$. Meanwhile, by the upper estimate (6.35) and the lower bound (6.37), we get

$$\frac{\epsilon_0^*}{4}\leq\widetilde{D}_h(\frac{3}{2}x^k-\frac{1}{2}x^{k-1})\leq\frac{3\tilde{C}^*}{2},\text{ for } k=n,n-1,\text{ so that }\frac{\epsilon_0^*}{4}\leq\xi^{(1)}\leq\frac{3\tilde{C}^*}{2}. \tag{6.43}$$

A substitution into (6.42), combined with the estimate (6.36), indicates the desired inequality, with $\tilde{C}_1=2(m-1)\max(\frac{4}{\epsilon_0^*},\frac{3\tilde{C}^*}{2})^{|m-2|}(C^*+1)$.

A similar application of intermediate value theorem reveals that

$$(S_h(W^n,W^{n-1}))^{m-1}-(S_h(x^n,x^{n-1}))^{m-1}=(m-1)(\xi^{(2)})^{m-2}\widetilde{D}_h\breve{\tilde{x}}^{n+1/2}, \tag{6.44}$$

and $\xi^{(2)}$ between $\widetilde{D}_h\left(\frac{3}{2}W^n-\frac{1}{2}W^{n-1}\right)$ and $\widetilde{D}_h\left(\frac{3}{2}x^n-\frac{1}{2}x^{n-1}\right)$. Using the same argument as in (6.43), we see that $\frac{\epsilon_0^*}{4}\leq\xi^{(2)}\leq\frac{3\tilde{C}^*}{2}$, so that

$$\|\mathcal{N}^{n+1/2}\|_\infty=\|(m-1)(\xi^{(2)})^{m-2}\|_\infty\leq\tilde{C}_2:=(m-1)\max(\frac{4}{\epsilon_0^*},\frac{3\tilde{C}^*}{2})^{|m-2|}, \tag{6.45}$$

which completes the proof of (6.40). Moreover, for two adjacent grid points x_i and x_{i+1} (with $1\leq i\leq M-2$), motivated by the fact that

$$\xi_i^{(2)}\text{ is between }\widetilde{D}_h\breve{W}_i^{n+1/2}\text{ and }\widetilde{D}_h\breve{x}_i^{n+1/2},\ \xi_{i+1}^{(2)}\text{ is between }\widetilde{D}_h\breve{W}_{i+1}^{n+1/2}\text{ and }\widetilde{D}_h\breve{x}_{i+1}^{n+1/2}, \tag{6.46}$$

we have the following observation:

$$
\begin{aligned}
|\xi_{i+1}^{(2)}-\xi_i^{(2)}| &\le \left|\widetilde{D}_h \breve{W}_{i+1}^{n+1/2}-\widetilde{D}_h \breve{W}_i^{n+1/2}\right|+\left|\widetilde{D}_h \breve{\tilde{x}}_i^{n+1/2}\right|+\left|\widetilde{D}_h \breve{\tilde{x}}_{i+1}^{n+1/2}\right| \\
&\le \frac{1}{2}h\Big(\left|D_h^2 \breve{W}_{i+1}^{n+1/2}\right|+\left|D_h^2 \breve{W}_i^{n+1/2}\right|\Big)+C\mathcal{C}(\Delta t^{\frac{5}{2}}+h^{\frac{5}{2}}) \\
&\le C^* h+h=(C^*+1)h,
\end{aligned} \tag{6.47}
$$

in which the preliminary estimate (6.34) and the regularity assumption (6.26) for the constructed profile have been used. On the other hand, with another application of intermediate value theorem:

$$
\begin{aligned}
\mathcal{N}_{i+1}^{n+1/2}-\mathcal{N}_i^{n+1/2} &= (m-1)(\xi_{i+1}^{(2)})^{m-2}-(m-1)(\xi_i^{(2)})^{m-2} \\
&= (m-1)(m-2)(\eta^{(1)})^{m-3}(\xi_{i+1}^{(2)}-\xi_i^{(2)}),
\end{aligned} \tag{6.48}
$$

with $\eta^{(1)}$ between $\xi_i^{(2)}$ and $\xi_{i+1}^{(2)}$, we get the desired estimate

$$
\left|\mathcal{N}_{i+1}^{n+1/2}-\mathcal{N}_i^{n+1/2}\right| \le (m-1)|m-2|\max(\frac{4}{\epsilon_0^*},\frac{3\tilde{C}^*}{2})^{|m-3|}(C^*+1)h. \tag{6.49}
$$

This completes the proof of (6.41), by taking $\tilde{C}_3=(m-1)|m-2|\max(\frac{4}{\epsilon_0^*},\frac{3\tilde{C}^*}{2})^{|m-3|}(C^*+1)$. ∎

Now we proceed with a rough error estimate. Taking a discrete inner product with (6.31) by $2\tilde{x}^{n+1}$ leads to

$$
\begin{aligned}
&\Big\langle \frac{(S_h(x^n,x^{n-1}))^{m-1}}{m(f_0(X))^{m-2}}\cdot\frac{\tilde{x}^{n+1}-\tilde{x}^n}{\Delta t},2\tilde{x}^{n+1}\Big\rangle+2A_0\Delta t\langle D_h(\tilde{x}^{n+1}-\tilde{x}^n),D_h\tilde{x}^{n+1}\rangle \\
&\quad -2\Big\langle f_0(X)\Big(\frac{\ln(D_hW^{n+1})-\ln(D_hW^n)}{D_h(W^{n+1}-W^n)}-\frac{\ln(D_hx^{n+1})-\ln(D_hx^n)}{D_h(x^{n+1}-x^n)}\Big),\tilde{x}^{n+1}\Big\rangle \\
&= -2\Big\langle\frac{(S_h(W^n,W^{n-1}))^{m-1}-(S_h(x^n,x^{n-1}))^{m-1}}{m(f_0(X))^{m-2}}\cdot\frac{W^{n+1}-W^n}{\Delta t},\tilde{x}^{n+1}\Big\rangle \\
&\quad -2\Delta t^2\Big\langle\frac{D_h\tilde{x}^{n+1}}{D_hW^{n+1}D_hx^{n+1}}-\frac{D_h\tilde{x}^n}{D_hW^nD_hx^n},D_h\tilde{x}^{n+1}\Big\rangle+2\langle\Delta t^n,\tilde{x}^{n+1}\rangle.
\end{aligned} \tag{6.50}
$$

For the temporal derivative term, we make use of the equality

$$
2\tilde{x}^{n+1}(\tilde{x}^{n+1}-\tilde{x}^n)=(\tilde{x}^{n+1})^2-(\tilde{x}^n)^2+(\tilde{x}^{n+1}-\tilde{x}^n)^2\ge(\tilde{x}^{n+1})^2-(\tilde{x}^n)^2, \tag{6.51}
$$

and get

$$
\begin{aligned}
&\Big\langle\frac{(S_h(x^n,x^{n-1}))^{m-1}}{m(f_0(X))^{m-2}}\cdot\frac{\tilde{x}^{n+1}-\tilde{x}^n}{\Delta t},2\tilde{x}^{n+1}\Big\rangle \\
&\ge \frac{1}{\Delta t}\Big(\Big\langle\frac{(S_h(x^n,x^{n-1}))^{m-1}}{m(f_0(X))^{m-2}},(\tilde{x}^{n+1})^2\Big\rangle-\Big\langle\frac{(S_h(x^n,x^{n-1}))^{m-1}}{m(f_0(X))^{m-2}},(\tilde{x}^n)^2\Big\rangle\Big) \\
&= \frac{1}{\Delta t}\Big(\Big\langle\frac{(S_h(x^n,x^{n-1}))^{m-1}}{m(f_0(X))^{m-2}},(\tilde{x}^{n+1})^2\Big\rangle-\Big\langle\frac{(S_h(x^{n-1},x^{n-2}))^{m-1}}{m(f_0(X))^{m-2}},(\tilde{x}^n)^2\Big\rangle\Big) \\
&\quad -\frac{1}{\Delta t}\Big\langle\frac{(S_h(x^n,x^{n-1}))^{m-1}-(S_h(x^{n-1},x^{n-2}))^{m-1}}{m(f_0(X))^{m-2}},(\tilde{x}^n)^2\Big\rangle.
\end{aligned} \tag{6.52}
$$

Furthermore, with an application of the preliminary estimate (6.39), the last term of (6.52) could be bounded as

$$
\frac{1}{\Delta t}\Big\langle\frac{(S_h(x^n,x^{n-1}))^{m-1}-(S_h(x^{n-1},x^{n-2}))^{m-1}}{m(f_0(X))^{m-2}},(\tilde{x}^n)^2\Big\rangle\le\tilde{C}_4\|\tilde{x}^n\|_2^2, \tag{6.53}
$$

$$\text{with } \tilde{C}_4 = \tilde{C}_1 \max\left(\frac{1}{\min_\Omega f_0(X)}, \max_\Omega f_0(X)\right)^{|m-2|}.$$

The second term on the left-hand side of (6.50) could be controlled by a standard inequality:

$$2A_0\Delta t\langle D_h(\tilde{x}^{n+1}-\tilde{x}^n), D_h\tilde{x}^{n+1}\rangle \geq A_0\Delta t(\|D_h\tilde{x}^{n+1}\|_2^2 - \|D_h\tilde{x}^n\|_2). \tag{6.54}$$

The Cauchy inequality could be applied to bound the term associated with the local truncation error:

$$2\langle \tau^n, \tilde{x}^{n+1}\rangle \leq \|\tau^n\|_2^2 + \|\tilde{x}^{n+1}\|_2^2. \tag{6.55}$$

For the second term on the right-hand side of (6.50), we make the following observation:

$$-2\Delta t^2\left\langle \frac{D_h\tilde{x}^{n+1}}{D_hW^{n+1}D_hx^{n+1}}, D_h\tilde{x}^{n+1}\right\rangle \leq 0, \tag{6.56}$$

$$\begin{aligned} 2\Delta t^2\left\langle \frac{D_h\tilde{x}^n}{D_hW^nD_hx^n}, D_h\tilde{x}^{n+1}\right\rangle &\leq 2\Delta t^2 \cdot \frac{1}{\frac{1}{2}(\epsilon_0^*)^2}\|D_h\tilde{x}^n\|_2 \cdot \|D_h\tilde{x}^{n+1}\|_2 \\ &\leq \Delta t^3\|D_h\tilde{x}^{n+1}\|_2^2 + 4\Delta t(\epsilon_0^*)^{-4}\|D_h\tilde{x}^n\|_2^2 \\ &\leq \tilde{C}_5\Delta t\|\tilde{x}^{n+1}\|_2^2 + 4\Delta t(\epsilon_0^*)^{-4}\|D_h\tilde{x}^n\|_2^2, \end{aligned} \tag{6.57}$$

in which (6.56) is based on the fact that $D_hW^{n+1} > 0$, $D_hx^{n+1} > 0$ (as given by the unique solvability analysis in Theorem 4.1), the first step of (6.57) comes from the separation property (6.25) (for the constructed profile W) and the preliminary estimate (6.35), and an inverse inequality has been applied at the last step.

For the first term on the right-hand side of (6.50), the preliminary estimate (6.40) and the regularity assumption (6.26) have to be applied:

$$\begin{aligned} &-2\left\langle \frac{(S_h(W^n,W^{n-1}))^{m-1} - (S_h(x^n,x^{n-1}))^{m-1}}{m(f_0(X))^{m-2}} \cdot \frac{W^{n+1}-W^n}{\Delta t}, \tilde{x}^{n+1}\right\rangle \\ &= -2\left\langle \frac{\mathcal{N}^{n+1/2}\widetilde{D}_h\breve{\tilde{x}}^{n+1/2}}{m(f_0(X))^{m-2}} \cdot \frac{W^{n+1}-W^n}{\Delta t}, \tilde{x}^{n+1}\right\rangle \\ &\leq \tilde{C}_6\|\widetilde{D}_h\breve{\tilde{x}}^{n+1/2}\|_2 \cdot \|\tilde{x}^{n+1}\|_2 \leq \frac{\tilde{C}_6}{2}(\|D_h\breve{\tilde{x}}^{n+1/2}\|_2^2 + \|\tilde{x}^{n+1}\|_2^2), \end{aligned} \tag{6.58}$$

$$\text{with } \tilde{C}_6 = \frac{2}{m}\tilde{C}_2C^* \cdot \max\left(\frac{1}{\min_\Omega f_0(X)}, \max_\Omega f_0(X)\right)^{|m-2|}.$$

Notice that the inequality $\|\widetilde{D}_hf\|_2 \leq \|D_hf\|_2$ has been used in the last step.

The rest of the analysis is focused on the term associated with the nonlinear diffusion part:

$$\mathcal{NLE}^{n+1/2} := -\frac{\ln(D_hW^{n+1}) - \ln(D_hW^n)}{D_h(W^{n+1}-W^n)} + \frac{\ln(D_hx^{n+1}) - \ln(D_hx^n)}{D_h(x^{n+1}-x^n)}. \tag{6.59}$$

The following lemma is needed in the nonlinear estimate; its proof will be provided in the Appendix.

LEMMA 6.2. *Fix $x_0 > 0$, and we define $q_1(x) := -\frac{\ln x - \ln x_0}{x - x_0}$ for $x > 0$. The following properties are valid:*

$$q_1'(x) > 0, \text{ for any } x > 0, \tag{6.60}$$

$$q_1(y) - q_1(x) = q_1'(\eta)(y-x), \text{ with } q_1'(\eta) \text{ between } \tfrac{1}{2y^2}, \tfrac{1}{2x^2} \text{ and } \tfrac{1}{2x_0^2}, \ \forall x>0, y>0, \tag{6.61}$$

$$q_1''(x) \leq 0, \text{ for any } x > 0, \tag{6.62}$$

$$\frac{q_1(y)-q_1(x)}{y-x} \text{ is a decreasing function of } x, \text{ for any fixed } y > 0. \tag{6.63}$$

Subsequently, the following point-wise nonlinear estimate becomes available for $\mathcal{NLE}^{n+1/2}$.

LEMMA 6.3. *At each numerical mesh cell (x_i, x_{i+1}), we have the following estimate*

$$\mathcal{NLE}^{n+1/2}_{i+1/2} = \xi^{(3)}_{i+1/2} D_h \tilde{x}^{n+1}_{i+1/2} + \xi^{(4)}_{i+1/2} D_h \tilde{x}^{n}_{i+1/2}, \tag{6.64}$$

$$\textit{with } \xi^{(3)}_{i+1/2} \geq \frac{1}{2\tilde{C}^*} h, \ \frac{1}{2(\tilde{C}^*)^2} \leq \xi^{(4)}_{i+1/2} \leq 2(\epsilon_0^*)^{-2}. \tag{6.65}$$

Proof. The following decomposition is performed for $\mathcal{NLE}$:

$$\begin{aligned}
\mathcal{NLE}^{n+1/2} &= \mathcal{NLE}^{n+1/2,(1)} + \mathcal{NLE}^{n+1/2,(2)}, \text{ with}\\
\mathcal{NLE}^{n+1/2,(1)} &= -\frac{\ln(D_h W^{n+1}) - \ln(D_h x^n)}{D_h W^{n+1} - D_h x^n} + \frac{\ln(D_h x^{n+1}) - \ln(D_h x^n)}{D_h x^{n+1} - D_h x^n}\\
&= q_1(D_h W^{n+1}) - q_1(D_h x^{n+1}), \text{ with fixed } x_0 = D_h x^n, \qquad (6.66)\\
\mathcal{NLE}^{n+1/2,(2)} &= -\frac{\ln(D_h W^{n+1}) - \ln(D_h W^n)}{D_h W^{n+1} - D_h W^n} + \frac{\ln(D_h W^{n+1}) - \ln(D_h x^n)}{D_h W^{n+1} - D_h x^n}\\
&= q_1(D_h W^n) - q_1(D_h x^n), \text{ with fixed } x_0 = D_h W^{n+1}. \qquad (6.67)
\end{aligned}$$

For the first part $\mathcal{NLE}^{n+1/2,(1)}$, we make use of the following bound for $D_h x^{n+1}$:

$$0 < (D_h x^{n+1})_{i+1/2} \leq \frac{1}{h}, \quad \text{since } 0 \leq x_k^{n+1} \leq 1, \, \forall 0 \leq k \leq M, \tag{6.68}$$

so that an application of property (6.63) implies that

$$\begin{aligned}
\frac{q_1(D_h W^{n+1}) - q_1(D_h x^{n+1})}{D_h W^{n+1} - D_h x^{n+1}} &\geq \frac{q_1(D_h W^{n+1}) - q_1(\frac{1}{h})}{D_h W^{n+1} - \frac{1}{h}}\\
&= \frac{-\frac{\ln \frac{1}{h} - \ln(D_h x^n)}{\frac{1}{h} - D_h x^n} - q_1(D_h W^{n+1})}{\frac{1}{h} - D_h W^{n+1}} \geq \frac{1}{2\tilde{C}^*} h, \qquad (6.69)
\end{aligned}$$

in which the last step is based on the preliminary estimate (6.35), as well as the fact that the value of $-q_1(D_h W^{n+1})$ is between $\frac{1}{D_h W^{n+1}}$ and $\frac{1}{D_h x^n}$. This inequality is equivalent to

$$\mathcal{NLE}^{n+1/2,(1)}_{i+1/2} = \xi^{(3)}_{i+1/2} D_h \tilde{x}^{n+1}_{i+1/2}, \quad \text{with } \xi^{(3)}_{i+1/2} \geq \frac{1}{2\tilde{C}^*} h. \tag{6.70}$$

For the first part $\mathcal{NLE}^{n+1/2,(2)}$, we apply property (6.61) so that

$$\mathcal{NLE}^{n+1/2,(2)} = q_1(D_h W^n) - q_1(D_h x^n) = q_1'(\eta) D_h \tilde{x}^n, \tag{6.71}$$

$$\text{with } q_1'(\eta) \text{ between } \frac{1}{2(D_h W^n)^2}, \ \frac{1}{2(D_h x^n)^2}, \text{ and } \frac{1}{2(D_h W^{n+1})^2}. \tag{6.72}$$

On the other hand, by the separation property (6.25), the regularity assumption (6.26) for W, combined with the preliminary estimate (6.35), we obtain the desired estimate:

$$\frac{1}{2(\tilde{C}^*)^2} \leq q_1'(\eta) \leq 2(\epsilon_0^*)^{-2}. \tag{6.73}$$

In other words, the second estimate in (6.64) becomes available:

$$\mathcal{NLE}^{n+1/2,(2)}_{i+1/2} = \xi^{(4)}_{i+1/2} D_h \tilde{x}^{n}_{i+1/2}, \quad \text{with } \frac{1}{2(\tilde{C}^*)^2} \leq \xi^{(4)}_{i+1/2} \leq 2(\epsilon_0^*)^{-2}. \tag{6.74}$$

This completes the proof of Lemma 6.3. ∎

As a consequence of this lemma, we analyze the nonlinear product at each cell (x_i, x_{i+1}):

$$
\begin{aligned}
& f_0(X_{i+1/2})\mathcal{NLE}^{n+1/2}_{i+1/2}\cdot 2D_h\tilde{x}^{n+1}_{i+1/2} \\
= & 2f_0(X_{i+1/2})\Big(\xi^{(3)}_{i+1/2}(D_h\tilde{x}^{n+1}_{i+1/2})^2+\xi^{(4)}_{i+1/2}D_h\tilde{x}^{n+1}_{i+1/2}\cdot D_h\tilde{x}^{n}_{i+1/2}\Big) \\
\geq & 2f_0(X_{i+1/2})\Big(\frac{1}{2\tilde{C}^*}h(D_h\tilde{x}^{n+1}_{i+1/2})^2-2(\epsilon_0^*)^{-2}D_h\tilde{x}^{n+1}_{i+1/2}\cdot D_h\tilde{x}^{n}_{i+1/2}\Big) \\
\geq & 2f_0(X_{i+1/2})\cdot\Big(-\frac{2\tilde{C}^*(\epsilon_0^*)^{-4}}{h}\cdot(D_h\tilde{x}^{n}_{i+1/2})^2\Big),
\end{aligned} \tag{6.75}
$$

in which the Cauchy inequality has been applied at the last step. A summation of this inequality yields

$$
\Big\langle f_0(X)\mathcal{NLE}^{n+1/2},2D_h\tilde{x}^{n+1}\Big\rangle\geq-\tilde{C}_7h^{-1}\|D_h\tilde{x}^n\|_2^2,\quad \text{with } \tilde{C}_7=4\tilde{C}^*(\epsilon_0^*)^{-4}\|f_0(X)\|_\infty. \tag{6.76}
$$

Finally, a substitution of (6.52), (6.53), (6.54), (6.55), (6.56), (6.57), (6.58) and (6.76) into (6.50) leads to

$$
\begin{aligned}
& \Big\langle\frac{(S_h(x^n,x^{n-1}))^{m-1}}{m(f_0(X))^{m-2}}-\Delta t\Big(\frac{\tilde{C}_6}{2}+1+\tilde{C}_5\Delta t\Big),(\tilde{x}^{n+1})^2\Big\rangle \\
\leq & \Big\langle\frac{(S_h(x^{n-1},x^{n-2}))^{m-1}}{m(f_0(X))^{m-2}},(\tilde{x}^n)^2\Big\rangle\Big)+\Delta t\tilde{C}_4\|\tilde{x}^n\|_2^2+(A_0+4(\epsilon_0^*)^{-4})\Delta t^2\|D_x\tilde{x}^n\|_2^2 \\
& +\Delta t\|\tau^n\|_2^2+\frac{\tilde{C}_6}{2}\Delta t\|D_h\breve{\tilde{x}}^{n+1/2}\|_2^2+\tilde{C}_7\cdot\frac{\Delta t}{h}\|D_h\tilde{x}^n\|_2^2.
\end{aligned} \tag{6.77}
$$

On the left-hand side, we observe the following point-wise lower bound, which comes from the preliminary estimate (6.37):

$$
\frac{(S_h(x^n,x^{n-1}))^{m-1}}{m(f_0(X))^{m-2}}\geq\tilde{C}_8:=\frac{1}{m}\Big(\frac{\epsilon_0^*}{4}\Big)^{m-1}\min\Big(\frac{1}{\max_\Omega f_0(X)},\min_\Omega f_0(X)\Big)^{|m-2|}, \tag{6.78}
$$

$$
\Delta t\Big(\frac{\tilde{C}_6}{2}+1+\tilde{C}_5\Delta t\Big)\leq\frac{\tilde{C}_8}{2},\quad \text{provided that } \Delta t \text{ is sufficiently small,} \tag{6.79}
$$

which in turn indicates that

$$
\Big\langle\frac{(S_h(x^n,x^{n-1}))^{m-1}}{m(f_0(X))^{m-2}}-\Delta t\Big(\frac{\tilde{C}_6}{2}+1+\tilde{C}_5\Delta t\Big),(\tilde{x}^{n+1})^2\Big\rangle\geq\frac{\tilde{C}_8}{2}\|\tilde{x}^{n+1}\|_2^2. \tag{6.80}
$$

On the right-hand side, the following estimates are available, based on the a-priori assumption (6.32) and the preliminary estimate (6.33):

$$
\Big\langle\frac{(S_h(x^{n-1},x^{n-2}))^{m-1}}{m(f_0(X))^{m-2}},(\tilde{x}^n)^2\Big\rangle\Big)=O((\Delta t^4+h^4)^2), \tag{6.81}
$$

$$
\Delta t\tilde{C}_4\|\tilde{x}^n\|_2^2+\Delta t\|\tau^n\|_2^2=O(\Delta t(\Delta t^4+h^4)^2), \tag{6.82}
$$

$$
(A_0+4(\epsilon_0^*)^{-4})\Delta t^2\|D_x\tilde{x}^n\|_2^2=O(\Delta t^2(\Delta t^3+h^3)^2), \tag{6.83}
$$

$$
\frac{\tilde{C}_6}{2}\Delta t\|D_h\breve{\tilde{x}}^{n+1/2}\|_2^2=O(\Delta t(\Delta t^3+h^3)^2), \tag{6.84}
$$

$$
\tilde{C}_7\cdot\frac{\Delta t}{h}\|D_h\tilde{x}^n\|_2^2\leq C\tilde{C}_7C_2\mathcal{C}^2(\Delta t^3+h^3)^2. \tag{6.85}
$$

Then we arrive at a rough estimate for the numerical error function at time step t^{n+1}:

$$
\frac{\tilde{C}_8}{2}\|\tilde{x}^{n+1}\|_2^2\leq\Big(C\tilde{C}_7C_2\mathcal{C}^2+1\Big)(\Delta t^3+h^3)^2,\quad \text{provided that } \Delta t \text{ and } h \text{ are sufficiently small,}
$$

$$\text{i.e. } \|\tilde{x}^{n+1}\|_2 \le \hat{C}(\Delta t^3 + h^3), \quad \text{with } \hat{C} := \Big(\frac{2(C\tilde{C}_7 C_2 \mathcal{C}^2 + 1)}{\tilde{C}_8}\Big)^{\frac{1}{2}}, \tag{6.86}$$

under the linear refinement requirement $C_1 h \le \Delta t \le C_2 h$. Subsequently, an application of 1-D inverse inequality implies that

$$\|D_h \tilde{x}^{n+1}\|_\infty \le \frac{C\|\tilde{x}^{n+1}\|_2}{h^{\frac{3}{2}}} \le \hat{C}_1(\Delta t^{\frac{3}{2}} + h^{\frac{3}{2}}), \quad \text{with } \hat{C}_1 = C\hat{C}, \tag{6.87}$$

under the same linear refinement requirement. Because of the accuracy order, we could take Δt and h sufficiently small so that $\hat{C}_1(\Delta t^{\frac{3}{2}} + h^{\frac{3}{2}}) \le \frac{\epsilon_0^*}{2}$, which in turn gives

$$\frac{\epsilon_0^*}{2} \le D_h x^{n+1} = D_h W^{n+1} - D_h \tilde{x}^{n+1} \le C^* + \frac{\epsilon_0^*}{2} = \tilde{C}^*, \tag{6.88}$$

in which the lower and upper bounds (6.25), (6.26) (for the constructed profile W) have been used again. Such a uniform $\|\cdot\|_{W_h^{1,\infty}}$ bound will play a very important role in the refined error estimate.

Remark 6.3. In the rough error estimate (6.86), we see that the accuracy order is lower than the one given by the a-priori-assumption (6.32). Therefore, such a rough estimate could not be used for a global induction analysis. Instead, the purpose of such an estimate is to establish a uniform $\|\cdot\|_{W_h^{1,\infty}}$ bound for the numerical solution at time step t^{n+1}, as well as its temporal derivative, via the technique of inverse inequality. With these bounds established for the numerical solution, the refined error analysis will yield much sharper estimates.

6.3. A further rough error estimate. Meanwhile, we have to derive a discrete $W_h^{1,\infty}$ bound for the second order temporal derivative of the numerical solution at time step t^{n+1}, which will be needed in the refined error estimate. In fact, such a bound could not be obtained by (6.87). To obtain such a bound, we have to perform a further rough error estimate.

We revisit the proof of Lemma 6.3 and discover that, $\xi_{i+1/2}^{(3)}$ has to be between $D_h W^{n+1}$ and $D_h x^{n+1}$, based on the representation (6.66) and the property (6.61). In more details, the regularity assumption (6.26) (for W) and the $W_h^{1\infty}$ bound (6.88) imply a similar bound for $\xi_{i+1/2}^{(3)}$:

$$\frac{1}{2(\tilde{C}^*)^2} \le \xi_{i+1/2}^{(4)} \le 2(\epsilon_0^*)^{-2}. \tag{6.89}$$

With such a bound at hand, we are able to rewrite the inner product in a more precise way:

$$\begin{aligned}
& f_0(X_{i+1/2}) \mathcal{NLE}_{i+1/2}^{n+1/2} \cdot 2D_h \tilde{x}_{i+1/2}^{n+1} \\
= & 2f_0(X_{i+1/2})\Big(\xi_{i+1/2}^{(3)}(D_h\tilde{x}_{i+1/2}^{n+1})^2 + \xi_{i+1/2}^{(4)} D_h \tilde{x}_{i+1/2}^{n+1} \cdot D_h \tilde{x}_{i+1/2}^{n}\Big) \\
\ge & 2f_0(X_{i+1/2})\Big(\frac{1}{2(\tilde{C}^*)^2}(D_h\tilde{x}_{i+1/2}^{n+1})^2 - 2(\epsilon_0^*)^{-2} D_h\tilde{x}_{i+1/2}^{n+1} \cdot D_h\tilde{x}_{i+1/2}^{n}\Big) \\
\ge & 2f_0(X_{i+1/2}) \cdot \Big(\frac{1}{4(\tilde{C}^*)^2}(D_h\tilde{x}_{i+1/2}^{n+1})^2 - 4(\tilde{C}^*)^2(\epsilon_0^*)^{-4} \cdot (D_h\tilde{x}_{i+1/2}^{n})^2\Big),
\end{aligned}$$

which in turn gives

$$\Big\langle f_0(X)\mathcal{NLE}^{n+1/2}, 2D_h\tilde{x}^{n+1}\Big\rangle \ge \tilde{C}_9\|\tilde{x}^{n+1}\|_2^2 - \tilde{C}_{10}\|D_h\tilde{x}^n\|_2^2, \tag{6.90}$$

$$\text{with } \tilde{C}_9 = \frac{1}{2(\tilde{C}^*)^2}\min_\Omega(f_0(X)),\ \tilde{C}_{10} = 8(\tilde{C}^*)^2(\epsilon_0^*)^{-4}\|f_0(X)\|_\infty.$$

A substitution of this updated estimate yields a rewritten inequality for (6.77):

$$\Big\langle \frac{(S_h(x^n, x^{n-1}))^{m-1}}{m(f_0(X))^{m-2}} - \Delta t\Big(\frac{\tilde{C}_6}{2} + 1 + \tilde{C}_5\Delta t\Big), (\tilde{x}^{n+1})^2\Big\rangle + \tilde{C}_9\Delta t\|D_h\tilde{x}^{n+1}\|_2^2$$

$$\le \Big\langle \frac{(S_h(x^{n-1},x^{n-2}))^{m-1}}{m(f_0(X))^{m-2}},(\tilde{x}^n)^2\Big\rangle\Big) + \Delta t\tilde{C}_4\|\tilde{x}^n\|_2^2 + (A_0+4(\epsilon_0^*)^{-4})\Delta t^2\|D_x\tilde{x}^n\|_2^2$$
$$+\Delta t\|\tau^n\|_2^2 + \frac{\tilde{C}_6}{2}\Delta t\|D_h\breve{\tilde{x}}^{n+1/2}\|_2^2 + \tilde{C}_{10}\Delta t\|D_h\tilde{x}^n\|_2^2. \tag{6.91}$$

Meanwhile, all other estimates (6.78)-(6.84) are still valid, then we arrive at

$$\begin{aligned}
&\frac{\tilde{C}_8}{2}\|\tilde{x}^{n+1}\|_2^2 + \tilde{C}_9\Delta t\|D_h\tilde{x}^{n+1}\|_2^2\\
&\le \frac{\tilde{C}_6}{2}\Delta t\|D_h\breve{\tilde{x}}^{n+1/2}\|_2^2 + \tilde{C}_{10}\Delta t\|D_h\tilde{x}^n\|_2^2 + O(\Delta t^2(\Delta t^3+h^3)^2)\\
&\le \tilde{C}_{11}\Delta t(\|D_h\tilde{x}^n\|_2^2 + \|D_h\tilde{x}^{n-1}\|_2^2) + O(\Delta t^2(\Delta t^3+h^3)^2)\\
&\le \tilde{C}_{12}\Delta t(\Delta t^3+h^3)^2,
\end{aligned} \tag{6.92}$$

with $\tilde{C}_{11} = \tilde{C}_9 + \frac{9\tilde{C}_6}{4}$, $\tilde{C}_{12} = C\tilde{C}_{11}\mathcal{C}^2 + 1$, provided that Δt and h are sufficiently small. This in turn results in a further rough estimate for $\tilde{x}^{n+1}$:

$$\|D_h\tilde{x}^{n+1}\|_2 \le \hat{C}_2(\Delta t^3+h^3), \quad \text{with } \hat{C}_2 := \Big(\frac{\tilde{C}_{12}}{\tilde{C}_9}\Big)^{\frac{1}{2}}. \tag{6.93}$$

As a consequence, an application of 1-D inverse inequality gives a sharper estimate for $\|D_h\tilde{x}^{n+1}\|_\infty$:

$$\|D_h\tilde{x}^{n+1}\|_\infty \le \frac{C\|D_h\tilde{x}^{n+1}\|_2}{h^{\frac{1}{2}}} \le \hat{C}_3(\Delta t^{\frac{5}{2}} + h^{\frac{5}{2}}), \quad \text{with } \hat{C}_3 = C\hat{C}_2, \tag{6.94}$$

$$\|D_hD_t^2\tilde{x}^n\|_\infty \le (\hat{C}_3+1)(\Delta t^{\frac{3}{2}} + h^{\frac{3}{2}}) \le \Delta t, \ D_t^2\tilde{x}^n := \frac{\tilde{x}^{n+1} - 2\tilde{x}^n + \tilde{x}^{n-1}}{\Delta t^2}, \tag{6.95}$$

in combination with (6.34) , under the same linear refinement requirement. This $\|\cdot\|_{W_h^{1,\infty}}$ bound for the second order temporal derivative will play a very important role in the refined error estimate.

6.4. The refined error estimate. Now we proceed with the refined error estimate. Taking a discrete inner product with (6.31) by $2\tilde{x}^{n+1/2} = \tilde{x}^{n+1} + \tilde{x}^n$ leads to

$$\begin{aligned}
&\Big\langle \frac{(S_h(x^n,x^{n-1}))^{m-1}}{m(f_0(X))^{m-2}}\cdot\frac{\tilde{x}^{n+1}-\tilde{x}^n}{\Delta t},\tilde{x}^{n+1}+\tilde{x}^n\Big\rangle + A_0\Delta t\langle D_h(\tilde{x}^{n+1}-\tilde{x}^n), D_h(\tilde{x}^{n+1}+\tilde{x}^n)\rangle\\
&\quad -\Big\langle f_0(X)\Big(\frac{\ln(D_hW^{n+1})-\ln(D_hW^n)}{D_h(W^{n+1}-W^n)} - \frac{\ln(D_hx^{n+1})-\ln(D_hx^n)}{D_h(x^{n+1}-x^n)}\Big), D_h(\tilde{x}^{n+1}+\tilde{x}^n)\Big\rangle\\
&= -\Big\langle \frac{(S_h(W^n,W^{n-1}))^{m-1} - (S_h(x^n,x^{n-1}))^{m-1}}{m(f_0(X))^{m-2}}\cdot\frac{W^{n+1}-W^n}{\Delta t}, 2\tilde{x}^{n+1/2}\Big\rangle\\
&\quad -\Delta t^2\Big\langle \frac{D_h\tilde{x}^{n+1}}{D_hW^{n+1}D_hx^{n+1}} - \frac{D_h\tilde{x}^n}{D_hW^nD_hx^n}, D_h(\tilde{x}^{n+1}+\tilde{x}^n)\Big\rangle + \langle\tau^n,\tilde{x}^{n+1}+\tilde{x}^n\rangle.
\end{aligned} \tag{6.96}$$

For the temporal derivative term, the equality $(\tilde{x}^{n+1}+\tilde{x}^n)(\tilde{x}^{n+1}-\tilde{x}^n) = (\tilde{x}^{n+1})^2 - (\tilde{x}^n)^2$ implies a similar estimate as in (6.52)-(6.53):

$$\begin{aligned}
&\Big\langle \frac{(S_h(x^n,x^{n-1}))^{m-1}}{m(f_0(X))^{m-2}}\cdot\frac{\tilde{x}^{n+1}-\tilde{x}^n}{\Delta t},\tilde{x}^{n+1}+\tilde{x}^n\Big\rangle\\
&= \frac{1}{\Delta t}\Big(\Big\langle \frac{(S_h(x^n,x^{n-1}))^{m-1}}{m(f_0(X))^{m-2}},(\tilde{x}^{n+1})^2\Big\rangle - \Big\langle \frac{(S_h(x^n,x^{n-1}))^{m-1}}{m(f_0(X))^{m-2}},(\tilde{x}^n)^2\Big\rangle\Big)\\
&\ge \frac{1}{\Delta t}\Big(\Big\langle \frac{(S_h(x^n,x^{n-1}))^{m-1}}{m(f_0(X))^{m-2}},(\tilde{x}^{n+1})^2\Big\rangle - \Big\langle \frac{(S_h(x^{n-1},x^{n-2}))^{m-1}}{m(f_0(X))^{m-2}},(\tilde{x}^n)^2\Big\rangle\Big) - \tilde{C}_4\|\tilde{x}^n\|_2^2.
\end{aligned} \tag{6.97}$$

For the second term on the left-hand side, the second term on the right-hand side of (6.96), and the local truncation error terms, the following bounds could be similarly derived:

$$A_0\Delta t\langle D_h(\tilde{x}^{n+1}-\tilde{x}^n), D_h(\tilde{x}^{n+1}+\tilde{x}_n)\rangle = A_0\Delta t(\|D_h\tilde{x}^{n+1}\|_2^2-\|D_h\tilde{x}^n\|_2), \tag{6.98}$$

$$2\langle\Delta t^n,\tilde{x}^{n+1}+\tilde{x}^n\rangle \le \|\Delta t^n\|_2^2+\|\tilde{x}^{n+1/2}\|_2^2 \le \|\Delta t^n\|_2^2+\frac{1}{2}(\|\tilde{x}^{n+1}\|_2^2+\|\tilde{x}^n\|_2^2), \tag{6.99}$$

$$\begin{aligned}
&-\Delta t^2\Big\langle \frac{D_h\tilde{x}^{n+1}}{D_hW^{n+1}D_hx^{n+1}}-\frac{D_h\tilde{x}^n}{D_hW^nD_hx^n}, D_h(\tilde{x}^{n+1}+\tilde{x}^n)\Big\rangle\\
&\le 2\Delta t^2\cdot\frac{1}{\frac{1}{2}(\epsilon_0^*)^2}\|D_h\tilde{x}^n\|_2\cdot\|D_h\tilde{x}^{n+1}\|_2-\Delta t^2\cdot\frac{1}{\frac{1}{2}(\epsilon_0^*)^2}\|D_h\tilde{x}^n\|_2^2\\
&\le 2\Delta t^2(\epsilon_0^*)^{-2}(\|D_h\tilde{x}^{n+1}\|_2^2+\|D_h\tilde{x}^n\|_2^2)+2\Delta t^2(\epsilon_0^*)^{-2}\|D_h\tilde{x}^n\|_2^2\\
&\le 2\Delta t^2(\epsilon_0^*)^{-2}(\|D_h\tilde{x}^{n+1}\|_2^2+2\|D_h\tilde{x}^n\|_2^2),
\end{aligned}\tag{6.100}$$

in which the separation property (6.25) (for the constructed profile W), the preliminary estimates (6.35), (6.88) (for x^n and x^{n+1}, respectively), have been used in (6.100).

For the first term on the right-hand side of (6.96), we cannot count on the estimate (6.58), since there is no stability control for $\|D_h\breve{\tilde{x}}^{n+1/2}\|_2^2=\|D_h(\frac{3}{2}\tilde{x}^n-\frac{1}{2}\tilde{x}^{n-1})\|_2^2$. To overcome this difficulty, a summation by parts formula is applied, due to the fact that $\tilde{x}_0^{n+1/2}=\tilde{x}_M^{n+1/2}=0$:

$$\begin{aligned}
&-2\Big\langle\frac{(S_h(W^n,W^{n-1}))^{m-1}-(S_h(x^n,x^{n-1}))^{m-1}}{m(f_0(X))^{m-2}}\cdot\frac{W^{n+1}-W^n}{\Delta t},\tilde{x}^{n+1/2}\Big\rangle\\
&=-2\Big\langle\frac{\mathcal{N}^{n+1/2}\widetilde{D}_h\breve{\tilde{x}}^{n+1/2}}{m(f_0(X))^{m-2}}\cdot\frac{W^{n+1}-W^n}{\Delta t},\tilde{x}^{n+1/2}\Big\rangle\\
&=2\Big\langle\widetilde{D}_h\Big(\frac{\mathcal{N}^{n+1/2}\frac{W^{n+1}-W^n}{\Delta t}}{m(f_0(X))^{m-2}}\cdot\tilde{x}^{n+1/2}\Big),\breve{\tilde{x}}^{n+1/2}\Big\rangle.
\end{aligned}\tag{6.101}$$

Meanwhile, the following observation is made in the finite difference space:

$$\begin{aligned}
&\Big\|\widetilde{D}_h\Big(\frac{\mathcal{N}^{n+1/2}\frac{W^{n+1}-W^n}{\Delta t}}{m(f_0(X))^{m-2}}\cdot\tilde{x}^{n+1/2}\Big)\Big\|_2\le\Big\|D_h\Big(\frac{\mathcal{N}^{n+1/2}\frac{W^{n+1}-W^n}{\Delta t}}{m(f_0(X))^{m-2}}\cdot\tilde{x}^{n+1/2}\Big)\Big\|_2\\
&\le\frac{1}{m}(\|\tilde{x}^{n+1/2}\|_2+\|D_h\tilde{x}^{n+1/2}\|_2)\cdot(\|\mathcal{N}^{n+1/2}\|_\infty+\|D_h\mathcal{N}^{n+1/2}\|_\infty)\\
&\quad\cdot\Big(\|\frac{W^{n+1}-W^n}{\Delta t}\|_\infty+\|\frac{D_h(W^{n+1}-W^n)}{\Delta t}\|_\infty\Big)\\
&\quad\cdot\Big(\|\frac{1}{(f_0(X))^{m-2}}\|_\infty+\|D_h\Big(\frac{1}{(f_0(X))^{m-2}}\Big)\|_\infty\Big)\\
&\le\frac{1}{m}(\tilde{C}_2+\tilde{C}_3)\tilde{C}_{13}C^*(\|\tilde{x}^{n+1/2}\|_2+\|D_h\tilde{x}^{n+1/2}\|_2),
\end{aligned}\tag{6.102}$$

$$\text{with } \tilde{C}_{13}:=\|\frac{1}{(f_0(X))^{m-2}}\|_\infty+\|D_h\Big(\frac{1}{(f_0(X))^{m-2}}\Big)\|_\infty,$$

in which the preliminary $W_h^{1,\infty}$ estimates (6.40), (6.41), and the regularity assumption (6.26) have been repeatedly used in the derivation. Then we arrive at

$$\begin{aligned}
&-2\Big\langle\frac{(S_h(W^n,W^{n-1}))^{m-1}-(S_h(x^n,x^{n-1}))^{m-1}}{m(f_0(X))^{m-2}}\cdot\frac{W^{n+1}-W^n}{\Delta t},\tilde{x}^{n+1/2}\Big\rangle\\
&\le 2\tilde{C}_{14}(\|\tilde{x}^{n+1/2}\|_2+\|D_h\tilde{x}^{n+1/2}\|_2)\|\breve{\tilde{x}}^{n+1/2}\|_2,\quad\Big(\text{with } \tilde{C}_{14}=\frac{1}{m}(\tilde{C}_2+\tilde{C}_3)\tilde{C}_9C^*\Big)\\
&\le\Big(\tilde{C}_{14}+\tilde{C}_{14}^2(\tilde{C}^*)^2\Big)\|\breve{\tilde{x}}^{n+1/2}\|_2^2+\tilde{C}_{14}\|\tilde{x}^{n+1/2}\|_2^2+(\tilde{C}^*)^{-2}\|D_h\tilde{x}^{n+1/2}\|_2^2.
\end{aligned}\tag{6.103}$$

Again, the rest of the work is focused on the error analysis associated with the nonlinear diffusion part, as given by (6.59). However, the point-wise estimate (6.64), (6.65) is not useful in the refined analysis any more. Instead, we begin with the following application of higher order Taylor expansion for $\ln x$, around $\frac{x+x_0}{2}$:

$$\frac{\ln x - \ln x_0}{x-x_0} = \frac{1}{\frac{x+x_0}{2}} + \frac{2}{3(\frac{x+x_0}{2})^3} \cdot \frac{(x-x_0)^2}{8} + \Big(\frac{1}{5(\xi^{(5)})^5} + \frac{1}{5(\xi^{(6)})^5} \Big) \cdot \frac{(x-x_0)^4}{32}, \tag{6.104}$$

with $\xi^{(5)}$ between $\frac{x+x_0}{2}$ and x, $\xi^{(6)}$ between $\frac{x+x_0}{2}$ and x_0. This in turn gives

$$\begin{aligned}\frac{\ln(D_h W^{n+1}) - \ln(D_h W^n)}{D_h(W^{n+1}-W^n)} &= \frac{1}{D_h(\frac{W^{n+1}+W^n}{2})} + \frac{2}{3(D_h \frac{W^{n+1}+W^n}{2})^3} \cdot \frac{(D_h(W^{n+1}-W^n))^2}{8} \\ &\quad + \Big(\frac{1}{5(\eta^{(1)})^5} + \frac{1}{5(\eta^{(2)})^5} \Big) \cdot \frac{(D_h(W^{n+1}-W^n))^4}{32}, \end{aligned} \tag{6.105}$$

$$\begin{aligned}\frac{\ln(D_h x^{n+1}) - \ln(D_h x^n)}{D_h(x^{n+1}-x^n)} &= \frac{1}{D_h(\frac{x^{n+1}+x^n}{2})} + \frac{2}{3(D_h \frac{x^{n+1}+x^n}{2})^3} \cdot \frac{(D_h(x^{n+1}-x^n))^2}{8} \\ &\quad + \Big(\frac{1}{5(\eta^{(3)})^5} + \frac{1}{5(\eta^{(4)})^5} \Big) \cdot \frac{(D_h(x^{n+1}-x^n))^4}{32}, \end{aligned} \tag{6.106}$$

with

$$\eta^{(1)} \text{ between } D_h(\frac{W^{n+1}+W^n}{2}) \text{ and } D_h W^{n+1}, \quad \eta^{(2)} \text{ between } D_h(\frac{W^{n+1}+W^n}{2}) \text{ and } D_h W^n,$$
$$\eta^{(3)} \text{ between } D_h(\frac{x^{n+1}+x^n}{2}) \text{ and } D_h x^{n+1}, \quad \eta^{(4)} \text{ between } D_h(\frac{x^{n+1}+x^n}{2}) \text{ and } D_h x^n.$$

Then we arrive at a decomposition for $\mathcal{NLE}^{n+1/2}$:

$$\mathcal{NLE}^{n+1/2} = \mathcal{NL}^{(1)} + \mathcal{NL}^{(2)} + \mathcal{NL}^{(3)}, \quad \text{with}$$

$$\mathcal{NL}^{(1)} = -\frac{1}{D_h(\frac{W^{n+1}+W^n}{2})} + \frac{1}{D_h(\frac{x^{n+1}+x^n}{2})} = \frac{D_h \tilde{x}^{n+1/2}}{D_h(\frac{W^{n+1}+W^n}{2}) \cdot D_h(\frac{x^{n+1}+x^n}{2})}, \tag{6.107}$$

$$\mathcal{NL}^{(2)} = \frac{1}{12} \Big(-\frac{(D_h(W^{n+1}-W^n))^2}{(D_h \frac{W^{n+1}+W^n}{2})^3} + \frac{(D_h(x^{n+1}-x^n))^2}{(D_h \frac{x^{n+1}+x^n}{2})^3} \Big), \tag{6.108}$$

$$\begin{aligned}\mathcal{NL}^{(3)} &= \frac{1}{160} \Big(-\Big(\frac{1}{(\eta^{(1)})^5} + \frac{1}{(\eta^{(2)})^5} \Big) (D_h(W^{n+1}-W^n))^4 \\ &\quad + \Big(\frac{1}{(\eta^{(3)})^5} + \frac{1}{(\eta^{(4)})^5} \Big) (D_h(x^{n+1}-x^n))^4 \Big). \end{aligned} \tag{6.109}$$

For the leading expansion $\mathcal{NL}^{(1)}$, the following nonlinear estimate is available:

$$\begin{aligned}\langle \mathcal{NL}^{(1)}, 2D_h \tilde{x}^{n+1/2} \rangle &= \Big\langle \frac{2}{D_h(\frac{W^{n+1}+W^n}{2}) \cdot D_h(\frac{x^{n+1}+x^n}{2})}, (D_h \tilde{x}^{n+1/2})^2 \Big\rangle \\ &\geq \frac{2}{(\tilde{C}^*)^2} \|D_h \tilde{x}^{n+1/2}\|_2^2, \end{aligned} \tag{6.110}$$

with repeated applications of (6.40), (6.41) and (6.26). The second expansion $\mathcal{NL}^{(2)}$ could have a further decomposition: $\mathcal{NL}^{(2)} = \mathcal{NL}^{(2),1} + \mathcal{NL}^{(2),2}$, with

$$\mathcal{NL}^{(2),1} = -\frac{D_h(W^{n+1}-W^n+x^{n+1}-x^n) \cdot D_h(\tilde{x}^{n+1}-\tilde{x}^n)}{(D_h W^{n+1/2})^3}, \tag{6.111}$$

$$\mathcal{NL}^{(2),2} = \mathcal{NLC}^{(2)} \cdot D_h \tilde{x}^{n+1/2} \cdot (D_h(x^{n+1}-x^n))^2, \tag{6.112}$$

$$\mathcal{NLC}^{(2)} = \frac{(D_h W^{n+1/2})^2 + (D_h W^{n+1/2})(D_h x^{n+1/2}) + (D_h x^{n+1/2})^2)}{(D_h W^{n+1/2})^3 (D_h x^{n+1/2})^3}. \tag{6.113}$$

We notice a bound for the nonlinear coefficient $\mathcal{NLC}^{(2)}$:

$$|\mathcal{NLC}^{(2)}| \le \frac{3(\tilde{C}^*)^2}{\frac{1}{2}(\epsilon_0^*)^2} = 6(\tilde{C}^*)^2(\epsilon_0^*)^{-2} := \tilde{C}_{15}, \quad \text{by (6.25), (6.26), (6.35), (6.88).} \tag{6.114}$$

This in turn yields an estimate for the term associated with $\mathcal{NL}^{(2),2}$:

$$\langle \mathcal{NL}^{(2),1}, 2D_h\tilde{x}^{n+1/2}\rangle \ge -2\tilde{C}_{16}\Delta t\|D_h\tilde{x}^{n+1/2}\|_2^2, \quad \text{with } \tilde{C}_{16} := \tilde{C}_{15}(C^*+1)^2, \tag{6.115}$$

in which a point-wise bound $\|D_x(x^{n+1}-x^n)\|_\infty \le (C^*+1)\Delta t$ comes from (6.26), (6.35) and (6.87). For $\mathcal{NL}^{(2),1}$, we introduce the following discrete function

$$\gamma^{n+1/2} := -\frac{\frac{D_h(W^{n+1}-W^n)}{\Delta t} + \frac{D_h(x^{n+1}-x^n)}{\Delta t}}{(D_hW^{n+1/2})^3}, \quad \text{so that} \tag{6.116}$$

$$\begin{aligned}\langle \mathcal{NL}^{(2),1}, D_h(\tilde{x}^{n+1}+\tilde{x}^n)\rangle &= \Delta t\langle \gamma^{n+1/2}, (D_h\tilde{x}^{n+1})^2 - (D_h\tilde{x}^n))^2\rangle \\ &= \Delta t(\langle \gamma^{n+1/2}, (D_h\tilde{x}^{n+1})^2\rangle - \langle \gamma^{n-1/2}, (D_h\tilde{x}^n))^2\rangle) \\ &\quad -\Delta t\langle \gamma^{n+1/2} - \gamma^{n-1/2}, (D_h\tilde{x}^n))^2\rangle.\end{aligned} \tag{6.117}$$

For the last correction term, the following observation is made:

$$\begin{aligned}\gamma^{n+1/2} - \gamma^{n-1/2} = &-\frac{\Delta t D_h(D_t^2W^n) + \Delta t D_h(D_t^2x^n)}{(D_hW^{n+1/2})^3} \\ &+\mathcal{NLC}^{(2)} \cdot \Big(\frac{D_h(W^{n+1}-W^n)}{\Delta t} + \frac{D_h(x^{n+1}-x^n)}{\Delta t}\Big) \cdot D_h\tilde{x}^{n+1/2}.\end{aligned} \tag{6.118}$$

Meanwhile, a $\|\cdot\|_{W_h^{1,\infty}}$ bound for $D_t^2x^n$ is available, as a result of the further rough estimate (6.95) and the regularity assumption (6.26) for W:

$$\|D_h(D_t^2x^n)\|_\infty \le \|D_h(D_t^2W^n)\|_\infty + \|D_h(D_t^2\tilde{x}^n)\|_\infty \le C^* + \frac{\epsilon_0^*}{2} = \tilde{C}^*. \tag{6.119}$$

This in turn implies an $O(\Delta t)$ estimate for the first part, in combination with (6.25)

$$\Delta t\Big|\frac{D_h(D_t^2W^n) + D_h(D_t^2x^n)}{(D_hW^{n+1/2})^3}\Big| \le \frac{C^*+\tilde{C}^*}{(\epsilon_0^*)^3}\Delta t. \tag{6.120}$$

A similar bound for the second part is also available, which comes from (6.26), (6.34), (6.94) and (6.114):

$$\begin{aligned}&\Big|\mathcal{NLC}^{(2)} \cdot \Big(\frac{D_h(W^{n+1}-W^n)}{\Delta t} + \frac{D_h(x^{n+1}-x^n)}{\Delta t}\Big) \cdot D_h\tilde{x}^{n+1/2}\Big| \\ &\le (2C^*+1)\tilde{C}_{15}(\hat{C}+\hat{C}_3)(\Delta t^{\frac{5}{2}} + h^{\frac{5}{2}}) \le \Delta t.\end{aligned} \tag{6.121}$$

Therefore, an $O(\Delta t)$ bound for $\gamma^{n+1/2} - \gamma^{n-1/2}$ is obtained:

$$\|\gamma^{n+1/2} - \gamma^{n-1/2}\|_\infty \le \tilde{C}_{17}\Delta t, \quad \text{with } \tilde{C}_{17} := (C^*+\tilde{C}^*)(\epsilon_0^*)^{-3} + 1, \tag{6.122}$$

so that the nonlinear inner product associated with $\mathcal{NL}^{(2),1}$ could be analyzed as follows:

$$\begin{aligned}\langle \mathcal{NL}^{(2),1}, D_h(\tilde{x}^{n+1}+\tilde{x}^n)\rangle \ge &\Delta t(\langle \gamma^{n+1/2}, (D_h\tilde{x}^{n+1})^2\rangle - \langle \gamma^{n-1/2}, (D_h\tilde{x}^n))^2\rangle) \\ &-\tilde{C}_{17}\Delta t^2\|D_h\tilde{x}^n\|_2^2.\end{aligned} \tag{6.123}$$

Its combination with (6.115) yields the nonlinear estimate for $\mathcal{NL}^{(2)}$:

$$\begin{aligned}\langle \mathcal{NL}^{(2)}, D_h(\tilde{x}^{n+1}+\tilde{x}^n)\rangle \ge &\Delta t(\langle \gamma^{n+1/2}, (D_h\tilde{x}^{n+1})^2\rangle - \langle \gamma^{n-1/2}, (D_h\tilde{x}^n))^2\rangle) \\ &-\tilde{C}_{18}\Delta t^2(\|D_h\tilde{x}^{n+1}\|_2^2 + \|D_h\tilde{x}^n\|_2^2), \quad \text{with } \tilde{C}_{18} = \tilde{C}_{16} + \tilde{C}_{17}\ .\end{aligned} \tag{6.124}$$

The analysis for $\mathcal{NL}^{(3)}$ is similar for that of $\mathcal{NL}^{(2),2}$. We are able to obtain the following estimate; the technical details are skipped for the sake of brevity.

$$\langle \mathcal{NL}^{(3)}, D_h(\tilde{x}^{n+1}+\tilde{x}^n)\rangle \geq -\tilde{C}_{19}\Delta t^2(\|D_h\tilde{x}^{n+1}\|_2^2+\|D_h\tilde{x}^n\|_2^2)-(\Delta t^4+h^4)^2. \quad (6.125)$$

A combination of (6.110), (6.124) and (6.125) results in an estimate for $\mathcal{NLE}^{n+1/2}$:

$$\begin{aligned}&\langle \mathcal{NLE}^{n+1/2}, 2D_h\tilde{x}^{n+1/2}\rangle\\ \geq{}& \frac{2}{(\tilde{C}^*)^2}\|D_h\tilde{x}^{n+1/2}\|_2^2+\Delta t(\langle\gamma^{n+1/2},(D_h\tilde{x}^{n+1})^2\rangle-\langle\gamma^{n-1/2},(D_h\tilde{x}^n))^2\rangle)\\ &-\tilde{C}_{20}\Delta t^2(\|D_h\tilde{x}^{n+1}\|_2^2+\|D_h\tilde{x}^n\|_2^2)-(\Delta t^4+h^4)^2, \end{aligned} \quad (6.126)$$

with $\tilde{C}_{20}=\tilde{C}_{18}+\tilde{C}_{19}$.

Finally, a substitution of (6.97)-(6.100), (6.103) and (6.126) into (6.96) results in

$$\begin{aligned}&\Big\langle\frac{(S_h(x^n,x^{n-1}))^{m-1}}{m(f_0(X))^{m-2}},(\tilde{x}^{n+1})^2\Big\rangle+\frac{1}{(\tilde{C}^*)^2}\Delta t\|D_h\tilde{x}^{n+1/2}\|_2^2\\ &+A_0\Delta t^2(\|D_h\tilde{x}^{n+1}\|_2^2-\|D_h\tilde{x}^n\|_2^2)+\Delta t^2(\langle\gamma^{n+1/2},(D_h\tilde{x}^{n+1})^2\rangle-\langle\gamma^{n-1/2},(D_h\tilde{x}^n))^2\rangle)\\ \leq{}&\Big\langle\frac{(S_h(x^{n-1},x^{n-2}))^{m-1}}{m(f_0(X))^{m-2}},(\tilde{x}^n)^2\Big\rangle\Big)+\tilde{C}_{22}\Delta t\|\tilde{x}^{n+1}\|_2^2+\tilde{C}_{23}\Delta t\|\tilde{x}^n\|_2^2+\frac{\tilde{C}_{21}}{2}\Delta t\|\tilde{x}^{n-1}\|_2^2\\ &+\tilde{C}_{24}\Delta t^3(\|D_h\tilde{x}^{n+1}\|_2^2+\|D_h\tilde{x}^n\|_2^2)+\Delta t(\|\tau^n\|_2^2+(\Delta t^4+h^4)^2),\end{aligned} \quad (6.127)$$

with $\tilde{C}_{21}=\tilde{C}_{14}+\tilde{C}_{14}^2(\tilde{C}^*)^2$, $\tilde{C}_{22}=\frac{\tilde{C}_{14}}{2}+\frac{\tilde{C}_6}{2}+2$, $\tilde{C}_{23}=\tilde{C}_4+\frac{1}{2}+\frac{\tilde{C}_{14}}{2}+\frac{9\tilde{C}_{21}}{2}$, $\tilde{C}_{24}=\tilde{C}_{20}+4(\epsilon_0^*)^{-2}$. Subsequently, a summation in time gives

$$\begin{aligned}&\Big\langle\frac{(S_h(x^n,x^{n-1}))^{m-1}}{m(f_0(X))^{m-2}},(\tilde{x}^{n+1})^2\Big\rangle+\frac{1}{(\tilde{C}^*)^2}\Delta t\sum_{k=0}^{n}\|D_h\tilde{x}^{k+1/2}\|_2^2\\ &+\Delta t^2\Big(A_0\|D_h\tilde{x}^{n+1}\|_2^2+\langle\gamma^{n+1/2},(D_h\tilde{x}^{n+1})^2\rangle\Big)\\ \leq{}&\tilde{C}_{25}\Delta t\sum_{k=0}^{n+1}\|\tilde{x}^k\|_2^2+2\tilde{C}_{24}\Delta t^3\sum_{k=0}^{n+1}\|D_h\tilde{x}^k\|_2^2+C(\Delta t^4+h^4)^2,\end{aligned} \quad (6.128)$$

with $\tilde{C}_{25}=\tilde{C}_{22}+\tilde{C}_{23}+\frac{\tilde{C}_{21}}{2}$. Meanwhile, by the definition of $\gamma^{n+1/2}$ (6.116), we have

$$\|\gamma^{n+1/2}\|_\infty\leq\frac{2C^*+1}{(\epsilon_0^*)^3},\quad \text{by (6.26), (6.35), (6.87).} \quad (6.129)$$

In turn, by taking $A_0=(2C^*+1)(\epsilon_0^*)^{-3}+1$, and making use of the inequality (6.78), we obtain

$$\begin{aligned}&\tilde{C}_8\|\tilde{x}^{n+1}\|_2^2+\frac{1}{(\tilde{C}^*)^2}\Delta t\sum_{k=0}^{n}\|D_h\tilde{x}^{k+1/2}\|_2^2+\Delta t^2\|D_h\tilde{x}^{n+1}\|_2^2\\ \leq{}&\tilde{C}_{25}\Delta t\sum_{k=0}^{n+1}\|\tilde{x}^k\|_2^2+2\tilde{C}_{24}\Delta t^3\sum_{k=0}^{n+1}\|D_h\tilde{x}^k\|_2^2+C(\Delta t^4+h^4)^2.\end{aligned} \quad (6.130)$$

Therefore, an application of discrete Gronwall inequality (in the integral form) leads to the desired higher order convergence estimate

$$\|\tilde{x}^{n+1}\|_2+\Big((\tilde{C}^*)^{-2}\Delta t\sum_{m=0}^{n}\|\frac{1}{2}D_h(\tilde{x}^{m+1}+\tilde{x}^m)\|_2^2\Big)^{1/2}\leq\hat{C}_4(\Delta t^4+h^4). \quad (6.131)$$

This completes the refined error estimate.

6.5. Recovery of the a-priori assumption (6.32). With the higher order error estimate (6.131) at hand, we conclude that the a-priori assumption in (6.32) is satisfied at the next time step t^{n+1}, since $\hat{C}_4$ takes the following form:

$$\hat{C}_4 \le C\exp\Big(\frac{(\tilde{C}_{25}+2\tilde{C}_{24})t^{n+1}}{\min(\tilde{C}_8,1)}\Big) \le \hat{C} := C\exp\Big(\frac{(\tilde{C}_{25}+2\tilde{C}_{24})T}{\min(\tilde{C}_8,1)}\Big). \tag{6.132}$$

We also notice that $\tilde{C}_8$, $\tilde{C}_{24}$ and $\tilde{C}_{25}$ are independent of $\mathcal{C}$. Therefore, the a-priori assumption in (6.32) is satisfied, so that an induction analysis could be applied. This finishes the higher order convergence analysis.

Finally, the convergence estimate (6.3) is a direct consequence of (6.131), combined with the definition (6.4) of the constructed approximate solution W. This completes the proof of Theorem 6.1.

7. Convergence analysis of Newton's iteration

In this section, we prove the convergence of damped Newton's iteration (3.11) in the convex set $\mathcal{Q}$, based on *self-concordant* [30]. The definition of *self-concordant* is given as the following:

DEFINITION 7.1. *Let $\mathcal{G}$ be a finite-dimensional real vector space, $\mathcal{Q}$ be an open nonempty convex subset of $\mathcal{G}$, $\Lambda : \mathcal{Q} \to \mathbb{R}$ be a function, $a > 0$. Λ is called self-concordant on $\mathcal{Q}$ with the parameter value a, if $\Lambda \in C^3$ is a convex function on $\mathcal{Q}$, and, for all $x \in \mathcal{Q}$ and all $u \in \mathcal{G}$, the following inequality holds:*

$$|D^3\Lambda(x)[u,u,u]| \le 2a^{-1/2}(D^2\Lambda(x)[u,u])^{3/2}$$

($D^k\Lambda(x)[u_1,\cdots,u_k]$ henceforth denotes the value of the k-th differential of Λ taken at x along the collection of directions $u_1,\cdots,u_k$) [30].

The *self-concordant* function has two typical characteristics [30]:

- Linear and (convex) quadratic functions are evidently self-concordant, since they have zero third derivative.
- A function $f : \mathcal{R}^n \to R$ is self-concordant if it is self-concordant along every line in its domain.

Then we review the definition of F, given by (4.3)-(4.6), in which $G(x,x_0)$ turns out to be the primitive function of $-\frac{\ln(1+x)-\ln x_0}{1+x-x_0}$, for a fixed x_0, as formulated in (4.7). The convexity of F_1, F_3 and F_4 (in terms of $\hat{x}$) is obvious. For the functional F_2, the inequality (4.8) is valid, in which the convexity of $-\ln(1+x)$ has been used. This fact implies the convexity of F_2. Therefore, we conclude that F is convex in terms of $\hat{x}$, provided that $D_h\hat{x} > -1$ at a point-wise level. Furthermore, F is strictly convex, because of the strict convexity of F_1.

THEOREM 7.1. *Suppose $f_0(X) \in \mathcal{E}_N$ is the initial distribution with a positive lower bound for $X \in \mathcal{Q}$ and $\sqrt{a} := \min\limits_{0 \le i \le M} f_0(X_i)\frac{hC_{Newton}}{2}$ with a positive constant C_{Newton}, then $F(\hat{x})$, defined in (4.3)-(4.6), is a self-concordant function and Newton's iteration is convergent in $\mathcal{Q}$.*

Proof. Since linear and quadratic functions have zero third derivative, $F_1(\hat{x})$ and $F_3(\hat{x})$ are self-concordant. We just need to prove $F_2(\hat{x})$ and $F_4(\hat{x})$ are self-concordant along every line in $\mathcal{Q}$.

Suppose $x^n \in \mathcal{Q}$ and let

$$\xi_{i-\frac{1}{2}} := \frac{D_h x^n_{i-\frac{1}{2}}}{1+D_h\hat{x}_{i-\frac{1}{2}}}, i = 1,\cdots,M.$$

Then $\xi_{i-\frac{1}{2}} > 0$, $i = 1,\cdots,M$. For $\forall i = 1,\cdots,M$, we can obtain

$$\frac{\partial F_2(\hat{x}_i)}{\partial \hat{x}_i} = \frac{f_0(X_{i-\frac{1}{2}})}{1+D_h\hat{x}_{i-\frac{1}{2}} - D_h x^n_{i-\frac{1}{2}}}\ln(\xi_{i-\frac{1}{2}}) - \frac{f_0(X_{i+\frac{1}{2}})}{1+D_h\hat{x}_{i+\frac{1}{2}} - D_h x^n_{i+\frac{1}{2}}}\ln(\xi_{i+\frac{1}{2}}),$$

$$\frac{\partial^2 F_2(\hat{x}_i)}{\partial \hat{x}_i^2} = -\frac{f_0(X_{i-\frac{1}{2}})}{h\Big(1+D_h\hat{x}_{i-\frac{1}{2}}-D_hx^n_{i-\frac{1}{2}}\Big)^2}[(1-\xi_{i-\frac{1}{2}})+\ln(\xi_{i-\frac{1}{2}})]$$
$$-\frac{f_0(X_{i+\frac{1}{2}})}{h\Big(1+D_h\hat{x}_{i+\frac{1}{2}}-D_hx^n_{i+\frac{1}{2}}\Big)^2}[(1-\xi_{i+\frac{1}{2}})+\ln(\xi_{i+\frac{1}{2}})],$$

$$\Big|\frac{\partial^3 F_2(\hat{x}_i)}{\partial \hat{x}_i^3}\Big| = \Big|\frac{2f_0(X_{i-\frac{1}{2}})}{h^2\Big(1+D_h\hat{x}_{i-\frac{1}{2}}-D_hx^n_{i-\frac{1}{2}}\Big)^3}\Big[(1-\xi_{i-\frac{1}{2}})+\ln(\xi_{i-\frac{1}{2}})+\frac{1}{2}(1-\xi_{i-\frac{1}{2}})^2\Big]$$
$$-\frac{2f_0(X_{i+\frac{1}{2}})}{h^2\Big(1+D_h\hat{x}_{i+\frac{1}{2}}-D_hx^n_{i+\frac{1}{2}}\Big)^3}\Big[(1-\xi_{i+\frac{1}{2}})+\ln(\xi_{i+\frac{1}{2}})+\frac{1}{2}(1-\xi_{i+\frac{1}{2}})^2\Big]\Big|$$
$$\leq \Big|\frac{2f_0(X_{i-\frac{1}{2}})}{h^2\Big(1+D_h\hat{x}_{i-\frac{1}{2}}-D_hx^n_{i-\frac{1}{2}}\Big)^3}\Big[(1-\xi_{i-\frac{1}{2}})+\ln(\xi_{i-\frac{1}{2}})+\frac{1}{2}(1-\xi_{i-\frac{1}{2}})^2\Big]$$
$$+\frac{2f_0(X_{i+\frac{1}{2}})}{h^2\Big(1+D_h\hat{x}_{i+\frac{1}{2}}-D_hx^n_{i+\frac{1}{2}}\Big)^3}\Big[(1-\xi_{i+\frac{1}{2}})+\ln(\xi_{i+\frac{1}{2}})+\frac{1}{2}(1-\xi_{i+\frac{1}{2}})^2\Big]\Big|.$$

Notice that

$$\ln t = \ln(1+(t-1)) = (t-1) - \frac{1}{2}(t-1)^2 + O(t-1)^3, \ \forall t>0,$$

hence there exists a constant $C_{Newton}>0$ such that

$$\Big[(1-\xi_{i-\frac{1}{2}})+\ln(\xi_{i-\frac{1}{2}})+\frac{1}{2}(1-\xi_{i-\frac{1}{2}})^2\Big]^2 \leq C_{Newton}\Big[-(1-\xi_{i-\frac{1}{2}})-\ln(\xi_{i-\frac{1}{2}})\Big]^3.$$

If the parameter $\sqrt{a} := \min_{0\leq i\leq M} f_0(X_i))\frac{hC_{Newton}}{2}$, we obtain

$$\left|\frac{2f_0(X_{i-\frac{1}{2}})}{h^2\Big(1+D_h\hat{x}_{i-\frac{1}{2}}-D_hx^n_{i-\frac{1}{2}}\Big)^3}\Big[(1-\xi_{i-\frac{1}{2}})+\ln(\xi_{i-\frac{1}{2}})+\frac{1}{2}(1-\xi_{i-\frac{1}{2}})^2\Big]\right|^2$$
$$\leq 2a^{-\frac{1}{2}}\left(-\frac{f_0(X_{i-\frac{1}{2}})}{h\Big(1+D_h\hat{x}_{i-\frac{1}{2}}-D_hx^n_{i-\frac{1}{2}}\Big)^2}\Big[(1-\xi_{i-\frac{1}{2}})+\ln(\xi_{i-\frac{1}{2}})\Big]\right)^3, \forall i=1,\cdots,M. \tag{7.1}$$

So $F_2(\hat{x})$ is self-concordant. By the same method, $F_4(\hat{x})$ is also self-concordant. Based on **Theorem 2.2.3** in [30], Newton's iteration is convergent in $\mathcal{Q}$. ☐

8. Numerical results

In this section, we present an example with a positive state to demonstrate the convergence rate of the numerical scheme.

Before that, we define the error of a numerical solution measured in the $\mathcal{L}^2$ and $\mathcal{L}^\infty$ norms as:

$$\|e_h\|_2^2 = \frac{1}{2}\left(e_{h_0}^2 h_{x_0} + \sum_{i=1}^{M-1} e_{h_i}^2 h_{x_i} + e_{h_M}^2 h_{x_M}\right), \tag{8.1}$$

and

$$\|e_h\|_\infty = \max_{0\leq i\leq M}\{|e_{h_i}|\}, \tag{8.2}$$

where $e_h = (e_{h_0}, e_{h_1}, \cdots, e_{h_M})$ and for the error of the density $f-f_h$,

$$h_{x_i} = x_{i+1}-x_{i-1}, \ \ 1\leq i\leq M-1; \quad h_{x_0} = x_1 - x_0; \ \ h_{x_M} = x_M - x_{M-1},$$

and for the error of the trajectory $x-x_h$,

$$h_{x_i}=2h,\quad 1\le i\le M-1,\quad h_{x_0}=h_{x_M}=h,$$

where h is the spatial step.

Example 1 (The problem with a positive initial state). Consider the problem (2.2)-(2.5) in dimension one with a smooth positive initial data

$$f_0(x)=0.5-(x-0.5)^2,\quad x\in\Omega:=(0,1). \tag{8.3}$$

Firstly, the trajectory Equation (2.11) with the initial and boundary condition (2.12)-(2.13) can be solved by the fully discrete scheme (3.10), and then the density function f in (2.8) can be approximated by (3.12). The reference "exact" solution is obtained numerically on a much finer mesh with $h=\frac{1}{10000}$, $\Delta t=\frac{1}{10000}$. By Theorem 7.1, the convergence of Newton's iteration is ensured for any choice of Δt and h. Meanwhile, we take $\Delta t=h$ in the practical computations for simplicity of presentation.

Table 8.1 shows the second order convergence for density f and trajectory x in the $\mathcal{L}^2$ and $\mathcal{L}^\infty$ norm with both $m=\frac{5}{3}$ and $m=3$ at time $t=0.05$. That verifies the optimal convergence rate of the numerical scheme. Figure 8.1 (a) and (b) present the density f for both values of m at time $t=0.05$ and $t=0.1$, respectively. That implies that the speed of diffusion increases as m increases. The reason is that the density f is bigger than 1 such that the diffusion become larger with the growth of m. Figure 8.2 (a) displays the decay of total energy and Figure 8.2 (b) shows that particles move outward at a finite speed for $m=2$.

				$m=5/3$					
h	Δt	$\Vert e_h^f\Vert_2$	Order	$\Vert e_h^f\Vert_\infty$	Order	$\Vert e_h^x\Vert_2$	Order	$\Vert e_h^x\Vert_\infty$	Order
1/200	1/200	1.506e-04		3.277e-04		7.593e-05		7.844e-05	
1/400	1/400	3.620e-05	2.056	8.421e-05	1.960	1.871e-05	2.021	1.934e-05	2.020
1/800	1/800	8.495e-06	2.092	2.033e-05	2.050	4.464e-06	2.067	4.617e-06	2.066
1/1600	1/1600	1.887e-06	2.170	4.695e-06	2.114	1.000e-06	2.158	1.036e-06	2.156
				$m=2$					
h	Δt	$\Vert e_h^f\Vert_2$	Order	$\Vert e_h^f\Vert_\infty$	Order	$\Vert e_h^x\Vert_2$	Order	$\Vert e_h^x\Vert_\infty$	Order
1/200	1/200	1.502e-04		3.279e-04		7.642e-05		7.902e-05	
1/400	1/400	3.599e-05	2.061	8.370e-05	1.970	1.873e-05	2.028	1.938e-05	2.028
1/800	1/800	8.431e-06	2.094	2.005e-05	2.061	4.458e-06	2.071	4.615e-06	2.070
1/1600	1/1600	1.853e-06	2.186	4.563e-06	2.136	9.871e-07	2.175	1.024e-06	2.172

[1] Δt is the time step and h is the space step.

TABLE 8.1. *Convergence rate of solution f and trajectory x at time $t=0.05$.*

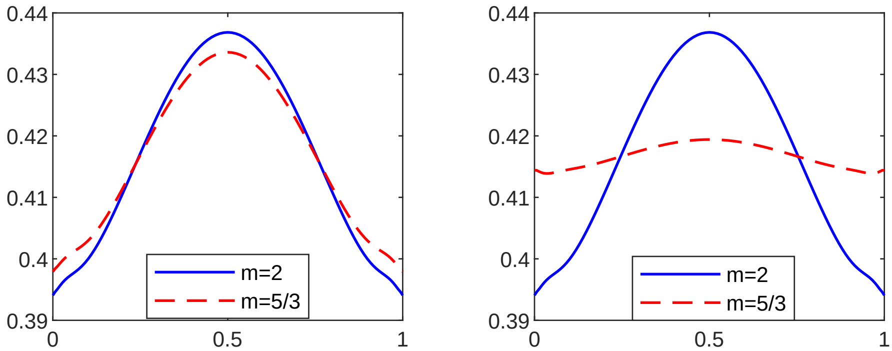
Example 2 (The problem with free boundaries). In this example, we consider the Barenblatt solution [13], which can be expressed as

$$B_m(x,t)=\begin{cases}(t+1)^{-k}\left(1-\dfrac{k(m-1)}{2m}\dfrac{|x|^2}{(t+1)^{2k}}\right)^{1/(m-1)}, & x\in[-\xi_B^t,\xi_B^t],\ t\ge 0,\\ 0, & \text{otherwise in } \Omega,\end{cases} \tag{8.4}$$

where $l_+=\max\{l,0\}$, $k=(m+1)^{-1}$ and

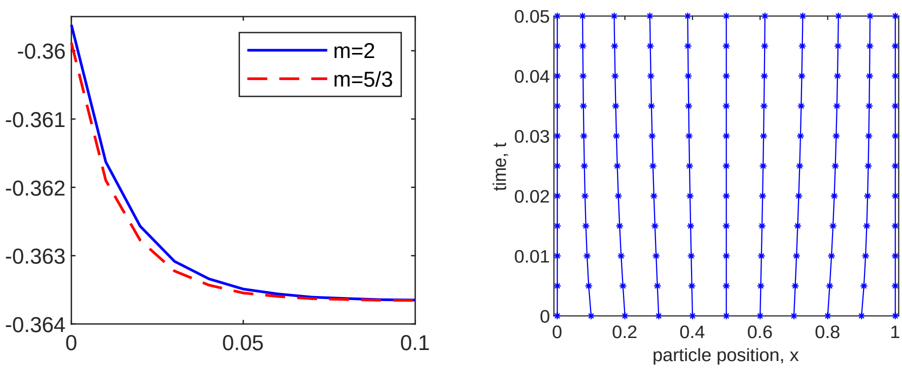
$$\xi_B^t=\sqrt{\frac{2m}{k(m-1)}}(t+1)^k.$$

Let $\Omega=(-10,10)$. We take $B_m(x,0)$ as the initial data and solve the problem with the numerical scheme (3.10) with (3.17). Figure 8.3 (a) displays the numerical density f_{num} and the exact solution f_{exact} at time $t=5$ and (b) gives a zoomed-in view near the right interface with the parameter $m=2$, the spatial resolution $M=100$ and the time step $\Delta t=1/100$. The results indicate that the numerical solution is an effective approximation to the exact solution. Moreover, Figure 8.4 (a) and (b) demonstrates the total energy decay and the total mass



(a) t=0.05 (b) t=0.1

Fig. 8.1. *The density f at time $t=0.05$ and $t=0.1$ ($h=1/100$, $\Delta t=1/100$).*



(a) Total energy (b) Particle Position for $m=2$

Fig. 8.2. *The evolution of total energy and particle position ($h=1/100$, $\Delta t=1/100$).*

conservation, as well as the finite movement of particles, respectively. Table 8.2 gives the $\mathcal{L}^2$ and $\mathcal{L}^\infty$ convergence rates of f at time $t=1$, and at $X=0$. Because of the low regularity of the exact solution in the free boundary set-up, only the first order convergence rates are observed in this case.

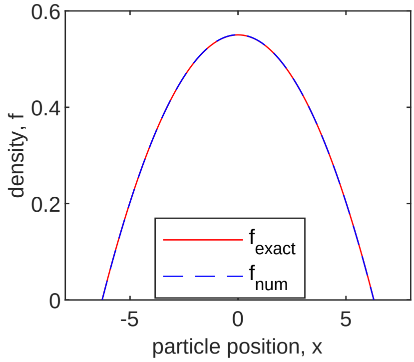
Example 3 (The waiting time phenomenon). The waiting time phenomenon is a classical and difficult problem in porous medium equation, which could occur for some initial state [39]. More detailed descriptions could be found in [13]. Now we consider the following set-up:

$$\Omega=(-5,5), \tag{8.5}$$

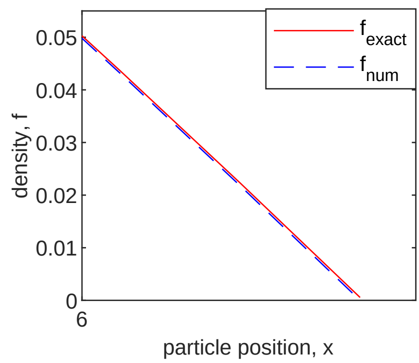
$$f_0(x)=\begin{cases}\left\{\frac{m-1}{m}[(1-\theta)\sin^2(x)+\theta\sin^4(x)\right\}^{1/(m-1)}, & x\in[-\pi,0],\\ 0, & \text{otherwise in } \Omega,\end{cases} \tag{8.6}$$

with $\theta\in[0,\frac{1}{4}]$. Then the waiting time is positive and the exact value is given by [2]

$$t^*_{exact}=\frac{1}{2(m+1)(1-\theta)}. \tag{8.7}$$

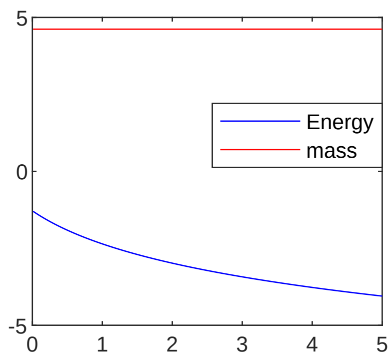


(a) t=5

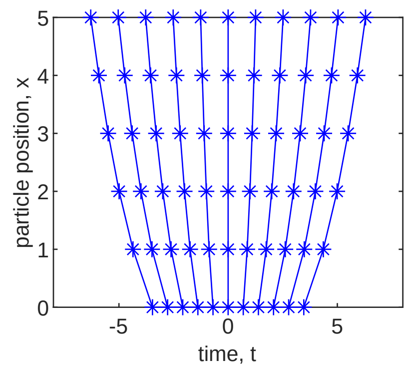


(b) Zoomed-in plot near right interface

FIG. 8.3. *The density f at time $t = 5$ and zoomed-in plot near right interface with the exact solution f_{exact} and the numerical solution f_{num} ($M = 100$, $\Delta t = 1/100$).*



(a) Total energy and mass



(b) Particle Position

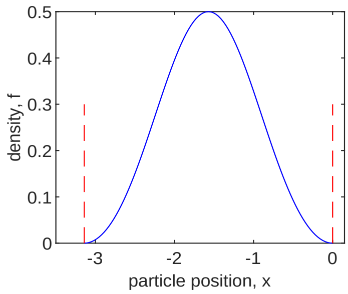
FIG. 8.4. *The evolution of total energy, total mass and particle position for $m = 2$ ($M = 100$, $\Delta t = 1/100$).*

In the computation, we take $m = 2$ and $\theta = \frac{1}{4}$. Figure 8.5 displays the evolution of f. It is observed that the interface of f remains static before $t = 0.2405$ (the waiting time), and then moves at a finite speed with $M = 100$ and $\Delta t = 1/2000$. Furthermore, Figure 8.6 (a) presents the total energy decay and the mass conservation of the system. The particle movement in Figure 8.6 (b) reveals that the particles at the interface do not move until $t = 0.2405$. In addition, Table 8.3 gives the convergence order for the waiting time. Although the convergence rate is less than first order, at least a positive accuracy order has been obtained in the simulation of this challenging problem. This example demonstrates another advantage of the proposed algorithm.

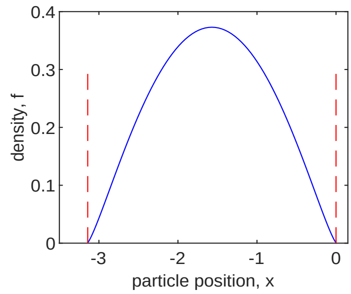
This numerical example suggests that the proposed scheme works for the case that $f_0(X)$ is degenerate in Ω_0. Such a performance comes from the subtle fact that, there is no logarithmic term in the Equation (3.1), while the energy formulation is always valid for non-negative $f_0(X)$ (by noticing that the energy density function $f \ln f$ is well-defined even if $f = 0$ in a certain region).

		$m=2$			
M	Δt	$\|e_h^f\|_2$	Order	$\|e_h^f\|_\infty$	Order
100	1/100	3.147e-03		2.381e-03	
200	1/200	1.573e-03	1.0005	1.191e-03	0.999
400	1/400	7.865e-04	1.0001	5.953e-04	1.000
800	1/800	3.932e-04	1.0000	2.976e-04	1.000

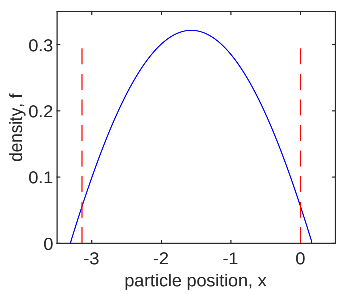
TABLE 8.2. **Example 2.** *The convergence rate of f at the finite time $T=1$.*



(a) t=0

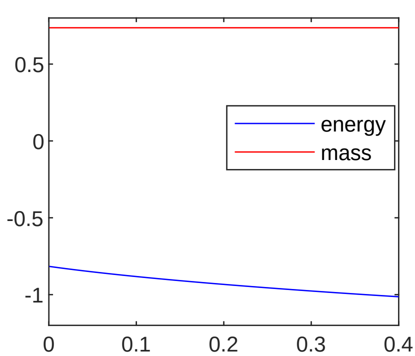


(b) t=0.2405

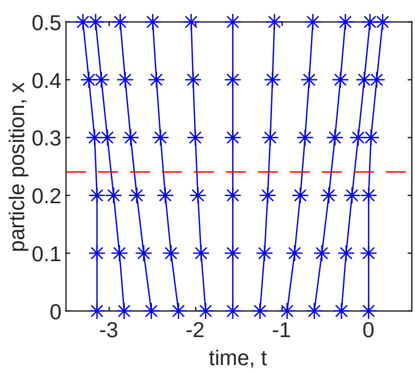


(c) t=0.5

FIG. 8.5. *The evolution of density f with $m=2$ ($M=100$, $\Delta t=1/2000$).*



(a) Total energy and mass



(b) Particle Position for $m=2$

FIG. 8.6. *The evolution of total energy, total mass and particle position ($M=100$, $\Delta t=1/2000$).*

M	Δt	t_w^*	$\lvert t_w^* - t_{w,e}^* \rvert$	Order
50	1/1000	0.2540	0.032	
100	1/2000	0.2405	0.0183	0.806
200	1/4000	0.2328	0.01053	0.7973
t_{exact}^*		0.2222		

TABLE 8.3. **Example 3**. *The convergence rate of waiting time ($m=2$, $\theta=\frac{1}{4}$).*

REMARK 8.1. In two or higher dimensions, the determinant of the deformation gradient, i.e., $\det \frac{\partial x}{\partial X}$, will appear in the trajectory equation, which becomes a fully nonlinear degenerate parabolic system. An efficient numerical method, which can satisfy the discrete energy dissipation law, has not been available in this area. A similar numerical method, based on evolving diffeomorphisms [5], will also encounter a similar difficulty. Because of this limitation, we have to focus on the one-dimensional equation in this article. The proposed numerical scheme (3.10) preserves the second order temporal accuracy, as well as all the nice theoretical properties, such as positivity-preservation, unique solvability, energy stability and optimal rate convergence analysis (provided that the exact PDE solution is sufficiently smooth). In addition, this numerical algorithm is able to capture complicated physical structures in an accurate way, such as waiting time phenomenon, in the practical computations, due to its Lagrange approach. These combined features make the proposed numerical method very attractive, at least for the one-dimensional case.

9. Concluding remarks

A second order accurate numerical scheme, both in time and space, is constructed and analyzed for the one-dimensional porous medium equation (PME) based on an energetic variational approach (EVA). A modified Crank-Nicolson temporal discretization is applied, combined with the finite difference over a uniform spatial mesh. Such a highly nonlinear numerical scheme is proved to be uniquely solvable on an admissible convex set, and an energy dissipation property is established, in which the convexity of the nonlinear implicit terms has played an important role. Moreover, an optimal rate convergence analysis is provided in this work, in which many highly non-standard estimates have to be involved. The higher order asymptotic expansion is performed to obtain higher order consistency estimate, the rough error estimate is needed so that an application of inverse inequality leads to an $W_h^{1,\infty}$ bound for the numerical variable. Subsequently, the refined error estimate is carried out to accomplish the desired convergence result, in which the $W_h^{1,\infty}$ bound for the numerical solution is applied. A few numerical results are also presented in this article, which demonstrates the robustness of the proposed numerical scheme.

On the other hand, one obvious limitation of this work is associated with the one-dimensional nature of the problem. In two or higher dimensions, the determinant of the deformation gradient, i.e., $\det \frac{\partial x}{\partial X}$, will arise in the trajectory equation, which is a complex nonlinear degenerate parabolic equation system. A suitable numerical method in multi-dimensional case, which can satisfy the discrete energy dissipation law, is still in the investigation process. Solving for multi-dimensional PME by this energetic method and the corresponding optimal error estimate will be left to the future works. Another limitation is the assumption of a positive initial condition ($f_0 > 0$), in which the convergence rate does not depend on the constant m. It is well known that if the initial state has a compact support, the convergent rate decreases with m. In this case, the trajectory equation with a free boundary makes the convergence analysis more difficult. This problem will also be considered in the future works.

Acknowledgments. This work is supported in part by the National Science Foundation (USA) grants NSF DMS-1759535, NSF DMS-1759536 (C. Liu), NSF DMS-2012669 (C. Wang), and NSF of China 11901109 (C. Duan), 11971342 (X. Yue), 11671098,91630309, and a 111 Project B08018 (W. Chen).

Appendix. Proof of Lemma 6.2. A direct calculation gives

$$q_1'(x) = -\frac{\frac{1}{x}(x-x_0) - (\ln x - \ln x_0)}{(x-x_0)^2} > 0, \tag{A.1}$$

in which the convexity of $-\ln x$ (for $x>0$) has been applied:

$$-\frac{1}{x}(x-x_0) + (\ln x - \ln x_0) > 0. \tag{A.2}$$

Meanwhile, a detailed Taylor expansion leads to

$$\ln x - \ln x_0 = \frac{1}{x}(x-x_0) + \frac{1}{2\zeta^2}(x-x_0)^2, \quad \text{with } \zeta \text{ between } x_0 \text{ and } x, \tag{A.3}$$

which in turn implies that

$$q_1'(x) = -\frac{\frac{1}{x}(x-x_0) - (\ln x - \ln x_0)}{(x-x_0)^2} = \frac{1}{2\zeta^2}, \quad \text{with } \zeta \text{ between } x_0 \text{ and } x. \tag{A.4}$$

Therefore, an application of the intermediate value theorem indicates that

$$q_1(y) - q_1(x) = q_1'(\eta)(y-x), \text{ with } \eta \text{ between } x \text{ and } y, \tag{A.5}$$

$$q_1'(\eta) = \frac{1}{2\zeta_\eta^2}, \quad \text{with } \zeta_\eta \text{ between } x_0 \text{ and } \eta. \tag{A.6}$$

Of course, a careful analysis implies that

$$q_1'(\eta) = \frac{1}{2\zeta_\eta^2}, \quad q_1'(\eta) \text{ is between } \frac{1}{2y^2}, \frac{1}{2x^2}, \text{and } \frac{1}{2x_0^2}, \forall x>0, y>0. \tag{A.7}$$

This completes the proof of (6.61).

A further calculation gives

$$q_1''(x) = -\frac{-\frac{1}{x^2}(x-x_0)^3 - 2(x-x_0)\left(\frac{1}{x}(x-x_0) - (\ln x - \ln x_0)\right)}{(x-x_0)^4} \le 0, \tag{A.8}$$

for any $x>0$, in which a higher order Taylor expansion has been applied:

$$\ln x - \ln x_0 = \frac{1}{x}(x-x_0) + \frac{1}{2x^2}(x-x_0)^2 + \frac{1}{3\zeta^3}(x-x_0)^3, \quad \text{with } \zeta \text{ between } x_0 \text{ and } x. \tag{A.9}$$

As a direct consequence, by an introduction of $q_2(x) := \frac{q_1(y)-q_1(x)}{y-x}$ for a fixed $y>0$, we get

$$q_2'(x) = \frac{-q_1'(x)(y-x) + (q_1(y)-q_1(x))}{(y-x)^2} \le 0, \tag{A.10}$$

$$\text{since } q_1(y) \text{ is concave: } -q_1'(x)(y-x) + (q_1(y)-q_1(x)) \le 0, \ \forall y>0, x>0.$$

This completes the proof of Lemma 6.2.

REFERENCES

[1] D.G. Aronson, *Regularity properties of flows through porous media*, SIAM J. Appl. Math., 17:461–467, 1969. 1

[2] D.G. Aronson, L.A. Caffarelli, and S. Kamin, *How an initially stationary interface begins to move in porous medium flow*, SIAM J. Math. Anal., 14(4):639–658, 1983. 8

[3] A. Baskaran, J.L. Lowengrub, C. Wang, and S.M. Wise, *Convergence analysis of a second order convex splitting scheme for the modified phase field crystal equation*, SIAM J. Numer. Anal., 51:2851–2873, 2013. 6.1

[4] J.A. Carrilo, B. Düring, D. Matthes, and D.S. McCormick, *A Lagrangian scheme for the solution of nonlinear diffusion equations using moving simplex meshes*, J. Sci. Comput., 75:1463–1499, 2018. 1

[5] J.A. Carrilo, H. Ranetbauer, and M.T. Wolfram, *Numerical simulation of nonlinear continuity equations by evolving diffeomorphisms*, J. Comput. Phys., 327:186–202, 2016. 8.1

[6] W. Chen, J. Jing, C. Wang, X. Wang, and S.M. Wise, *A modified Crank-Nicolson scheme for the Flory-Huggins Cahn-Hilliard model*, Commun. Comput. Phys., 31(1):60–93, 2022. 3.1

[7] W. Chen, C. Wang, X. Wang, and S.M. Wise, *A linear iteration algorithm for a second-order energy stable scheme for a thin film model without slope selection*, J. Sci. Comput., 59:574–601, 2014. 3.1

[8] K. Cheng, C. Wang, S.M. Wise, and X. Yue, *A second-order, weakly energy-stable pseudo-spectral scheme for the Cahn-Hilliard equation and its solution by the homogeneous linear iteration method*, J. Sci. Comput., 69:1083–1114, 2016. 3.1

[9] E. DiBenedetto and D. Hoff, *An interface tracking algorithm for the porous medium equation*, Trans. Amer. Math. Soc., 284:463–500, 1984. 1

[10] A. Diegel, C. Wang, X. Wang, and S.M. Wise, *Convergence analysis and error estimates for a second order accurate finite element method for the Cahn-Hilliard-Navier-Stokes system*, Numer. Math., 137:495–534, 2017. 3.1

[11] A. Diegel, C. Wang, and S.M. Wise, *Stability and convergence of a second order mixed finite element method for the Cahn-Hilliard equation*, IMA J. Numer. Anal., 36:1867–1897, 2016. 3.1

[12] L. Dong, W. Feng, C. Wang, S.M. Wise, and Z. Zhang, *Convergence analysis and numerical implementation of a second order numerical scheme for the three-dimensional phase field crystal equation*, Comput. Math. Appl., 75(6):1912–1968, 2018. 3.1

[13] C. Duan, C. Liu, C. Wang, and X. Yue, *Numerical methods for porous medium equation by an energetic variational approach*, J. Comput. Phys., 385:13–32, 2019. 1, 2.1, 2, 3.3, 8, 8

[14] C. Duan, C. Liu, C. Wang, and X. Yue, *Convergence analysis of a numerical scheme for the porous medium equation by an energetic variational approach*, Numer. Math. Theor. Meth. Appl., 13:1–18, 2020. 1, 2.1, 6.1

[15] W. E and J.G. Liu, *Projection method I: Convergence and numerical boundary layers*, SIAM J. Numer. Anal., 32:1017–1057, 1995. 1, 6.1

[16] W. E and J.G. Liu, *Projection method. III. Spatial discretization on the staggered grid*, Math. Comput., 71:27–47, 2002. 1, 6.1

[17] J.L. Graveleau and P. Jamet, *A finite difference approach to some degenerate nonlinear parabolic equations*, SIAM J. Appl. Math., 410:199–223, 1971. 1

[18] Y. Gu and J. Shen, *Bound preserving and energy dissipative schemes for porous medium equation*, J. Comput. Phys., 410:109378, 2020. 1, 3.3

[19] Z. Guan, J.L. Lowengrub, and C. Wang, *Convergence analysis for second order accurate schemes for the periodic nonlocal Allen-Cahn and Cahn-Hilliard equations*, Math. Meth. Appl. Sci., 40:6836–6863, 2017. 1, 6.1

[20] Z. Guan, J.L. Lowengrub, C. Wang, and S.M. Wise, *Second order convex splitting schemes for periodic nonlocal Cahn-Hilliard and Allen-Cahn equations*, J. Comput. Phys., 277:48–71, 2014. 3.1

[21] Z. Guan, C. Wang, and S.M. Wise, *A convergent convex splitting scheme for the periodic nonlocal Cahn-Hilliard equation*, Numer. Math., 128:377–406, 2014. 1, 6.1

[22] J. Guo, C. Wang, S.M. Wise, and X. Yue, *An H^2 convergence of a second-order convex-splitting, finite difference scheme for the three-dimensional Cahn-Hilliard equation*, Commun. Math. Sci., 14:489–515, 2016. 3.1

[23] J. Guo, C. Wang, S.M. Wise, and X. Yue, *An improved error analysis for a second-order numerical scheme for the Cahn-Hilliard equation*, J. Comput. Appl. Math., 388:113300, 2021. 3.1

[24] Z. Hu, S.M. Wise, C. Wang, and J. Lowengrub, *Stable and efficient finite-difference nonlinear-multigrid schemes for the phase field crystal equation*, J. Comput. Phys., 228:5323–5339, 2009. 3.1

[25] S. Jin, L. Pareschi, and G. Toscani, *Diffusive relaxation schemes for multi-scale discrete-velocity kinetic equations*, SIAM J. Numer. Anal., 35(6):2405–2439, 1998. 1

[26] A.S. Kalašnikov, *Formation of singularities in solutions of the equation of nonstationary filtration*, Ž. Vyčisl. Mat. Mat. Fiz., 7:440–444, 1967. 1

[27] X. Li, Z. Qiao, and C. Wang, *Convergence analysis for a stabilized linear semi-implicit numerical scheme for the nonlocal Cahn-Hilliard equation*, Math. Comput., 90:171–188, 2021. 1, 6.1

[28] C. Liu, C. Wang, S. Wise, X. Yue, and S. Zhou, *A positivity-preserving, energy stable and convergent numerical scheme for the Poisson-Nernst-Planck system*, Math. Comput., 90:2071–2106, 2021. 1, 6.1

[29] C. Liu and Y. Wang, *On Lagrangian schemes for porous medium type generalized diffusion equations: A discrete energetic variational approach*, J. Comput. Phys., 417:109566, 2020. 1

[30] Y. Nesterov and A. Nemirovskii, *Interior-Point Polynomial Algorithms in Convex Programming*, SIAM, 13, 1994. 7, 7.1, 7, 7

[31] C. Ngo and W.Z. Huang, *A study on moving mesh finite element solution of the porous medium equation*, J. Comput. Phys., 331:357–380, 2017. 1

[32] O.A. Oleĭnik, A.S. Kalašinkov, and Y. Čžou, *The Cauchy problem and boundary problems for equations of the type of non-stationary filtration*, Izv. Akad. Nauk SSSR, Ser. Mat., 22:667–704, 1958. 1

[33] R. Samelson, R. Temam, C. Wang, and S. Wang, *Surface pressure Poisson equation formulation of the primitive equations: Numerical schemes*, SIAM J. Numer. Anal., 41:1163–1194, 2003. 1, 6.1

[34] R. Samelson, R. Temam, C. Wang, and S. Wang, *A fourth order numerical method for the planetary geostrophic equations with inviscid geostrophic balance*, Numer. Math., 107:669–705, 2007. 6.1

[35] J. Shen, C. Wang, X. Wang, and S.M. Wise, *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, 2009. 3.1

[36] S.I. Shmarev, *Interfaces in multidimensional diffusion equations with absorption terms*, Nonlinear Anal., 53:791–828, 2003. 1

[37] S. Shmarev, *Interfaces in solutions of diffusion-absorption equations in arbitrary space dimension*, in J.F. Rodrigues, G. Seregin, and J.M. Urbano (eds.), Trends in Partial Differential Equations of Mathematical Physics, Birkhäuser, Basel, 257–273, 2005. 1

[38] R. Temam, *Navier-Stokes Equations: Theory and Numerical Analysis*, Amer. Math. Soc., Providence, Rhode Island, 2001. 6.1

[39] J.L. Vázquez, *The Porous Medium Equation*, Oxford University Press, Oxford, 2007. 1, 8

[40] C. Wang, J.G. Liu, and H. Johnston, *Analysis of a fourth order finite difference method for incompressible Boussinesq equations*, Numer. Math., 97:555–594, 2004. 6.1

[41] C. Wang and J.G. Liu, *Analysis of finite difference schemes for unsteady Navier-Stokes equations in vorticity formulation*, Numer. Math., 91:543–576, 2002. 6.1

[42] C. Wang and J.G. Liu, *Convergence of gauge method for incompressible flow*, Math. Comput., 69:1385–1407, 2000. 1, 6.1

[43] L. Wang, W. Chen, and C. Wang, *An energy-conserving second order numerical scheme for nonlinear hyperbolic equation with an exponential nonlinear term*, J. Comput. Appl. Math., 280:347–366, 2015. 6.1

[44] M. Westdickenberg and J. Wilkening, *Variational particle schemes for the porous medium equation and for the system of isentropic Euler equations*, ESAIM Math. Model. Numer. Anal., 44(1):133–166, 2010. 1

[45] Q. Zhang and Z.L. Wu, *Numerical simulation for porous medium equation by local discontinuous Galerkin finite element method*, J. Sci. Comput., 38(2):127–148, 2009. 1