

## Convergence Analysis of a Numerical Scheme for the Porous Medium Equation by an Energetic Variational Approach

Chenghua Duan<sup>1,2</sup>, Chun Liu<sup>3</sup>, Cheng Wang<sup>4</sup> and Xingye Yue<sup>1,\*</sup>

<sup>1</sup> Department of Mathematics, Soochow University, Suzhou 215006, China

<sup>2</sup> Shanghai Center for Mathematical Sciences, Fudan University, Shanghai 200438, China

<sup>3</sup> Department of Applied Mathematics, Illinois Institute of Technology, Chicago, IL 60616, USA

<sup>4</sup> Department of Mathematics, University of Massachusetts, Dartmouth, North Dartmouth, MA 02747-2300, USA

Received 11 May 2019; Accepted (in revised version) 1 September 2019

---

**Abstract.** The porous medium equation (PME) is a typical nonlinear degenerate parabolic equation. We have studied numerical methods for PME by an energetic variational approach in [C. Duan et al., J. Comput. Phys., 385 (2019), pp. 13–32], where the trajectory equation can be obtained and two numerical schemes have been developed based on different dissipative energy laws. It is also proved that the nonlinear scheme, based on  $f \log f$  as the total energy form of the dissipative law, is uniquely solvable on an admissible convex set and preserves the corresponding discrete dissipation law. Moreover, under certain smoothness assumption, we have also obtained the second order convergence in space and the first order convergence in time for the scheme. In this paper, we provide a rigorous proof of the error estimate by a careful higher order asymptotic expansion and two step error estimates. The latter technique contains a rough estimate to control the highly nonlinear term in a discrete  $W^{1,\infty}$  norm and a refined estimate is applied to derive the optimal error order.

**AMS subject classifications:** 65M12, 65M75

**Key words:** Energetic variational approach, porous medium equation, trajectory equation, optimal rate convergence analysis.

---

\*Corresponding author. *Email addresses:* xyue@suda.edu.cn (X. Y. Yue), chduan@fudan.edu.cn (C. H. Duan), cwang1@umassd.edu (C. Wang), cliu124@iit.edu (C. Liu)

## 1. Introduction and background

One of the typical nonlinear degenerate parabolic equations is the porous medium equation (PME):

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

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 been applied 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 [18], etc.

It is well known that the PME is degenerate at points where  $f = 0$ . In turn, the PME has many special features: the finite speed of propagation, the free boundary, a possible waiting time phenomenon [5, 18]. Various numerical methods have been studied for the PME, such as finite difference approach [8], tracking algorithm method [3], a local discontinuous Galerkin finite element method [24], Variational Particle Scheme (VPS) [23] and an adaptive moving mesh finite element method [13]. Many theoretical analyses have been derived in the existing literature [1, 12, 14, 16–18], etc.

Relevant detailed descriptions can be found in a recent paper [5], in which the numerical methods for the PME were constructed by an Energetic Variational Approach (EnVarA) to naturally keep the physical laws, such as the conservation of mass, energy dissipation and force balance. Meanwhile, based on different dissipative energy laws, two different numerical schemes have been studied. In more details, based on the total energy form  $f \log f$  and  $\frac{1}{2f}$ , a fully discrete nonlinear scheme and a linear numerical scheme could be appropriately designed for the trajectory equation, respectively. It has also been proved that the former one is uniquely solvable on an admissible convex set and both schemes preserve the corresponding discrete dissipation law. Numerical experiments have demonstrated that both schemes have yielded 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. In addition, under certain smoothness assumption, the second order convergence in space and the first order convergence in time have been reported for both schemes in [5]. The aim of the paper is to provide a rigorous proof of the optimal rate convergence analysis for the nonlinear scheme. On the other hand, the highly nonlinear nature of the trajectory equation makes the convergence analysis every challenging. To overcome these difficulties, we use 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 [6, 7, 15, 21], 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. 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.

This paper is organized as follows. The trajectory equation of the PME and the numerical scheme are outlined in Section 2.1 and Section 2.2, respectively. Subsequently, the proof of optimal rate convergence analysis is provided in Section 3. Finally we present a simple numerical example to demonstrate the convergence rate of the numerical scheme in Section 4.

## 2. Trajectory equation and the numerical scheme

In this section, we review the trajectory equation and the corresponding numerical scheme.

### 2.1. Trajectory equation of the PME

In this part, the one-dimensional trajectory equation will be reviewed, derived by an Energetic Variational Approach [5]. We solve the following initial-boundary problem:

$$\partial_t f + \partial_x(f\mathbf{v}) = 0, \quad x \in \Omega \subset \mathbb{R}, \quad t > 0, \quad (2.1a)$$

$$f\mathbf{v} = -\partial_x(f^m), \quad x \in \Omega, \quad m > 1, \quad (2.1b)$$

$$f(x, 0) = f_0(x) \geq 0, \quad x \in \Omega, \quad (2.1c)$$

$$\partial_x f = 0, \quad x \in \partial\Omega, \quad t > 0, \quad (2.1d)$$

where  $\Omega$  is a bounded domain,  $f$  is a non-negative function,  $t$  is the time,  $x$  is the particle position and  $\mathbf{v}$  is the velocity of particle .

The following lemma is available.

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

$$\frac{d}{dt} \int_{\Omega} f \ln f dx = - \int_{\Omega} \frac{f}{m f^{m-1}} |\mathbf{v}|^2 dx. \quad (2.2)$$

*Proof.* We first prove the energy dissipation law (2.2) if  $f$  is the solution of (2.1a)-(2.1d). Multiplying by  $(1 + \ln f)$  and integrating on both sides of (2.1a), 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 (2.1b), we have

$$\frac{d}{dt} \int_{\Omega} f \ln f dx = \int_{\Omega} \frac{\partial_x f}{f} (f \mathbf{v}) dx = - \int_{\Omega} \frac{f}{m f^{m-1}} |\mathbf{v}|^2 dx \leq 0. \quad (2.3)$$

Subsequently, we are also able to derive (2.1b) from the energy dissipation law (2.2) by EnVarA.

In addition, (2.1a) is the conservation law. In the Lagrangian coordinate, its solution can be expressed by:

$$f(x(X, t), t) = \frac{f_0(X)}{\frac{\partial x(X, t)}{\partial X}}, \quad (2.4)$$

where  $f_0(X)$  is the positive initial data and  $\partial_X x := \frac{\partial x(X, t)}{\partial X}$  is the *deformation gradient* in one dimension.

Based on an Energetic Variational Approach, we can obtain the trajectory equation.

### Energy Dissipation Law.

The **total energy** of the PME is

$$E^{total} := \int_{\Omega} f \ln f dx. \quad (2.5)$$

### Least Action Principle Step.

With (2.4), the action functional in Lagrangian coordinate becomes

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

where  $T^* > 0$  is a given terminal time and  $\mathcal{H}$  is the free energy depending on  $x$ . Thus for any test function  $y(X, t) = \tilde{y}(x(X, t), t) \in C_0^\infty(\Omega \times (0, T^*))$  and  $\varepsilon \in \mathbb{R}$ , taking the variational of  $\mathcal{A}(x)$  with respect to  $x$ , we have

$$\begin{aligned} \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \mathcal{A}(x + \varepsilon y) &= \int_0^{T^*} \int_{\Omega} \frac{f_0(X)}{\partial_X x} \cdot \partial_X y dX dt \\ &= - \int_0^{T^*} \int_{\Omega} \partial_x f \cdot \tilde{y} dx dt. \end{aligned}$$

Then the conservation force turns out to be

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

in the Eulerian coordinate, and

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

in the Lagrangian coordinate.

**Maximal Dissipation Principle Step.**

Define the entropy production

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

Taking the variational of  $\frac{1}{2}\Delta$  with respect to the velocity  $\mathbf{v}$  and  $x_t$ , we obtain the dissipation force

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

in the Eulerian coordinate and

$$F_{dis} := \frac{\delta \frac{1}{2}\Delta}{\delta(\partial_t x)} = \frac{f_0(X)}{m\left(\frac{f_0(X)}{\partial_X x}\right)^{m-1}} \partial_t x,$$

in the Lagrangian coordinate.

**Force balance Step.**

Based on the Newton's force balance law, we get

$$\frac{f_0(X)}{m\left(\frac{f_0(X)}{\partial_X x}\right)^{m-1}} \partial_t x = -\partial_X \left( \frac{f_0(X)}{\partial_X x} \right), \quad (2.6)$$

in the Lagrangian coordinate and the Darcy's Law in the Eulerian coordinate

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

Thus, we complete the proof.  $\square$

It is noticed that, there is an assumption that the value of initial state  $f_0(x)$  is positive in  $\Omega$  to make  $\int_{\Omega} f \ln f dx$  well-defined in (2.2). More details can be found in [5].

Since then, we should first settle the initial and boundary conditions for (2.6). From (2.1a) and (2.1d), we have  $x_t|_{\partial\Omega} = 0$ , for  $t > 0$ . This means that the particles lying on boundary will stay there forever, so a Dirichlet boundary condition should be subject to as  $x|_{\partial\Omega} = X|_{\partial\Omega}$ , for  $t \geq 0$ . As a result, the trajectory problem becomes

$$\frac{f_0(X)}{m\left(\frac{f_0(X)}{\partial_X x}\right)^{m-1}} \partial_t x = -\partial_X \left( \frac{f_0(X)}{\partial_X x} \right), \quad X \in \Omega, \quad t > 0, \quad (2.7a)$$

$$x|_{\partial\Omega} = X|_{\partial\Omega}, \quad t > 0, \quad (2.7b)$$

$$x(X, 0) = X, \quad X \in \Omega. \quad (2.7c)$$

Finally, with a substitution of (2.7a) into (2.4), we obtain the solution  $f(x, t)$  to (2.1a)-(2.1d).

## 2.2. Numerical scheme of the trajectory equation

Let  $\tau = \frac{T}{N}$ , where  $N \in \mathbb{N}^+$ ,  $T$  is the final time and the grid points are given by  $t_n = n\tau$ ,  $n = 0, \dots, N$ . Let  $X_0$  be the left point of  $\Omega$  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  $\{X_i | i = 0, \dots, M\}$  and  $\{X_{i-\frac{1}{2}} | i = 1, \dots, M\}$ , respectively. In component form, these functions are identified via  $l_i = l(X_i)$ ,  $i = 0, \dots, M$ , for  $l \in \mathcal{E}_M$  and  $\phi_{i-\frac{1}{2}} = \phi(X_{i-\frac{1}{2}})$ ,  $i = 1, \dots, M$  for  $\phi \in \mathcal{C}_M$ .

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

$$(D_h l)_{i-\frac{1}{2}} = (l_i - l_{i-1})/h, \quad i = 1, \dots, M, \quad (2.8a)$$

$$(d_h \phi)_i = (\phi_{i+\frac{1}{2}} - \phi_{i-\frac{1}{2}})/h, \quad i = 1, \dots, M-1, \quad (2.8b)$$

$$(\tilde{D}_h l)_i = (l_{i+1} - l_{i-1})/2h, \quad i = 1, \dots, M-1, \quad (2.8c)$$

$$(\tilde{D}_h l)_i = (4l_{i+1} - l_{i+2} - 3l_i)/2h, \quad i = 0, \quad (2.8d)$$

$$(\tilde{D}_h l)_i = (l_{i-2} - 4l_{i-1} + 3l_i)/2h, \quad i = M, \quad (2.8e)$$

respectively.

Let  $\mathcal{Q} := \{l \in \mathcal{E}_M | l_{i-1} < l_i, 1 \leq i \leq M; l_0 = X_0, l_M = X_M\}$  with its boundary set  $\partial\mathcal{Q} := \{l \in \mathcal{E}_M | l_{i-1} \leq l_i, 1 \leq i \leq M; l_0 = X_0, l_M = X_M; \exists i \in \{1, \dots, M\}, s.t. l_{i-1} = l_i\}$ . Then  $\bar{\mathcal{Q}} := \mathcal{Q} \cup \partial\mathcal{Q}$  is a closed convex set. Its physical meaning indicates that particles are arranged in the order without twisting or exchanging in  $\mathcal{Q}$ .

A few more notations have to be introduced. Let  $l, g \in \mathcal{E}_M$  and  $\phi, \varphi \in \mathcal{C}_M$ . We define the *inner product* on space  $\mathcal{E}_M$  and  $\mathcal{C}_M$  respectively as:

$$\langle l, g \rangle := h \left( \frac{1}{2} l_0 g_0 + \sum_{i=1}^{M-1} l_i g_i + \frac{1}{2} l_M g_M \right), \quad (2.9a)$$

$$\langle \phi, \varphi \rangle_e := h \sum_{i=0}^{M-1} \phi_{i+\frac{1}{2}} \varphi_{i+\frac{1}{2}}. \quad (2.9b)$$

The following summation by parts formula is available:

$$\langle l, d_h \phi \rangle = - \langle D_h l, \phi \rangle_e \quad \text{with } l_0 = l_M = 0, \quad \phi \in \mathcal{C}_M, \quad l \in \mathcal{E}_M. \quad (2.10)$$

The inverse inequality is given by:

$$\|l\|_\infty \leq C_m \frac{\|l\|_2}{h^{1/2}}, \quad \forall l \in \mathcal{E}_M \quad \text{with} \quad \|l\|_\infty := \max_{0 \leq i \leq M} \{l_i\}, \quad \|l\|_2^2 := \langle l, l \rangle. \quad (2.11)$$

The **fully discrete scheme** is formulated as follows: Given the positive initial state  $f_0(X) \in \mathcal{E}_M$  and the particle position  $x^n \in \mathcal{Q}$ , find  $x^{n+1} = (x_0^{n+1}, \dots, x_M^{n+1}) \in \mathcal{Q}$  such that

$$\frac{f_0(X_i)}{m \left( \frac{f_0(X)}{\tilde{D}_h x^n} \right)_i^{m-1}} \cdot \frac{x_i^{n+1} - x_i^n}{\tau} = -d_h \left[ \left( \frac{f_0(X)}{\tilde{D}_h x^{n+1}} \right) \right]_i, \quad 1 \leq i \leq M-1, \quad (2.12)$$

with  $x_0^{n+1} = X_0$  and  $x_M^{n+1} = X_M$ ,  $n = 0, \dots, N-1$ .

It is noticed that (2.12) is still a nonlinear system which can be solved by Newton's iteration method [5]. Then we obtain the numerical solution  $f(x_i, t^n) := f_i^n$  by

$$f_i^n = \frac{f_0(X)}{\tilde{D}_h x_i^n}, \quad 0 \leq i \leq M, \quad (2.13)$$

which is the discrete scheme of (2.4).

### 3. Convergence analysis

In this section, the second order spatial convergence and the first order temporal convergence will be theoretically justified for the numerical scheme (2.12). We first introduce a higher order approximate expansion of the exact solution, since a consistency estimate (second order in space and first order in time) is not able to control the discrete  $W^{1,\infty}$  norm of the numerical solution. Also see the related works in the earlier literature [2, 6, 7, 9, 10, 15, 19–22], etc.

**Lemma 3.1.** *Assume a higher order approximate solution of the exact solution  $x_e$ :*

$$W := x_e + \tau w_\tau^{(1)} + \tau^2 w_\tau^{(2)} + h^2 w_h, \quad (3.1)$$

where  $w_\tau^{(1)}$ ,  $w_\tau^{(2)}$ ,  $w_h \in C^\infty(\Omega; 0, T)$ . Then there exists a small  $\tau_0 > 0$ , such that  $\forall \tau, h \leq \tau_0$ ,  $\tilde{D}_h W > 0$ , i.e.,  $W \in \mathcal{Q}$ , where  $\tau$  and  $h$  are the time step and the spatial mesh sizes, respectively.

*Proof.* Because of a point-wise condition for the exact solution,  $x_e \in \mathcal{Q}$ , i.e.,  $\exists \varepsilon_0 > 0$ , such that  $D_h x_e > \varepsilon_0 > 0$ . For small  $\tau_0$ , such that

$$\|\tau D_h w_\tau^{(1)}\|_{L^\infty} \leq \frac{1}{9} \varepsilon_0, \quad \|\tau^2 D_h w_\tau^{(2)}\|_{L^\infty} \leq \frac{1}{9} \varepsilon_0 \quad \text{and} \quad \|h^2 D_h w_h\|_{L^\infty} \leq \frac{1}{9} \varepsilon_0,$$

for  $\forall \tau, h \leq \tau_0$ . As a consequence, for  $\forall \tau, h \leq \tau_0$ , we have

$$D_h W \geq \frac{1}{3} \varepsilon_0 > 0, \quad (3.2)$$

which in turn implies that  $W \in \mathcal{Q}$ . □

**Theorem 3.1.** Assume that the initial function  $f_0(X)$  is positive and bounded, i.e.,  $0 < b_f \leq f_0(X) \leq B_f$ . Denote  $x_e \in \Omega$  as the exact solution to the original PDE (2.7a) (with enough regularity) and  $x_h \in \mathcal{Q}$  as the numerical solution to (2.12). The numerical error function is defined at a point-wise level:

$$e_i^n = x_{e_i}^n - x_{h_i}^n, \quad (3.3)$$

where  $x_{e_i}^n, x_{h_i}^n \in \mathcal{Q}$ ,  $0 \leq i \leq N$ ,  $n = 0, \dots, M$ .

Then we have

- $e^n = (e_0^n, \dots, e_M^n)$  satisfies

$$\|e^n\|_2 := \langle e^n, e^n \rangle \leq C(\tau + h^2).$$

- $\tilde{D}_h e^n = (\tilde{D}_h e_0^n, \dots, \tilde{D}_h e_M^n)$  satisfies

$$\|\tilde{D}_h e^n\|_2 \leq C(\tau + h^2).$$

Moreover, the error between the numerical solution  $f_h^n$  and the exact solution  $f_e^n$  of Eqs. (2.1a)-(2.1d) can be estimated by:

$$\|f_h^n - f_e^n\|_2 \leq C(\tau + h^2),$$

where  $C$  is a positive constant,  $h$  is the spatial step,  $\tau$  is the time step and  $n = 0, \dots, N$ .

*Proof.* A careful Taylor expansion of the exact solution in both time and space, in terms of the numerical scheme (2.7a), gives that

$$\begin{aligned} & \frac{f_0(X_i)}{m \left( \frac{f_0(X_i)}{D_h x_{e_i}^n} \right)^{m-1}} \frac{x_{e_i}^{n+1} - x_{e_i}^n}{\tau} \\ &= -d_h \left( \frac{f_0(X)}{D_h x_e^{n+1}} \right)_i + \tau l_i^{(1)} + \tau^2 l_i^{(2)} + \tau^3 l_i^{(3)} + h^2 g_i^{(1)} + h^4 g_i^{(2)}, \quad 1 \leq i \leq M-1, \\ & \text{with } x_{e_0}^{n+1} = X_0, \quad x_{e_M}^{n+1} = X_M, \end{aligned} \quad (3.4)$$

where  $\|l^{(1)}\|_2, \|l^{(2)}\|_2, \|l^{(3)}\|_2, \|g^{(1)}\|_2, \|g^{(2)}\|_2 \leq C_e$ , with  $C_e$  only dependent on the exact solution.

To perform a higher order consistency analysis for an approximate solution of the exact solution, we have to construct the approximation  $W$  as in (3.1).

The term  $w_\tau^{(1)} \in C^\infty(\Omega; 0, T)$  is given by the following linear equation:

$$\frac{f_0(X)}{m \left( \frac{f_0(X)}{\partial_X x_e} \right)^{m-1}} \partial_t w_\tau^{(1)} + \frac{m-1}{m \left( \frac{f_0(X)}{\partial_X x_e} \right)^{m-2}} \partial_t x_e \cdot \partial_X w_\tau^{(1)} = \partial_X \left( \frac{f_0(X)}{(\partial_X x_e)^2} \partial_X w_\tau^{(1)} \right) - l^{(1)}, \quad (3.5a)$$

$$w_\tau^{(1)}|_{\partial\Omega} = 0, \quad w_\tau^{(1)}(\cdot, 0) = 0. \quad (3.5b)$$

The term  $w_\tau^{(2)} \in C^\infty(\Omega; 0, T)$  is given by the following linear equation:

$$\begin{aligned} & \frac{f_0(X)}{m\left(\frac{f_0(X)}{\partial_X x_e}\right)^{m-1}} \partial_t w_\tau^{(2)} + \frac{m-1}{m\left(\frac{f_0(X)}{\partial_X x_e}\right)^{m-2}} \partial_t x_e \cdot \partial_X w_\tau^{(2)} \\ & + \frac{(m-1)(m-2)}{2m\left(\frac{f_0(X)}{\partial_X x_e}\right)^{m-2} \partial_X x_e} (\partial_X w_\tau^{(1)})^2 \cdot \partial_t x_e + \frac{(m-1)}{m\left(\frac{f_0(X)}{\partial_X x_e}\right)^{m-2}} \partial_t w_\tau^{(1)} \cdot \partial_X w_\tau^{(1)} \\ & = \partial_X \left( \frac{f_0(X)}{(\partial_X x_e)^2} \partial_X w_\tau^{(2)} \right) - \partial_X \left( \frac{f_0(X)}{(\partial_X x_e)^3} (\partial_X w_\tau^{(1)})^2 \right) - l^{(2)}, \end{aligned} \quad (3.6a)$$

$$w_\tau^{(2)}|_{\partial\Omega} = 0, \quad w_\tau^{(2)}(\cdot, 0) = 0. \quad (3.6b)$$

The term  $w_h \in C^\infty(\Omega; 0, T)$  is given by the following linear equation:

$$\frac{f_0(X)}{m\left(\frac{f_0(X)}{\partial_X x_e}\right)^{m-1}} \partial_t w_h + \frac{(m-1)\partial_t x_e}{m\left(\frac{f_0(X)}{\partial_X x_e}\right)^{m-2}} \partial_X w_h = \partial_X \left( \frac{f_0(X)}{(\partial_X x_e)^2} \partial_X w_h \right) - g^{(1)}, \quad (3.7a)$$

$$w_h|_{\partial\Omega} = 0, \quad w_h(\cdot, 0) = 0. \quad (3.7b)$$

Since  $w_\tau^{(1)}$ ,  $w_\tau^{(2)}$ ,  $w_h$  are dependent only on  $W$  and  $x_e$ , we have the following estimate:

$$\|W - x_e\|_{H^m} = \tau \|w_\tau^{(1)}\|_{H^m} + \tau^2 \|w_\tau^{(1)}\|_{H^m} + h^2 \|w_h\|_{H^m} \leq C'(\tau + h^2). \quad (3.8)$$

With these expansion terms, the constructed approximation  $W \in \mathcal{Q}$  satisfies the numerical scheme with a higher order truncation error:

$$\begin{aligned} & \frac{f_0(X_i)}{m\left(\frac{f_0(X)}{D_h W^n}\right)_i^{m-1}} \cdot \frac{W_i^{n+1} - W_i^n}{\tau} = -d_h \left( \frac{f_0(X)}{D_h W^{n+1}} \right)_i + \tau^3 l_i^* + h^4 g_i^*, \quad 1 \leq i \leq M-1, \\ & \text{with } W_0^{n+1} = X_0, \quad W_M^{n+1} = X_M, \quad n = 0, 1, \dots, N-1, \end{aligned} \quad (3.9)$$

where  $l^*$ ,  $g^*$  are dependent only on  $l^{(1)}$ ,  $l^{(2)}$ ,  $l^{(3)}$ ,  $g^{(1)}$ ,  $g^{(2)}$  and the derivatives of  $w_\tau^{(1)}$ ,  $w_\tau^{(2)}$ ,  $w_h$ .

Then we define  $\tilde{e}_i^n := W_i^n - x_{h_i}^n$ ,  $0 \leq i \leq M$ ,  $n = 0, 1, \dots, N$ . In other words, instead of a direct comparison between the numerical solution and exact PDE solution, we evaluate the numerical error between the numerical solution and the constructed solution  $W$ . The higher order truncation error enables us to obtain a required  $W_h^{1,\infty}$  of the numerical solution, which is necessary in the nonlinear convergence analysis.

Note that the discrete  $L^2$  norm  $\|\tilde{e}^0\|_2 = 0$  at time step  $t^0$ . We make the following a-priori assumption at time step  $t^n$ :

$$\|\tilde{e}^n\|_2 \leq (\tau^{\frac{11}{4}} + h^{\frac{7}{2}}). \quad (3.10)$$

In turn, the following estimates become available, by making use of inverse inequalities:

$$\|\tilde{D}_h \tilde{e}^n\|_2 \leq C(\tau^{\frac{7}{4}} + h^{\frac{5}{2}}), \quad (3.11a)$$

$$\|\tilde{D}_h \tilde{e}^n\|_\infty \leq CC_m \frac{\|\tilde{D}_h \tilde{e}^n\|_2}{h^{1/2}} \leq CC_m(\tau^{\frac{5}{4}} + h^2), \quad \text{if } h = \mathcal{O}(\tau), \quad (3.11b)$$

$$\|\tilde{D}_h x_h^n\|_\infty = \|\tilde{D}_h W^n - \tilde{D}_h \tilde{e}^n\|_\infty \leq C^* + 1 := C_0^*, \quad (3.11c)$$

$$\text{with } C^* := \|\tilde{D}_h W^n\|_\infty, \quad \text{if } CC_m(\tau^{\frac{5}{4}} + h^2) \leq 1, \quad (3.11d)$$

$$\left\| \frac{\tilde{D}_h x_h^n - \tilde{D}_h x_h^{n-1}}{\tau} \right\|_\infty = \left\| \frac{\tilde{D}_h W^n - \tilde{D}_h W^{n-1}}{\tau} - \frac{\tilde{D}_h \tilde{e}^n - \tilde{D}_h \tilde{e}^{n-1}}{\tau} \right\|_\infty \leq \tilde{C}_t^* + 1, \quad (3.11e)$$

$$\text{with } \tilde{C}_t^* := \left\| \frac{\tilde{D}_h W^n - \tilde{D}_h W^{n-1}}{\tau} \right\|_\infty, \quad \text{if } CC_m(\tau^{\frac{1}{4}} + h) \leq 1. \quad (3.11f)$$

For  $x_h, W \in \mathcal{Q}$ , i.e.,  $\exists \delta_0 > 0$ , such that  $\tilde{D}_h W_i^n \geq \delta_0$ , then  $\tilde{D}_h x_{h_i}^n \geq \frac{\delta_0}{2} > 0$ ,  $0 \leq i \leq M$ , if  $C_m \gamma(\tau^{\frac{5}{4}} + h^2) \leq \frac{\delta_0}{2}$ .

In turn, subtracting (3.9) from the numerical scheme (2.12) yields

$$\begin{aligned} & \frac{f_0(X_i)}{m \left( \frac{f_0(X)}{\tilde{D}_h x_h^n} \right)_i^{m-1}} \cdot \frac{\tilde{e}_i^{n+1} - \tilde{e}_i^n}{\tau} \\ & + \frac{f_0(X_i)}{m [f_0(X_i)]^{m-1}} \cdot \frac{W_i^{n+1} - W_i^n}{\tau} \cdot [(\tilde{D}_h W^n)_i^{m-1} - (\tilde{D}_h x_h^n)_i^{m-1}] \\ = & d_h \left( \frac{f_0(X)}{D_h W_i^{n+1} D_h x_h^{n+1}} D_h \tilde{e}^{n+1} \right)_i + \tau^3 l_i^* + h^4 g_i^*, \quad 1 \leq i \leq M-1, \\ & \text{with } \tilde{e}_0^{n+1} = \tilde{e}_M^{n+1} = 0, \end{aligned} \quad (3.12)$$

in which the form of the left term comes from the following identity:

$$\begin{aligned} & \frac{f_0(X_i)}{m \left( \frac{f_0(X)}{\tilde{D}_h W^n} \right)_i^{m-1}} \frac{W_i^{n+1} - W_i^n}{\tau} - \frac{f_0(X)}{m \left( \frac{f_0(X)}{\tilde{D}_h x_h^n} \right)_i^{m-1}} \frac{x_{h_i}^{n+1} - x_{h_i}^n}{\tau} \\ = & \frac{f_0(X_i)}{\tau m [f_0(X_i)]^{m-1}} [(\tilde{D}_h W^n)_i^{m-1} (W_i^{n+1} - W_i^n) - (\tilde{D}_h x_h^n)_i^{m-1} (x_{h_i}^{n+1} - x_{h_i}^n)] \\ & + (\tilde{D}_h x_h^n)_i^{m-1} (W_i^{n+1} - W_i^n) - (\tilde{D}_h x_h^n)_i^{m-1} (W_i^{n+1} - W_i^n) \\ = & \frac{f_0(X_i)}{m [f_0(X_i)]^{m-1}} \cdot \frac{W_i^{n+1} - W_i^n}{\tau} \cdot [(\tilde{D}_h W^n)_i^{m-1} - (\tilde{D}_h x_h^n)_i^{m-1}] \\ & + \frac{f_0(X_i)}{m \left( \frac{f_0(X)}{\tilde{D}_h x_h^n} \right)_i^{m-1}} \cdot \frac{\tilde{e}_i^{n+1} - \tilde{e}_i^n}{\tau}. \end{aligned}$$

Based on the preliminary results, taking a discrete inner product with (3.12) by  $2\tilde{e}^{n+1}$  gives

$$\begin{aligned} & 2\langle \alpha_n(\tilde{e}^{n+1} - \tilde{e}^n), \tilde{e}^{n+1} \rangle - 2\tau \left\langle d_h \left( \frac{f_0(X)}{D_h W^{n+1} D_h x_h^{n+1}} D_h \tilde{e}^{n+1} \right), \tilde{e}^{n+1} \right\rangle \\ &= -2\tau \left\langle \frac{f_0(X)}{m[f_0(X)]^{m-1}} \cdot \frac{W^{n+1} - W^n}{\tau} \cdot [(\tilde{D}_h W^n)^{m-1} - (\tilde{D}_h x_h^n)^{m-1}], \tilde{e}^{n+1} \right\rangle \\ & \quad + 2\tau \langle \tau^3 f^* + h^4 g^*, \tilde{e}^{n+1} \rangle, \end{aligned} \quad (3.13)$$

where

$$\alpha_n := \frac{f_0(x)}{m \left( \frac{f_0(X)}{D_h x_h^n} \right)^{m-1}}. \quad (3.14)$$

For the first term of the left side, we get

$$\begin{aligned} 2\langle \alpha_n(\tilde{e}^{n+1} - \tilde{e}^n), \tilde{e}^{n+1} \rangle &= \alpha_n \|\tilde{e}^{n+1}\|_2^2 + \alpha_n \|\tilde{e}^{n+1} - \tilde{e}^n\|_2^2 - \alpha_n \|\tilde{e}^n\|_2^2 \\ &\geq \alpha_n \|\tilde{e}^{n+1}\|_2^2 - \alpha_n \|\tilde{e}^n\|_2^2. \end{aligned} \quad (3.15)$$

For the second term of the left side, we see that

$$\begin{aligned} & -2\tau \left\langle d_h \left( \frac{f_0(X)}{D_h W^{n+1} D_h x_h^{n+1}} D_h \tilde{e}^{n+1} \right), \tilde{e}^{n+1} \right\rangle \\ &= 2\tau \left\langle \frac{f_0(X)}{D_h W^{n+1} D_h x_h^{n+1}} D_h \tilde{e}^{n+1}, D_h \tilde{e}^{n+1} \right\rangle_e \geq 0, \end{aligned} \quad (3.16)$$

in which the summation by parts formula (2.10) is applied with  $\tilde{e}_0^{n+1} = \tilde{e}_N^{n+1} = 0$ .

For the right side term, we have

$$\begin{aligned} & -2\tau \left\langle \frac{f_0(X)}{m[f_0(X)]^{m-1}} \cdot \frac{W^{n+1} - W^n}{\tau} \cdot [(\tilde{D}_h W^n)^{m-1} - (\tilde{D}_h x_h^n)^{m-1}], \tilde{e}^{n+1} \right\rangle \\ &= -2\tau \left\langle \frac{f_0(X)}{m[f_0(X)]^{m-1}} \cdot \frac{W^{n+1} - W^n}{\tau} \cdot [(m-1)(\tilde{D}_h \zeta^n)^{m-2} \tilde{D}_h \tilde{e}^n], \tilde{e}^{n+1} \right\rangle \\ &\leq 2\tau C_1 \|\tilde{D}_h \tilde{e}^n\|_2 \|\tilde{e}^{n+1}\|_2 \quad \left( C_1 := \frac{(m-1)B_f C_t^* C_\zeta^{m-2}}{m b_f^{(m-1)}} \right) \\ &\leq \tau C_1 \|\tilde{D}_h \tilde{e}^n\|_2^2 + \tau C_1 \|\tilde{e}^{n+1}\|_2^2, \end{aligned} \quad (3.17)$$

in which  $C_t^* = \|W_t\|_\infty$ ,  $\tilde{D}_h \zeta$  is between  $\tilde{D}_h x_h^n$  and  $\tilde{D}_h W^n$ ,  $\|\tilde{D}_h \zeta\|_\infty \leq C_\zeta$ , with

$$C_\zeta := \begin{cases} C_0^*, & m \geq 2, \\ \delta_0/2, & m < 2. \end{cases}$$

The local truncation error term could be bounded by the standard Caught inequality:

$$\begin{aligned}
& 2\tau \langle \tau^3 l^* + h^4 g^*, \tilde{e}^{n+1} \rangle \\
& \leq \tau \|\tau^3 l^* + h^4 g^*\|_2^2 + \tau \|\tilde{e}^{n+1}\|_2^2 \\
& \leq \tau C(\tau^3 + h^4)^2 + \tau \|\tilde{e}^{n+1}\|_2^2.
\end{aligned} \tag{3.18}$$

Next we estimate  $\|D_h x_h^{n+1}\|_\infty$  roughly. Based on (3.14),  $\alpha_n$  can be estimated by

$$C_\alpha := \frac{b_f}{m(\frac{B_f}{\delta_0/2})^{m-1}} \leq \|\alpha\|_n \leq \frac{B_f}{m b_f^{m-1}} (C_0^*)^{m-1} := \bar{C}_\alpha.$$

A substitution of (3.15)-(3.18) into (3.13), in combination with (3.11a), leads to

$$\begin{aligned}
& (\alpha_n - \tau(1 + C_1)) \|\tilde{e}^{n+1}\|_2^2 \\
& \leq \alpha_n \|\tilde{e}^n\|_2^2 + \tau C_1 \|\tilde{D}_h \tilde{e}^n\|_2^2 + \tau C(\tau^3 + h^4)^2 \\
& \leq \tau \bar{C}(\tau^{\frac{7}{4}} + h^{\frac{5}{2}})^2,
\end{aligned}$$

where  $\bar{C}$  is dependent on  $C$ ,  $C_1$  and  $\bar{C}_\alpha$ . Then we get

$$\|\tilde{e}^{n+1}\|_2^2 \leq \tilde{C}^2 \tau (\tau^{\frac{7}{4}} + h^{\frac{5}{2}})^2, \quad \text{i.e., } \|\tilde{e}^{n+1}\|_2 \leq \tilde{C} \tau^{\frac{1}{2}} (\tau^{\frac{7}{4}} + h^{\frac{5}{2}}), \tag{3.19}$$

with

$$\tilde{C} := \left( \frac{\bar{C}}{C_\alpha/2} \right)^{\frac{1}{2}}, \quad \text{if } \tau(1 + C_1) \leq C_\alpha/2.$$

Based on the inverse inequality (2.11), we obtain that, by choosing  $h = \mathcal{O}(\tau)$ ,

$$\|\tilde{e}^{n+1}\|_\infty \leq \frac{C_m \|\tilde{e}^{n+1}\|_2}{h^{\frac{1}{2}}} \leq C_m \tilde{C} (\tau^{\frac{7}{4}} + h^{\frac{5}{2}}). \tag{3.20}$$

Then we have

$$\begin{aligned}
\|D_h x_h^{n+1}\|_\infty &= \|D_h W^{n+1} - D_h \tilde{e}^{n+1}\|_\infty \\
&\leq C^* + C_m \tilde{C} (\tau^{\frac{3}{4}} + h^{\frac{3}{2}}) \leq C^* + 1 := C_0^*,
\end{aligned} \tag{3.21}$$

if  $C_m \tilde{C} (\tau + h^2) \leq 1$ .

As a result, (3.16) can be re-estimated as follows:

$$2\tau \left\langle \frac{f_0(X)}{D_h W^{n+1} D_h x_h^{n+1}} D_h \tilde{e}^{n+1}, D_h \tilde{e}^{n+1} \right\rangle_e \geq 2\tau C_2 \|D_h \tilde{e}^{n+1}\|_2^2, \tag{3.22}$$

with  $C_2 := \frac{b_f}{C^* C_0^*}$ .

As a consequence, a substitution of (3.15)-(3.18) with (3.22) into (3.13) leads to

$$\begin{aligned} & \alpha_n \|\tilde{e}^{n+1}\|_2^2 - \alpha_n \|\tilde{e}^n\|_2^2 + \tau C_2 \|\tilde{D}_h \tilde{e}^{n+1}\|_2^2 \\ & \leq \tau \left(1 + \frac{C_1^2}{C_2}\right) \|\tilde{e}^{n+1}\|_2^2 + \tau C(\tau^3 + h^4)^2, \end{aligned}$$

where the following estimates are applied:  $\|\tilde{D}_h x_h^n\|_2 \leq \|D_h x_h^n\|_2$  and

$$2\tau C_1 \|\tilde{D}_h \tilde{e}^n\|_2 \|\tilde{e}^{n+1}\|_2 \leq \tau \frac{C_1^2}{C_2} \|\tilde{e}^{n+1}\|_2^2 + \tau C_2 \|\tilde{D}_h \tilde{e}^n\|_2^2. \quad (3.23)$$

Subsequently, a summation in time shows that

$$\begin{aligned} & \alpha_n \|\tilde{e}^{n+1}\|_2^2 + \tau C_2 \sum_{k=1}^{n+1} \|\tilde{D}_h \tilde{e}^k\|_2^2 \\ & \leq \tau \sum_{k=1}^n \frac{(\alpha_k - \alpha_{k-1})}{\tau} \|\tilde{e}^k\|_2^2 + \tau \left(\frac{C_1^2}{C_2} + 1\right) \sum_{k=1}^{n+1} \|\tilde{e}^k\|_2^2 + CT(\tau^3 + h^4)^2, \\ & \|\tilde{e}^{n+1}\|_2^2 + \tau \frac{C_2}{C_\alpha} \sum_{k=1}^{n+1} \|\tilde{D}_h \tilde{e}^k\|_2^2 \\ & \leq \frac{\tau}{C_\alpha} \left(\frac{C_1^2}{C_2} + 1 + \tilde{C}_\alpha\right) \sum_{k=1}^{n+1} \|\tilde{e}^k\|_2^2 + \frac{CT}{C_\alpha} (\tau^3 + h^4)^2, \end{aligned}$$

where we have used the estimate

$$\begin{aligned} \left\| \frac{\alpha^k - \alpha^{k-1}}{\tau} \right\|_\infty &= \left\| \frac{f_0(X)}{m[f_0(X)]^{m-1}} \cdot \frac{(\tilde{D}_h x_h^k)^{m-1} - (\tilde{D}_h x_h^{k-1})^{m-1}}{\tau} \right\|_\infty \\ &= \left\| \frac{f_0(X)}{m[f_0(X)]^{m-1}} (m-1) (\tilde{D}_h \vartheta)^{m-2} \frac{\tilde{D}_h x_h^k - \tilde{D}_h x_h^{k-1}}{\tau} \right\|_\infty \\ &\leq \frac{(m-1)B_f}{mb_f^{m-1}} (C_\vartheta)^{m-2} (\tilde{C}_t^* + 1) := \tilde{C}_\alpha. \end{aligned}$$

It is noticed that  $T$  is the terminal time, (3.11f) is applied and  $\tilde{D}_h \vartheta$  is between  $\tilde{D}_h x_h^k$  and  $\tilde{D}_h x_h^{k-1}$  with

$$\|\tilde{D}_h \vartheta\|_\infty \leq C_\vartheta := \begin{cases} C_0^*, & m \geq 2, \\ \delta_0/2, & m < 2. \end{cases}$$

In turn, an application of discrete Gronwall inequality yields the desired convergence result:

$$\|\tilde{e}^{n+1}\|_2^2 + \tau \frac{C_2}{C_\alpha} \sum_{k=1}^{n+1} \|\tilde{D}_h \tilde{e}^k\|_2^2 \leq e^{TC_0} \frac{CT}{C_\alpha} (\tau^3 + h^4)^2,$$

i.e.,

$$\|\tilde{e}^{n+1}\|_2 \leq \gamma(\tau^3 + h^4),$$

where  $C_0 := \frac{1}{C_\alpha}(\frac{C_1^2}{C_2} + \tilde{C}_\alpha + 1)$  and

$$\gamma := \left(\frac{CT}{C_\alpha}\right)^{\frac{1}{2}} e^{\frac{C_0 T}{2}}. \quad (3.24)$$

Therefore, the a-priori assumption (3.10) is also valid at  $t^{n+1}$ :

$$\|\tilde{e}^{n+1}\|_2 \leq \gamma(\tau^3 + h^4) \leq \tau^{\frac{11}{4}} + h^{\frac{7}{2}}, \quad (3.25)$$

provided that  $\tau \leq \gamma^{-4}$ ,  $h \leq \gamma^{-2}$ .

Based on the following estimate

$$\|\tilde{D}_h \tilde{e}^{n+1}\|_2 = \|\tilde{D}_h x_h^{n+1} - \tilde{D}_h W^{n+1}\|_2 \leq C\gamma(\tau^2 + h^3), \quad (3.26)$$

we obtain

$$\|\tilde{D}_h x_h^{n+1} - \tilde{D}_h x_e^{n+1}\|_2 \leq C(\tau + h^2). \quad (3.27)$$

Finally, we estimate the error between the numerical solution  $f_h^{n+1}$  and the exact solution  $f_e^{n+1}$  of the problem (2.1a)-(2.1d):

$$\begin{aligned} \|f_e^{n+1} - f_h^{n+1}\|_2 &= \left\| \frac{f_0(X)}{\partial_X x_e^{n+1}} - \frac{f_0(X)}{\tilde{D}_h x_h^{n+1}} \right\|_2 \\ &= \left\| \frac{f_0(X)}{\partial_X x_e^{n+1}} - \frac{f_0(X)}{\tilde{D}_h x_e^{n+1}} + \frac{f_0(X)}{\tilde{D}_h x_e^{n+1}} - \frac{f_0(X)}{\tilde{D}_h x_h^{n+1}} \right\|_2 \leq C(\tau + h^2). \end{aligned}$$

Thus, we complete the proof.  $\square$

#### 4. Numerical results

In this section, we present some numerical results 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:

$$\|e_h\|_2^2 = \frac{1}{2} \left( e_{h_0}^2 h_{x_0} + \sum_{i=1}^{M-1} e_{h_i}^2 h_{x_i} + e_{h_M}^2 h_{x_M} \right), \quad (4.1a)$$

$$\|e_h\|_\infty = \max_{0 \leq i \leq M} \{ |e_{h_i}| \}, \quad (4.1b)$$

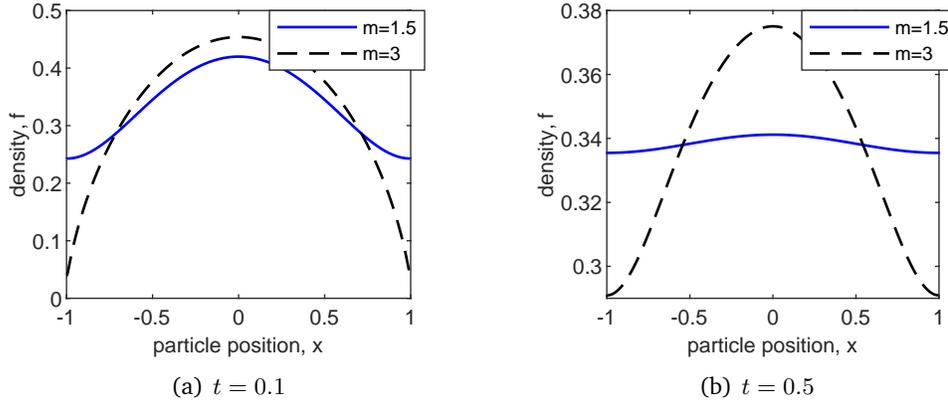


Figure 1: The evolution of  $f$  ( $h = 1/1000$ ,  $\tau = 1/1000$ ).

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

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

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

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

where  $h$  is the spatial step.

Consider the problem (2.1a)-(2.1d) in dimension one with a smooth positive initial data

$$f_0(x) = \frac{1}{2}(-x^2 + 1.01), \quad x \in \Omega := [-1, 1]. \quad (4.2)$$

Firstly, the trajectory equation (2.7a) with the initial and boundary condition (2.7b)-(2.7c) can be solved by the fully discrete scheme (2.12). And then the density function  $f$  in (2.4) can be approximated by (2.13). The reference “exact” solution is obtained numerically on a much finer mesh with  $h = \frac{1}{10000}$ ,  $\tau = \frac{1}{10000}$ .

Table 1 shows the convergence rate with  $m = 1.5$  and  $m = 3$  at time  $T = 0.5$ . The rate for density  $f$  and trajectory  $x$  in the  $\mathcal{L}^2$  and  $\mathcal{L}^\infty$  norm is second order in space and first order in time without dependence on  $m$ . Fig. 1 presents the density  $f$  at time  $t = 0.1$  and  $t = 0.5$  for both values of  $m$ . The results imply that the speed of diffusion decreases as  $m$  increases. Fig. 2 displays the evolution of particles whose initial positions are  $X = -0.001, 0.000, 0.001$ , respectively. We see that particles move outward at a finite speed. However, the speed is lower as  $m$  increases, except for the center point which remains stationary. As shown in Fig. 3, the total energy decays as time evolves for both values of  $m$  and the decreasing rate is slowed down as  $m$  increases.

More interesting examples can be found in [5], such as a free boundary problem with an exact Barenblatt solution, the waiting time phenomenon and the problem with two support sets at the initial state.

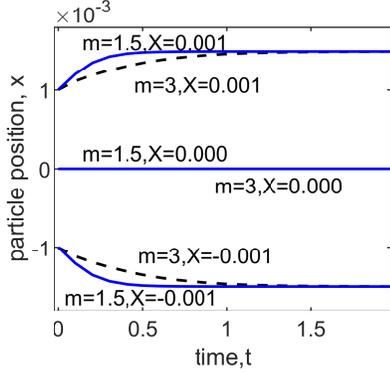


Figure 2: The evolution of particle with initial position  $X = -0.001, 0.000, 0.001$  ( $h = 1/1000$ ,  $\tau = 1/1000$ ).

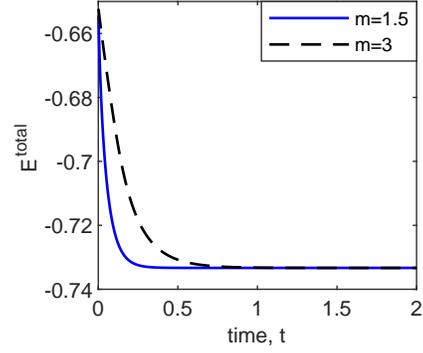


Figure 3: The evolution of total energy ( $h = 1/1000$ ,  $\tau = 1/1000$ ).

Table 1: Convergence rate of solution  $f$  and trajectory  $x$  at time  $T = 0.5$ .

$m = 1.5$									
$h$	$\tau$	$\mathcal{L}^2$ -error ( $f$ )	Order	$\mathcal{L}^\infty$ -error( $f$ )	Order	$\mathcal{L}^2$ -error ( $x$ )	Order	$\mathcal{L}^\infty$ -error ( $x$ )	Order
1/10	1/10	1.5683e-02		1.1680e-01		1.5122e-03		5.2129e-03	
1/20	1/40	2.8389e-03	2.7621	3.8639e-02	1.5114	1.4812e-03	2.3315	1.1415e-03	2.2833
1/40	1/160	4.8401e-04	2.9327	1.0515e-02	1.8373	3.1765e-04	2.1501	2.6587e-04	2.1468
1/80	1/640	9.2853e-05	2.6063	2.6850e-03	1.9582	7.3867e-05	2.1474	6.1999e-05	2.1441
$m = 3$									
$h$	$\tau$	$\mathcal{L}^2$ -error ( $f$ )	Order	$\mathcal{L}^\infty$ -error( $f$ )	Order	$\mathcal{L}^2$ -error ( $x$ )	Order	$\mathcal{L}^\infty$ -error ( $x$ )	Order
1/10	1/10	1.7044e-02		1.1471e-01		1.9900e-03		7.1606e-03	
1/20	1/40	3.0278e-03	2.8146	3.5746e-02	1.6045	4.9937e-04	1.9924	1.8319e-03	1.6045
1/40	1/160	5.5190e-04	2.7431	9.6637e-03	1.8495	1.2332e-04	2.0247	4.5340e-04	1.8495
1/80	1/640	1.1139e-04	2.4773	2.4584e-03	1.9655	2.9237e-05	2.1089	1.0757e-04	1.9655

$\mathcal{L}^2$ -error and  $\mathcal{L}^\infty$ -error is defined by (4.1a) and (4.1b), respectively.  
 $\tau$  is the time step and  $h$  is the space step.

## 5. Conclusions

The numerical methods of the PME based on EnVarA has been proposed and studied in [5], while a theoretical justification for optimal convergence analysis has not been available. In this paper, we prove the second order spatial convergence and the first order temporal convergence for the nonlinear numerical scheme. A careful asymptotic expansion for the exact solution in terms of the numerical scheme is applied to obtain higher order consistency. Furthermore, we use two step error estimates: a rough estimate to control a discrete  $W^{1,\infty}$  bound of the numerical solution and a refined estimate to derive the desired convergence result.

One obvious limitation of this work is associated with the one-dimensional nature of the problem. In two or higher dimension, 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 that 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.

**Acknowledgements** The work of Yue is supported in part by NSF of China under the grants No. 11971342. Chun Liu and Cheng Wang are partially supported by NSF grants DMS-1216938, DMS-1418689, respectively.

## Reference

- [1] D. G. ARONSON, *Regularity properties of flows through porous media*, SIAM J. Appl. Math., 17 (1969), pp. 461–467.
- [2] 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 (2013), pp. 2851–2873.
- [3] E. DiBENEDETTO AND D. HOFF, *An interface tracking algorithm for the porous medium equation*, Trans. Am. Math. Soc., 284 (1984), pp. 463–500.
- [4] C. DUAN, C. LIU, C. WANG AND X. YUE, *Numerical complete solution for random genetic drift by energetic variational approach*, ESAIM: M2AN, 53 (2019), pp. 615–634.
- [5] C. DUAN, C. LIU, C. WANG AND X. YUE, *Numerical methods for Porous medium equation by an energetic variational approach*, J. Comput. Phys., 385 (2019), pp. 13–32.
- [6] W. E AND J. G. LIU, *Projection method I: convergence and numerical boundary layers*, SIAM J. Numer. Anal., 32 (1995), pp. 1017–1057.
- [7] W. E AND J. G. LIU, *Projection method III. spatial discretization on the staggered grid*, Math. Comput., 71 (2002), pp. 27–47.
- [8] J. L. GRAVELEAU AND P. JAMET, *A finite difference approach to some degenerate nonlinear parabolic equations*, SIAM J. Appl. Math., 20 (1971), pp. 199–223.
- [9] 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. Model. Appl. Sci., 40 (2017), pp. 6836–6863.
- [10] Z. GUAN, C. WANG AND S. M. WISE, *A convergent convex splitting scheme for the periodic nonlocal Cahn-Hilliard equation*, Numer. Math., 128 (2014), pp. 377–406.
- [11] S. JIN, L. PARESCHI AND G. TOSCANI, *Diffusive relaxation schemes for multi-scale discrete-velocity kinetic equations*, SIAM J. Numer. Anal., 35(6) (1998), pp. 2405–2439.
- [12] A. S. KALAŠNIKOV, *Formation of singularities in solutions of the equation of nonstationary filtration*, Ž. Vyčisl. Mat. Mat. Fiz., 7 (1967), pp. 440–444.
- [13] C. NGO AND W. Z. HUANG, *A study on moving mesh finite element solution of the porous medium equation*, J. Comput. Phys., 331 (2017), pp. 357–380.

- [14] 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 (1958), pp. 667–704.
- [15] 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 (2003), pp. 1163–1194.
- [16] S. I. SHMAREV, *Interfaces in multidimensional diffusion equations with absorption terms*, *Non-linear Anal.*, 53 (2003), pp. 791–828.
- [17] S. SHMAREV, *Interfaces in solutions of diffusion-absorption equations in arbitrary space dimension*, in: *Trends in Partial Differential Equations of Mathematical Physics*, in: *Progr. Nonlinear Differential Equations Appl.*, Birkhäuser, Basel, (2005), pp. 257–273.
- [18] J. L. VÁZQUEZ, *The Porous Medium Equation*, Oxford University Press, Oxford, 2007.
- [19] C. WANG, J. G. LIU AND H. JOHNSTON, *Analysis of a fourth order finite difference method for incompressible Boussinesq equations*, *Numer. Math.*, 97 (2004), pp. 555–594.
- [20] C. WANG AND J. G. LIU, *Analysis of finite difference schemes for unsteady Navier-Stokes equations in vorticity formulation*, *Numer. Math.*, 91 (2002), pp. 543–576.
- [21] C. WANG AND J. G. LIU, *Convergence of gauge method for incompressible flow*, *Math. Comput.*, 69 (2000), pp. 1385–1407.
- [22] 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 (2015), pp. 347–366.
- [23] M. WESTDICKENBERG AND J. WILKENING, *Variational particle schemes for the porous medium equation and for the system of isentropic Euler equations*, *ESAIM: M2AN.*, 44(1) (2010), pp. 133–166.
- [24] Q. ZHANG AND Z. L. WU, *Numerical simulation for porous medium equation by local discontinuous Galerkin finite element method*, *J. Sci. Comput.*, 38(2) (2009), pp. 127–148.