The variational formulation of the fully parabolic Keller-Segel system with degenerate diffusion

退化拡散項を持つ完全放物型 Keller-Segel 系に対する変分的定式化

三村　与士文

The variational formulation of the fully parabolic Keller-Segel system with degenerate diffusion

Yoshifumi MIMURA
Graduate School of Mathematical Sciences
The University of Tokyo, Japan

Abstract

We prove the time global existence of solutions of the degenerate Keller-Segel system, under the assumption that the mass of the first component is below a certain critical value. What we deal with is the full parabolic-parabolic system rather than the simplified parabolic-elliptic system. Our approach is to formulate the problem as a gradient flow on the Wasserstein space — a new approach for the parabolic-parabolic system.

1 Introduction

1.1 Description of the problem

We consider the following degenerate parabolic system:

$$\begin{cases} \partial_t u = \nabla\cdot(\nabla u^m - \chi u\nabla v), & x\in\Omega,\ t>0, \\ \varepsilon\partial_t v = \Delta v - \gamma v + \alpha u, & x\in\Omega,\ t>0, \\ u(x,0)=u_0(x),\ \varepsilon v(x,0)=\varepsilon v_0(x), & x\in\Omega, \end{cases} \tag{1}$$

where $\alpha,\chi,\gamma,\varepsilon,m$ are constants satisfying $\alpha,\chi>0,\ \gamma\geq 0,\ \varepsilon>0, m\geq 2-\frac{2}{d}, d>2$ and Ω is a bounded domain in $\mathbb{R}^d$ with smooth boundary. We impose the following boundary conditions:

$$\frac{\partial u^m}{\partial\nu} - \chi u\frac{\partial v}{\partial\nu} = v = 0,\quad x\in\partial\Omega,\ t>0. \tag{2}$$

The aim of this thesis is to prove the time global existence of solutions of the system (1) under the assumption that the initial mass $\int_\Omega u_0\,dx$ is below a certain critical mass and for an arbitrary $v_0\geq 0$.

Our approach is first to formulate (1) as a gradient flow on a certain metric space, then to apply the techniques of [3] to prove the time global existence. Note that the system (1) does not have a gradient structure in standard function spaces such as L^2 because of the presence of the drift term $\nabla\cdot(\chi u\nabla v)$. This is why we use the Wasserstein formulation.

The system (1) was proposed by Keller and Segel [13] in 1970 to describe an aggregation phenomenon of certain microorganisms called "*slime molds*", which have a characteristic property called chemotaxis. Chemotaxis is a motion toward higher concentration of a chemical substance. This kind of microorganisms, when put in a nutrition-poor environment, produces a chemical substance that attracts other individuals within the same population. This leads to formation of an aggregate, which produces spores. In this way, the slime molds propagate next generation. In equations (1), u stands for the density of slime molds and v stands for the concentration of the chemical substance, hence we are interested in non-negative solutions of (1). From the mathematical aspect, the aggregation phenomenon can be interpreted as a blow-up phenomenon of the solution of (1), that is, the density of slime molds singularly concentrates at some point.

Mathematically, it is known that whether the above aggregation phenomenon occurs or not depends largely on the mass of u. Notice that (1) preserves the mass $\int_\Omega u\,dx$. In particular, the case $m = 1$ is well-understood and it is known that the following sharp threshold M_c exists provided that $d = 2$ (which is the case with the classical Keller-Segel model): If $\|u_0\|_{L^1} < M_c$ then the solution exists globally in time, while for any $M > M_c$ there exists a solution with $\|u_0\|_{L^1} = M$ that blows up in finite time [15, 8, 7, 5]. These results hold true both for the case $\varepsilon > 0$ (parabolic-parabolic system) and for the case $\varepsilon = 0$ (parabolic-elliptic system).

In the case $m > 1$, the situation is more complicated. It is known, at least formally and partially rigorously, that the above-mentioned sharp threshold mass M_c can exist (if it ever does) only when $m = 2 - \frac{2}{d}$ (hence $m = 1$ if $d = 2$). Early results in this direction were given by Sugiyama [19], who showed that there exist constants $0 < M_1 < M_2$ (whose values are specified in [19]) such that if $0 < \|u_0\|_{L^1} < M_1$ the solutions exist globally in time, while there exist a solution with $\|u_0\|_{L^1} > M_2$ that blows up in finite time. Later, under the additional assumption that $\gamma = 0$, Blanchet, Carrillo and Laurençot [6] showed the existence of a sharp threshold mass M_c as in the case $m = 1, d = 2$, which we have mentioned above.

At present, the existence of such threshold phenomena is known only for the parabolic-elliptic case with $m = 2 - 2/d$. As for the parabolic-parabolic case, the recent work of Ishida and Yokota [11] proves the time global existence under the assumption that both u_0 and Δv_0 are relatively small – a condition that is too restrictive compared with what we have mentioned for the parabolic-elliptic system. So far, there has been no result that suggests the existence of a threshold mass for the parabolic-parabolic system (1).

In the case $m > 2 - \frac{2}{d}$, it is shown by Sugiyama [20] for $\varepsilon = 0$ and by Ishida and Yokota [10] for $\varepsilon = 1$ that solutions of the system (1) exist globally in time without any restriction on the size of the mass. That is, the solutions never blow-up.

As mentioned earlier, our approach is to formulate (1) as a gradient flow in a certain metric space. One of the advantages of this approach is that it gives us better understanding of the relation between the time global existence of (1) and the variational properties of the Lyapunov functional ϕ_m, which is to be defined in Section 1.3. More precisely, our approach shows that the lower boundedness of the Lyapunov functional guarantees the time global existence of the solution of (1). Our approach is, in its spirit, similar to that of Blanchet, Calvez and Carrillo [5], who formulated the non-degenerate parabolic-elliptic Keller-Segel system (where one sets $m = 1$ and $\varepsilon = \gamma = 0$ in (1)) as a gradient flow in a Wasserstein space. In the present case, where $\varepsilon > 0$, one cannot reduce (1) to a single non-local equation. Nonetheless, we can still formulate it partly in the framework of the Wasserstein space, as we will show later. The Wasserstein techniques for the present type of evolution PDE's were developed in the pioneering work of Otto [16, 17], Jordan, Kinderlehrer and Otto [12], and other related papers such as [1, 3, 14].

A common strategy in the above-mentioned papers [1, 3, 5, 12, 14, 16] is first to approximate the evolution equation by a time-discrete problem, which consists of solving a certain minimization problem at each time step. One then proves the convergence of the approximate solution to a weak solution of the original evolution equation as the time mesh size tends to 0. In proving the convergence, one needs some compactness properties of the time-discrete solution, but the minimizing nature of the time discretization simplifies the compactness argument significantly.

There are two approaches for proving the convergence of discretized solution. One is to use the Euler-Lagrange equation associated with the minimization problem at each time step. This Euler-Lagrange equation is written explicitly in the form of a backward Euler difference scheme for the original evolution equation with some penalty term. The other approach, found in [3, 14], uses the concept of "curves of maximal slope", which is formulated in the framework of abstract metric spaces. The former approach based on the Euler-Lagrange equation is visually more explicit than the latter, but this latter approach based on the notion of curves of maximal slope requires milder compactness properties in the convergence proof, which is an advantage. The present thesis adopts this latter approach.

Note that, in this latter approach, the subdifferentials of the functional play a crucial role. In [3], existence of subdifferentials having certain good properties is shown for what they call "regular" functionals. However, in our present problem, it is not easy to check if our Lyapunov functional is regular. We therefore will use a different argument that guarantees the existence of subdifferentials of the Lyapunov functional at every point where the approximate discrete solution curves pass, which is sufficient for the convergence proof. More precisely, we use a two-step time discretization scheme, in which the solution of the next time level is given by solving a minimization problem for u (in the Wasserstein space) and one for v (in L^2) alternately rather than simultaneously. With this new scheme one can obtain the regularity of the discrete solution, which is sufficient to exist the subdifferentials of our Lyapunov functional.

The main difference between our results and the earlier results by Ishida-Yokota [11] is that our variational approach makes it possible to obtain global existence more directly from the Lyapunov functional, which is known to play a fundamental role in determining the sharp threshold mass for the parabolic-elliptic system. More precisely, our results show the time-global existence of solutions of (1) under a rather mild condition $\|u_0\|_{L^1} < M_*$ (and for an arbitrary v_0), where M_* is the critical mass for the Lyapunov functional to be bounded from below. Furthermore, the constant M_* coincides with the sharp threshold mass M_c given by Blanchet *et al.* [6] for the parabolic-elliptic case. Considering that the threshold does not depend on ε in the case of $m = 1$, we suspect that M_* is the sharp threshold even for our parabolic-parabolic system with $m = 2 - \frac{2}{d}$. Note that uniqueness is not known for the initial value problem (1), expect that some partial results are given by Sugiyama [22] for the parabolic-elliptic case $\varepsilon = 0$.

1.2 Main Results

Now we state our main results. The following functional ϕ_m is known as a Lyapunov functional associated with the Keller-Segel system (1):

$$\phi_m(u, v) := \frac{1}{m-1} \int_\Omega u^m \, dx - \chi \int_\Omega uv \, dx + \frac{\chi}{2\alpha} \int_\Omega |\nabla v|^2 + \gamma v^2 \, dx.$$

We consider the functional ϕ_m in the space

$$X_M := \Big\{(u, v) \in \big(L^1 \cap L^m(\Omega)\big) \times H_0^1(\Omega) \, ; \, \|u\|_{L^1} = M, u \geq 0, v \geq 0\Big\}.$$

We define μ_M and M_* by

$$\mu_M := \inf_{(u,v)\in X_M} \phi_m(u, v),$$

$$M_* := \sup\{M \geq 0 \, ; \, \mu_M > -\infty\}. \tag{3}$$

Theorem 1.1 (well-definedness and properties of M_*). *M_* has the following properties:*

(i) $M_* > 0$

(ii) We have $\mu_M > -\infty$ for every $M \leq M_*$, while we have $\mu_M = -\infty$ for every $M > M_*$.

Particularly, we have $M_* = \infty$ for $m > 2 - \frac{2}{d}$ and $M_* < \infty$ for $m = 2 - \frac{2}{d}$. That is, in the case of $m = 2 - \frac{2}{d}$, M_* is the threshold of the lower bounds of Lyapunov functional ϕ_m. Moreover, M_* depends on α, χ, d and we have $M_*(\alpha, \chi, d) = (\alpha\chi)^{-\frac{d}{2}} M_*(1, 1, d)$.

Next, we state the time-global existence of solutions of the system (1). Here the meaning of solutions of (1) is understood in the sense of (16).

Theorem 1.2 (time-global existence). *For any $u_0 \in L^2(\Omega)$ and $v_0 \in H_0^1(\Omega)$ with $u_0, v_0 \geq 0$, there exists a non-negative solution of (1) with this initial data that exists globally for all $t \geq 0$, provided that u_0 satisfies*

$$\int_\Omega u_0 \, dx < M_*.$$

1.3 Formal derivation

In this subsection we will explain how to formulate the system (1) as a gradient flow for the following functional:

$$\phi_m(u,v) := \frac{1}{m-1}\int_\Omega u^m\,dx - \chi\int_\Omega uv\,dx + \frac{\chi}{2\alpha}\int_\Omega |\nabla v|^2 + \gamma v^2\,dx.$$

The arguments below are not rigorous but may highlight the underlying ideas that will come in the later sections.

Let us first consider the following system of equations:

$$\begin{cases} \partial_t u + \nabla\cdot(u\xi) = 0 & \text{in } \Omega\times(0,T) \\ \partial_t v + \eta = 0 & \text{in } \Omega\times(0,T) \end{cases} \tag{4}$$

with the boundary conditions: $(\boldsymbol{\nu}\cdot\xi)u = 0$ and $v = 0$ on $\partial\Omega, t > 0$. Here the vector field $\xi : \Omega \to \mathbb{R}^d$ and the scalar field $\eta : \Omega \to \mathbb{R}$ are to be specified later. We assume that $w := (u,v)$ is sufficiently smooth. The time derivative of $\phi_m(w(t))$ for $w(t) = (u(t), v(t))$ is then given by

$$\begin{aligned} \frac{d}{dt}\phi_m(w(t)) &= \int_\Omega \left(\frac{m}{m-1}u^{m-1} - \chi v\right)\partial_t u\,dx - \frac{\chi}{\alpha}\int_\Omega (\Delta v - \gamma v + \alpha u)\partial_t v\,dx \\ &= \int_\Omega \langle g_1(w), \xi\rangle u\,dx + \frac{\varepsilon\chi}{\alpha}\int_\Omega g_2(w)\eta\,dx, \end{aligned} \tag{5}$$

where $\langle\cdot,\cdot\rangle$ denotes standard inner product in $\mathbb{R}^d$ and

$$g(w) = (g_1(w), g_2(w)) = \left(\nabla\left(\frac{m}{m-1}u^{m-1} - \chi v\right),\ \frac{\Delta v - \gamma v + \alpha u}{\varepsilon}\right).$$

We introduce the following inner product $\langle\cdot,\cdot\rangle_w$ at the point $w = (u,v)$: for a vector field $\xi_i : \Omega \to \mathbb{R}^d$ and a scalar field $\eta_i : \Omega \to \mathbb{R}$, $i = 1,2$, we define

$$\langle\zeta_1,\zeta_2\rangle_w := \int_\Omega \langle\xi_1,\xi_2\rangle u\,dx + \frac{\varepsilon\chi}{\alpha}\int_\Omega \eta_1\eta_2\,dx \quad \text{for } \zeta_i = (\xi_i, \eta_i).$$

Using this notation, (5) can be rewritten as

$$\frac{d}{dt}\phi_m(w(t)) = \langle g(w), \zeta\rangle_w, \quad \zeta = (\xi, \eta). \tag{6}$$

Therefore by the Cauchy-Schwarz inequality and the Young inequality, we have

$$\begin{aligned} \frac{d}{dt}\phi_m(w(t)) &= \langle g(w), \zeta\rangle_w \\ &\ge -\|g(w)\|_w\|\zeta\|_w \\ &\ge -\frac{1}{2}\|g(w)\|_w^2 - \frac{1}{2}\|\zeta\|_w^2. \end{aligned}$$

In particular, all the equalities hold if and only if $\zeta = -g(w)$. In this case, (4) coincides with (1), and the following identity holds:

$$\frac{d}{dt}\phi_m(w(t)) = -\frac{1}{2}\|g(w)\|_w^2 - \frac{1}{2}\|\zeta\|_w^2. \tag{7}$$

Conversely, if (4), (6) and (7) are satisfied, then $w = (u, v)$ is a solution of (1).

In this thesis, using the theory of minimizing movements [3], we will prove the existence of a curve w satisfying the relations (4), (6) and (7) in a weak sense. As a consequence, we obtain the time global existence of solutions of (1).

1.4 Notation

$\mathscr{L}^d$	d-dimensional Lebesgue measure
$\boldsymbol{t}_{\#}\mu$	push-forward of the measure μ through the map $\boldsymbol{t}$
$\boldsymbol{t}_\mu^\nu$	optimal transport map from a measure μ to a measure ν
d_W	Wasserstein distance
$D(\phi)$	effective domain of functional ϕ
$\|v'\|$	metric derivative of w in a metric space $\mathscr{S}$
$\|\partial\phi\|(v)$	metric slope of functional ϕ at v
$\mathscr{P}_2(\Omega)$	probability measures on Ω with finite second moment
$\mathscr{P}_2^r(\Omega)$	regular measures in $\mathscr{P}_2(\Omega)$ with finite second moment
$L^p(\Omega)$	p-summable functions on $\Omega \subset \mathbb{R}^d$ with respect to $\mathscr{L}^d$
$L_\mu^2(\Omega;\mathbb{R}^d)$	$\mathbb{R}^d$-valued 2-summable functions on Ω with respect to μ
$\|\cdot\|_{L^2(\mu)}$	the norm in $L_\mu^2(\Omega;\mathbb{R}^d)$
$AC(a,b;\mathscr{S})$	absolutely continuous curves in a metric space $\mathscr{S}$
$W^{k,p}(\Omega)$	Sobolev space over Ω

2 Preliminaries

In this section, we collect some results on the Wasserstein metric. We refer to the book [3] by Ambrosio, Gigli and Savaré.

$\mathscr{P}(\mathbb{R}^d)$ denotes the space of probability measures on $\mathbb{R}$ endowed with the following topology.

Definition 2.1 (narrow convergence). We say that a sequence $(\mu_n) \subset \mathscr{P}(\mathbb{R}^d)$ is *narrowly convergent* to $\mu \in \mathscr{P}(\mathbb{R}^d)$ as $n \to \infty$ if

$$\lim_{n\to\infty} \int_{\mathbb{R}^d} f(x)\, d\mu_n(x) = \int_{\mathbb{R}^d} f(x)\, d\mu(x)$$

for every function $f \in C_b(\mathbb{R}^d)$, the space of continuous and bounded real functions defined on $\mathbb{R}^d$.

We define the subset $\mathscr{P}_2(\mathbb{R}^d)$ of $\mathscr{P}(\mathbb{R}^d)$ by

$$\mathscr{P}_2(\mathbb{R}^d) := \left\{ \mu \in \mathscr{P}(\mathbb{R}^d) : \int_{\mathbb{R}^d} |x|^2 \, d\mu(x) < \infty \right\}, \tag{8}$$

and $\mathscr{P}_2^r(\mathbb{R}^d)$ denotes the subset of $\mathscr{P}_2(\mathbb{R}^d)$ whose element is absolutely continuous with respect to the Lebesgue measure.

For $\Omega \subset \mathbb{R}^d$, we identify $\mathscr{P}_2(\Omega)$ with the set of measures $\mu \in \mathscr{P}_2(\mathbb{R}^d)$ such that $\mu(\mathbb{R}^d \setminus \Omega) = 0$. If Ω is bounded, then $\mathscr{P}_2(\Omega)$ coincides with $\mathscr{P}(\Omega)$.

Definition 2.2 (push-forward)**.** Let $\mu, \nu \in \mathscr{P}(\mathbb{R}^d)$. If, for a μ-measurable map $\boldsymbol{t} : \mathbb{R}^d \to \mathbb{R}^d$ and for every $f \in C_b(\mathbb{R}^d)$, it holds that

$$\int_{\mathbb{R}^d} f(y) \, d\nu(y) = \int_{\mathbb{R}^d} f(\boldsymbol{t}(x)) \, d\mu(x),$$

then we say that ν is a *push-forward* of μ through $\boldsymbol{t}$ and write $\nu = \boldsymbol{t}_{\#}\mu$.

Definition 2.3 (Wasserstein distance)**.** *Wasserstein distance* d_W is defined by

$$d_W^2(\mu, \nu) = \inf_{p \in \Gamma(\mu,\nu)} \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^2 \, dp(x, y) \tag{9}$$

where the set $\Gamma(\mu, \nu)$ of transport plans between μ and ν is defined by

$$\Gamma(\mu, \nu) := \{ p \in \mathscr{P}(\mathbb{R}^d \times \mathbb{R}^d) : (\pi_1)_{\#}p = \mu \text{ and } (\pi_2)_{\#}p = \nu \}$$

with $\pi_1(x, y) = x$ and $\pi_2(x, y) = y$, that is,

$$\int_{\mathbb{R}^d \times \mathbb{R}^d} f(x) \, dp = \int_{\mathbb{R}^d} f(x) \, d\mu, \quad \int_{\mathbb{R}^d \times \mathbb{R}^d} f(y) \, dp = \int_{\mathbb{R}^d} f(y) \, d\nu$$

for every $f \in C_b(\mathbb{R}^d)$, where $\mathscr{P}(\mathbb{R}^d \times \mathbb{R}^d)$ denotes the set of probability measures on $\mathbb{R}^d \times \mathbb{R}^d$.

The space $(\mathscr{P}_2(\mathbb{R}^d), d_W)$ is a complete metric space and is called the "Wasserstein space" [3, 23].

Theorem 2.4 (Brenier's theorem)**.** *Let $\mu, \nu \in \mathscr{P}_2(\mathbb{R}^d)$. If μ is absolutely continuous with respect to the Lebesgue measure $\mathscr{L}^d$, then there exist the optimal transport plan p_0 and the optimal transport map $\mathbf{t}_\mu^\nu$ such that*

$$\begin{aligned} d_W(\mu, \nu)^2 &= \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^2 \, dp_0(x, y) \\ &= \int_{\mathbb{R}^d} |x - \boldsymbol{t}_\mu^\nu(x)|^2 \, d\mu(x). \end{aligned}$$

Moreover, the map $\boldsymbol{t}_\mu^\nu$ coincides μ-a.e. with the gradient of a convex function φ_0.

Before stating the next theorem, we recall the definitions of absolutely continuous curves in a complete metric space $(\mathscr{S}, d)$ and the metric derivative.

Definition 2.5 (absolutely continuous curves). Let $(\mathscr{S}, d)$ be a complete metric space and (a, b) a finite interval in $\mathbb{R}$. We say that a curve $v : (a, b) \to \mathscr{S}$ is *absolutely continuous* if there exists $m \in L^1(a, b)$ such that

$$d(v(t), v(s)) \leq \int_s^t m(r)\, dr \quad \forall (s, t) \subset (a, b).$$

We write $v \in AC(a, b; \mathscr{S})$ to mean that v is absolutely continuous in the above sense. If, in addition, $m \in L^p(a, b)$ with $p > 1$, then we write $v \in AC^p(a, b; \mathscr{S})$.

Definition 2.6 (metric derivative). For $v \in AC^p(a, b; \mathscr{S})$, $p \geq 1$, we define the *metric derivative* $|v'|$ as

$$|v'|(t) := \lim_{s \to t} \frac{d(v(t), v(s))}{|t - s|} \quad \text{for } t \in (a, b).$$

The metric derivative is the equivalent of the normed value of a tangent vector in Hilbert space. It is known that for $v \in AC^p(a, b; \mathscr{S})$, $p \geq 1$, the metric derivative exists for $\mathscr{L}^1$-a.e. $t \in (a, b)$ and that $|v'| \in L^p(a, b)$ (see for instance [3]).

Theorem 2.7 (Representing formula for absolutely continuous curves). *If $\mu \in AC(a, b; \mathscr{P}_2(\mathbb{R}^d))$ then there exists a unique vector field $\boldsymbol{\xi} : \mathbb{R}^d \times (a, b) \to \mathbb{R}^d$ such that*

$$\partial_t \mu_t + \nabla \cdot (\boldsymbol{\xi}_t \mu_t) = 0 \quad \text{in } \mathcal{D}'(\mathbb{R}^d \times (a, b))$$
$$\text{and } |\mu'|(t) = \|\boldsymbol{\xi}_t\|_{L^2(\mu(t))} \quad \text{for } \mathscr{L}^1\text{-a.e. } t \in (a, b),$$

where $\|\cdot\|_{L^2(\mu)}$ stands for the norm in $L^2_\mu(\mathbb{R}^d; \mathbb{R}^d)$.

3 Reformulation of the main results

In this section, we reformulate our main result Theorem 1.2 in an equivalent statement in a Wasserstein framework. To do so, we first need to normalize u to satisfy $\int_\Omega u\, dx = 1$, since the Wasserstein space is a space of probability measures. For this purpose, we make the following change of variables:

$$\tilde{u} = \frac{u}{M}, \quad \tilde{t} = M^{m-1} t, \quad \text{where } M := \int_\Omega u\, dx,$$

along with the new parameters:

$$\tilde{\chi} = \frac{\chi}{M^{m-1}}, \quad \tilde{\varepsilon} = M^{m-1}\varepsilon, \quad \tilde{\alpha} = \alpha M. \tag{10}$$

Then $(\tilde{u}, v)$ satisfies the same equations as (1) with the above new parameters; furthermore, we have $\int_\Omega \tilde{u}\, dx = 1$. Therefore, in what follows it suffices to consider only the solution of (1) that satisfy

$$\int_\Omega u_0\, dx = 1. \tag{11}$$

We consider a partition of the time interval $[0, +\infty)$ and we identify a partition with a sequence $\boldsymbol{\tau} := (\tau_1, \tau_2, \ldots, \tau_k, \ldots)$ by the relation

$$\begin{cases} 0 = t_{\boldsymbol{\tau}}^0 < t_{\boldsymbol{\tau}}^1 < t_{\boldsymbol{\tau}}^2 < \cdots < t_{\boldsymbol{\tau}}^k < \cdots \\ t_{\boldsymbol{\tau}}^k = t_{\boldsymbol{\tau}}^{k-1} + \tau_k \end{cases} \tag{12}$$

and let $|\boldsymbol{\tau}| := \max_{k\geq 1} \tau_k$.

Let $(u_{\boldsymbol{\tau}}^0, v_{\boldsymbol{\tau}}^0) = (u_0, v_0) \in (\mathscr{P}(\Omega) \cap L^2(\Omega)) \times H_0^1(\Omega)$ be initial data and for $k = 1, 2, 3, \ldots$, we recursively define $(u_{\boldsymbol{\tau}}^k, v_{\boldsymbol{\tau}}^k)$ by

$$\begin{aligned} v_{\boldsymbol{\tau}}^k &\in \underset{v\in H_0^1(\Omega)}{\operatorname{argmin}} \left\{ \phi_m(u_{\boldsymbol{\tau}}^{k-1}, v) + \frac{\varepsilon\chi}{2\alpha\tau_k} \|v - v_{\boldsymbol{\tau}}^{k-1}\|_{L^2}^2 \right\}, \\ u_{\boldsymbol{\tau}}^k &\in \underset{u\in \mathscr{P}(\Omega)}{\operatorname{argmin}} \left\{ \phi_m(u, v_{\boldsymbol{\tau}}^k) + \frac{1}{2\tau_k} d_W^2(u, u_{\boldsymbol{\tau}}^{k-1}) \right\}, \end{aligned} \tag{13}$$

that is, $v_{\boldsymbol{\tau}}^k$ minimizes

$$v \mapsto \phi_m(u_{\boldsymbol{\tau}}^{k-1}, v) + \frac{\varepsilon\chi}{2\alpha\tau_k} \|v - v_{\boldsymbol{\tau}}^{k-1}\|_{L^2}^2 \quad \text{in } H_0^1(\Omega)$$

and $u_{\boldsymbol{\tau}}^k$ minimizes

$$u \mapsto \phi_m(u, v_{\boldsymbol{\tau}}^k) + \frac{1}{2\tau_k} d_W^2(u, u_{\boldsymbol{\tau}}^{k-1}) \quad \text{in } \mathscr{P}(\Omega).$$

Notice that in the Euclidian case, the variational scheme

$$x_{\boldsymbol{\tau}}^k \in \underset{x\in\mathbb{R}^d}{\operatorname{argmin}} \left\{ f(x) + \frac{1}{2\tau_k} \left| x - x_{\boldsymbol{\tau}}^{k-1} \right|^2 \right\} \tag{14}$$

leads to the implicit Euler scheme

$$\frac{x_{\boldsymbol{\tau}}^k - x_{\boldsymbol{\tau}}^{k-1}}{\tau_k} = -\nabla f(x_{\boldsymbol{\tau}}^k)$$

where f is a function on $\mathbb{R}^d$.

Proposition 3.1. *If $M_* = M_*(\alpha, \chi, d) > 1$ and $(u_0, v_0) \in (\mathscr{P}(\Omega) \cap L^m(\Omega)) \times H_0^1(\Omega)$, then $(u_{\boldsymbol{\tau}}^k, v_{\boldsymbol{\tau}}^k)$ are well-defined for all $k \in \mathbb{N}$.*

Definition 3.2 (discrete solutions)**.** If the above minimization problems (13) are well-defined for every $k \in \mathbb{N}$, then we define the piecewise constant interpolation

$$\begin{cases} \overline{u}_{\boldsymbol{\tau}}(t) := u_{\boldsymbol{\tau}}^k & \text{for } t \in (t_{\boldsymbol{\tau}}^{k-1}, t_{\boldsymbol{\tau}}^k], \\ \overline{v}_{\boldsymbol{\tau}}(t) := v_{\boldsymbol{\tau}}^k & \text{for } t \in (t_{\boldsymbol{\tau}}^{k-1}, t_{\boldsymbol{\tau}}^k]. \end{cases} \tag{15}$$

We call $(\overline{u}_{\boldsymbol{\tau}}, \overline{v}_{\boldsymbol{\tau}})$ a *discrete solution.*

Theorem 3.3 (Lyapunov solution). *Assume* (11) *and* $M_* > 1$. *If* $(u_0, v_0) \in (\mathscr{P}(\Omega) \cap L^2(\Omega)) \times H_0^1(\Omega)$, *then a subsequence of discrete solution* $(\overline{u}_\tau, \overline{v}_\tau)$ *converges to* (u, v) *weakly in* $L^m(\Omega) \times H_0^1(\Omega)$ *for every* $t \in [0, \infty)$, *which satisfies the system:*

$$\begin{cases} \partial_t u_t + \nabla \cdot (u_t \xi_t) = 0 & \text{in } \mathscr{D}'(\Omega \times (0, +\infty)), \\ \varepsilon \partial_t v_t + \eta_t = 0 & \text{in } \mathscr{D}'(\Omega \times (0, +\infty)), \\ -u_t \xi_t = \nabla u_t^m - \chi u_t \nabla v_t & \text{in } \Omega \text{ for } \mathscr{L}^1\text{-a.e. } t > 0, \\ \eta_t = -\Delta v_t + \gamma v_t - \alpha u_t & \text{in } \Omega \text{ for } \mathscr{L}^1\text{-a.e. } t > 0, \end{cases} \tag{16}$$

with the boundary conditions:

$$\frac{\partial u^m}{\partial \nu} - \chi u \frac{\partial v}{\partial \nu} = v = 0, \quad x \in \partial\Omega, \ t > 0,$$

where the meaning of these boundary conditions are understood in the sense of Remark 3.4 below.

Moreover, the energy inequality

$$\phi_m(w(a)) - \phi_m(w(t)) \geq \frac{1}{2} \int_a^t \int_\Omega |\xi(x,s)|^2 u(x,s) dx ds + \frac{\chi}{2\alpha\varepsilon} \int_a^t \int_\Omega |\eta(x,s)|^2 \, dx ds$$

holds for every $t \in [0, +\infty)$ *and* $a \in [0, t) \backslash \mathcal{N}$, $\mathcal{N}$ *being a* $\mathscr{L}^1$*-negligible subset of* $(0, +\infty)$.

Remark 3.4 (boundary conditions). Since $u \in \mathscr{P}_2(\Omega)$, we have

$$\begin{aligned} 0 = \frac{d}{dt} \int_\Omega u \, dx &= \int_\Omega \partial_t u \, dx \\ &= -\int_\Omega \nabla \cdot (u_t \xi) \, dx \\ &= \int_{\partial\Omega} \left(\frac{\partial u^m}{\partial \nu} - \chi u \frac{\partial v}{\partial \nu} \right) dS. \end{aligned}$$

On the other hand, since $v \in H_0^1(\Omega)$ it holds

$$v = 0 \quad \text{on } \partial\Omega.$$

Remark 3.5. Note that the uniqueness of the solution of (1) (for each given initial data) is not known.

Remark 3.6. If initial data $w_0 = (u_0, v_0) \in (\mathscr{P}(\Omega) \cap L^2(\Omega)) \times H_0^1(\Omega)$ satisfies $v_0 \geq 0$, then the above limit functions u and v are a non-negative weak solution of system (1) (see remark 4.4).

4 Variational analysis

In this section, we give a proof of Theorem 1.1 and Proposition 3.1.

Proof of Theorem 1.1. First, we consider the case of $m > 2 - \frac{2}{d}$ and show that $M_* = \infty$ in this case. By the Hölder inequality, the Sobolev inequality, and the interpolation inequality, we have

$$\begin{aligned}\left|\int uv\,dx\right| &\leq \|u\|_{L^{\frac{2d}{d+2}}}\|v\|_{L^{\frac{2d}{d-2}}} \\ &\leq C_s\|u\|_{L^{\frac{2d}{d+2}}}\|\nabla v\|_{L^2} \\ &\leq C_s\|u\|_{L^1}^{1-\theta}\|u\|_{L^m}^{\theta}\|\nabla v\|_{L^2}\end{aligned} \tag{17}$$

where C_s denotes the Sobolev constant and $\theta = \frac{m(d-2)}{2d(m-1)}$. Therefore, for any $\delta > 0$ we have

$$\left|\int uv\,dx\right| \leq \frac{(\alpha+\delta)C_s^2}{2}\|u\|_{L^1}^{2(1-\theta)}\|u\|_{L^m}^{2\theta} + \frac{1}{2(\alpha+\delta)}\|\nabla v\|_{L^2}^2$$

Hence, it holds that

$$\phi_m \geq \frac{1}{m-1}\|u\|_{L^m}^m - \frac{(\alpha+\delta)\chi C_s^2}{2}\|u\|_{L^1}^{2(1-\theta)}\|u\|_{L^m}^{2\theta} + \frac{\delta\chi}{2\alpha(\alpha+\delta)}\|\nabla v\|_{L^2}^2 \tag{18}$$

This means that $M_* = \infty$ for $m > 2 - \frac{2}{d}$, since $2\theta < m$.

Next, we consider the case of $m = 2 - \frac{2}{d}$. By the estimate (18), we immediately have

$$M_* > \left(\frac{2}{\alpha\chi(m-1)C_s^2}\right)^{d/2}.$$

In order to obtain the optimal estimate, we define the constant C_* by

$$C_* := \sup_{(u,v)\in X_M}\frac{\langle u,v\rangle_{L^2}}{\|u\|_{L^1}^{\frac{1}{d}}\|u\|_{L^m}^{\frac{m}{2}}\|\nabla v\|_{L^2}}$$

and set

$$M_1 := \left(\frac{2}{\alpha\chi(m-1)C_*^2}\right)^{d/2}.$$

These are well-defined by the estimate (17). We show that $M_1 = M_*$. By using M_1, Lyapunov functional ϕ_m is estimated as follows,

$$\phi_m \geq \frac{\alpha\chi C_*^2}{2}\left(M_1^{\frac{2}{d}} - \frac{\alpha+\delta}{\alpha}\|u\|_{L^1}^{\frac{2}{d}}\right)\|u\|_{L^m}^m + \frac{\delta\chi}{2\alpha(\alpha+\delta)}\|\nabla v\|_{L^2}, \tag{19}$$

for any $\delta > 0$. This means $M_* \geq M_1$.

On the other hand, by the definition of the constant C_*, for any $\delta > 0$, there exists a pair $(u_\delta, v_\delta) \in X_M$ such that

$$\langle u_\delta, v_\delta\rangle_{L^2} \geq (C_* - \delta)\|u_\delta\|_{L^1}^{\frac{1}{d}}\|u_\delta\|_{L^m}^{\frac{m}{2}}\|\nabla v_\delta\|_{L^2}.$$

Hence for $\alpha = \dfrac{\|\nabla v_\delta\|_{L^2}}{(C_* - \delta)\|u_\delta\|_{L^1}^{\frac{1}{d}}\|u_\delta\|_{L^m}^{\frac{m}{2}}}$, we have

$$\langle u_\delta, v_\delta \rangle_{L^2} \geq \frac{\alpha (C_* - \delta)^2 \|u_\delta\|_{L^1}^{\frac{2}{d}} \|u_\delta\|_{L^m}^m}{2} + \frac{1}{2\alpha}\|\nabla v_\delta\|_{L^2}^2.$$

Therefore we have

$$\phi_m(u_\delta, v_\delta) - \frac{\chi\gamma}{2\alpha}\int v_\delta^2\, dx \leq \frac{\alpha\chi C_*^2}{2}\left(M_1^{\frac{2}{d}} - \frac{(C_* - \delta)^2}{C_*^2}\|u_\delta\|_{L^1}^{\frac{2}{d}}\right)\|u_\delta\|_{L^m}^m.$$

Now we define the functions (u_λ, v_λ) by

$$u_\lambda(x) := \begin{cases} \lambda^d u_\delta(\lambda x) & \lambda x \in \Omega \\ 0 & \lambda x \notin \Omega, \end{cases} \qquad v_\lambda(x) := \begin{cases} \lambda^{d-2} v_\delta(\lambda x) & \lambda x \in \Omega \\ 0 & \lambda x \notin \Omega, \end{cases}$$

and then, we have

$$\begin{aligned} \phi_m(u_\lambda, v_\lambda) &= \lambda^{d-2}\left(\phi_m(u_\delta, v_\delta) - \frac{\chi\gamma}{2\alpha}\int v_\delta^2\, dx\right) + \frac{\chi\gamma}{2\alpha}\int v_\lambda^2\, dx \\ &\leq \lambda^{d-2}\frac{\alpha\chi C_*^2}{2}\left(M_1^{\frac{2}{d}} - \frac{(C_* - \delta)^2}{C_*^2}\|u_\delta\|_{L^1}^{\frac{2}{d}}\right)\|u_\delta\|_{L^m}^m + \lambda^{d-4}\frac{\chi\gamma}{2\alpha}\int v_\delta^2\, dx. \end{aligned}$$

Therefore, if $\|u_\delta\|_{L^1} = M > M_1$, then we can choose $\delta > 0$ such that

$$\phi_m(u_\lambda, v_\lambda) \to -\infty, \quad (\lambda \to \infty).$$

Thus we have $M_1 = M_*$. □

Remark 4.1. By the change of variable

$$u_\lambda(x) := \begin{cases} \lambda^d u(\lambda x) & \lambda x \in \Omega \\ 0 & \lambda x \notin \Omega, \end{cases} \qquad v_\lambda(x) := \begin{cases} \lambda^{d-2} v(\lambda x) & \lambda x \in \Omega \\ 0 & \lambda x \notin \Omega, \end{cases}$$

we see that C_* does not depend on Ω. Therefore M_* also does not depend on Ω.

Remark 4.2. M_* coincides with the sharp threshold mass M_c for the parabolic-elliptic case given by Blanchet, Carrillo and Laurençot [6].

Indeed, we define the functional $A(u, v)$ by

$$A(u, v) := \frac{\|\nabla v\|_{L^2}^2}{|\langle u, v \rangle_{L^2}|^2}.$$

Then

$$
\begin{aligned}
C_*^2 &= \sup_{(u,v)\in X_M} \frac{|\langle u, v\rangle_{L^2}|^2}{\|u\|_{L^1}^{\frac{2}{d}} \|u\|_{L^m}^{m} \|\nabla v\|_{L^2}^2} \\
&= \sup_u \left\{ \frac{1}{\|u\|_{L^1}^{\frac{2}{d}} \|u\|_{L^m}^{m}} \sup_v \left(\frac{1}{A(u,v)} \right) \right\} \\
&= \sup_u \left(\frac{1}{\|u\|_{L^1}^{\frac{2}{d}} \|u\|_{L^m}^{m} \inf\limits_v A(u,v)} \right).
\end{aligned}
$$

Here we can characterize the function v which attains $\inf\limits_v A(u,v)$ by the Euler-Lagrange equation. The function v which attains $\inf\limits_v A(u,v)$ is given by

$$
-\Delta v = \langle u, v\rangle_{L^2} A(u,v) u.
$$

Therefore, using the fundamental solution $\mathcal{K}$ of the Laplace operator $-\Delta$, we can write

$$
\frac{1}{\inf\limits_v A(u,v)} = \langle \mathcal{K} * u, u\rangle_{L^2}.
$$

As a consequence we obtain

$$
C_*^2 = \sup_u \frac{\langle \mathcal{K} * u, u\rangle_{L^2}}{\|u\|_{L^1}^{\frac{2}{d}} \|u\|_{L^m}^{m}}.
$$

Therefore the constant M_* equals the threshold mass M_c for the parabolic-elliptic case (see [6]).

Lemma 4.3 (lower semicontinuity of ϕ_m). *Let $u_n \rightharpoonup u$ weakly in $L^m(\Omega)$ and $v_n \rightharpoonup v$ weakly in $H_0^1(\Omega)$. Then we have*

$$
\phi_m(u,v) \leq \liminf_n \phi_m(u_n, v_n).
$$

Proof. It suffices to show that $\langle u, v\rangle_{L^2}$ is continuous with respect to the weak topology in $L^m(\Omega) \times H_0^1(\Omega)$, because it is well known that $\phi_m(u,v) - \chi\langle u, v\rangle_{L^2}$ is lower semicontinuous with respect to this weak topology. By Rellich's compactness theorem, we can extract a subsequence still denoted by v_n such that $v_n \to v$ strongly in $L^p(\Omega)$ for $p := \frac{2(d-\varepsilon)}{d-2}$ $(0 < \varepsilon < 1)$. On the other hand, for the conjugate exponent p' of p we have $m - p' = \frac{2(d-2)(1-\varepsilon)}{d(d+2(1-\varepsilon))} > 0$. Hence the interpolation inequality assures the uniform bounds of $\|u_n\|_{L^{p'}}$. Therefore, taking into account that $\langle u, v\rangle_{L^2}$ is a bilinear form in $L^{p'}(\Omega) \times L^p(\Omega)$, we have

$$
\begin{aligned}
\lim_{n\to\infty} \left| \int_\Omega (u_n v_n - uv)\, dx \right| &= \lim_{n\to\infty} \left| \int_\Omega (u - u_n)(v - v_n)\, dx \right| \\
&\leq \lim_{n\to\infty} \|u - u_n\|_{L^{p'}} \|v - v_n\|_{L^p} = 0.
\end{aligned} \tag{20}
$$

□

Remark 4.4 (non-negativity). For any τ and $k \in \mathbb{N}$, u_τ^k and v_τ^k are non-negative. Indeed, the non-negativity of u_τ^k is clear because u_τ^k belongs to $\mathscr{P}(\Omega)$. Hence we have

$$\phi_m(u_\tau^k, v_\tau^k) \geq \phi_m(u_\tau^k, |v_\tau^k|)$$

Therefore, if the initial data v_0 is non-negative, then v_τ^k is non-negative.

Proof of Proposition 3.1. Let v_j be a minimizing sequence of

$$v \mapsto \phi_m(u_\tau^{k-1}, v) + \frac{\varepsilon\chi}{2\alpha\tau_k}\|v - v_\tau^{k-1}\|_{L^2}^2,$$

and I_k be the infimum. Then, from the lower bounds of ϕ_m, the infimum I_k is a finite value, and the coercivity (18)-(19) induces the weak compactness in $H_0^1(\Omega)$. Therefore, there exists a subsequence of v_j still denoted by v_j and $v_\infty \in H_0^1(\Omega)$ such that v_j converges v_∞ weakly in $H_0^1(\Omega)$. By Lemma 4.3, we have

$$\phi_m(u_\tau^{k-1}, v_\infty) + \frac{\varepsilon\chi}{2\alpha\tau_k}\|v_\infty - v_\tau^{k-1}\|_{L^2}^2 \leq I_k.$$

Since the opposite inequality is obvious, we have

$$I_k = \phi_m(u_\tau^{k-1}, v_\infty) + \frac{\varepsilon\chi}{2\alpha\tau_k}\|v_\infty - v_\tau^{k-1}\|_{L^2}^2.$$

Therefore, we can define v_τ^k by v_∞. Taking into account the lower semicontinuity of the Wasserstein distance with respect to the weak convergence in $L^m(\Omega)$, which is stronger than the narrow convergence (see Ambrosio *et al.* [3]), by a similar argument, u_τ^k is well-defined. □

5 Subdifferential calculus

In this section, we introduce the notion of metric slope and investigate its properties. The metric slope is the equivalent of the normed value of a gradient of a functional in Hilbert space. Therefore, the concept of the metric slope plays a crucial role in the construction of gradient flows in metric spaces.

Let $X = \mathscr{P}_2(\Omega) \times L^2(\Omega)$ be a metric space endowed with the following distance

$$d^2(w_1, w_2) := d_1^2(\mu_1, \mu_2) + d_2^2(v_1, v_2) \quad \text{for } w_i = (\mu_i, v_i) \in X,$$

where the distances d_1 and d_2 are defined by

$$\begin{aligned} d_1(\mu, \tilde{\mu}) &:= d_W(\mu, \tilde{\mu}) \quad \text{for } \mu, \tilde{\mu} \in \mathscr{P}(\Omega) \subset \mathscr{P}_2(\mathbb{R}^d), \\ d_2(v, \tilde{v}) &:= \sqrt{\frac{\varepsilon\chi}{\alpha}}\|v - \tilde{v}\|_{L^2} \quad \text{for } v, \tilde{v} \in L^2(\Omega), \end{aligned} \tag{21}$$

and d_W denotes the Wasserstein space introduced in §2 and $\|\cdot\|_{L^2}$ denotes the usual $L^2(\Omega)$ norm. We define the subset $K \subset X$ by

$$K := \Big\{(\mu, v) \in X\,;\, \mu = u\mathscr{L}^d \text{ with } u \in L^m(\Omega) \text{ and } v \in H^1(\Omega)\Big\}, \tag{22}$$

and the functional ϕ_m on X as follows: for $w = (\mu, v) \in X$,

$$\phi_m(w) := \begin{cases} \dfrac{1}{m-1}\displaystyle\int_\Omega u^m\,dx - \chi\int_\Omega uv\,dx + \frac{\chi}{2\alpha}\int_\Omega |\nabla v|^2 + \gamma v^2 dx & \text{if } w \in K, \\ +\infty & \text{otherwise,} \end{cases}$$

where $\mathscr{L}^d$ denotes d-dimensional Lebesgue measure and u denotes the density function of μ. The effective domain of ϕ_m is defined by

$$D(\phi_m) := \Big\{(\mu, v) \in X\,;\, \phi_m(\mu, v) < +\infty\Big\}. \tag{23}$$

We see that $D(\phi_m) = K$. We often identify probability measure $\mu = u\mathscr{L}^d$ with its density u and write $(u, v) \in D(\phi_m)$ instead of $(\mu, v) \in D(\phi_m)$.

Let us recall the definition of the metric slope.

Definition 5.1 (metric slope). The *metric slope* $|\partial\phi_m|(w)$ of ϕ_m at $w \in D(\phi_m)$ is defined by

$$|\partial\phi_m|(w) := \limsup_{\tilde{w}\to w} \frac{\big(\phi_m(w) - \phi_m(\tilde{w})\big)^+}{d(w, \tilde{w})}.$$

For $w = (\mu, v)$, we also define $|\partial_1\phi_m|(w)$ and $|\partial_2\phi_m|(w)$ as

$$\begin{aligned} |\partial_1\phi_m|(w) &:= \limsup_{\tilde{\mu}\to\mu} \frac{\big(\phi_m(\mu, v) - \phi_m(\tilde{\mu}, v)\big)^+}{d_1(\mu, \tilde{\mu})}, \\ |\partial_2\phi_m|(w) &:= \limsup_{\tilde{v}\to v} \frac{\big(\phi_m(\mu, v) - \phi_m(\mu, \tilde{v})\big)^+}{d_2(v, \tilde{v})}. \end{aligned} \tag{24}$$

In order to investigate the properties of the metric slope, we first give definitions of the inner product $\langle\cdot,\cdot\rangle_w$ on $L^2_\mu(\Omega; \mathbb{R}^d) \times L^2(\Omega)$ and subdifferentials.

Definition 5.2 (inner product). For vector fields $\xi_i : \Omega \to \mathbb{R}^d$ and scalar fields $\eta_i : \Omega \to \mathbb{R}$, $(i = 1, 2)$ we define their *inner product* by

$$\langle\zeta_1, \zeta_2\rangle_w := \int_\Omega \langle\xi_1, \xi_2\rangle\,d\mu + \frac{\varepsilon\chi}{\alpha}\int_\Omega \eta_1\eta_2\,dx, \quad \zeta_i = (\xi_i, \eta_i).$$

Note that this inner product depends on a point $w = (\mu, v) \in X$.

Definition 5.3 (subdifferential $\partial\phi_m(w)$). We say that $g = (g_1, g_2) \in L^2_\mu(\Omega : \mathbb{R}^d) \times L^2(\Omega)$ belongs to the *subdifferential* of ϕ_m at $w = (\mu, v) \in D(\phi_m)$ if for any $\tilde{w} = (\tilde{\mu}, \tilde{v}) \in D(\phi_m)$

$$\phi_m(\tilde{w}) - \phi_m(w) \geq \langle g, \zeta\rangle_w + o(d(w, \tilde{w})),$$

where $\zeta = (\boldsymbol{t}_\mu^{\tilde{\mu}} - \boldsymbol{id}, \tilde{v} - v)$ and $\boldsymbol{t}_\mu^{\tilde{\mu}}$ is the optimal transport map from μ to $\tilde{\mu}$. We denote by $\partial\phi_m(w)$ the set of all the subdifferentials of ϕ_m at w.

In order to calculate the subdifferential of ϕ_m, we consider a curve $w(t) \in X$ that passes through a given point $w = (\mu, v) \in D(\phi_m)$ at $t = 0$ with a none-zero metric derivative $|w'|(0)$ and calculate the differential of $\phi_m \circ w(t)$. Since $D(\phi_m) \subset \mathscr{P}_2^r(\Omega) \times H_0^1(\Omega)$, for every $(\mu, v) \in D(\phi_m)$ and $\nu \in \mathscr{P}_2(\Omega)$ there exists the optimal transport map r such that

$$\nu = \boldsymbol{r}_{\#}\mu \quad \text{and} \quad d_1^2(\mu, \nu) = \int_{\mathbb{R}^d} |x - r(x)|^2 \, d\mu. \tag{25}$$

Let

$$\boldsymbol{r}_t := \boldsymbol{id} + t(\boldsymbol{r} - \boldsymbol{id}) \quad \text{and} \quad \mu_t := \boldsymbol{r}_{t\#}\mu \tag{26}$$

and

$$v_t := v + \eta, \quad \eta \in C_c^\infty(\Omega). \tag{27}$$

Lemma 5.4. *If v_n converges to v in $L^2(\Omega)$, then for every $t \in (0, T)$, $v_n(\boldsymbol{r}_t)$ converges to $v(\boldsymbol{r}_t)$ in $L^2(\Omega)$. In addition, It holds*

$$\|v_n(\boldsymbol{r}_t) - v(\boldsymbol{r}_t)\|_{L^2} \leq C_T \|v_n - v\|_{L^2} \quad (0 \leq t \leq T),$$

where C_T is a positive constant depending T.

Proof.

$$\begin{aligned}\int_\Omega |v_n(\boldsymbol{r}_t(x)) - v(\boldsymbol{r}_t(x))|^2 \, dx &= \int_\Omega |v_n(y) - v(y)|^2 \det(D\boldsymbol{r}_t^{-1}(y)) \, dy \\ &\leq \sup_y \Big(\det(\tilde{D}\boldsymbol{r}_t^{-1}(y)) \Big) \|v_n - v\|_{L^2}.\end{aligned}$$

Here $\tilde{D}$ denotes the approximate differential. See [3] for a definition of the approximate differential and on approximate differentiability of the optimal transport map. □

Lemma 5.5. *$g = (g_1, g_2)$ belongs to $\partial\phi_m(w)$ if $\mu = u\mathscr{L}^d$ with $u^m \in W^{1,1}(\Omega)$ and*

$$g_1 = \left(\frac{m}{m-1} \nabla u^{m-1} - \chi \nabla v \right) \in L_\mu^2(\Omega; \mathbb{R}^d), \quad \varepsilon g_2 = -\Delta v + \gamma v - \alpha u \in L^2(\Omega).$$

Proof. The subdifferential of ϕ_m splits into the following three parts.

$$\begin{aligned}\mathcal{F}[\mu] &= \frac{1}{m-1} \int_\Omega u^m \, dx, \quad I(\mu, v) = -\chi \int_\Omega uv \, dx, \\ \mathcal{G}[v] &= \frac{\chi}{2\alpha} \int_\Omega |\nabla v|^2 + \gamma v^2 \, dx.\end{aligned}$$

Subdifferentials of $\mathcal{F}[\mu]$ and $\mathcal{G}[v]$ are well-known because they are subdifferentials in Wasserstein space and L^2-space, respectively. Therefore we will focus on the subdifferential of I.

Let $\mu_t = u_t \mathscr{L}^d$ and v_t be curves defined by (26) and (27), respectively. Considering the relation $u_t(r_t(x)) \det \tilde{D} r_t(x) = u(x)$, it easily follows that

$$\lim_{t\to 0} \frac{\mathcal{F}[\mu_t] - \mathcal{F}[\mu]}{t} = \lim_{t\to 0} \frac{1}{(m-1)t} \int_\Omega \left(\frac{1}{(\det \tilde{D} r_t(x))^{m-1}} - 1 \right) u^m(x)\, dx$$
$$= - \int_\Omega u^m(x) \operatorname{tr} \tilde{D}(r(x) - x)\, dx.$$

Here, since u belongs to $W^{1,1}(\Omega)$, using a weak integration by parts (see Thm 10.4.5 of [3]) we can write

$$\lim_{t\to 0} \frac{\mathcal{F}[\mu_t] - \mathcal{F}[\mu]}{t} \geq \int_\Omega \left\langle \frac{\nabla u^m}{u}, r(x) - x \right\rangle u\, dx.$$

Moreover, it easily follows that

$$\lim_{t\to 0} \frac{I[\mu_t, v_t] - I[\mu_t, v]}{t} = -\chi \lim_{t\to 0} \int_\Omega u_t \eta dx = -\chi \int u\eta\, dx,$$

and if $\Delta v \in L^2(\Omega)$ then the following holds:

$$\lim_{t\to 0} \frac{\mathcal{G}[v_t] - \mathcal{G}[v]}{t} = \frac{\chi}{\alpha} \int_\Omega \nabla v \cdot \nabla \eta + \gamma v \eta\, dx$$
$$= \frac{\chi}{\alpha} \int_\Omega (-\Delta v + \gamma v)\eta\, dx.$$

As we shall see below note that by the Sobolev embedding theorem, u belongs to $L^2(\Omega)$ and then Δv also belongs to $L^2(\Omega)$.

Let $(v_n)_{n\in\mathbb{N}} \subset C_c^\infty(\Omega)$ be a sequence converging to v in $H_0^1(\Omega)$. We define the following functions:

$$I(t) := -\chi \int_\Omega v(r_t(x)) u(x)\, dx,$$

$$I_n(t) := -\chi \int_\Omega v_n(r_t(x)) u(x)\, dx.$$

By Lemma 5.4, $I_n(t)$ converges to $I(t)$ uniformly in $[0, T]$. Moreover $I_n'(t)$ converges to

$$\tilde{I}'(t) := -\chi \int_\Omega \langle \nabla v(r_t(x)), r(x) - x \rangle u(x)\, dx \quad \text{uniformly in } [0, T].$$

In fact, by the assumption $u^m \in W^{1,1}(\Omega)$ and the Sobolev embedding theorem, $u \in L^2(\Omega)$. Considering the definition of the push-forward and $\mu, \nu \in \mathscr{P}(\Omega)$, we see that $r(x)$ belongs to $\overline{\Omega}$ for μ-a.e. $x \in \Omega$. Since Ω is bounded, $r - \boldsymbol{id}$ belongs to $L_\mu^\infty(\Omega)$. Therefore by Lemma 5.4, $I_n'(t)$ uniformly converges to $\tilde{I}'(t)$ and hence $I(t)$ is differentiable in $(0, T)$ and we have

$$\lim_{t\to 0} \frac{I[\mu_t, v] - I[\mu, v]}{t} = I'(0) = -\chi \int_\Omega \langle \nabla v(x), r(x) - x \rangle u(x)\, dx.$$

Combining all the relations above, we have

$$\lim_{t\to 0}\frac{\phi_m(\mu_t,v_t)-\phi_m(\mu,v)}{t}$$
$$\geq \int_\Omega \left\langle \frac{\nabla u^m}{u}-\chi\nabla v, \boldsymbol{r}(x)-x\right\rangle u\,dx+\frac{\varepsilon\chi}{\alpha}\int_\Omega \frac{-\Delta v+\gamma v-\alpha u}{\varepsilon}\eta\,dx. \quad (28)$$

Finally, we notice that for $w=(\mu,v)$ and $w_t=(\mu_t,v_t)$, we have $d(w,w_t)=O(t)$. In fact,

$$\begin{aligned} d^2(w,w_t) &= d_1^2(\mu,\mu_t)+d_2^2(v,v_t)\\ &\leq \int_{\mathbb{R}^d}|x-\boldsymbol{r}_t(x)|^2\,d\mu(x)+\frac{\varepsilon\chi}{\alpha}\|t\eta\|_{L^2}\\ &= t^2\left(d_1^2(\mu,\nu)+\frac{\varepsilon\chi}{\alpha}\|\eta\|_{L^2}^2\right). \end{aligned}$$

Therefore (28) means that $g=(g_1,g_2)\in\partial\phi_m(w)$. □

Lemma 5.6. *If $(u,v)\in D(|\partial\phi_m|)$ and $u\in L^2(\Omega)$, then $\partial\phi_m(u,v)$ is not empty and we have*

$$|\partial\phi_m|^2(w)=\left\|\nabla\left(\frac{m}{m-1}u^{m-1}-\chi v\right)\right\|_{L^2(\mu)}^2+\frac{\varepsilon\chi}{\alpha}\left\|\frac{\Delta v-\gamma v+\alpha u}{\varepsilon}\right\|_{L^2}^2,$$

where $w=(\mu,v)$ with $\mu=u\mathscr{L}^d$.

Proof. By the definition of subdifferentials, for every $g\in\partial\phi_m(w)$, $|\partial\phi_m|(w)\leq\|g\|_w$ obviously holds. We prove that the opposite inequality holds for the subdifferential given in Lemma 5.5. Let $\boldsymbol{\xi}\in C_c^\infty(\Omega;\mathbb{R}^d)$ and X_t be the solution of the following equation:

$$\begin{cases} \dot{X}_t=\xi(X_t),\\ X_0=\boldsymbol{id}. \end{cases} \quad (29)$$

We define $\mu_t=(X_t)_\#\mu$ and $v_t=v+\eta$, $\eta\in C_c^\infty(\Omega)$. Then we have

$$\begin{aligned} d_1^2(\mu,\mu_t) &\leq \int_{\mathbb{R}^d}|x-X_t(x)|^2\,d\mu(x)\\ &= \int_{\mathbb{R}^d}\left|\int_0^t \xi(X_s(X))\,ds\right|^2 d\mu(x)\\ &\leq t\int_0^t\int_{\mathbb{R}^d}|\xi(X_s(x))|^2\,d\mu ds \end{aligned} \quad (30)$$

and hence

$$\lim_{t\to 0}\frac{d_1(\mu,\mu_t)}{t}=\|\xi\|_{L^2(\mu)}.$$

Therefore for $w_t=(\mu,v)$ and $w_t=(\mu_t,v_t)$ we have

$$\begin{aligned} \lim_{t\to 0}\frac{\phi_m(w)-\phi_m(w_t)}{t} &\leq |\partial\phi_m|(w)\lim_{t\to 0}\frac{d(w,w_t)}{t}\\ &= |\partial\phi_m|(w)\|\zeta\|_w \quad \text{where } \zeta=(\xi,\eta). \end{aligned} \quad (31)$$

On the other hand, we have

$$\lim_{t\to 0}\frac{\phi_m(w)-\phi_m(w_t)}{t}=\int_\Omega u^m \operatorname{div}\xi+\chi\langle\nabla v,\xi\rangle u\,dx \\ +\frac{\varepsilon\chi}{\alpha}\int_\Omega\frac{\Delta v-\gamma v+\alpha u}{\varepsilon}\eta\,dx. \quad (32)$$

Hence combining (31) and (32), by the density of $C_c^\infty(\Omega;\mathbb{R}^d)\times C_c^\infty(\Omega)$ in $L^2_\mu(\Omega;\mathbb{R}^d)\times L^2(\Omega)$ and the duality in $L^2_\mu(\Omega;\mathbb{R}^d)\times L^2(\Omega)$, we obtain

$$g_1=\left(\frac{m}{m-1}\nabla u^{m-1}-\chi\nabla v\right)\in L^2_\mu(\Omega;\mathbb{R}^d),\quad \varepsilon g_2=-\Delta v+\gamma v-\alpha u\in L^2(\Omega),$$

and for $g=(g_1,g_2)$

$$\|g\|_w\le|\partial\phi_m|(w).$$

Finally, we complete the proof by proving $g\in\partial\phi_m(w)$. By the Cauchy-Schwarz inequality, we have

$$\|\nabla u^m-\chi u\nabla v\|_{L^1}\le\left(\int_\Omega\left|\nabla\left(\frac{m}{m-1}u^{m-1}-\chi v\right)\right|^2u\,dx\right)^{\frac12}\left(\int_\Omega u(x)\,dx\right)^{\frac12} \\ \lesssim|\partial\phi_m|(u,v).$$

By the assumption $u\in L^2(\Omega)$, we obtain

$$\|\nabla u^m\|_{L^1}\lesssim|\partial\phi_m|(u,v)+\chi\|u\|_{L^2}\|\nabla v\|_{L^2}.$$

Hence by Lemma 5.5, g belongs to $\partial\phi_m(w)$. □

Remark 5.7. From the proof above, we see that

$$|\partial_1\phi_m|(w)=\|g_1\|_{L^2(\mu)}\quad\text{and}\quad|\partial_2\phi_m|(w)=\sqrt{\frac{\varepsilon\chi}{\alpha}}\|g_2\|_{L^2}.$$

By the convexity of $\partial\phi_m(w)$ and the relation $|\partial\phi_m|(w)\le\|g\|_w$ for every $g\in\partial\phi_m(w)$, a element of $\partial\phi_m(w)$ satisfying $|\partial\phi_m|(w)=\|g\|_w$ is unique.

Definition 5.8. We denote by $\operatorname{grad}_X\phi_m(w)$ the unique element of $\partial\phi_m(w)$ which satisfies $\|\operatorname{grad}_X\phi_m(w)\|_w=|\partial\phi_m|(w)$.

Lemma 5.9. *If $(u,v)\in D(|\partial_1\phi_m|)$ and $v\in W^{2,2}(\Omega)$, then u belongs to $L^2(\Omega)$. Conversely, if $(u,v)\in D(|\partial_2\phi_m|)$ and $u\in L^2(\Omega)$, then v belongs to $W^{2,2}(\Omega)$.*

Proof. By the Cauchy-Schwarz inequality, we have

$$\|\nabla u^m-\chi u\nabla v\|_{L^1}\le\left(\int_\Omega\left|\nabla\left(\frac{m}{m-1}u^{m-1}-\chi v\right)\right|^2u\,dx\right)^{\frac12}\left(\int_\Omega u(x)\,dx\right)^{\frac12} \\ \le|\partial_1\phi_m|(u,v).$$

By the Sobolev embedding theorem, if $v \in W^{2,2}(\Omega)$, then $\nabla v \in L^{\frac{2d}{d-2}}(\Omega)$. Therefore we have $\nabla v \in L^{m'}(\Omega)$ $(1/m + 1/m' = 1)$, since $m' \leq \frac{2(d-1)}{d-2} < \frac{2d}{d-2}$, and then $u\nabla v$ belongs to $L^1(\Omega)$. Hence we have

$$\|\nabla u^m\|_{L^1} \leq |\partial_1\phi_m|(u,v) + \chi\|u\nabla v\|_{L^1}.$$

On the other hand, by the interpolation inequality, for $\theta = \frac{1}{2}\frac{md}{(m-1)d+1} \in (0,1)$ and $p = \frac{m}{\theta} > 1$, we have

$$\begin{aligned}\|u\|_{L^2}^p &\leq \left(\|u\|_{L^1}^{1-\theta}\|u\|_{L^{\frac{md}{d-1}}}^{\theta}\right)^p \\ &= \|u^m\|_{L^{\frac{d}{d-1}}}.\end{aligned} \tag{33}$$

Moreover, by the Sobolev inequality, there exists a constant C_1 depending only on d and Ω such that

$$\|u^m\|_{L^{\frac{d}{d-1}}} \leq C_1\left(\|u^m\|_{L^1} + \|\nabla u^m\|_{L^1}\right).$$

Therefore we have $u \in L^2(\Omega)$ and

$$\|u\|_{L^2}^p \leq C_1\left(\|u\|_{L^m}^m + |\partial_1\phi_m|(u,v) + \chi\|u\nabla v\|_{L^1}\right). \tag{34}$$

The second assertion follows from the estimate

$$\sqrt{\frac{\chi}{\alpha\varepsilon}}\|\Delta v - \gamma v + \alpha u\|_{L^2} \leq |\partial_2\phi_m|(u,v),$$

and the L^2-estimate $\|v\|_{W^{2,2}} \leq C\|\Delta v\|_{L^2}$ for $v \in W^{2,2}(\Omega) \cap H_0^1(\Omega)$. □

Lemma 5.10 (L^2-estimate). *If $u_n \in L^2(\Omega)$ and*

$$\sup_n |\partial_1\phi_m|(u_n,v_n) < +\infty, \quad \sup_n \|u_n\|_{L^m} < +\infty, \quad \sup_n \|\nabla v_n\|_{L^2} < +\infty,$$

then we have $\sup_n \|u_n\|_{L^2} < +\infty$ and $\sup_n \|\nabla(u_n)^m\|_{L^1} < +\infty$.

Proof. From the estimate (34), we have

$$\|u_n\|_{L^2}^p \leq C_1\left(\sup_n \|u_n\|_{L^m}^m + \sup_n |\partial_1\phi_m|(u_n,v_n) + \chi\left(\sup_n \|\nabla v_n\|_{L^2}\right)\|u_n\|_{L^2}\right).$$

Since $p > 1$, this estimate means that

$$\sup_n \|u_n\|_{L^2} < +\infty,$$

and then we obtain

$$\sup_n \|\nabla(u_n)^m\|_{L^1} \leq \sup_n\left(|\partial_1\phi_m|(u_n,v_n) + \chi\|u_n\|_{L^2}\|\nabla v_n\|_{L^2}\right). \tag{35}$$

□

Lemma 5.11 (lower semicontinuity of $|\partial\phi_m|$). *Assume that $u_n \rightharpoonup u$ weakly in $L^m(\Omega)$ and $v_n \rightharpoonup v$ weakly in $H_0^1(\Omega)$ with*

$$\sup_n \|u_n\|_{L^2} < +\infty, \quad \sup_n \|v_n\|_{W^{2,2}} < +\infty.$$

Then we have

$$\begin{aligned} |\partial_1\phi_m|(u,v) &\le \liminf_n |\partial_1\phi_m|(u_n,v_n), \\ |\partial_2\phi_m|(u,v) &\le \liminf_n |\partial_2\phi_m|(u_n,v_n). \end{aligned} \tag{36}$$

Proof. Take any sequence $w_n = (u_n, v_n) \in D(|\partial\phi_m|)$ such that $u_n \rightharpoonup u$ weakly in $L^m(\Omega)$ and $v_n \rightharpoonup v$ weakly in $H_0^1(\Omega)$ with

$$\sup_n \|u_n\|_{L^2} < +\infty, \quad \sup_n \|v_n\|_{W^{2,2}} < +\infty.$$

In order to prove the lower semicontinuity, it suffices to consider the case where the metric slope at w_n remains bounded, that is,

$$|\partial\phi_m|^2(w_n) = \|g_1(w_n)\|^2_{L^2(u_n)} + \frac{\varepsilon\chi}{\alpha}\|g_2(w_n)\|^2_{L^2} \le C, \quad (n = 1,2,3,\ldots)$$

for some constant C, where

$$u_n g_1(w_n) = \nabla u_n^m - \chi u_n \nabla v_n, \quad \varepsilon g_2(w_n) = -\Delta v_n + \gamma v_n - \alpha u_n.$$

The estimate (35) and the boundedness of $\|u_n\|_{L^m}$ imply that u_n^m is bounded in $W^{1,1}(\Omega)$, hence in $\mathrm{BV}(\Omega)$. By the compactness theorem for BV functions (see Thm.3.23 of [2]), a bounded sequence in $\mathrm{BV}(\Omega)$ has a subsequence that is weakly convergent in $\mathrm{BV}(\Omega)$, thus strongly convergent in $L^1(\Omega)$. Therefore, there exists a function $L \in \mathrm{BV}(\Omega)$ such that a subsequence of u_n^m converges to L in $L^1(\Omega)$. We can extract a further subsequence still denoted by u_n such that $u_n^m(x) \to L(x)$ for $\mathscr{L}^d$-a.e. $x \in \Omega$ and then, $u_n(x) \to L^{\frac{1}{m}}(x)$ for $\mathscr{L}^d$-a.e. $x \in \Omega$. Let $\rho_n := |L^{\frac{1}{m}} - u_n|$. Then the Lebesgue dominated convergence theorem yields

$$\int_\Omega \frac{\rho_n}{1+\rho_n}\,dx \to 0 \quad (n \to \infty).$$

By the Hölder inequality we have

$$\begin{aligned} \int_\Omega \rho_n\,dx &= \int_\Omega \left(\frac{\rho_n}{1+\rho_n}\right)^{\frac{m-1}{m}} \left(\rho_n(1+\rho_n)^{m-1}\right)^{\frac{1}{m}} dx \\ &\le \left(\int_\Omega \frac{\rho_n}{1+\rho_n}\,dx\right)^{\frac{m-1}{m}} \left(\int_\Omega \rho_n(1+\rho_n)^{m-1}\,dx\right)^{\frac{1}{m}}. \end{aligned}$$

The second factor of the right hand side is bounded since Ω is bounded and $u_n, L^{\frac{1}{m}} \in L^m(\Omega)$ with $\sup_n \|u_n\|_{L^m} < +\infty$. Hence we have $u_n \to L^{\frac{1}{m}}$ in $L^1(\Omega)$. It thereby follows that $L^{\frac{1}{m}} = u$ and then

$$\begin{aligned} u_n &\to u && \text{in } L^1(\Omega), \\ u_n^m &\to u^m && \text{in } L^1(\Omega). \end{aligned} \tag{37}$$

Therefore the compactness of BV functions yields

$$\int_\Omega \langle \nabla u_n^m, \xi \rangle \, dx \to \int_\Omega \langle \nabla u^m, \xi \rangle \, dx \quad \text{for every } \xi \in C_c^\infty(\Omega; \mathbb{R}^d).$$

On the other hand, by Rellich's compactness theorem, we can extract a subsequence still denoted by v_n such that $v_n \to v$ strongly in $H_0^1(\Omega)$. Thus we have

$$-\chi \int_\Omega \langle \nabla v_n, \xi \rangle u_n \, dx \to -\chi \int_\Omega \langle \nabla v, \xi \rangle u \, dx.$$

By the Cauchy-Schwarz inequality, we have

$$\begin{aligned} \langle g_1(w_n), \xi \rangle_{L^2(u_n)} &\leq \|g_1(w_n)\|_{L^2(u_n)} \|\xi\|_{L^2(u_n)} \\ &= |\partial_1 \phi_m|(w_n) \|\xi\|_{L^2(u_n)}. \end{aligned} \tag{38}$$

Therefore, passing to the limit as $n \to +\infty$, we have

$$\langle g_1(u, v), \xi \rangle_{L^2(u)} \leq \liminf_{n \to \infty} |\partial_1 \phi_m|(w_n) \|\xi\|_{L^2(u)}. \tag{39}$$

Hence by duality, we obtain

$$|\partial_1 \phi_m|(w) = \|g_1(w)\|_{L^2(u)} \leq \liminf_{k \to \infty} |\partial_1 \phi_m|(w_{n_k}).$$

The relation $|\partial_2 \phi_m|(u, v) \leq \liminf_n |\partial_2 \phi_m|(u_n, v_n)$ follows from

$$(\Delta v_{n_k} - \gamma v_{n_k} + \alpha u_{n_k}) \rightharpoonup (\Delta v - \gamma v + \alpha u) \quad \text{weakly in } L^2(\Omega)$$

for a subsequence still denoted by (u_n, v_n) and the weak lower semicontinuity of the norm in L^2-space. □

Lemma 5.12 (continuity of ϕ_m). *Assume that $u_n \rightharpoonup u$ weakly in $L^m(\Omega)$ and $v_n \rightharpoonup v$ weakly in $H_0^1(\Omega)$ with*

$$\sup_n |\partial \phi_m|(u_n, v_n) < +\infty, \quad \sup_n \|u_n\|_{L^2} < +\infty, \quad \sup_n \|v_n\|_{W^{2,2}} < +\infty.$$

Then there exists a subsequence (u_{n_k}, v_{n_k}) such that

$$\lim_{k \to \infty} \phi_m(u_{n_k}, v_{n_k}) = \phi_m(u, v).$$

Proof. From (37), it follows

$$u_{n_k} \to u \quad \text{in } L^m(\Omega).$$

On the other hand, by Rellich's compactness theorem, we can extract a subsequence v_{n_k} such that $v_{n_k} \to v$ strongly in $H_0^1(\Omega)$. Therefore, it follows that

$$\lim_{k \to \infty} \phi_m(u_{n_k}, v_{n_k}) = \phi_m(u, v).$$

□

6 Energy identity of discrete solutions

In this section, we apply some well-known results of minimizing movement scheme to our case. It is natural to treat our case in the metric space $X = \mathscr{P}(\Omega) \times L^2(\Omega)$ endowed with distance d defined by

$$d^2(w_1, w_2) := d_W^2(u_1, u_2) + \frac{\varepsilon\chi}{\alpha}\|v_2 - v_1\|_{L^2}^2 \quad \text{for } w_i = (u_i, v_i) \in X, (i = 1, 2).$$

That is, we should consider the following variational scheme

$$w_\tau^k \in \underset{w \in \mathscr{P}(\Omega) \times H_0^1(\Omega)}{\operatorname{argmin}} \left\{ \phi_m(w) + \frac{1}{2\tau_k} d^2(w, w_\tau^{k-1}) \right\}. \tag{40}$$

However, by this formulation, we encounter difficulties concening the existence of subdifferentials of ϕ_m. To avoid these difficulties, we apply minimizing movement scheme to $\phi_m(\cdot, v)$ in the metric space $X_1 := \mathscr{P}(\Omega)$ endowed with distance d_1 defined by

$$d_1(u, \tilde{u}) := d_W(u, \tilde{u}) \quad \text{for } u, \tilde{u} \in \mathscr{P}(\Omega) \subset \mathscr{P}_2(\mathbb{R}^d),$$

and to $\phi_m(u, \cdot)$ in the metric space $X_2 := L^2(\Omega)$ endowed with distance d_2 defined by

$$d_2(v, \tilde{v}) := \sqrt{\frac{\varepsilon\chi}{\alpha}}\|v - \tilde{v}\|_{L^2} \quad \text{for } v, \tilde{v} \in L^2(\Omega),$$

respectively. In this section, we refer to [3] and [18].

Remark 6.1 (implicit Euler method). Assume that we define $w_\tau^k = (u_\tau^k, v_\tau^k)$ by the variational scheme (40). Then, we consider the following perturbation: for any $\varphi \in C_c^\infty(\Omega)$,

$$w_h := (u_h, v_h) = ((\nabla\varphi_h)_\# u_\tau^k, v_\tau^k + h\varphi),$$

where $\nabla\varphi_h = \boldsymbol{id} + h\nabla\varphi$ and $h \in \mathbb{R}$. Then we have

$$\frac{d}{dh}\left\{ \phi_m(w_h) + \frac{1}{2\tau_k} d^2(w_h, w_\tau^{k-1}) \right\}\Bigg|_{h=0} = 0.$$

This relation yields the following Euler-Lagrange equations:

$$\begin{aligned} \left\langle \frac{u_\tau^k - u_\tau^{k-1}}{\tau_k}, \varphi \right\rangle_{L^2} &= -\left\langle \nabla(u_\tau^k)^m + \chi u_\tau^k \nabla v_\tau^k, \nabla\varphi \right\rangle_{L^2} + O(d_W^2(u_\tau^{k-1}, u_\tau^k)), \\ \varepsilon\left\langle \frac{v_\tau^k - v_\tau^{k-1}}{\tau_k}, \varphi \right\rangle_{L^2} &= -\langle \nabla v_\tau^k, \nabla\varphi \rangle_{L^2} + \langle \alpha u_\tau^k - \gamma v_\tau^k, \varphi \rangle_{L^2}. \end{aligned} \tag{41}$$

Note that the above equations coincide with the weak form of the implicit Euler method for (1) except for the penalty term $O(d_W^2(u_\tau^{k-1}, u_\tau^k))$ in the first equation. In order to prove the existence of a weak solution of (1), one could possibly argue the convergence of (41) as $|\tau| \to 0$. Such an approach works well in the parabolic-elliptic case ($\varepsilon = 0$) as shown in [5]. However, in the present problem it is not easy to show the convergence of (41) directly. This is why we are using another approach due to Ambrosio *et al.* [3].

Definition 6.2 (Moreau-Yosida approximation). *The Moreau-Yosida approximations* $a_{1,\tau}(u,v)$ and $a_{2,\tau}(u,v)$ of ϕ_m are respectively defined by

$$\begin{aligned} a_{1,\tau}(u,v) &:= \inf_{\tilde{u} \in \mathscr{P}(\Omega)} \left\{ \phi_m(\tilde{u}, v) + \frac{1}{2\tau} d_1^2(\tilde{u}, u) \right\}, \\ a_{2,\tau}(u,v) &:= \inf_{\tilde{v} \in H_0^1(\Omega)} \left\{ \phi_m(u, \tilde{v}) + \frac{1}{2\tau} d_2^2(v, \tilde{v}) \right\}. \end{aligned} \tag{42}$$

We also define the sets of minimizers $J_{1,\tau}[u,v]$ and $J_{2,\tau}[u,v]$ as follows:

$$\begin{aligned} u_\tau \in J_{1,\tau}[u,v] &\Leftrightarrow a_{1,\tau}(u,v) = \phi_m(u_\tau, v) + \frac{1}{2\tau} d_1^2(u_\tau, u), \\ v_\tau \in J_{2,\tau}[u,v] &\Leftrightarrow a_{2,\tau}(u,v) = \phi_m(u, v_\tau) + \frac{1}{2\tau} d_2^2(v, v_\tau) \end{aligned} \tag{43}$$

Furthermore, we set

$$\begin{aligned} d_{1,\tau}^+(u,v) &:= \sup \{ d_1(u, u_\tau) \mid u_\tau \in J_{1,\tau}[u,v] \}, \\ d_{1,\tau}^-(u,v) &:= \inf \{ d_1(u, u_\tau) \mid u_\tau \in J_{1,\tau}[u,v] \}, \\ d_{2,\tau}^+(u,v) &:= \sup \{ d_2(v, v_\tau) \mid v_\tau \in J_{2,\tau}[u,v] \}, \\ d_{2,\tau}^-(u,v) &:= \inf \{ d_2(v, v_\tau) \mid v_\tau \in J_{2,\tau}[u,v] \}. \end{aligned} \tag{44}$$

Lemma 6.3 (continuity of the Moreau-Yosida approximation). *The map $(\tau, u) \mapsto a_{1,\tau}(u,v)$ is continuous in $(0,\infty) \times \mathscr{P}(\Omega)$ for every $v \in H_0^1(\Omega)$, and $(\tau, v) \mapsto a_{2,\tau}(u,v)$ is continuous in $(0,\infty) \times L^2(\Omega)$ for every $u \in L^m(\Omega)$. If $0 < \tau_0 < \tau_1$ and $u_{\tau_0} \in J_{1,\tau_0}[u,v], u_{\tau_1} \in J_{1,\tau_1}[u,v]$, then the following holds:*

$$\begin{aligned} &\phi_m(u,v) \geq a_{1,\tau_0}(u,v) \geq a_{1,\tau_1}(u,v), \quad d_1(u, u_{\tau_0}) \leq d_1(u, u_{\tau_1}), \\ &\phi_m(u,v) \geq \phi_m(u_{\tau_0}, v) \geq \phi_m(u_{\tau_1}, v), \quad d_{1,\tau_0}^+(u,v) \leq d_{1,\tau_1}^-(u,v), \end{aligned} \tag{45}$$

$$\lim_{\tau \downarrow 0} a_{1,\tau}(u,v) = \lim_{\tau \downarrow 0} \left(\inf_{u_\tau \in J_{1,\tau}[u,v]} \phi_m(u_\tau, v) \right) = \phi_m(u,v),$$

Moreover, if $(u,v) \in D(\phi_m)$, then it holds that $\lim_{\tau \downarrow 0} d_{1,\tau}^+(u,v) = 0$. In particular, there exists an (at most) countable set $\mathscr{N}_{u,v}$ such that

$$d_{1,\tau}^+(u,v) = d_{1,\tau}^-(u,v) \quad \forall \tau \in (0,\infty) \setminus \mathscr{N}_{u,v}.$$

The same relations hold for $v_\tau \in J_{2,\tau}[u,v]$ and $a_{2,\tau}(u,v)$.

Proof. Let $\tau_n \to \tau$, $d_1(u_n, u) \to 0$ and $\tilde{u}_n \in J_{1,\tau_n}[u_n, v]$. Then we have

$$
\begin{aligned}
\limsup_{n\to\infty} a_{1,\tau_n}(u_n, v) &= \limsup_{n\to\infty} \left\{ \phi_m(\tilde{u}_n, v) + \frac{1}{2\tau_n} d_1^2(\tilde{u}_n, u_n) \right\} \\
&\leq \limsup_{n\to\infty} \left\{ \phi_m(\tilde{u}, v) + \frac{1}{2\tau_n} d_1^2(\tilde{u}, u_n) \right\} \\
&= \phi_m(\tilde{u}, v) + \frac{1}{2\tau} d_1^2(\tilde{u}, u), \quad \forall \tilde{u} \in \mathscr{P}(\Omega).
\end{aligned}
\tag{46}
$$

Taking the infimum, we get

$$
\limsup_{n\to\infty} a_{1,\tau_n}(u_n, v) \leq a_{1,\tau}(u, v).
$$

On the other hand,

$$
\begin{aligned}
\liminf_{n\to\infty} a_{1,\tau_n}(u_n, v) &= \liminf_{n\to\infty} \left\{ \phi_m(\tilde{u}_n, v) + \frac{1}{2\tau_n} d_1^2(\tilde{u}_n, u_n) \right\} \\
&\geq \liminf_{n\to\infty} \left\{ \phi_m(\tilde{u}_n, v) + \frac{1}{2\tau_n} \left(d_1(\tilde{u}_n, u) - d_1(u, u_n)\right)^2 \right\} \\
&\geq \liminf_{n\to\infty} \left\{ \phi_m(\tilde{u}_n, v) + \frac{1}{2\tau} d_1^2(\tilde{u}_n, u) \right\} \\
&\quad + \liminf_{n\to\infty} \left\{ \frac{\tau - \tau_n}{2\tau_n \tau} d_1^2(\tilde{u}_n, u) - \frac{1}{\tau_n} d_1(\tilde{u}_n, u) d_1(u_n, u) \right\} \\
&\geq a_{1,\tau}(u, v).
\end{aligned}
\tag{47}
$$

Therefore we have

$$
\lim_{n\to\infty} a_{1,\tau_n}(u_n, v) = a_{1,\tau}(u, v).
$$

Hence $(\tau, u) \mapsto a_{1,\tau}(u, v)$ is continuous in $(0, \infty) \times \mathscr{P}(\Omega)$.

By the monotonicity $\tau \mapsto \tau^{-1}$, it easily follows that

$$
\begin{aligned}
\phi_m(u_{\tau_1}, v) + \frac{1}{2\tau_1} d_1^2(u_{\tau_1}, u) &\leq \phi_m(u_{\tau_0}, v) + \frac{1}{2\tau_1} d_1^2(u_{\tau_0}, u) \\
&\leq \phi_m(u_{\tau_0}, v) + \frac{1}{2\tau_0} d_1^2(u_{\tau_0}, u) \\
&\leq \phi_m(u, v),
\end{aligned}
\tag{48}
$$

all the equalities hold if and only if $u_{\tau_0} = u_{\tau_1} = u$. Therefore we have the first inequality of (45) and

$$
\lim_{\tau\downarrow 0} a_{1,\tau}(u, v) = \lim_{\tau\downarrow 0} \left(\inf_{u_\tau \in J_{1,\tau}[u,v]} \phi_m(u_\tau, v) \right) = \phi_m(u, v).
$$

The second inequality follows from

$$\begin{aligned}\phi_m(u_{\tau_0},v)+\frac{1}{2\tau_0}d_1^2(u_{\tau_0},u)&\le\phi_m(u_{\tau_1},v)+\frac{1}{2\tau_0}d_1^2(u_{\tau_1},v)\\&=\phi_m(u_{\tau_1},v)+\frac{1}{2\tau_1}d_1^2(u_{\tau_1},u)+\frac{\tau_1-\tau_0}{2\tau_0\tau_1}d_1^2(u_{\tau_1},u)\\&\le\phi_m(u_{\tau_0},v)+\frac{1}{2\tau_1}d_1^2(u_{\tau_0},u)+\frac{\tau_1-\tau_0}{2\tau_0\tau_1}d_1^2(u_{\tau_1},u).\end{aligned}$$

From the definition of u_{τ_1}, we have

$$\phi_m(u_{\tau_1},v)+\frac{1}{2\tau_1}d_1^2(u_{\tau_1},u)\le\phi_m(u_{\tau_0},v)+\frac{1}{2\tau_1}d_1^2(u_{\tau_0},u). \tag{49}$$

Combining with the second inequality of (45), we have

$$\phi_m(u_{\tau_1},v)\le\phi_m(u_{\tau_0},v).$$

Let $u_\tau\in J_{1,\tau}[u,v]$. Then for arbitrary $\tilde{u}$ such that $(\tilde{u},v)\in D(\phi_m)$, we have

$$d_1^2(u_\tau,u)\le d_1^2(\tilde{u},u)+2\tau(\phi_m(\tilde{u},v)-\phi_m(u_\tau,v)).$$

Taking the supremum, we have

$$(d_{1,\tau}^+(u,v))^2\le d_1^2(\tilde{u},u)+2\tau(\phi_m(\tilde{u},v)-\inf\phi_m),$$

and then

$$\limsup_{\tau\downarrow 0}(d_{1,\tau}^+(u,v))^2\le d_1^2(\tilde{u},u).$$

Here we can choose $\tilde{u}=u$, since $(u,v)\in D(\phi_m)$. Therefore we have

$$\lim_{\tau\downarrow 0}d_{1,\tau}^+(u,v)=0.$$

We omit the proof of the last assertion. See [18] to complete the proof. □

Lemma 6.4 (slope estimate). *If $u_\tau\in J_{1,\tau}[u,v]$, then $(u_\tau,v)\in D(|\partial_1\phi_m|)$ and it holds that*

$$|\partial_1\phi_m|(u_\tau,v)\le\frac{d_1(u,u_\tau)}{\tau}.$$

On the other hand, if $v_\tau\in J_{2,\tau}[u,v]$, then $(u,v_\tau)\in D(|\partial_2\phi_m|)$ and

$$|\partial_2\phi_m|(u,v_\tau)\le\frac{d_2(v,v_\tau)}{\tau}.$$

Proof. Let $u_\tau\in J_{1,\tau}[u,v]$, then for every $\tilde{u}\in\mathscr{P}(\Omega)$, we have

$$\phi_m(u_\tau,v)+\frac{1}{2\tau}d_1^2(u_\tau,u)\le\phi_m(\tilde{u},v)+\frac{1}{2\tau}d_1^2(\tilde{u},u),$$

and then

$$\begin{aligned}\phi_m(u_\tau, v) - \phi_m(\tilde{u}, v) &\le \frac{1}{2\tau} d_1^2(\tilde{u}, u) - \frac{1}{2\tau} d_1^2(u_\tau, u) \\ &= \frac{1}{2\tau}\big(d(\tilde{u}, u) - d_1(u_\tau, u)\big)\big(d_1(\tilde{u}, u) + d_1(u, u_\tau)\big) \\ &\le \frac{1}{2\tau} d(u_\tau, \tilde{u})\big(d_1(\tilde{u}, u) + d_1(u, u_\tau)\big).\end{aligned} \tag{50}$$

Hence we have

$$\begin{aligned}|\partial_1 \phi_m|(u_\tau, v) &= \limsup_{\tilde{u}\to u_\tau} \frac{\big(\phi_m(u_\tau, v) - \phi_m(\tilde{u}, v)\big)^+}{d_1(u_\tau, \tilde{u})} \\ &\le \limsup_{\tilde{u}\to u_\tau} \frac{1}{2\tau}\big(d_1(\tilde{u}, u) + d_1(u_\tau, u)\big) \\ &= \frac{d_1(u_\tau, u)}{\tau}.\end{aligned} \tag{51}$$

Similarly, we can obtain

$$|\partial_2 \phi_m|(u, v_\tau) \le \frac{d_2(v, v_\tau)}{\tau} \quad \text{for } v_\tau \in J_{2,\tau}[u, v].$$

□

Lemma 6.5 (derivative of Moreau-Yosida approximations). *Moreau-Yosida approximation $\tau \mapsto a_{i,\tau}(u, v)$ is differentiable $\mathscr{L}^1$-a.e. $\tau > 0$ and we have*

$$\frac{d}{d\tau} a_{i,\tau}(u, v) = -\frac{\big(d_{1,\tau}^{\pm}(u, v)\big)^2}{2\tau^2} \quad \textit{for } \mathscr{L}^1\textit{-a.e. } \tau > 0, (i = 1, 2),$$

$$\frac{d_1^2(u, u_\tau)}{2\tau} + \int_0^\tau \frac{\big(d_{1,r}^{\pm}(u, v)\big)^2}{2r^2}\, dr = \phi_m(u, v) - \phi_m(u_\tau, v) \quad \forall u_\tau \in J_{1,\tau}[u, v],$$

$$\frac{d_2^2(v, v_\tau)}{2\tau} + \int_0^\tau \frac{\big(d_{2,r}^{\pm}(u, v)\big)^2}{2r^2}\, dr = \phi_m(u, v) - \phi_m(u, v_\tau) \quad \forall v_\tau \in J_{2,\tau}[u, v].$$

Proof. Let $0 < \tau_0 < \tau_1$ and $u_{\tau_0} \in J_{1,\tau_0}[u, v], u_{\tau_1} \in J_{1,\tau_1}[u, v]$. Then, we easily get the following estimate,

$$\begin{aligned}a_{1,\tau_0}(u, v) - a_{1,\tau_1}(u, v) &= \phi_m(u_{\tau_0}, v) + \frac{d_1^2(u, u_{\tau_0})}{2\tau_0} - \phi_m(u_{\tau_1}, v) - \frac{d_1^2(u, u_{\tau_1})}{2\tau_1} \\ &\le \phi_m(u_{\tau_1}, v) + \frac{d_1^2(u, u_{\tau_1})}{2\tau_0} - \phi_m(u_{\tau_1}, v) - \frac{d_1^2(u, u_{\tau_1})}{2\tau_1} \\ &= \frac{\tau_1 - \tau_0}{2\tau_0\tau_1} d_1^2(u, u_{\tau_1}).\end{aligned} \tag{52}$$

Similarly, we also have

$$a_{1,\tau_1}(u,v) - a_{1,\tau_0}(u,v) \le \frac{\tau_0-\tau_1}{2\tau_0\tau_1} d_1^2(u,u_{\tau_0}).$$

Therefore it holds that

$$0 \le \frac{\tau_1-\tau_0}{2\tau_0\tau_1} d_1^2(u,u_{\tau_0}) \le a_{1,\tau_0}(u,v) - a_{1,\tau_1}(u,v) \le \frac{\tau_1-\tau_0}{2\tau_0\tau_1} d_1^2(u,u_{\tau_1}), \tag{53}$$

and then, for $0<\tau_0<\tau_1$ we have

$$\frac{1}{2\tau_0\tau_1}(d_{1,\tau_0}^+(u,v))^2 \le \frac{a_{1,\tau_0}(u,v) - a_{1,\tau_1}(u,v)}{\tau_1-\tau_0} \le \frac{1}{2\tau_0\tau_1}(d_{1,\tau_1}^-(u,v))^2. \tag{54}$$

This implies that the map $\tau \mapsto a_{1,\tau}(u,v)$ is locally Lipschitz. Passing to the limit as $\tau_0 \uparrow \tau$ and $\tau_1 \downarrow \tau$, we obtain

$$\frac{d}{d\tau} a_{1,\tau}(u,v) = -\frac{\left(d_{1,\tau}^\pm(u,v)\right)^2}{2\tau^2} \qquad \text{for } \mathscr{L}^1\text{-a.e. } \tau > 0.$$

□

Recall the definition of the discrete solution. Let $(u_\tau^0, v_\tau^0) = (u_0, v_0)$ be the initial data. For $k = 1,2,3,\ldots$, we recursively define v_τ^k and u_τ^k by

$$\begin{aligned} v_\tau^k &\in J_{2,\tau_k}[u_\tau^{k-1}, v_\tau^{k-1}], \\ u_\tau^k &\in J_{1,\tau_k}[u_\tau^{k-1}, v_\tau^k]. \end{aligned} \tag{55}$$

Then the discrete solution $\overline{w}_\tau = (\overline{u}_\tau, \overline{v}_\tau)$ is defined by

$$\overline{w}_\tau(t) := (u_\tau^k, v_\tau^k) \qquad \text{for } t \in (t_\tau^{k-1}, t_\tau^k].$$

Furthermore, we also define $\underline{u}_\tau$ by

$$\underline{u}_\tau(t) := u_\tau^{k-1} \qquad \text{for } t \in (t_\tau^{k-1}, t_\tau^k]. \tag{56}$$

Definition 6.6 (De Giorgi variational interpolation). Let $\{(u_\tau^k, v_\tau^k)\}_{k=0}^\infty$ be a solution of the variational scheme (55). We define the De Giorgi variational interpolation $(\tilde{u}_\tau, \tilde{v}_\tau)$ by

$$\begin{aligned} \tilde{u}_\tau(t) &:= \tilde{u}_\tau(t_\tau^{k-1}+\delta) \in J_{1,\delta}[u_\tau^{k-1}, v_\tau^k] &&\text{for } t = t_\tau^{k-1}+\delta \in (t_\tau^{k-1}, t_\tau^k], \\ \tilde{v}_\tau(t) &:= \tilde{v}_\tau(t_\tau^{k-1}+\delta) \in J_{2,\delta}[u_\tau^{k-1}, v_\tau^{k-1}] &&\text{for } t = t_\tau^{k-1}+\delta \in (t_\tau^{k-1}, t_\tau^k]. \end{aligned} \tag{57}$$

We also define two functions G_τ and $|w_\tau'|$ by

$$G_\tau(t) := \sqrt{\left(\frac{d_{1,\delta}^+(u_\tau^{k-1}, v_\tau^k)}{\delta}\right)^2 + \left(\frac{d_{2,\delta}^+(u_\tau^{k-1}, v_\tau^{k-1})}{\delta}\right)^2} \qquad \text{for } t \in (t_\tau^{k-1}, t_\tau^k],$$

$$\begin{aligned} |w_\tau'|(t) &:= \frac{d(w_\tau^{k-1}, w_\tau^k)}{t_\tau^k - t_\tau^{k-1}} \\ &= \sqrt{\left(\frac{d_1(u_\tau^{k-1}, u_\tau^k)}{t_\tau^k - t_\tau^{k-1}}\right)^2 + \left(\frac{d_2(v_\tau^{k-1}, v_\tau^k)}{t_\tau^k - t_\tau^{k-1}}\right)^2} \qquad \text{for } t \in (t_\tau^{k-1}, t_\tau^k]. \end{aligned}$$

Lemma 6.7. *The function $G_\tau(t)$ satisfies the following inequality:*

$$|\partial_1\phi_m|^2(\tilde{u}_\tau(t), \overline{v}_\tau(t)) + |\partial_2\phi_m|^2(\underline{u}_\tau(t), \tilde{v}_\tau(t)) \le G_\tau^2(t). \tag{58}$$

Proof. By Lemma 6.4 and the definitions of $d_{1,\delta}^+(u,v)$ and $\tilde{u}_\tau$, we have

$$\begin{aligned} |\partial_1\phi_m|(\tilde{u}_\tau(t_\tau^{k-1}+\delta), v_\tau^k) &\le \frac{d_1(\tilde{u}_\tau(t_\tau^{k-1}+\delta), u_\tau^{k-1})}{\delta} \\ &\le \frac{d_{1,\delta}^+(u_\tau^{k-1}, u_\tau^k)}{\delta}. \end{aligned} \tag{59}$$

Similarly, we have

$$|\partial_2\phi_m|(u_\tau^{k-1}, \tilde{v}_\tau(t_\tau^{k-1}+\delta)) \le \frac{d_{2,\delta}^+(u_\tau^{k-1}, v_\tau^{k-1})}{\delta}.$$

Therefore, it holds that

$$|\partial_1\phi_m|^2(\tilde{u}_\tau(t), \overline{v}_\tau(t)) + |\partial_2\phi_m|^2(\underline{u}_\tau(t), \tilde{v}_\tau(t)) \le G_\tau^2(t).$$

□

Proposition 6.8 (energy identity of discrete solutions). *The discrete solution $\overline{w}_\tau = (\overline{u}_\tau, \overline{v}_\tau)$ satisfies the following identity:*

$$\frac{1}{2}\int_{t_\tau^{k-1}}^{t_\tau^k} |w_\tau'|^2(r)\,dr + \frac{1}{2}\int_{t_\tau^{k-1}}^{t_\tau^k} G_\tau^2(r)\,dr = \phi_m(w_\tau(t_\tau^{k-1})) - \phi_m(w_\tau(t_\tau^k)). \tag{60}$$

Proof. Since $u_\tau^k \in J_{1,\tau}[u_\tau^{k-1}, v_\tau^k]$ and $v_\tau^k \in J_{2,\tau}[u_\tau^{k-1}, v_\tau^{k-1}]$, by Lemma 6.5, we have

$$\frac{d_1^2(u_\tau^{k-1}, u_\tau^k)}{2\tau_k} + \int_0^{\tau_k} \frac{\left(d_{1,r}^{\pm}(u_\tau^{k-1}, v_\tau^k)\right)^2}{2r^2}\,dr = \phi_m(u_\tau^{k-1}, v_\tau^k) - \phi_m(u_\tau^k, v_\tau^k),$$

$$\frac{d_2^2(v_\tau^{k-1}, v_\tau^k)}{2\tau_k} + \int_0^{\tau_k} \frac{\left(d_{2,r}^{\pm}(u_\tau^{k-1}, v_\tau^{k-1})\right)^2}{2r^2}\,dr = \phi_m(u_\tau^{k-1}, v_\tau^{k-1}) - \phi_m(u_\tau^{k-1}, v_\tau^k).$$

By adding both equalities, and then taking into account the definitions of $|w_\tau'|$ and G_τ, (60) holds true. □

Lemma 6.9. *The discrete solution $(\overline{u}_\tau, \overline{v}_\tau)$ and the De Giorgi variational interpolation $(\tilde{u}_\tau, \tilde{v}_\tau)$ satisfy the following inequalities:*

$$\begin{aligned} d_1^2(\underline{u}_\tau(t), \overline{u}_\tau(t)) &\le 2|\boldsymbol{\tau}|(\phi_m(u_0, v_0) - \inf\phi_m), \\ d_1^2(\tilde{u}_\tau(t), \overline{u}_\tau(t)) &\le 8|\boldsymbol{\tau}|(\phi_m(u_0, v_0) - \inf\phi_m), \\ d_2^2(\tilde{v}_\tau(t), \overline{v}_\tau(t)) &\le 8|\boldsymbol{\tau}|(\phi_m(u_0, v_0) - \inf\phi_m). \end{aligned} \tag{61}$$

Proof. Let $t \in (t_{\tau}^{k-1}, t_{\tau}^{k}]$. From the definition of u_{τ}^{k}, we have

$$\phi_m(u_{\tau}^{k}, v_{\tau}^{k}) + \frac{1}{2\tau_k} d_1^2(u_{\tau}^{k-1}, u_{\tau}^{k}) \le \phi_m(u_{\tau}^{k-1}, v_{\tau}^{k}).$$

On the other hand, from the definition of v_{τ}^{k}, we have

$$\phi_m(u_{\tau}^{k-1}, v_{\tau}^{k}) + \frac{1}{2\tau_k} d_2^2(v_{\tau}^{k-1}, v_{\tau}^{k}) \le \phi_m(u_{\tau}^{k-1}, v_{\tau}^{k-1}).$$

Therefore, for every $k \in \mathbb{N}$,

$$\phi_m(u_{\tau}^{k}, v_{\tau}^{k}) \le \phi_m(u_{\tau}^{k-1}, v_{\tau}^{k}) \le \phi_m(u_{\tau}^{k-1}, v_{\tau}^{k-1}) \le \cdots \le \phi_m(u_0, v_0). \tag{62}$$

Hence we have

$$\begin{aligned} d_1^2(\underline{u}_{\tau}(t), \overline{u}_{\tau}(t)) = d_1^2(u_{\tau}^{k-1}, u_{\tau}^{k}) &\le 2\tau_k(\phi_m(u_{\tau}^{k-1}, v_{\tau}^{k}) - \phi_m(u_{\tau}^{k}, v_{\tau}^{k})) \\ &\le 2|\boldsymbol{\tau}|(\phi_m(u_0, v_0) - \inf \phi_m). \end{aligned} \tag{63}$$

By the definition of $\tilde{u}_{\tau}(t)$ and Lemma 6.3, we have

$$d_1(\tilde{u}_{\tau}(t), u_{\tau}^{k-1}) \le d_1(u_{\tau}^{k-1}, u_{\tau}^{k}).$$

Therefore we have

$$\begin{aligned} d_1^2(\tilde{u}_{\tau}(t), \overline{u}_{\tau}(t)) &\le \left(d_1(\tilde{u}_{\tau}(t), u_{\tau}^{k-1}) + d_1(u_{\tau}^{k-1}, u_{\tau}^{k})\right)^2 \\ &\le 4d_1^2(u_{\tau}^{k-1}, u_{\tau}^{k}) \\ &\le 8|\boldsymbol{\tau}|(\phi_m(u_0, v_0) - \inf \phi_m). \end{aligned} \tag{64}$$

Similarly, we have

$$d_2^2(\tilde{v}_{\tau}(t), \overline{v}_{\tau}(t)) \le 8|\boldsymbol{\tau}|(\phi_m(u_0, v_0) - \inf \phi_m).$$

□

Definition 6.10 (generalized minimizing movements). We say that a curve $w = (u, v) : [0, +\infty) \to L^m(\Omega) \times H^1(\Omega)$ is a *Generalized Minimizing Movement* starting from $w_0 = (u_0, v_0)$, if there exists a subsequence of partitions $\boldsymbol{\tau}_n$ with $|\boldsymbol{\tau}_n| \downarrow 0$ and a corresponding sequence of discrete solutions $(\overline{u}_{\boldsymbol{\tau}_n}, \overline{v}_{\boldsymbol{\tau}_n})$ defined by Definition 3.2 such that

$$\begin{aligned} \overline{u}_{\boldsymbol{\tau}_n}(t) &\rightharpoonup u(t) \text{ weakly in } L^m(\Omega) \quad \forall t \in [0, +\infty), \\ \overline{v}_{\boldsymbol{\tau}_n}(t) &\rightharpoonup v(t) \text{ weakly in } H_0^1(\Omega) \quad \forall t \in [0, +\infty). \end{aligned} \tag{65}$$

The set of all the generalized minimizing movements starting from (u_0, v_0) will be denoted by $(u, v) \in \mathrm{GMM}(u_0, v_0)$ or $w \in \mathrm{GMM}(w_0)$.

Proposition 6.11 (non-emptiness of GMM(w_0)). *If $w_0 \in L^m(\Omega)\times H_0^1(\Omega)$, then there exists a subsequence $w_{\tau_n} = (u_{\tau_n}, v_{\tau_n})$ of discrete solutions with $|\tau_n| \downarrow 0$ and $w = (u,v) \in AC^2_{\rm loc}([0,+\infty); X)$ such that*

$$\begin{aligned} \overline{u}_{\tau_n}(t) &\rightharpoonup u(t) \text{ weakly in } L^m(\Omega), \quad \forall t \geq 0, \\ \overline{v}_{\tau_n}(t) &\rightharpoonup v(t) \text{ weakly in } H_0^1(\Omega), \quad \forall t \geq 0. \end{aligned} \tag{66}$$

Moreover, $\underline{u}_{\tau_n}$ and $\tilde{u}_{\tau_n}$ also converge to u weakly in $L^m(\Omega)$. In particular, it holds that

$$\int_s^t |w'|^2(r)\,dr \leq \liminf_{n\to\infty} \int_s^t |w'_{\tau_n}|^2(r)\,dr \quad \textit{for } \forall (s,t) \subset [0,\infty).$$

Proof. From Proposition 6.8, we have

$$\int_0^\infty |w'_\tau|^2(r)\,dr \leq 2(\phi_m(w_0) - \inf \phi_m). \tag{67}$$

Therefore we can extract a sequence (τ_n) such that

$$|w'_{\tau_n}| \rightharpoonup A \quad \text{weakly in } L^2(0,\infty),$$

for some function $A \in L^2(0,\infty)$. For $0 \leq s < t$, there exist $s(n) \in \mathbb{N}$ and $t(n) \in \mathbb{N}$ such that

$$t_{\tau_n}^{s(n)-1} < s \leq t_{\tau_n}^{s(n)}, \quad t_{\tau_n}^{t(n)-1} < t \leq t_{\tau_n}^{t(n)},$$

$$\lim_{n\to\infty} t_{\tau_n}^{s(n)} = s, \quad \lim_{n\to\infty} t_{\tau_n}^{t(n)} = t.$$

Taking into account that

$$\overline{w}_{\tau_n}(s) = \overline{w}_{\tau_n}(t_{\tau_n}^{s(n)}) \quad \text{and} \quad \overline{w}_{\tau_n}(t) = \overline{w}_{\tau_n}(t_{\tau_n}^{t(n)}),$$

we obtain

$$\begin{aligned} d(\overline{w}_{\tau_n}(s), \overline{w}_{\tau_n}(t)) &\leq \sum_{k=s(n)}^{t(n)-1} d(\overline{w}_{\tau_n}(t_{\tau_n}^k), \overline{w}_{\tau_n}(t_{\tau_n}^{k+1})) \\ &= \int_{t_{\tau_n}^{s(n)}}^{t_{\tau_n}^{t(n)}} |w'_{\tau_n}|(r)\,dr, \end{aligned} \tag{68}$$

and therefore

$$\limsup_{n\to\infty} d(\overline{w}_{\tau_n}(s), \overline{w}_{\tau_n}(t)) \leq \int_s^t A(r)\,dr. \tag{69}$$

Moreover, for each fixed t, from the coercivity (18)-(19), the sequence $\{\overline{w}_{\tau_n}(t)\}$ is relatively weakly compact in $L^m(\Omega) \times H_0^1(\Omega)$. Hence by a standard diagonal argument we can find a subsequence still denoted by τ_n and $w : \mathbb{Q} \cap [0,\infty) \to$

$L^m(\Omega) \times H_0^1(\Omega)$ such that $\overline{w}_{\tau_n}(t) \rightharpoonup w(t)$ weakly in $L^m(\Omega) \times H_0^1(\Omega)$ for all $t \in \mathbb{Q} \cap [0,\infty)$ and

$$\begin{aligned} d(w(s), w(t)) &\le \limsup_{n\to\infty} d(\overline{w}_{\tau_n}(s), \overline{w}_{\tau_n}(t)) \le \int_s^t A(r)\,dr \\ &\le \sqrt{t-s}\int_0^\infty A^2(r)\,dr \quad \text{for all } s,t \in \mathbb{Q}\cap[0,\infty),\ s<t. \end{aligned} \tag{70}$$

Moreover, by the inequality (70), we can extend w to an element of $AC^2_{\mathrm{loc}}([0,+\infty);X)$ and

$$|w'|(t) \le A(t) \quad \text{for } \mathscr{L}^1\text{-a.e. } t>0.$$

By Fatou's lemma and weak lower semicontinuity of the norm in L^2-space, we get

$$\int_s^t |w'|^2(r)\,dr \le \int_s^t A^2(r)\,dr \le \liminf_{n\to\infty}\int_s^t |w'_{\tau_n}|^2(r)\,dr \quad \text{for all } (s,t)\subset[0,\infty).$$

□

7 Proof of Theorem 3.3

In this final section, we show that $w \in \mathrm{GMM}(w_0)$ is a time global weak solution of the system (1). Our strategy comes from the formal derivation given in §1.3.

Proposition 7.1 (tangent velocity vector). *If $w = (\mu, v) \in AC^2(0,T;X)$ then there exist a unique vector field $\boldsymbol{\xi} : t \in [0,T] \mapsto \boldsymbol{\xi}_t \in L^2_\mu(\mathbb{R}^d;\mathbb{R}^d)$ and scalar field $\eta : t\in[0,T] \mapsto \eta_t \in L^2(\mathbb{R}^d)$ such that*

$$\begin{cases} \dfrac{d}{dt}\mu_t + \nabla\cdot(\boldsymbol{\xi}_t\mu_t) = 0 & \text{in } \mathcal{D}'(\Omega\times(0,T)) \\ \dfrac{d}{dt}v_t + \eta_t = 0 & \text{in } \mathcal{D}'(\Omega\times(0,T)) \end{cases} \tag{71}$$

$$\text{and} \quad |w'|(t) = \sqrt{\|\boldsymbol{\xi}_t\|^2_{L^2(\mu(t))} + \frac{\varepsilon\chi}{\alpha}\|\eta_t\|^2_{L^2}}.$$

Proof. By the definitions of the distance d and its metric derivative, we have

$$|w'|(t) := \lim_{h\to 0}\frac{d(w(t+h),w(t))}{h} = \sqrt{|\mu'|^2(t) + \frac{\varepsilon\chi}{\alpha}\|\dot v(t)\|^2_{L^2}},$$

where $|\mu'|(t)$ and $\|\dot v(t)\|_{L^2}$ denote the metric derivatives in $\mathscr{P}_2(\Omega)$ and $L^2(\Omega)$, respectively. Therefore if $w\in AC^2(0,T;X)$ then we see that

$$\mu\in AC^2(0,T;\mathscr{P}_2(\Omega)) \text{ and } v\in AC^2(0,T;L^2(\Omega)).$$

By Theorem 2.7 there exists $\boldsymbol{\xi}$ such that the first equation of (71) holds and $\|\boldsymbol{\xi}_t\|_{L^2(\mu(t))} = |\mu'|(t)$. On the other hand, It is known that $v\in AC^2(0,T;L^2(\Omega))$ if and only if $t\mapsto v(t)$ is differentiable in $L^2(\Omega)$ and its derivative $\dot v$ belongs to $L^2(0,T;L^2(\Omega))$ (see [4]). It thereby follows the proposition. □

Definition 7.2 (tangent velocity vector). We denote by w' the tangent velocity vector $(\boldsymbol{\xi}, \eta)$ of $w = (\mu, v) \in AC^2(0, T; X)$, which satisfies the equations (71) in the sense of distributions.

Lemma 7.3 (chain rule). *Let $w : (a, b) \mapsto w_t \in D(\phi_m)$ be an absolutely continuous curve in X with tangent velocity vector w'_t. Let $\Lambda \subset (a, b)$ be the set of points $t \in (a, b)$ such that*

(a) $|\partial\phi_m|(w_t) < +\infty$,

(b) *$\phi_m \circ w_t$ is equal to a function $\psi(t)$ with finite pointwise variation,*

(c) *differential coefficient $\psi'(t)$ exists,*

(d) $\operatorname{grad}_X \phi_m(w_t)$ *exists.*

Then we have

$$\psi'(t) = \Big\langle \operatorname{grad}_X \phi_m(w(t)), w'(t) \Big\rangle_{w(t)} \qquad \textit{for } t \in \Lambda.$$

Proof. Let $t \in \Lambda$. Observing that

$$\lim_{h \to 0} \frac{\boldsymbol{t}_{\mu_t}^{\mu_{t+h}} - \boldsymbol{id}}{h} = \boldsymbol{\xi}_t \text{ in } L^2_{\mu_t}(\mathbb{R}; \mathbb{R}) \quad \text{and} \quad \lim_{h \to 0} \frac{v_{t+h} - v_t}{h} = \eta_t \text{ in } L^2(\Omega)$$

(see Prop.8.4.6 of [3]), from the definition of subdifferentials, we have

$$\phi_m(w_{t+h}) - \phi_m(w_t) \geq \Big\langle \operatorname{grad}_X \phi_m(w_t), \zeta(w_t, w_{t+h}) \Big\rangle_{w_t} + o(d(w_t, w_{t+h})),$$

where $\zeta(w_t, w_{t+h}) = (\boldsymbol{t}_{\mu_t}^{\mu_{t+h}} - \boldsymbol{id},\ v_{t+h} - v_t)$. Hence for $\mathscr{L}^1$-a.e. h,

$$\psi(t+h) - \psi(t) \geq \Big\langle \operatorname{grad}_X \phi_m(w_t), \zeta(w_t, w_{t+h}) \Big\rangle_{w_t} + o(d(w_t, w_{t+h})).$$

Dividing by h and passing to the limit as $h \downarrow 0$ and $h \uparrow 0$, we get

$$\psi'(t) = \Big\langle \operatorname{grad}_X \phi_m(w_t), w'_t \Big\rangle_{w_t}.$$

□

Lemma 7.4 (regularity of discrete solutions). *If $u_0 \in L^2(\Omega)$ and $(u_0, v_0) \in D(\phi_m)$, then we have*

$$\begin{aligned} \overline{u}_\tau(t) &\in L^2(\Omega) \quad && \forall t > 0; \\ \overline{v}_\tau(t) &\in W^{2,2}(\Omega) \quad && \forall t > 0. \end{aligned} \tag{72}$$

Proof. By Lemma 6.4, we have $(u_0, v_\tau^1) \in D(|\partial_2\phi_m|)$ and by Lemma 5.9 we have $v_\tau^1 \in W^{2,2}(\Omega)$. Again, by Lemmas 6.4and 5.9, we have $(u_\tau^1, v_\tau^1) \in D(|\partial_1\phi_m|)$ and $u_\tau^1 \in L^2(\Omega)$. Repeating this argument, we obtain $\overline{u}_\tau(t) \in L^2(\Omega)$ and $\overline{v}_\tau(t) \in W^{2,2}(\Omega)$ for every $t > 0$. □

Remark 7.5. We can obtain the same results under the assumption $\nabla v_0 \in L^{m'}(\Omega)$ instead of $u_0 \in L^2(\Omega)$, where $1/m + 1/m' = 1$. Then, we recursively define u_τ^k and v_τ^k by

$$\begin{aligned} u_\tau^k &\in J_{1,\tau_k}[u_\tau^{k-1}, v_\tau^{k-1}], \\ v_\tau^k &\in J_{2,\tau_k}[u_\tau^k, v_\tau^{k-1}]. \end{aligned} \tag{73}$$

Lemma 7.6 (finiteness of the slope). *The following relations hold:*

$$\liminf_{|\tau|\to 0} |\partial_1\phi_m|(\overline{u}_\tau(t), \overline{v}_\tau(t)) < +\infty \quad \textit{for } \mathscr{L}^1\textit{-a.e. } t > 0, \tag{74}$$

$$\liminf_{|\tau|\to 0} |\partial_2\phi_m|(\underline{u}_\tau(t), \overline{v}_\tau(t)) < +\infty \quad \textit{for } \mathscr{L}^1\textit{-a.e. } t > 0, \tag{75}$$

$$\liminf_{|\tau|\to 0} |\partial_1\phi_m|(\tilde{u}_\tau(t), \overline{v}_\tau(t)) < +\infty \quad \textit{for } \mathscr{L}^1\textit{-a.e. } t > 0, \tag{76}$$

$$\liminf_{|\tau|\to 0} |\partial_2\phi_m|(\underline{u}_\tau(t), \tilde{v}_\tau(t)) < +\infty \quad \textit{for } \mathscr{L}^1\textit{-a.e. } t > 0, \tag{77}$$

Proof. It easily follows from Proposition 6.8 and Lemma 6.7. □

Lemma 7.7. *Let* $(u, v) \in \mathrm{GMM}(u_0, v_0)$. *Let us suppose that* $(\overline{v}_{\tau_n}(t), \overline{v}_{\tau_n}(t))$ *and* $(\tilde{v}_{\tau_n}(t), \tilde{v}_{\tau_n}(t))$ *converge to* $(u(t), v(t))$ *weakly in* $L^m(\Omega) \times H_0^1(\Omega)$ *for any* $t \geq 0$ *and* $\underline{u}_{\tau_n}(t)$ *converges to* $u(t)$ *weakly in* $L^m(\Omega)$ *for any* $t \geq 0$. *Then the following inequalities hold:*

$$\begin{aligned} |\partial_1\phi_m|(u(t), v(t)) &\leq \liminf_{|\tau_n|\to 0} |\partial_1\phi_m|(\tilde{u}_{\tau_n}(t), \overline{v}_{\tau_n}(t)) \quad \textit{for } \mathscr{L}^1\textit{-a.e. } t > 0, \\ |\partial_2\phi_m|(u(t), v(t)) &\leq \liminf_{|\tau_n|\to 0} |\partial_2\phi_m|(\underline{u}_{\tau_n}(t), \tilde{v}_{\tau_n}(t)) \quad \textit{for } \mathscr{L}^1\textit{-a.e. } t > 0. \end{aligned} \tag{78}$$

Proof. In order to fix the ideas, let us suppose that

$$\liminf_{|\tau|\to 0} |\partial_1\phi_m|(\tilde{u}_\tau(t_0), \overline{v}_\tau(t_0)) = \lim_{n\to\infty} |\partial_1\phi_m|(\tilde{u}_{\tau_n}(t_0), \overline{v}_{\tau_n}(t_0)) < +\infty.$$

Then, we can further assume $\sup_n |\partial_1\phi_m|(\tilde{u}_{\tau_n}(t_0), \overline{v}_{\tau_n}(t_0)) < +\infty$. Therefore by Lemmas 7.4 and 5.9, we have $\tilde{u}_{\tau_n}(t_0) \in L^2(\Omega)$. Moreover, by Lemmas 5.10 and 5.11, we have

$$|\partial_1\phi_m|(u(t_0), v(t_0)) \leq \lim_{n\to\infty} |\partial_1\phi_m|(\tilde{u}_{\tau_n}(t_0), \overline{v}_{\tau_n}(t_0)).$$

By Lemma 7.6, this estimate holds for $\mathscr{L}^1$-a.e. $t_0 > 0$.

Next, we prove the second relation. Let us suppose that

$$\begin{aligned} \liminf_{|\tau|\to 0} |\partial_2\phi_m|(\underline{u}_\tau(t_0), \overline{v}_\tau(t_0)) &= \lim_{n\to\infty} |\partial_2\phi_m|(\underline{u}_{\tau_n}(t_0), \overline{v}_{\tau_n}(t_0)) < +\infty, \\ \liminf_{|\tau|\to 0} |\partial_2\phi_m|(\underline{u}_\tau(t_0), \tilde{v}_\tau(t_0)) &= \lim_{n\to\infty} |\partial_2\phi_m|(\underline{u}_{\tau_n}(t_0), \tilde{v}_{\tau_n}(t_0)) < +\infty, \end{aligned} \tag{79}$$

and

$$\sup_n |\partial_2\phi_m|(\underline{u}_{\tau_n}(t_0), \overline{v}_{\tau_n}(t_0)) < +\infty, \tag{80}$$

$$\sup_n |\partial_2\phi_m|(\underline{u}_{\tau_n}(t_0), \tilde{v}_{\tau_n}(t_0)) < +\infty. \tag{81}$$

Then, by (80), Lemmas 7.4 and 5.10, we have

$$\sup_n \|\underline{u}_{\tau_n}(t_0)\|_{L^2} < +\infty.$$

Since (81) means that

$$\sup_n \|\Delta\tilde{v}_{\tau_n}(t_0) - \gamma\tilde{v}_{\tau_n}(t_0) + \alpha\underline{u}_{\tau_n}(t_0)\|_{L^2} < +\infty,$$

and since (18)-(19) and the Poincaré inequality imply the uniform boundedness of $\|\tilde{v}_{\tau_n}(t_0)\|_{L^2}$, applying L^2-estimate, we obtain

$$\sup_n \|\tilde{v}_{\tau_n}(t_0)\|_{W^{2,2}} < +\infty.$$

Therefore by Lemma 5.11, we have

$$|\partial_2\phi_m|(u(t_0), v(t_0)) \leq \lim_{n\to\infty} |\partial_2\phi_m|(\underline{u}_{\tau_n}(t_0), \tilde{v}_{\tau_n}(t_0)).$$

□

Corollary 7.8. *If $u_0 \in L^2(\Omega)$ and $v_0 \in H_0^1(\Omega)$, then for $w \in GMM(u_0, v_0)$, $\mathrm{grad}_X\, \phi_m(w(t))$ exists for $\mathscr{L}^1$-a.e. $t > 0$.*

Now we give a proof of Theorem 3.3.

Proof of Theorem 3.3. From Lemmas 6.7 and 7.7 and Fatou's lemma, one has that

$$\int_s^t |\partial\phi_m|^2(w(r))\, dr \leq \liminf_{|\tau_n|\to 0} \int_s^t G^2_{\tau_n}(r)\, dr \quad \text{for } 0 \leq s < t.$$

From the monotonicity (62) and Helly's selection theorem, there exists a non-increasing function ψ and a subsequence, still denoted by (τ_n), such that

$$\lim_{n\to\infty} \phi_m(\overline{w}_{\tau_n}(t)) = \psi(t) \quad \forall t \geq 0.$$

For $0 \leq s < t < +\infty$, there exist $s(n) \in \mathbb{N}$ and $t(n) \in \mathbb{N}$ such that

$$t^{s(n)-1}_{\tau_n} < s \leq t^{s(n)}_{\tau_n}, \quad t^{t(n)-1}_{\tau_n} < t \leq t^{t(n)}_{\tau_n},$$

$$\lim_{n\to\infty} t^{s(n)-1}_{\tau_n} = s, \quad \lim_{n\to\infty} t^{t(n)}_{\tau_n} = t.$$

Hence we have

$$
\begin{aligned}
\frac{1}{2}\int_s^t |w'|^2(r)\,dr + \frac{1}{2}\int_s^t |\partial\phi_m|^2(w(r))\,dr \\
&\le \liminf_{|\tau_n|\to 0}\left(\frac{1}{2}\int_{t_{\tau_n}^{s(n)}}^{t_{\tau_n}^{t(n)}} |w'_{\tau_n}|^2(r)\,dr + \frac{1}{2}\int_{t_{\tau_n}^{s(n)}}^{t_{\tau_n}^{t(n)}} G_{\tau_n}^2(r)\,dr\right) \\
&= \liminf_{|\tau_n|\to 0}\big(\phi_m(\overline{w}_{\tau_n}(t_{\tau_n}^{s(n)})) - \phi_m(\overline{w}_{\tau_n}(t_{\tau_n}^{t(n)}))\big) \\
&= \liminf_{|\tau_n|\to 0}\big(\phi_m(\overline{w}_{\tau_n}(s)) - \phi_m(\overline{w}_{\tau_n}(t))\big) \\
&= \psi(s) - \psi(t).
\end{aligned}
\tag{82}
$$

Therefore, taking into account the differentiability of a monotone function (see for instance Cor. 3.29 of [2]), it follows that

$$
\psi'(t) \le -\frac{1}{2}|w'|^2(t) - \frac{1}{2}|\partial\phi_m|^2(w(t)) \quad \text{for } \mathscr{L}^1\text{-a.e. } t > 0. \tag{83}
$$

Considering the definition of $\operatorname{grad}_X \phi_m$ and Proposition 7.1, we can write

$$
\psi'(t) \le -\frac{1}{2}\|w'\|_w^2 - \frac{1}{2}\|\operatorname{grad}_X \phi_m(w(t))\|_w^2 \quad \text{for } \mathscr{L}^1\text{-a.e. } t > 0. \tag{84}
$$

Moreover, from Lemmas 7.6 and 5.12,

$$
\lim_{|\tau_n|\to 0} \phi_m(\overline{w}_{\tau_n}(t)) = \phi_m(w(t)) \quad \text{for } \mathscr{L}^1\text{-a.e. } t \ge 0.
$$

Therefore we have

$$
\psi(t) = \phi_m(w(t)) \quad \mathscr{L}^1\text{-a.e. } t \ge 0.
$$

On the other hand, by Lemma 7.3, the Cauchy-Schwarz inequality and the Young inequality, for $\mathscr{L}^1$-a.e. $t > 0$, we have

$$
\begin{aligned}
\psi'(t) = \langle \operatorname{grad}_X \phi_m(w), w' \rangle_w &\ge -\|\operatorname{grad}_X \phi_m(w)\|_w \|w'\|_w \\
&\ge -\frac{1}{2}\|\operatorname{grad}_X \phi_m(w)\|_w^2 - \frac{1}{2}\|w'\|_w^2.
\end{aligned}
\tag{85}
$$

In particular, the equality holds if and only if

$$
w' = -\operatorname{grad}_X \phi_m(w). \tag{86}
$$

Therefore, combining (84) and (85), we obtain (86) for $\mathscr{L}^1$-a.e. $t > 0$. Finally, considering Proposition 7.1, we obtain the results of Theorem 3.3. $\square$

Acknowledgement. The author would like to thank Professor Hiroshi Matano for his valuable advise and continued encouragement, and also Professor Filippo Santambrogio for stimulating discussions.

References

[1] M. Agueh, *Existence of solutions to degenerate parabolic equations via the Monge-Kantrovich theory*, adv. differential Equations, 10 (2005), 309-360.

[2] L. Ambrosio, N. Fusco and D. Pallara, *Functions of bounded variation and free discontinuity problems*, Oxford Mathematical Monographs. The Clarendon Press Oxford University Press, New York (2000).

[3] L. Ambrosio, N. Gigli and G. Savaré, *Gradient flows in metric spaces and in the space of probability measures*, Lectures in Mathematics, Birkhäuser, (2005).

[4] L. Ambrosio and B. Kirchheim, *Rectifiable sets in metric and Bnanach spaces*, Math. Ann., 318 (2000), pp. 527-555.

[5] A. Blanchet, V. Calvez, and J. A. Carrillo, *Convergence of the mass-transport steepest descent scheme for the subcritical Patlak-Keller-Segel model*, SIAM J. Numer. Anal. 46 (2008), pp.691-721.

[6] A. Blanchet, J. A. Carrillo, and PH.Laurençot, *Critical mass for a Patlak-Keller-Segel model with degenerate diffusion in higher dimensions*, Calc. Var. Partial Differential Equations 35 (2009), pp.133-168.

[7] A. Blanchet, J. Dolbeault and B. Perthame, *Two-dimensional Keller-Segel model: optimal critical mass and qualitative properties of the solutions*, Electron. J. Differential Equations 44 (2006), 32 pp. (electronic).

[8] H. Gajewski and K. Zacharias, *Global behavior of a reaction-diffusion system modeling chemotaxis*, Math. Nachr. 195 (1998), 77-114.

[9] M. A. Herrero and Velázquez, *Chemotaxis collapse for Keller-Segel model*, J. Math. Biol. 35 (1996), 177-194.

[10] S. Ishida and T. Yokota, *Global existence of weak solutions to quasilinear degenerate Keller-Segel systems of parabolic-parabolic type*, J. Differential Equations 252 (2012) 1421-1440.

[11] S. Ishida and T. Yokota, *Global existence of weak solutions to quasilinear degenerate Keller-Segel systems of parabolic-parabolic type with small data*, J. Differential Equations 252 (2012) 2469-2491.

[12] R. Jordan, D. Kinderlehrer and F. Otto, *The variational formulation of the Fokker-Planck equation*, SIAM J. Math. Anal. 29 (1998) 1-17 (electronic).

[13] E. F. Keller and L. A. Segel, *Initiation of slime mold aggregation viewed as an instability*, J. Theor. Biol. 26 (1970), 399-415.

[14] S. Lisini, *Nonlinear diffusion equations with variable coefficients as gradient flows in Wasserstein spaces*, ESAIM: COCV 15, (2009), 712-740.

[15] T.Nagai, T.Senba and K.Yoshida, *Application of the Trudinger-Moser inequality to a parabolic system of chemotaxis*, Funckcial. Ekvac. 40 (1997), 411-433.

[16] F. Otto, *Doubly degenerate diffusion equations as steepest descent*, Manuscript (1996).

[17] F. Otto, *The geometry of dissipative evolution equations: the porous medium equation*, Comm. Partial Diff. Eq. 26 (2001) 101-174.

[18] R. Rossi and G. Savaré, *Gradient flows of non convex functionals in Hilbert spaces and applications*, ESAIM Control Optim. Calc. Var., 12 (2006), pp. 564-614 (electronic).

[19] Y. Sugiyama, *Application of the best constant of the Sobolev inequality to degenerate Keller-Segel models* Adv. Differential Equations 12(2007), pp. 121-144.

[20] Y. Sugiyama, *Global existence in the sub-critical cases and finite time blow-up in the super-critical cases to degenerate Keller-Segel systems*, Differential Integral Equations 19 (2006) 841-876.

[21] Y. Sugiyama, *Time global existence and asymptotic behavior of solutions to degenerate quasi-linear parabolic systems of chemotaxis*, Differential Integral Equations 20 (2007) 133-180.

[22] Y. Sugiyama and Y. Yahagi, *Uniqueness and continuity of solution for the initial data in the scaling invariant class of the degenerate Keller-Segel system* J. Evol. Equ. 11 (2011), 319-337.

[23] C. Villani, *Topics in optimal transportation*, Graduate Studies in Mathematics 58, American Mathematical Society, Providence, RI (2003).