ARTICLE IN PRESS
YJDEQ:8759

Available online at www.sciencedirect.com

ScienceDirect

J. Differential Equations ••• (••••) •••–•••

Journal of Differential Equations

www.elsevier.com/locate/jde

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

Yoshifumi Mimura

Graduate School of Mathematical Sciences, The University of Tokyo, 3-8-1, Komaba, Meguro-ku, Tokyo, 153-8914, Japan

Received 29 June 2012; revised 24 November 2016

Abstract

We prove the time-global existence of solutions of the degenerate Keller–Segel system in higher dimensions, 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. We first consider a time-discretized problem, in which the values of the solution are determined iteratively by solving a certain minimizing problem at each time step. Here we use a new minimizing scheme at each time level, which gives the time-discretized solutions favorable regularity properties. As a consequence, it becomes relatively easy to prove that the time-discretized solutions converge to a weak solution of the original system as the time step size tends to zero.

MSC: 35K65; 35B33; 47J30; 35K40

Keywords: Keller–Segel; Chemotaxis; Degenerate diffusion; Wasserstein distance; Gradient flows

E-mail address: ymimura@ms.u-tokyo.ac.jp.

http://dx.doi.org/10.1016/j.jde.2017.03.020
0022-0396/

Please cite this article in press as: Y. Mimura, The variational formulation of the fully parabolic Keller–Segel system with degenerate diffusion, J. Differential Equations (2017), http://dx.doi.org/10.1016/j.jde.2017.03.020

1. Introduction

1.1. Description of the problem

We consider the following degenerate parabolic system:

$$\begin{cases} \partial_t u = \nabla \cdot (\nabla u^m - \chi u \nabla v), & x \in \Omega,\ t > 0, \\ \varepsilon \partial_t v = \Delta v - \gamma v + \alpha u, & x \in \Omega,\ t > 0, \\ u(x,0) = u_0(x),\ \varepsilon v(x,0) = \varepsilon v_0(x), & x \in \Omega, \end{cases} \tag{1.1}$$

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

$$\frac{\partial u^m}{\partial \boldsymbol{\nu}} - \chi u \frac{\partial v}{\partial \boldsymbol{\nu}} = v = 0, \quad x \in \partial\Omega,\ t > 0, \tag{1.2}$$

where $\boldsymbol{\nu}$ is the outward unit normal to Ω. Notice that (1.1) preserves the mass $\int_\Omega u\, dx$.

The aim of this paper is to prove the time-global existence of solutions of the system (1.1) 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 to formulate (1.1) as a gradient flow on a certain metric space, then to apply the variational method to prove the time-global existence. Note that the system (1.1) does not have a gradient flow 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 where the Wasserstein distance comes in, as we will explain later.

Recall that (1.1) is a version of the celebrated Keller–Segel chemotaxis model featuring a nonlinear diffusion. The Keller–Segel model was proposed by Keller and Segel [1] 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 microorganism, 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 the next generation. In equations (1.1), 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 (1.1). From a mathematical point of view, the aggregation phenomenon can be interpreted as a blow-up phenomenon of the solution of (1.1), that is, the density of slime molds singularly concentrates at some point.

The case $\varepsilon = 0$ is the so-called "parabolic–elliptic system", while the case $\varepsilon > 0$ is the so-called "parabolic–parabolic system". In both cases, the exponent $m = 2 - 2/d$, $d \geq 2$ has been identified as a critical exponent separating two different behaviors described below.

(i) **sub-critical case** $m > 2 - 2/d$
All solutions exist global in time [2–6]. We can find related results for a system with non-degenerate diffusion [7].

(ii) **super-critical case** $m < 2 - 2/d$
There are global solutions starting from suitably small initial data [8,9,4–6], while, if $\varepsilon = 0$, there are solutions that blow up in finite time [10,11,4]. The latter phenomenon is expected to take place also when $\varepsilon > 0$ but is not yet proved. However, the existence of blow-up

Please cite this article in press as: Y. Mimura, The variational formulation of the fully parabolic Keller–Segel system with degenerate diffusion, J. Differential Equations (2017), http://dx.doi.org/10.1016/j.jde.2017.03.020

solutions has been shown for a related system which includes a non-degenerate diffusion term $\nabla \cdot (A(u)\nabla u)$ with $A(u) \le u^m$ $(u \ge 1)$ instead of Δu^m [12].

(iii) **critical case** $m = 2 - 2/d$
It is expected that there is a critical mass M_c of the L^1-norm $\|u_0\|_{L^1}$ such that
(iii-a) all solutions exist global in time if $\|u_0\|_{L^1} < M_c$
(iii-b) for any $M > M_c$, there exists a solution with $\|u_0\|_{L^1} = M$ that blows up in finite time.

When $\varepsilon = 0$, result (iii) is known to be true for $d \ge 2$ [13–16]. For $\varepsilon > 0$, the global existence assertion (iii-a) is known only for $d = 2$ [17–19]. (See also the "Note after submission" at the end of this section.) As for $d > 2$ and $\varepsilon > 0$, the recent work of Ishida and Yokota [9] proves the time-global existence under the assumption that both u_0 and Δv_0 are relatively small – a condition that is restrictive compared with (iii-a). When $\varepsilon > 0$, the assertion (iii-b) is not yet proved except for some partial results. More precisely, Herrero and Velàzquez [20] construct an example of radially symmetric solution that blows up in finite time with a mass larger than $8\pi/\alpha\chi$ when Ω is a disc domain. On the other hand, Horstmann and Wang [21] prove that when Ω is a smooth bounded domain in $\mathbb{R}^2$, for any mass M larger than $4\pi/\alpha\chi$ and not equal to an integer multiple of $4\pi/\alpha\chi$, unbounded solutions with mass M exist. However, this result does not clarify whether the existence interval is finite or infinite.

The present paper gives a proof of (i) and (iii-a) when $\varepsilon > 0$ and $d > 2$. We first find a candidate $M_* > 0$ of the threshold mass in (iii) by investigating variational properties of the Lyapunov functional associated with (1.1). Then we prove the time-global existence of solutions of (1.1) under the assumption that $\|u_0\|_{L^1} < M_*$. In particular, we have $M_* = +\infty$ if $m > 2 - 2/d$ and $M_* < +\infty$ if $m = 2-2/d$. We thus obtain the results (i) and (iii-a). The main difference between our results and the earlier results by Ishida–Yokota [2,9] 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 (1.1) under a rather mild condition $\|u_0\|_{L^1} < M_*$ (and for an arbitrary $v_0 \in H_0^1(\Omega)$), where M_* is the threshold mass for the Lyapunov functional to be bounded from below. We expect that M_* is the threshold value in (iii) for our parabolic–parabolic system with $m = 2 - 2/d$, $d > 2$.

As mentioned earlier, our approach is to formulate (1.1) as a gradient flow in a certain metric space. One of the advantages of this approach is that it gives us a better understanding of the relation between the time-global existence of solutions of (1.1) 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 (1.1). Our approach is similar in spirit to that of Blanchet, Calvez and Carrillo [22] and Blanchet, Carlen and Carrillo [23], who interpreted the parabolic–elliptic Keller–Segel system (where one sets $m = 1$ and $\varepsilon = \gamma = 0$ in (1.1)) as a gradient flow in the Wasserstein space. However, in the present case, where $\varepsilon > 0$, one cannot reduce (1.1) to a single non-local equation and hence cannot interpret system (1.1) as a gradient flow in the Wasserstein space. Nonetheless, we can still formulate it partly in the framework of the Wasserstein space, as we will show later. The applications of the Wasserstein distance to the present type of evolution PDE's has been developed in the pioneering work of Otto [24,25], Jordan, Kinderlehrer and Otto [26], and other related works [27–32].

A common strategy in the above-mentioned works [27,28,22,23,26,29,30,24] is first to approximate the evolution equation by a time-discrete problem, which consists of solving a certain

Please cite this article in press as: Y. Mimura, The variational formulation of the fully parabolic Keller–Segel system with degenerate diffusion, J. Differential Equations (2017), http://dx.doi.org/10.1016/j.jde.2017.03.020

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 step size tends to 0. In proving the convergence, one needs some compactness properties of the time-discretized solutions, but the minimizing nature of the time-discretization significantly simplifies the compactness argument.

There are two approaches for proving that the limit of the discretized solutions is indeed a weak solution of the original equation. 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 [28,29], 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 more direct and simpler than the latter, while the latter approach based on the notion of curves of maximal slope gives a framework to deal with the problem systematically. The present paper adopts the former approach to prove the existence of the solution of (1.1). We also employ some techniques of the latter approach to establish the energy dissipative inequality, which confirms that our Lyapunov functional ϕ_m indeed decreases along the weak solution.

Note that, in this latter approach, the subdifferentials of the functional play a crucial role. In [28], existence of subdifferentials having certain good properties is shown for what are called "regular" functionals. However, in our present problem, it is not clear if our Lyapunov functional is regular, therefore the known results cannot be applied directly. Neither is the former approach so straightforward. In order to make this approach work, we need to show that the time-discretized solution possesses adequate regularity properties as otherwise the Euler–Lagrange equation would not make sense, and this is one of the main difficulties to overcome in the former approach.

In order to solve this regularity issue in the Euler–Lagrange approach, we make the following crucial modification in the time-discretization. Rather than solving for u and v simultaneously, we alternate between solving for u and v. 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 it becomes relatively easier to obtain sufficient regularity of the time-discretized solutions for deriving the Euler–Lagrange equations rigorously, thus establishing the convergence of time-discretized solutions to a solution of (1.1).

Let us comment briefly the boundary conditions. One can find many papers that consider Keller–Segel system on whole space $\mathbb{R}^d$ [13,22,23,14,15,10,8,2,9,11,4–6] or that on a bounded domain of $\mathbb{R}^d$ with homogeneous Neumann boundary conditions for both u and v [12,17,20,21,3,18,33,7]. On the other hand, there are also some papers that deal with boundary conditions similar to ours, namely, the no-flux boundary condition for u and the homogeneous Dirichlet boundary condition for v [34–36,33]. One of the advantages of our boundary conditions is that we can systematically deal with the problem on the whole space $\mathbb{R}^d$ and that on a bounded domain of $\mathbb{R}^d$. More precisely, under the boundary conditions (1.2), our critical mass M_* is independent of the choice of Ω and coincides with the critical mass for $\mathbb{R}^d$ (see Remark 4.1 and Proposition 4.2).

Uniqueness results are obtained by Bedrossian, Rodríguez and Bertozzi [13], and Sugiyama and Yahagi [37] for parabolic–elliptic system with $m = 2 - 2/d$, and by Kowalczyk and Szymańska [3] for parabolic–parabolic system with $m > 3 - 4/d$. However, we do not know if uniqueness holds for our case. This is partly due to the fact that our weak solutions do not have adequate regularity and that our Lyapunov functional ϕ_m is not λ-convex in the sense of [28, §2.4] (see also [28, §4] for uniqueness results).

This paper is constructed as follows. In the rest of this section we present our main results and give a list of symbols that will be used later. In section 2, we explain the basic concept of the Wasserstein distance. In section 3, we reformulate our problem in the framework of the Wasserstein space and define time-discretized solutions, which are given by solving a certain minimizing problem at each time level. Section 4 is devoted to variational analysis. Among other things, we prove basic properties of the threshold M_* and show that our M_* coincides with the threshold mass M_c introduced in [14], despite the apparent difference in the two definitions. We also prove that the minimizing problem associated with the time-discretization of (1.1) is solvable, which guarantees the existence of time-global discrete solution for each time step size τ. In section 5, we prove that the time-discretized solution converges to a weak solution of (1.1) as $\tau \to 0$, which establishes Theorem 1.4. Finally, in section 6, we prove the energy dissipative inequality (Theorem 1.5).

Note added after submission: After submitting this paper, the author was informed of the following article that had been announced in arXiv three months prior to the submission of the present paper:

Blanchet and Laurençot, *The parabolic–parabolic Keller–Segel system with critical diffusion as a gradient flow in* $\mathbb{R}^d$*,* $d \geq 3$, arXiv:1203.3573v1 [math.AP], 15 Mar 2012.

This paper deals with the same problem as ours for $m = 2 - 2/d$, $d > 2$, on $\mathbb{R}^d$. Furthermore, it interprets (1.1) as a gradient flow in the Wasserstein space and proves the time-global existence of solutions of (1.1) under the assumption that $\|u_0\|_{L^1} < M_c$, M_c being the threshold mass for the parabolic–elliptic case introduced in [14]. As we will show in section 4, this constant M_c coincides with our threshold mass M_*, therefore their result is quite similar to our Theorem 1.4. The main difference between the two papers lies in the method of time-discretization. As a result, the two papers take different approaches for proving the regularity of time-discretized solutions, which is a cornerstone property for proving convergence of the discretized solution to a weak solution of the original system. More precisely, in the preprint of Blanchet and Laurençot, an elaborate technique developed by [30] combined with an ingenious choice of a certain auxiliary functional is used to obtain the regularity of the time-discretized solution. On the other hand, the present paper uses a different discretization scheme (3.7) (which we may call the splitting scheme) that allows us to obtain regularity without sophisticated techniques. Incidentally, the present paper proves the energy dissipative inequality for weak solutions (see section 6), which may be of independent interest. The author would like to thank the anonymous referee for pointing out the above paper to him.

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 (1.1):

$$\phi_m(u, v) := \frac{1}{m-1} \int_\Omega u^m \, dx - \chi \int_\Omega uv \, dx + \frac{\chi}{2\alpha} \int_\Omega |\nabla v|^2 + \gamma v^2 \, dx, \tag{1.3}$$

where Ω is a bounded domain in $\mathbb{R}^d$, $d > 2$, with smooth boundary. We consider the functional ϕ_m in the space

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

We define $\mu_M(\Omega)$ and M_* by

$$\begin{aligned} \mu_M(\Omega) &:= \inf_{(u,v)\in X_M(\Omega)} \phi_m(u, v), \\ M_*(\Omega) &:= \sup\{M \geq 0\ ;\ \mu_M(\Omega) > -\infty\}. \end{aligned} \tag{1.4}$$

We have the following theorem on the properties of M_*. The first theorem, which is concerned with the critical case $m = 2 - 2/d$, shows that M_* is indeed a threshold mass for the lower boundedness of ϕ_m.

Theorem 1.1 *(Properties of M_* for $m = 2 - 2/d$). Let $m = 2 - 2/d$. Then:*

(i) $0 < M_*(\Omega) < +\infty$;
(ii) $\mu_M(\Omega) \geq 0$ *for* $0 < M \leq M_*(\Omega)$, *while* $\mu_M(\Omega) = -\infty$ *for* $M > M_*(\Omega)$.

Furthermore, $M_*(\Omega)$ depends only on α, χ, d and is independent of γ and the choice of Ω, and it holds that

$$M_*(\alpha, \chi, d) = (\alpha\chi)^{-\frac{d}{2}} M_*(1, 1, d).$$

According to Theorem 1.1, we write just M_* without the dependence upon Ω hereafter. The next theorem is concerned with the case $m > 2 - 2/d$:

Theorem 1.2 *(Properties of M_* for $m > 2 - 2/d$). Let $m > 2 - 2/d$. Then: $M_* = +\infty$. In other words, $\mu_M(\Omega) > -\infty$ for every $M > 0$.*

Remark 1.1. The case $m > 2 - 2/d$ is often called the 'sub-critical' case. The justification of this meaning for case $\varepsilon = 0$ can be found in [4].

Next, we state the definition of weak solutions and the time-global existence of solutions of the system (1.1).

Definition 1.3 *(Weak solutions).* We say that a pair (u, v) of non-negative functions is a weak solution of (1.1) on the time interval $[0, T]$ if

(i) $(u, v) \in L^\infty(0, T; L^1(\Omega) \cap L^m(\Omega)) \times L^\infty(0, T; H_0^1(\Omega))$ and $\|u(\cdot, t)\|_{L^1} = \|u_0\|_{L^1}$.
(ii) $u \in L^{2p}(0, T; L^2(\Omega))$ for $p = 2m/(d(2 - m)) \geq m$ and $u^m \in L^2(0, T; W^{1,1}(\Omega))$, $v \in L^2(0, T; W^{2,2}(\Omega))$.
(iii) $\lim_{t\downarrow 0} d_W(u(t), u_0) = 0$ and $\lim_{t\downarrow 0} \|v(t) - v_0\|_{L^2} = 0$, where d_W denotes the Wasserstein distance to be introduced in §2.

(iv) (u, v) has the following additional regularities:

$$\int_0^T \int_\Omega \frac{|\nabla u^m - \chi u \nabla v|^2}{u} \, dx dt < +\infty,$$

$$\int_0^T \int_\Omega |\Delta v - \gamma v + \alpha u|^2 \, dx dt < +\infty.$$

(v) (u, v) satisfies

$$\begin{cases} \displaystyle\int_\Omega (u(b) - u(a))\varphi \, dx + \int_a^b \int_\Omega \langle \nabla u^m - \chi u \nabla v, \nabla \varphi \rangle \, dx dt = 0, \\ \displaystyle\varepsilon \int_\Omega (v(b) - v(a))\psi \, dx = \int_a^b \int_\Omega (\Delta v - \gamma v + \alpha u)\psi \, dx dt, \end{cases} \tag{1.5}$$

for any $a, b \in [0, T]$ and for all $\varphi \in C_c^\infty(\mathbb{R}^d)$ and for all $\psi \in C_c^\infty(\Omega)$.

Remark 1.2 *(Weak formulation of boundary conditions).* Since the test function φ in (1.5) belongs to $C_c^\infty(\mathbb{R}^d)$ rather than $C_c^\infty(\Omega)$, the identities (1.5) imply the following natural boundary condition in the weak sense:

$$\frac{\partial u^m}{\partial \boldsymbol{\nu}} - \chi u \frac{\partial v}{\partial \boldsymbol{\nu}} = 0 \text{ on } \partial\Omega \times (0, T).$$

In view of this and the fact that $v \in H_0^1(\Omega)$, one sees that the boundary conditions (1.2) are automatically satisfied by our weak solution (u, v).

Theorem 1.4 (Time-global existence). *Let $m \geq 2 - 2/d$. For any $u_0 \in L^2(\Omega) \cap L^m(\Omega)$ and $v_0 \in H_0^1(\Omega)$ with $u_0, v_0 \geq 0$, there exists a weak solution (u, v) of (1.1) with this initial data that exists globally for all $t \geq 0$, provided that u_0 satisfies*

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

Recall that $M_* = +\infty$ if $m > 2 - 2/d$. Thus, in this case, a time-global weak solution exists for any $u_0 \in L^2(\Omega) \cap L^m(\Omega)$, $v_0 \in H_0^1(\Omega)$ with $u_0, v_0 \geq 0$. As a matter of fact, as for the case $m > 2 - 2/d$, Theorem 1.4 has already been known in a slightly different setting. In fact Ishida and Yokota [2] obtained the same result for $\Omega = \mathbb{R}^d$, except that their definition of weak solutions is slightly different from ours. The main novelty of the present paper is the result for the case $m = 2 - 2/d$. As we mentioned earlier, the constant M_* coincides with the threshold mass obtained in Blanchet et al. [14] for $\varepsilon = \gamma = 0$.

Please cite this article in press as: Y. Mimura, The variational formulation of the fully parabolic Keller–Segel system with degenerate diffusion, J. Differential Equations (2017), http://dx.doi.org/10.1016/j.jde.2017.03.020

Theorem 1.4 will be proved by constructing time-discretized solutions as defined in subsection 3.2 and showing that this discretized solution converges to a weak solution of (1.1) as the time step size tends to 0. As we see below, any weak solution constructed in this way satisfies the energy dissipative inequality:

Theorem 1.5 *(Energy dissipative inequality). Let (u, v) be a weak solution of (1.1) that is given by a limit function of discrete solutions to be defined in Definition 3.2. Then*

$$\phi_m(u(a), v(a)) - \phi_m(u(b), v(b)) \geq \int_a^b\int_\Omega \frac{|\nabla u^m - \chi u \nabla v|^2}{u}\,dxdt + \frac{\chi}{\alpha\varepsilon}\int_a^b\int_\Omega |\Delta v - \gamma v + \alpha u|^2\,dxdt \tag{1.6}$$

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

Here, the meaning of the first term in the right-hand side of (1.6) is that the integrand is equal to $\frac{|\nabla u^m - \chi u \nabla v|^2}{u}$ on the subset of Ω where u is positive and zero otherwise. The above theorem implies, in particular, that ϕ_m is indeed a Lyapunov functional for (1.1) even for weak solutions. A similar inequality has been known for the parabolic–elliptic systems (see [13,14]), but as far as the author knows it is new in the case of parabolic–parabolic system.

1.3. 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
$\mathcal{D}(\phi)$	effective domain of functional ϕ
$\lvert\partial\phi\rvert(v)$	metric slope of functional ϕ at v
$\mathscr{P}_2(\Omega)$	probability measures on Ω with finite second moment
$L^p(\Omega)$	p-summable functions on $\Omega \subset \mathbb{R}^d$ with respect to $\mathscr{L}^d$
$L^2_u(\Omega; \mathbb{R}^d)$	$\mathbb{R}^d$-valued 2-summable functions on Ω with respect to $u\mathscr{L}^d$
$\lVert\cdot\rVert_{L^2(u)}$	the norm in $L^2_u(\Omega; \mathbb{R}^d)$ *i.e.* $\lVert\boldsymbol{\xi}\rVert_{L^2(u)} = \lVert\boldsymbol{\xi}\sqrt{u}\rVert_{L^2(\Omega)}$ with $u \geq 0$
$C_c^\infty(\Omega)$	compactly supported smooth functions on Ω
$C_b(\Omega)$	continuous and bounded functions on Ω
$W^{k,p}(\Omega)$	Sobolev space over Ω
$H_0^1(\Omega)$	the closure of $C_c^\infty(\Omega)$ in $W^{1,2}(\Omega)$
$\dot{H}^1(\Omega)$	the closure of $C_c^\infty(\Omega)$ in the seminorm $\lVert\nabla\cdot\rVert_{L^2}$

2. Preliminaries

In this section, we collect some results on the Wasserstein distance. We refer to the books [28,31,32] and the handbook [38].

Please cite this article in press as: Y. Mimura, The variational formulation of the fully parabolic Keller–Segel system with degenerate diffusion, J. Differential Equations (2017), http://dx.doi.org/10.1016/j.jde.2017.03.020

$\mathscr{P}(\Omega)$ denotes the space of probability measures on Ω. For any $\mu \in \mathscr{P}(\Omega)$, we identify μ with a measure on $\mathbb{R}^d$ by setting $\mu(\mathbb{R}^d \setminus \Omega) = 0$. Thus we have $\mathscr{P}(\Omega) \subset \mathscr{P}(\mathbb{R}^d)$. In what follows we will mainly work on $\mathscr{P}(\mathbb{R}^d)$. We endow $\mathscr{P}(\mathbb{R}^d)$ with the following topology:

Definition 2.1 *(Narrow convergence [28, §5.1]).* 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

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

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 [28, §5.2]).* 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 denote it by $\nu = \boldsymbol{t}_{\#}\mu$.

Remark 2.1. Note that, in the special case where μ, ν possess density functions $d\mu(x) := u(x)\, dx$, $d\nu(y) := v(y)\, dy$, and if $\boldsymbol{t}$ is a diffeomorphism, then we have

$$v(\boldsymbol{t}(x)) \det(D\boldsymbol{t}(x)) = u(x).$$

Hence,

$$\boldsymbol{t}_{\#}(u(x)dx) = u(\boldsymbol{t}^{-1}(x)) \det(D\boldsymbol{t}^{-1}(x))\, dx.$$

Definition 2.3 *(Wasserstein distance [28, §7.1], [31, Def. 6.1]).* The *Wasserstein distance* d_W is defined by

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

Please cite this article in press as: Y. Mimura, The variational formulation of the fully parabolic Keller–Segel system with degenerate diffusion, J. Differential Equations (2017), http://dx.doi.org/10.1016/j.jde.2017.03.020

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

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

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

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

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" (see [28, Prop. 7.15], [31, Def. 6.4, Thm. 6.18]). We say that a set $\mathcal{S} \subset \mathscr{P}(\mathbb{R}^d)$ is "tight" if for any $\varepsilon > 0$, there exists a compact set $K_\varepsilon \subset \mathbb{R}^d$ such that

$$\mu(\mathbb{R}^d \setminus K_\varepsilon) \le \varepsilon \quad \text{ for all } \mu \in \mathcal{S}.$$

It is known that if $\mathcal{S}$ is tight, then $\mathcal{S}$ is relatively compact in $\mathscr{P}(\mathbb{R}^d)$ with respect to the narrow topology (Prokhorov's compactness theorem; see for instance [39, III-59]). If Ω is a bounded domain in $\mathbb{R}^d$ and if $\mathcal{S} \subset \mathscr{P}(\Omega)$ $(\subset \mathscr{P}(\mathbb{R}^d))$, then it is clear that $\mathcal{S}$ is tight.

Proposition 2.4 *(Lower semicontinuity of d_W, [28, Lem. 7.1.4]). Let $(\mu_n), (\nu_n) \subset \mathscr{P}_2(\mathbb{R}^d)$ be two tight sequences narrowly converging to μ, ν in $\mathscr{P}(\mathbb{R}^d)$. Then*

$$d_W(\mu,\nu) \le \liminf_{n\to\infty} d_W(\mu_n,\nu_n).$$

Proposition 2.5 *(Brenier's theorem [38, Thm. 2.3], [32, Thm. 2.12]). 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 $p_0 = (\boldsymbol{id} \times \boldsymbol{t}_\mu^\nu)_\# \mu$ and*

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

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

In the rest of the this paper, we identify the probability measure $\mu = u\mathscr{L}^d$ with its density u and write u instead of $u\mathscr{L}^d$.

Please cite this article in press as: Y. Mimura, The variational formulation of the fully parabolic Keller–Segel system with degenerate diffusion, J. Differential Equations (2017), http://dx.doi.org/10.1016/j.jde.2017.03.020

3. Reformulation of the problem

In this section we reformulate the system (1.1) as a gradient flow in the Wasserstein space and define a time-discretization scheme for this new system.

3.1. Gradient flows in the Wasserstein space

Since the pioneering work of Jordan, Kinderlehrer and Otto [26], there have been many works that interpret the following type of evolution equation as a gradient flow on the Wasserstein space:

$$\partial_t u = \nabla \cdot \left(u \nabla \left(\frac{\delta \phi}{\delta u} \right) \right) \quad \text{in } \mathbb{R}^d \times (0, \infty), \tag{3.1}$$

where $\delta\phi/\delta u$ denotes the first variation of a functional ϕ and ∇ denotes differentiation with respect to the space variable x. Since (3.1) conserves the mass $\int_{\mathbb{R}^d} u\,dx$, by normalization one can assume

$$\int_{\mathbb{R}^d} u\,dx = 1;$$

hence u can be regarded as a probability measure, and thus there is room for interpreting (3.1) as an evolution equation on the space of probability measures – more specifically the Wasserstein space. However, it is not easy to directly interpret (3.1) as a gradient flow, as the Wasserstein space does not possess adequate differential structure. Given such difficulty, a usual strategy is first to consider time-discretization of (3.1) based on the following steepest descent scheme: Let $\tau > 0$ be the time step size and u_τ^0 be the initial data. One then recursively defines u_τ^k by

$$u_\tau^k \in \operatorname*{argmin}_{u \in \mathscr{P}_2(\mathbb{R}^d)} \left\{ \phi(u) + \frac{1}{2\tau} d_W^2(u, u_\tau^{k-1}) \right\} \quad \text{for } k = 1, 2, 3, \ldots, \tag{3.2}$$

where "argmin" denotes the set of minimizers. In order for this minimizing problem to be solvable, one assumes that $\phi(u)$ is coercive and lower semicontinuous, and possesses some sort of compactness properties. One then considers the limit of $(u_\tau^k)_{k\in\mathbb{N}}$ as $\tau \to 0$, which clearly exists for some sequence $\tau_n \to 0$ because of the compactness. The central issue is how to prove that the limit function is a weak solution of (3.1).

There are two different approaches for this last step. One approach, found in [27,22,23,26,30,24,25], is to use the Euler–Lagrange equation associated with (3.2). We will explain more about this approach below. The other approach, found in [28,29], uses the concept of "curves of maximal slope" on abstract metric spaces and proves first that the limit of $(u_\tau^k)_{k\in\mathbb{N}}$ is indeed a curve of maximal slope in the Wasserstein space. Then, characterization of curves of maximal slope in terms of subdifferentials shows that this limit function satisfies (3.1) in the weak sense. Both approaches have their advantages, but we will adopt the former one in the present paper.

Now let us explain more about this former approach. For each $k \in \mathbb{N}$, let $I_k(u)$ denote the functional on the right-hand side of (3.2). Then, for each $\zeta \in C_c^\infty(\mathbb{R}^d)$, $h \in \mathbb{R}$, we consider a perturbation of $u \in \mathscr{P}_2(\mathbb{R}^d)$ in the form of the push-forward $(\boldsymbol{id} + h\nabla\zeta)_\# u$. Here $|h|$ is chosen sufficiently small (say, $|h| \le \delta_0$) so that $(\boldsymbol{id} + h\nabla\zeta)$ is a diffeomorphism on $\mathbb{R}^d$. Then, since $(\boldsymbol{id} + h\nabla\zeta)_\# u \in \mathscr{P}_2(\mathbb{R}^d)$, by the minimality (3.2) we have

$$I_k(u_\tau^k) = \min_{|h| \le \delta_0} I_k((\boldsymbol{id} + h\nabla\zeta)_\# u_\tau^k).$$

Differentiating the functional I_k on the right-hand side with respect to h and setting $h = 0$, we obtain, at least formally, the following identity (cf. Lemmas 5.5 and 5.8):

$$\int_\Omega \frac{u_\tau^k - u_\tau^{k-1}}{\tau} \zeta \, dx + \int_\Omega \left\langle \nabla \left(\frac{\delta\phi}{\delta u} \right), \nabla\zeta \right\rangle u_\tau^k \, dx = o(1).$$

This is the Euler–Lagrange equation. From its form it is clear that this equation converges formally to a weak form of (3.1) as $h \to 0$. As $(u_\tau^k)_{k\in\mathbb{N}}$ is defined by a steepest descent scheme, this gives an (indirect) interpretation of (3.1) as a gradient flow.

Now let us come back to our problem (1.1). As is easily seen, (1.1) is formally written as follows:

$$\begin{cases} \partial_t u = \nabla \cdot \left(u \nabla \left(\frac{\delta\phi_m}{\delta u} \right) \right), \\ \varepsilon \partial_t v = -\frac{\alpha}{\chi} \left(\frac{\delta\phi_m}{\delta v} \right). \end{cases} \tag{3.3}$$

By an analogy with (3.1), this expression suggests that (1.1) may be interpreted (at least formally) as a gradient flow in the product space of the Wasserstein space and L^2-space. We will see that this interpretation can be justified rigorously.

3.2. Construction of time-discretized solutions

In this subsection, we construct a time-discretized solution associated with the system (1.1). We begin by normalizing u. We make the 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:

$$\tilde{\chi} = \frac{\chi}{M^{m-1}}, \quad \tilde{\varepsilon} = M^{m-1}\varepsilon, \quad \tilde{\alpha} = \alpha M. \tag{3.4}$$

Then $(\tilde{u}, v)$ satisfies the same equations as (1.1) with the above new parameters and the new time variable $\tilde{t}$; furthermore, we have $\int_\Omega \tilde{u} \, dx = 1$. Therefore, in what follows it suffices to consider only the solution of (1.1) that satisfies

$$\int_\Omega u_0 \, dx = 1. \tag{3.5}$$

Remark 3.1. When $m = 2 - 2/d$, by Theorem 1.1, we have

$$M_*(\tilde{\alpha}, \tilde{\chi}, d) = (\tilde{\alpha}\tilde{\chi})^{-\frac{d}{2}} M_*(1, 1, d) = \frac{M_*(\alpha, \chi, d)}{M} \tag{3.6}$$

Therefore, whether or not the mass M is below the threshold value M_* in the original parameters is equivalent to whether or not $M_*(\tilde{\alpha}, \tilde{\chi}, d) > 1$ in the new parameters. In the rest of this subsection, the new parameters $\tilde{\alpha}, \tilde{\chi}, \tilde{\varepsilon}, \tilde{u}_0, \tilde{u}$ will be denoted by $\alpha, \chi, \varepsilon, u_0, u$ for notational simplicity.

Now, we divide the time interval $[0, \infty)$ into sub-intervals with length $\tau > 0$. Let $(u_\tau^0, v_\tau^0) = (u_0, v_0)$ be initial data and we recursively define (u_τ^k, v_τ^k) for $k = 1, 2, 3, \dots$ by

$$\begin{aligned} v_\tau^k &\in \underset{v \in H_0^1(\Omega)}{\operatorname{argmin}} \left\{ \phi_m(u_\tau^{k-1}, v) + \frac{\varepsilon\chi}{2\alpha\tau} \|v - v_\tau^{k-1}\|_{L^2(\Omega)}^2 \right\}, \\ u_\tau^k &\in \underset{u \in \mathscr{P}_2(\Omega)}{\operatorname{argmin}} \left\{ \phi_m(u, v_\tau^k) + \frac{1}{2\tau} d_W^2(u, u_\tau^{k-1}) \right\}, \end{aligned} \tag{3.7}$$

that is, v_τ^k minimizes

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

and u_τ^k minimizes

$$u \mapsto \phi_m(u, v_\tau^k) + \frac{1}{2\tau} d_W^2(u, u_\tau^{k-1}) \ \text{ in } \mathscr{P}_2(\Omega).$$

Thus the recursive procedure $(u_\tau^{k-1}, v_\tau^{k-1}) \mapsto (u_\tau^k, v_\tau^k)$ at each time step consists of minimizing v and u alternately instead of minimizing them at once. We may therefore call (3.7) a "splitting scheme". The reason for adopting this splitting scheme is that it makes each minimizing procedure easier to handle as we only need to consider minimization in one variable rather than two.

Proposition 3.1. *Let $u_0 \in L^2(\Omega) \cap L^m(\Omega)$ with $u_0 \geq 0$ and $v_0 \in H_0^1(\Omega)$ with $v_0 \geq 0$. Assume (3.5). Then u_τ^k and v_τ^k are uniquely defined and non-negative for all $k \in \mathbb{N}$. Moreover, if $M_* = M_*(\alpha, \chi, m, d) > 1$, then there exists a constant $C_0 := C_0(\alpha, \chi, m, d, u_0, v_0)$ such that*

$$\|u_\tau^k\|_{L^m(\Omega)} + \|\nabla v_\tau^k\|_{L^2(\Omega)} \leq C_0 \quad \text{for all } k \in \mathbb{N}. \tag{3.8}$$

Definition 3.2 *(Discrete solutions).* Given a solution $(u_\tau^k, v_\tau^k)_{k \in \mathbb{N}}$ of (3.7), we define the piecewise constant interpolation

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

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

Please cite this article in press as: Y. Mimura, The variational formulation of the fully parabolic Keller–Segel system with degenerate diffusion, J. Differential Equations (2017), http://dx.doi.org/10.1016/j.jde.2017.03.020

Theorem 1.4 and Theorem 1.5 are consequences of the following theorem:

Theorem 3.3 *(Lyapunov solution). Let $u_0 \in L^2(\Omega) \cap L^m(\Omega)$ with $u_0 \geq 0$ and $v_0 \in H_0^1(\Omega)$ with $v_0 \geq 0$. Assume (3.5) and $M_* = M_*(\alpha, \chi, m, d) > 1$. Then there exists a sequence τ_n with $\tau_n \downarrow 0$ and a limit function (u, v) such that*

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

The limit function (u, v) is a time-global weak solution of (1.1) and satisfies the energy dissipative inequality (1.6) in Theorem 1.5.

Remark 3.2. Note that the uniqueness of the solution of (1.1) (for each given initial data) is not known.

We prove Proposition 3.1 in §4 and Theorem 3.3 in §5 except the energy dissipative inequality (1.6), which will be proved in §6.

4. Variational analysis

In this section, we give a proof of Theorem 1.1, Theorem 1.2 and Proposition 3.1. Here and throughout the rest of the paper, the symbol $\|\cdot\|_{L^p}$ stands for the norm $\|\cdot\|_{L^p(\Omega)}$, unless otherwise stated. We begin with some preliminary estimates.

By the Hölder inequality, the Sobolev inequality, and the interpolation inequality, we have

$$\begin{aligned} \left| \int_\Omega uv\,dx \right| &\leq \|u\|_{L^{\frac{2d}{d+2}}} \|v\|_{L^{\frac{2d}{d-2}}} \\ &\leq C_s \|u\|_{L^{\frac{2d}{d+2}}} \|\nabla v\|_{L^2} \\ &\leq C_s \|u\|_{L^1}^{1-\theta} \|u\|_{L^m}^{\theta} \|\nabla v\|_{L^2}. \end{aligned} \tag{4.1}$$

Here we have $\theta = \frac{m(d-2)}{2d(m-1)} \in (0,1)$ since $d > 2$ implies $1 < \frac{2d}{d+2} < \frac{2(d-1)}{d} \leq m$. The constant C_s denotes the best Sobolev constant for the embedding $\dot{H}^1(\Omega) \hookrightarrow L^{\frac{2d}{d-2}}(\Omega)$. We define C_m and C_* by

$$C_m := \begin{cases} C_s & \text{if } m > 2 - \frac{2}{d}, \\ C_* := \displaystyle\sup_{\substack{(u,v) \in \mathcal{Y}(\Omega) \\ u,v \neq 0}} \frac{\|uv\|_{L^1}}{\|u\|_{L^1}^{1/d} \|u\|_{L^m}^{m/2} \|\nabla v\|_{L^2}} & \text{if } m = 2 - \frac{2}{d}, \end{cases} \tag{4.2}$$

where $\mathcal{Y}(\Omega) := \{(u, v) \in \left(L^1(\Omega) \cap L^m(\Omega)\right) \times \dot{H}^1(\Omega)\}$. Note that $\theta = \frac{m}{2}$ when $m = 2 - 2/d$ and that $\dot{H}^1(\Omega)$ coincides with $H_0^1(\Omega)$ if Ω is bounded. It is obvious that $C_* \leq C_s$ from the inequality (4.1). Using this constant C_m and Young's inequality, for any $\beta > 0$ we have

$$\int_\Omega uv\,dx \le C_m \|u\|_{L^1}^{1-\theta}\|u\|_{L^m}^{\theta}\|\nabla v\|_{L^2}$$
$$\le \frac{\beta C_m^2}{2}\|u\|_{L^1}^{2(1-\theta)}\|u\|_{L^m}^{2\theta} + \frac{1}{2\beta}\|\nabla v\|_{L^2}^2.$$

Hence, it holds that

$$\phi_m(u,v) \ge \frac{1}{m-1}\|u\|_{L^m}^m - \frac{\beta\chi C_m^2}{2}\|u\|_{L^1}^{2(1-\theta)}\|u\|_{L^m}^{2\theta} + \frac{\chi}{2}\left(\frac{1}{\alpha}-\frac{1}{\beta}\right)\|\nabla v\|_{L^2}^2. \tag{4.3}$$

Particularly, if $m = 2-2/d$, then letting $\beta = \alpha+\delta$ with $\delta \ge 0$, (4.3) reduces to

$$\phi_m(u,v) \ge \frac{\alpha\chi C_*^2}{2}\left(M_1^{2/d} - \frac{\alpha+\delta}{\alpha}\|u\|_{L^1}^{2/d}\right)\|u\|_{L^m}^m + \frac{\delta\chi}{2\alpha(\alpha+\delta)}\|\nabla v\|_{L^2}^2, \tag{4.4}$$

where M_1 is the constant defined by

$$M_1 := \left(\frac{2}{\alpha\chi(m-1)C_*^2}\right)^{d/2}. \tag{4.5}$$

Now we are ready to prove Theorem 1.1 and Theorem 1.2.

Proof of Theorem 1.2. The condition $m > 2-2/d$ implies $2\theta < m$ in (4.3). Consequently, the sum of the first term and the second term in the right-hand side of (4.3) is estimated from below by $\frac{1}{2(m-1)}\|u\|_{L^m}^m - C$ if the constant $C>0$ is chosen sufficiently large. Therefore, letting $\beta = 2\alpha$ in (4.3), we have

$$\phi_m(u,v) \ge \frac{1}{2(m-1)}\|u\|_{L^m}^m + \frac{\chi}{4\alpha}\|\nabla v\|_{L^2}^2 - C > -\infty \tag{4.6}$$

on $X_M(\Omega)$ for any $M = \|u\|_{L^1}$. This implies $M_* = +\infty$ in this case. $\square$

Next, we consider the case of $m = 2-2/d$.

Proof of Theorem 1.1. From the estimate (4.4) with $\delta = 0$, it is clear that ϕ_m is bounded from below as long as $\|u\|_{L^1} \le M_1$. Thus we have $M_* \ge M_1$.

Next we show that $M_1 \ge M_*$. Without loss of generality, we can assume that Ω contains the origin. We define the operator $\mathcal{T}_{\lambda_1,\lambda_2} : (u,v) \in X_M(\Omega) \mapsto (U,V) \in X_M(\mathbb{R}^d)$ by

$$U(x) := \begin{cases} \lambda_1^d u(\lambda_1 x) & \text{if } \lambda_1 x \in \Omega, \\ 0 & \text{if } \lambda_1 x \in \mathbb{R}^d \setminus \Omega, \end{cases} \qquad V(x) := \begin{cases} \lambda_2^{d-2} v(\lambda_1 x) & \text{if } \lambda_1 x \in \Omega, \\ 0 & \text{if } \lambda_1 x \in \mathbb{R}^d \setminus \Omega, \end{cases}$$

for $\lambda_1, \lambda_2 > 0$. We identify $(u,v) \in X_M(\Omega)$ with $\mathcal{T}_{1,1}(u,v) \in X_M(\mathbb{R}^d)$. Then we can express ϕ_m as

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

One can easily check that for $(U,V)=\mathcal{T}_{\kappa,1}(u,v)$ and $\kappa>0$,

$$\begin{gathered}\phi_m(U,V;\alpha_\kappa,\chi_\kappa,\gamma_\kappa) = \kappa^{d-2}\phi_m(u,v;\alpha,\chi,\gamma) \text{ and } \|U\|_{L^1(\mathbb{R}^d)} = \|u\|_{L^1(\Omega)} \\ \text{where}\quad \alpha_\kappa = \frac{\alpha}{\kappa^{d-2}},\quad \chi_\kappa = \kappa^{d-2}\chi,\quad \gamma_\kappa = \kappa^2\gamma.\end{gathered} \tag{4.7}$$

That is, the lower boundedness of $\phi_m(u,v;\alpha,\chi,\gamma)$ is equivalent to that of $\phi_m(U,V;\alpha_\kappa,\chi_\kappa,\gamma_\kappa)$ for fixed κ.

We now investigate the lower boundedness of $\phi_m(U,V;\alpha_\kappa,\chi_\kappa,\gamma_\kappa)$. From the definition of the constant C_*, for any $\delta>0$, there exists a pair $(u_\delta,v_\delta)\in X_M(\Omega)$ such that

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

Then the pair $(U_\delta,V_\delta) := \mathcal{T}_{\kappa,1}(u_\delta,v_\delta)\in X_M(\mathbb{R}^d)$ also satisfies

$$\|U_\delta V_\delta\|_{L^1(\mathbb{R}^d)} \geq (C_*-\delta)\|U_\delta\|_{L^1(\mathbb{R}^d)}^{1/d}\|U_\delta\|_{L^m(\mathbb{R}^d)}^{m/2}\|\nabla V_\delta\|_{L^2(\mathbb{R}^d)}.$$

Let us assume that $M := \|u_\delta\|_{L^1(\Omega)} = \|U_\delta\|_{L^1(\mathbb{R}^d)} > M_1$ and we choose δ such that

$$M_1^{2/d} - \frac{(C_*-\delta)^2}{C_*^2}M^{2/d} < 0.$$

Let

$$\begin{gathered}\kappa^{d-2} = \left(\frac{\alpha(C_*-\delta)\|U_\delta\|_{L^1(\mathbb{R}^d)}^{1/d}\|U_\delta\|_{L^m(\mathbb{R}^d)}^{m/2}}{\|\nabla V_\delta\|_{L^2(\mathbb{R}^d)}}\right) \\ \text{or equivalently}\quad \alpha_\kappa = \frac{\|\nabla V_\delta\|_{L^2(\mathbb{R}^d)}}{(C_*-\delta)\|U_\delta\|_{L^1(\mathbb{R}^d)}^{1/d}\|U_\delta\|_{L^m(\mathbb{R}^d)}^{m/2}}.\end{gathered}$$

Then we have

$$\begin{aligned}\|U_\delta V_\delta\|_{L^1(\mathbb{R}^d)} &\geq (C_*-\delta)\|U_\delta\|_{L^1(\mathbb{R}^d)}^{1/d}\|U_\delta\|_{L^m(\mathbb{R}^d)}^{m/2}\|\nabla V_\delta\|_{L^2(\mathbb{R}^d)} \\ &= \frac{\alpha_\kappa(C_*-\delta)^2\|U_\delta\|_{L^1(\mathbb{R}^d)}^{2/d}\|U_\delta\|_{L^m(\mathbb{R}^d)}^{m}}{2} + \frac{1}{2\alpha_\kappa}\|\nabla V_\delta\|_{L^2(\mathbb{R}^d)}^2.\end{aligned}$$

Consequently, it holds that

Please cite this article in press as: Y. Mimura, The variational formulation of the fully parabolic Keller–Segel system with degenerate diffusion, J. Differential Equations (2017), http://dx.doi.org/10.1016/j.jde.2017.03.020

$$\begin{aligned}
&\phi_m(U_\delta, V_\delta; \alpha_\kappa, \chi_\kappa, \gamma_\kappa) - \frac{\chi_\kappa \gamma_\kappa}{2\alpha_\kappa} \int_{\mathbb{R}^d} V_\delta^2 \, dx \\
&\qquad\qquad \leq \frac{\alpha_\kappa \chi_\kappa C_*^2}{2} \left(M_1^{2/d} - \frac{(C_* - \delta)^2}{C_*^2} M^{2/d} \right) \|U_\delta\|_{L^m(\mathbb{R}^d)}^m .
\end{aligned}$$

Letting $(\tilde{U}(\lambda), \tilde{V}(\lambda)) := \mathcal{T}_{\lambda,\lambda}(U_\delta, V_\delta)$, we have

$$\begin{aligned}
\phi_m(\tilde{U}(\lambda), \tilde{V}(\lambda)) &= \lambda^{d-2} \left(\phi_m(U_\delta, V_\delta) - \frac{\chi_\kappa \gamma_\kappa}{2\alpha_\kappa} \int_{\mathbb{R}^d} V_\delta^2 \, dx \right) + \frac{\chi_\kappa \gamma_\kappa}{2\alpha_\kappa} \int_{\mathbb{R}^d} \tilde{V}^2(\lambda) \, dx \\
&\leq \lambda^{d-2} \frac{\alpha_\kappa \chi_\kappa C_*^2}{2} \left(M_1^{2/d} - \frac{(C_* - \delta)^2}{C_*^2} M^{2/d} \right) \|U_\delta\|_{L^m(\mathbb{R}^d)}^m + \lambda^{d-4} \frac{\chi_\kappa \gamma_\kappa}{2\alpha_\kappa} \|V_\delta\|_{L^2(\mathbb{R}^d)}^2 \\
&\to -\infty, \quad (\lambda \to +\infty).
\end{aligned}$$

Therefore, we obtain $M_1 \geq M_*$ and hence

$$M_* = M_1 = \left(\frac{2}{\alpha \chi (m-1) C_*^2} \right)^{d/2} . \tag{4.8}$$

Finally, we show that M_* is independent of the choice of Ω. For that purpose, we denote the value C_* in (4.2) by $C_*(\Omega)$, in order to clarify its dependence on the domain. Then, from (1.4), it suffices to prove that $C_*(\Omega)$ is independent of the choice of Ω. We denote by B_R an open ball in $\mathbb{R}^d$ with center at 0 and radius R. We can choose $R_1, R_2 > 0$ such that $B_{R_1} \subset \Omega \subset B_{R_2}$ since Ω is bounded and contains the origin. It is obvious that $C_*(B_{R_1}) \leq C_*(\Omega) \leq C_*(B_{R_2})$ by zero-extensions. Define

$$\mathcal{J}(u, v) := \frac{\|uv\|_{L^1(\mathbb{R}^d)}}{\|u\|_{L^1(\mathbb{R}^d)}^{1/d} \|u\|_{L^m(\mathbb{R}^d)}^{m/2} \|\nabla v\|_{L^2(\mathbb{R}^d)}} .$$

From the definition of $C_*(B_{R_2})$, for any $\delta > 0$, there exists a pair $(u_\delta, v_\delta) \in \mathcal{Y}(B_{R_2})$ such that

$$\mathcal{J}(u_\delta, v_\delta) \geq C_*(B_{R_2}) - \delta.$$

Taking into account that $\mathcal{J}(u, v)$ is invariant under the operator $\mathcal{T}_{\lambda,\lambda}$ and that $\mathcal{T}_{\lambda,\lambda}(u_\delta, v_\delta) \in \mathcal{Y}(B_{R_1})$ if $\lambda > R_2/R_1$, we have

$$C_*(B_{R_1}) \geq \mathcal{J}(\mathcal{T}_{\lambda,\lambda}(u_\delta, v_\delta)) = \mathcal{J}(u_\delta, v_\delta) \geq C_*(B_{R_2}) - \delta \quad \text{for } \lambda > R_2/R_1.$$

This implies that $C_*(B_{R_1}) \geq C_*(B_{R_2})$ and thus we have $C_*(B_{R_1}) = C_*(B_{R_2}) = C_*(\Omega)$. Therefore, M_* is independent of the choice of Ω. □

Proposition 4.1. *It holds that $C_*(\Omega) = C_*(\mathbb{R}^d)$ for any bounded domain $\Omega \subset \mathbb{R}^d$.*

Please cite this article in press as: Y. Mimura, The variational formulation of the fully parabolic Keller–Segel system with degenerate diffusion, J. Differential Equations (2017), http://dx.doi.org/10.1016/j.jde.2017.03.020

Proof. $C_*(\Omega) \le C_*(\mathbb{R}^d)$ is trivial by definition. We show the opposite relation. By definition, for any $\varepsilon > 0$ there exists a pair $(U_\varepsilon, V_\varepsilon) \in (L^1 \cap L^m)(\mathbb{R}^d) \times \dot{H}(\mathbb{R}^d)$ such that $U_\varepsilon \not\equiv 0$, $V_\varepsilon \not\equiv 0$ and

$$J(U_\varepsilon, V_\epsilon) + \frac{\varepsilon}{2} > C_*(\mathbb{R}^d).$$

On the other hand, since $C_c^\infty(\mathbb{R}^d)$ is dense in $(L^1 \cap L^m)(\mathbb{R}^d)$ and in $\dot{H}^1(\mathbb{R}^d)$, there exists a sequence $\{(u_n, v_n)\} \subset C_c^\infty(\mathbb{R}^d) \times C_c^\infty(\mathbb{R}^d)$ such that

$$u_n \to U_\delta \text{ in } (L^1 \cap L^m)(\mathbb{R}^d) \text{ and } \nabla v_n \to \nabla V_\delta \text{ in } L^2(\mathbb{R}^d).$$

Then if n is sufficiently large we have

$$|J(U_\delta, V_\delta) - J(u_n, v_n)| < \frac{\varepsilon}{2}.$$

In fact, by using the estimate (4.1),

$$\begin{aligned}\left|\int_{\mathbb{R}^d} U_\varepsilon V_\varepsilon \, dx - \int_{\mathbb{R}^d} u_n v_n \, dx\right| &\le \int_{\mathbb{R}^d} (|U_\varepsilon - u_n||V_\varepsilon| + |U_n||V_\varepsilon - v_n|) \, dx \\ &\le C_s \left(\|U_\varepsilon - u_n\|_{L^1}^{1/d} \|U_\varepsilon - u_n\|_{L^m}^{m/2} \|\nabla V_\varepsilon\|_{L^2} + \|U_n\|_{L^1}^{1/d} \|u_n\|_{L^m}^{m/2} \|\nabla V_\varepsilon - \nabla v_n\|_{L^2} \right)\end{aligned}$$

and we can deduce that the numerator of $J(u_n, v_n)$ converges to that of $J(U_\delta, V_\delta)$. It is clear that the denominator of $J(u_n, v_n)$ converges to that of $J(U_\delta, V_\delta)$. Hence $J(u_n, v_n)$ converges to $J(U_\delta, V_\delta)$. Since the support of u_n is bounded we already seen that

$$C_*(\Omega) \ge J(u_n, v_n)$$

for any bounded domain $\Omega \subset \mathbb{R}^d$. Thus we have

$$C_*(\Omega) \ge J(u_n, v_n) \ge C_*(\mathbb{R}^d) - \varepsilon,$$

which implies that $C_*(\Omega) \ge C_*(\mathbb{R}^d)$. $\square$

The next proposition implies that M_* is equal to the sharp critical mass M_c for the parabolic–elliptic case (see Blanchet et al. [14]).

Proposition 4.2 *(Alternative definitions of C_*). Let $\mathcal{K}$ be the fundamental solution of $-\Delta$ in $\mathbb{R}^d$, $d > 2$, that is,*

$$\mathcal{K}(x) := c_d \frac{1}{|x|^{d-2}} \quad \text{with} \quad c_d := \frac{\Gamma\left(\frac{d}{2}\right)}{2(d-2)\pi^{d/2}}.$$

Then it holds that

Please cite this article in press as: Y. Mimura, The variational formulation of the fully parabolic Keller–Segel system with degenerate diffusion, J. Differential Equations (2017), http://dx.doi.org/10.1016/j.jde.2017.03.020

$$C_*^2 = \sup_{u\in(L^1(\mathbb{R}^d)\cap L^m(\mathbb{R}^d))\setminus\{0\}} \left\{ \iint\limits_{\mathbb{R}^d\times\mathbb{R}^d} \left(\frac{\mathcal{K}(x-y)u(x)u(y)}{\|u\|_{L^1(\mathbb{R}^d)}^{2/d}\|u\|_{L^m(\mathbb{R}^d)}^m} \right) dxdy \right\}. \tag{4.9}$$

Proof. Throughout this proof, the norm $\|\cdot\|_{L^p}$ stands for $\|\cdot\|_{L^p(\mathbb{R}^d)}$. From Remark 4.1, we can redefine C_* by

$$C_*^2 := \sup_{\substack{(u,v)\in\mathcal{Y}(\mathbb{R}^d)\\ u,v\neq 0}} \frac{\int_{\mathbb{R}^d} uv\,dx}{\|u\|_{L^1}^{2/d}\|u\|_{L^m}^m\|\nabla v\|_{L^2}^2},$$

where

$$\mathcal{Y}(\mathbb{R}^d) := \left\{(u,v)\in \left(L^1(\mathbb{R}^d)\cap L^m(\mathbb{R}^d)\right)\times \dot{H}^1(\mathbb{R}^d);\, u\geq 0, v\geq 0\right\}.$$

We define the functional $A_u(v)$ by $A_u(v) := \|\nabla v\|_{L^2}^2/(\int_{\mathbb{R}^d} uv\,dx)^2$. Then

$$\begin{aligned}
C_*^2 &= \sup_{\substack{(u,v)\in\mathcal{Y}(\mathbb{R}^d)\\ u,v\neq 0}} \frac{(\int_{\mathbb{R}^d} uv\,dx)^2}{\|u\|_{L^1}^{2/d}\|u\|_{L^m}^m\|\nabla v\|_{L^2}^2} \\
&= \sup_{u\in(L^1(\mathbb{R}^d)\cap L^m(\mathbb{R}^d))\setminus\{0\}} \left\{ \frac{1}{\|u\|_{L^1}^{2/d}\|u\|_{L^m}^m} \sup_{v\in\dot{H}^1(\mathbb{R}^d)\setminus\{0\}} \left(\frac{1}{A_u(v)}\right)\right\} \\
&= \sup_{u\in(L^1(\mathbb{R}^d)\cap L^m(\mathbb{R}^d))\setminus\{0\}} \left(\frac{1}{\|u\|_{L^1}^{2/d}\|u\|_{L^m}^m \inf\limits_{v\in\dot{H}^1(\mathbb{R}^d)\setminus\{0\}} A_u(v)}\right).
\end{aligned}$$

Here one can characterize the function v which attains $\inf\limits_v A_u(v)$ by the Euler–Lagrange equation. Let $(v_j)\subset \dot{H}^1(\mathbb{R}^d)$ be a minimizing sequence of $A_u(v)$. Then $(V_j) := (v_j/\|\nabla v_j\|_{L^2})$ is also a minimizing sequence of $A_u(v)$ since $A_u(V_j) = A_u(v_j)$. So $\|\nabla V_j\|_{L^2} = 1$, there exists a subsequence, still denoted by V_j and V, such that

$$\nabla V_j \rightharpoonup \nabla V \text{ weakly in } L^2(\mathbb{R}^d), \quad V_j \rightharpoonup V \text{ weakly in } L^{\frac{2d}{d-2}}(\mathbb{R}^d).$$

It is clear that $\|\nabla V\|_{L^2} \leq \liminf_{j\to\infty}\|\nabla V_j\|_{L^2}$. By the interpolation inequality, $u\in L^1(\mathbb{R}^d)\cap L^m(\mathbb{R}^d)$ belongs to $L^{\frac{2d}{d+2}}(\mathbb{R}^d)$, the dual space of $L^{\frac{2d}{d-2}}(\mathbb{R}^d)$. Therefore, we obtain

$$\lim_{j\to\infty}\int\limits_{\mathbb{R}^d} uV_j\,dx = \int\limits_{\mathbb{R}^d} uV\,dx$$

and hence $A_u(V)\leq \liminf_{j\to\infty} A_u(V_j)$. This implies the existence of a minimizer of $A_u(v)$. Thanks to the calculus of variation, the function v_* which attains $\min\limits_v A_u(v)$ is given by

Please cite this article in press as: Y. Mimura, The variational formulation of the fully parabolic Keller–Segel system with degenerate diffusion, J. Differential Equations (2017), http://dx.doi.org/10.1016/j.jde.2017.03.020

$$-\Delta v_* = \left(\int_{\mathbb{R}^d} u v_* \, dx \right) A_u(v_*) u.$$

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

$$v_* = \left(\int_{\mathbb{R}^d} u v_* \, dx \right) A_u(v_*) \mathcal{K} * u.$$

Multiplying both sides by u and integrating on $\mathbb{R}^d$, we see that

$$\frac{1}{\inf_v A_u(v)} = \int_{\mathbb{R}^d \times \mathbb{R}^d} \mathcal{K}(x-y) u(x) u(y) \, dx dy.$$

As a consequence, (4.9) holds. $\square$

Proof of Proposition 3.1. In what follows, all the parameters are understood to be normalized ones as in Subsection 3.2. In particular, (3.5) holds. First, we show that v_τ^k is solvable under the assumption that u_τ^{k-1} and v_τ^{k-1} are solvable. Let (v_j) be a minimizing sequence of

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

and I_k be the infimum of the right-hand side. Choosing β to be $\alpha < \beta$ in (4.3), we see that $\|\nabla v_j\|_{L^2}$ is bounded and that I_k has a finite value. Therefore, there exist a subsequence of (v_j), still denoted by (v_j), and a function $v_\infty \in H_0^1(\Omega)$ such that (v_j) converges to v_∞ weakly in $H_0^1(\Omega)$ and hence in $L^{\frac{2d}{d-2}}(\Omega)$. Taking into account that $\int_\Omega u_\tau^{k-1} v_j \, dx$ converges to $\int_\Omega u_\tau^{k-1} v_\infty \, dx$ since $u_\tau^{k-1} \in (L^1(\Omega) \cap L^m(\Omega)) \subset L^{\frac{2d}{d+2}}(\Omega)$, one sees that

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

Since the opposite inequality is obvious, we have

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

Therefore, one can define v_τ^k by v_∞.

Next, we show that u_τ^k is solvable by a similar argument. Let u_j be a minimizing sequence of

$$u \mapsto \phi_m(u, v_\tau^k) + \frac{1}{2\tau} d_W^2(u, u_\tau^{k-1}),$$

and J_k be the infimum of the right-hand side. If $m > 2 - 2/d$, it is clear from (4.6) that $\|u_j\|_{L^m}$ is bounded and J_k has a finite value. In the case $m = 2 - 2/d$, choosing β to be $\beta < 2/(\chi(m-1) C_*^2 \|u\|_{L^1}^{2/d})$ in (4.3), we see that $\|u_j\|_{L^m}$ is bounded and that J_k has a finite value. Therefore, there exist a subsequence, still denoted by (u_j), and a function $u_\infty \in L^m(\Omega)$ such that (u_j) converges to u_∞ weakly in $L^m(\Omega)$. Taking into account the lower semicontinuity of the Wasserstein distance with respect to the narrow convergence (Proposition 2.4) which is weaker than the weak convergence in $L^m(\Omega)$, we see that u_∞ is indeed a minimizer of the above functional. Thus, one can define u_τ^k by u_∞.

In the case $m > 2 - 2/d$, the estimate (3.8) is clear from (4.6). In the case $m = 2 - 2/d$, from (4.4), we see that (3.8) holds if $M_* := M_1 > 1$.

The uniqueness of u_τ^k and v_τ^k follows from the strict convexity of $v \mapsto \phi_m(u_\tau^{k-1}, v)$, $v \mapsto \|v - v_\tau^{k-1}\|_{L^2}^2$ and $u \mapsto \phi_m(u, v_\tau^k)$, and the convexity of $u \mapsto d_W^2(u, u_\tau^{k-1})$ (see for instance [24, Prop. A.1]).

Finally, we prove that for any $\tau > 0$ and $k \in \mathbb{N}$, the functions u_τ^k and v_τ^k are non-negative. The non-negativity of u_τ^k is clear because u_τ^k belongs to $\mathscr{P}_2(\Omega)$. Consequently, it holds that

$$\phi_m(u_\tau^{k-1}, v_\tau^k) \geq \phi_m(u_\tau^{k-1}, |v_\tau^k|). \tag{4.10}$$

We prove $v_\tau^k \geq 0$ by induction in k. For $k = 0$, we have $v_\tau^0 = v_0 \geq 0$ by the assumption. By the triangle inequality, $\big||v_\tau^k| - |v_\tau^{k-1}|\big| \leq |v_\tau^k - v_\tau^{k-1}|$. Therefore, if $v_\tau^{k-1} \geq 0$, then we have

$$\big\||v_\tau^k| - v_\tau^{k-1}\big\|_{L^2} \leq \|v_\tau^k - v_\tau^{k-1}\|_{L^2}. \tag{4.11}$$

Combining (4.10) and (4.11), we obtain

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

Since v_τ^k is the unique minimizer of the functional

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

we have $|v_\tau^k| = v_\tau^k$. Therefore, if the initial data v_0 is non-negative, then v_τ^k is also non-negative for any $\tau > 0$ and $k \in \mathbb{N}$. □

5. Time-discretization method

We saw in the previous section that the discrete solutions $(\overline{u}_\tau, \overline{v}_\tau)$ defined in (3.2) exist for any $\tau > 0$. In this section, we prove that there exists a sequence (τ_n) with $\tau_n \downarrow 0$ such that the sequence $(\overline{u}_{\tau_n}, \overline{v}_{\tau_n})$ of discrete solutions converges to a time-global solution of (1.1), provided that $M_* > 1$.

Here and in what follows the parameters $\alpha, \chi, \varepsilon$ and u, u_0, t are understood to be normalized as in Subsection 3.2. Thus, throughout the rest of the paper, we assume that (3.5) and $M_* > 1$. By Theorem 1.1, Theorem 1.2 and Proposition 3.1, these assumptions induce the lower boundedness of ϕ_m and the estimate (3.8).

5.1. *Some fundamental estimates for discrete solutions*

In this subsection, we establish fundamental estimates for the discrete solutions and show that the discrete solution $(\overline{u}_\tau, \overline{v}_\tau)$ converges weakly in $L^m(\Omega) \times H_0^1(\Omega)$ for some sequence $\tau = \tau_n \downarrow 0$.

Let us recall the definition of the metric slope. The metric slope is the equivalent of the normed value of a gradient of a functional in Hilbert space.

Definition 5.1 *(Metric slopes [28, Def. 1.2.4]).* Define the distances d_1 and d_2 by

$$
\begin{aligned}
d_1(u_1, u_2) &:= d_W(u_1, u_2) \qquad \text{for } u_1, u_2 \in \mathscr{P}_2(\Omega) \subset \mathscr{P}_2(\mathbb{R}^d), \\
d_2(v_1, v_2) &:= \sqrt{\frac{\varepsilon\chi}{\alpha}} \|v_1 - v_2\|_{L^2} \qquad \text{for } v_1, v_2 \in L^2(\Omega).
\end{aligned}
$$

The *metric slopes* $|\partial_1\phi_m|(u, v)$ and $|\partial_2\phi_m|(u, v)$ of ϕ_m at $(u, v) \in \mathcal{D}(\phi_m)$ is defined by

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

where $\mathcal{D}(\phi_m)$ denotes the *effective domain* of ϕ_m, that is,

$$
\mathcal{D}(\phi_m) := \{(u, v)\,;\, \phi_m(u, v) < +\infty\}.
$$

We also define $\mathcal{D}(|\partial_1\phi_m|)$ and $\mathcal{D}(|\partial_2\phi_m|)$ likewise.

Lemma 5.2 *(Fundamental estimates [28, Lem. 3.1.3, Lem. 3.2.2], [32, §8.4.1]). Our recursive scheme (3.7) yields the following estimates:*

(i) Slope estimate: *for* $k = 1, 2, 3, \ldots$

$$
\begin{cases}
|\partial_1\phi_m|(u_\tau^k, v_\tau^k) \le \dfrac{d_1(u_\tau^k, u_\tau^{k-1})}{\tau}, \\[2ex]
|\partial_2\phi_m|(u_\tau^{k-1}, v_\tau^k) \le \dfrac{d_2(v_\tau^k, v_\tau^{k-1})}{\tau}.
\end{cases}
\tag{5.1}
$$

(ii) Energy estimate:

$$
\sup_{k\ge 0} \phi_m(u_\tau^k, v_\tau^k) \le \phi_m(u_0, v_0).
$$

Please cite this article in press as: Y. Mimura, The variational formulation of the fully parabolic Keller–Segel system with degenerate diffusion, J. Differential Equations (2017), http://dx.doi.org/10.1016/j.jde.2017.03.020

(iii) Total square distance estimate: *for any* $N \in \mathbb{N}$,

$$\sum_{k=1}^{N} \Big[d_1^2(u_\tau^k, u_\tau^{k-1}) + d_2^2(v_\tau^k, v_\tau^{k-1}) \Big] \le 2\tau(\phi_m(u_0, v_0) - \inf \phi_m).$$

(iv) Gradient energy estimate: *for any* $T > 0$,

$$\int_0^T |\partial_1 \phi_m|^2(\overline{u}_\tau(t), \overline{v}_\tau(t))\, dt + \int_0^T |\partial_2 \phi_m|^2(\underline{u}_\tau(t), \overline{v}_\tau(t))\, dt$$
$$\le 2(\phi_m(\overline{u}_0, \overline{v}_0) - \phi_m(\overline{u}_\tau(T), \overline{v}_\tau(T))),$$

where $(\overline{u}_\tau, \overline{v}_\tau)$ *is a discrete solution defined in Definition 3.2 and* $\underline{u}_\tau(t) := u_\tau^{k-1}$ *for* $t \in ((k-1)\tau, k\tau]$, *i.e.,* $\underline{u}_\tau(t) = \overline{u}_\tau(t - \tau)$.

Proof. We first show the estimate (i). Let $(u_*, v_*) \in \mathcal{D}(\phi_m)$ be given. If $u_\tau \in \operatorname*{argmin}_{u \in \mathscr{P}_2(\Omega)} \Big\{ \phi_m(u, v_*) + \dfrac{1}{2\tau} d_1^2(u, u_*) \Big\}$, then it follows that

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

for all $u \in \mathscr{P}_2(\Omega)$, and then

$$\begin{aligned} \phi_m(u_\tau, v_*) - \phi_m(u, v_*) &\le \frac{1}{2\tau} d_1^2(u, u_*) - \frac{1}{2\tau} d_1^2(u_\tau, u_*) \\ &= \frac{1}{2\tau} \big(d_1(u, u_*) - d_1(u_\tau, u_*) \big) \big(d_1(u, u_*) + d_1(u_\tau, u_*) \big) \\ &\le \frac{1}{2\tau} d_1(u_\tau, u) \big(d_1(u, u_*) + d_1(u_\tau, u_*) \big). \end{aligned}$$

Hence we have

$$\begin{aligned} |\partial_1 \phi_m|(u_\tau, v_*) &= \limsup_{u \to u_\tau} \frac{\big(\phi_m(u_\tau, v_*) - \phi_m(u, v_*) \big)^+}{d_1(u_\tau, u)} \\ &\le \limsup_{u \to u_\tau} \frac{1}{2\tau} \big(d_1(u, u_*) + d_1(u_\tau, u_*) \big) = \frac{d_1(u_\tau, u_*)}{\tau}. \end{aligned} \tag{5.2}$$

Similarly, for $v_\tau \in \operatorname*{argmin}_{v \in H_0^1(\Omega)} \Big\{ \phi_m(u_*, v) + \dfrac{1}{2\tau} d_2^2(v, v_*) \Big\}$, we obtain

$$|\partial_2 \phi_m|(u_*, v_\tau) \le \frac{d_2(v_\tau, v_*)}{\tau}. \tag{5.3}$$

Therefore, by the definitions of u_τ^k and v_τ^k, the estimates (5.1) hold.

Next, from the definition of u_τ^k and v_τ^k, it is obvious that

$$\begin{cases} \phi_m(u_\tau^k, v_\tau^k) + \dfrac{1}{2\tau} d_1^2(u_\tau^k, u_\tau^{k-1}) \le \phi_m(u_\tau^{k-1}, v_\tau^k), \\ \phi_m(u_\tau^{k-1}, v_\tau^k) + \dfrac{1}{2\tau} d_2^2(v_\tau^k, v_\tau^{k-1}) \le \phi_m(u_\tau^{k-1}, v_\tau^{k-1}). \end{cases}$$

Combining both inequalities, we obtain

$$\phi_m(u_\tau^k, v_\tau^k) + \frac{1}{2\tau}\left[d_1^2(u_\tau^k, u_\tau^{k-1}) + d_2^2(v_\tau^k, v_\tau^{k-1})\right] \le \phi_m(u_\tau^{k-1}, v_\tau^{k-1}).$$

The statements (ii) and (iii) follow from this inequality.

Finally, from (i) and (iii), we have

$$\begin{aligned} &\int_0^T |\partial_1 \phi_m|^2(\overline{u}_\tau, \overline{v}_\tau)\, dt + \int_0^T |\partial_2 \phi_m|^2(\underline{u}_\tau, \overline{v}_\tau)\, dt \\ &\qquad \le \tau \sum_{k=1}^N \left(|\partial_1 \phi_m|^2(u_\tau^k, v_\tau^k) + |\partial_2 \phi_m|^2(u_\tau^{k-1}, v_\tau^k)\right) \\ &\qquad \le \frac{1}{\tau} \sum_{k=1}^N \left(d_1^2(u_\tau^k, u_\tau^{k-1}) + d_2^2(v_\tau^k, v_\tau^{k-1})\right) \\ &\qquad \le 2(\phi_m(u_0, v_0) - \phi_m(\overline{u}_\tau(T), \overline{v}_\tau(T))). \qquad \square \end{aligned}$$

Lemma 5.3 *(Convergence of discrete solutions [28, Prop. 3.3.1], [32, §8.4.1]). There exist a sequence $(\overline{u}_{\tau_n}, \overline{v}_{\tau_n})$ of discrete solutions with $\tau_n \downarrow 0$ and a function $(u, v) \in L^m(\Omega) \times H_0^1(\Omega)$ such that*

$$\begin{aligned} &\overline{u}_{\tau_n}(t) \rightharpoonup u(t) \text{ weakly in } L^m(\Omega), \quad \forall t \ge 0, \\ &\overline{v}_{\tau_n}(t) \rightharpoonup v(t) \text{ weakly in } H_0^1(\Omega), \quad \forall t \ge 0. \end{aligned} \tag{5.4}$$

Proof. Let $\overline{w}_\tau := (\overline{u}_\tau, \overline{v}_\tau)$ and define the distance d by $d^2 := d_1^2 + d_2^2$. First, we show that $d(\overline{w}_\tau(t), \overline{w}_\tau(s)) \le C\sqrt{|t-s|}$. Without loss of generality, we can assume $s < t$. There exist $\ell_1, \ell_2 \in \mathbb{N}$ such that

$$(\ell_1 - 1)\tau < s \le \ell_1 \tau, \quad (\ell_2 - 1)\tau < t \le \ell_2 \tau.$$

By the triangle inequality, the Cauchy–Schwarz inequality and Lemma 5.2-(iii), we have

$$\begin{aligned} d(\overline{w}_\tau(t), \overline{w}_\tau(s)) &\le \sum_{k=\ell_1+1}^{\ell_2} d(w_\tau^k, w_\tau^{k-1}) \le \sqrt{t-s}\left(\sum_{\ell_1+1}^{\ell_2} \frac{d^2(w_\tau^k, w_\tau^{k-1})}{\tau}\right)^{1/2} \\ &\le \sqrt{2(\phi_m(u_0, v_0) - \inf \phi_m)(t-s)}. \end{aligned} \tag{5.5}$$

Please cite this article in press as: Y. Mimura, The variational formulation of the fully parabolic Keller–Segel system with degenerate diffusion, J. Differential Equations (2017), http://dx.doi.org/10.1016/j.jde.2017.03.020

From the estimate (3.8), for each fixed t, $\{\overline{w}_\tau(t)\}_{\tau>0}$ is relatively weakly compact in $L^m(\Omega) \times H_0^1(\Omega)$. Hence by [28, Prop. 3.3.1], our assertion holds true. □

5.2. Regularity of discrete solutions

The aim of this subsection is to derive the Euler–Lagrange equations satisfied by discrete solutions at each time step. By our minimizing scheme (3.7), the discrete solutions possess a certain regularity, which enables us to derive the Euler–Lagrange equations.

First, given $\boldsymbol{\xi} \in C_c^\infty(\mathbb{R}^d; \mathbb{R}^d)$, we calculate the derivative of ϕ_m with respect to the perturbation $(\boldsymbol{id} + t\boldsymbol{\xi})_\# u$ as $t \to 0$. Similar perturbations can be found in [26] and [28, §10.4.1]. Since the range of $(\boldsymbol{id} + t\boldsymbol{\xi})$ lies on $\mathbb{R}^d$ rather than Ω, this perturbation requires the extension of the domain of the functions u and v. Therefore, we extend the domain of u and v from Ω to $\mathbb{R}^d$. More precisely, we extend u by setting $u = 0$ outside Ω. As for v, we use a standard extension operator from $W^{2,2}(\Omega)$ to $W^{2,2}(\mathbb{R}^d)$ (see for instance [40, Thm. 4.26]).

Lemma 5.4. *Let $1 \le p < \infty$ and $\boldsymbol{\xi} \in C_c^\infty(\mathbb{R}^d; \mathbb{R}^d)$. If v_n converges to v in $L^p(\mathbb{R}^d)$, then for $\delta > 0$ small enough and for every $t \in [0, \delta]$, $v_n(\boldsymbol{id} + t\boldsymbol{\xi})$ converges to $v(\boldsymbol{id} + t\boldsymbol{\xi})$ in $L^p(\mathbb{R}^d)$. In addition, it holds that*

$$\|v_n(\boldsymbol{id} + t\boldsymbol{\xi}) - v(\boldsymbol{id} + t\boldsymbol{\xi})\|_{L^p(\mathbb{R}^d)} \le C_\delta \|v_n - v\|_{L^p(\mathbb{R}^d)} \quad t \in [0, \delta],$$

where C_δ is a positive constant depending on δ and $\boldsymbol{\xi}$.

Proof. Let $\boldsymbol{r}_t(x) := x + t\boldsymbol{\xi}(x)$. Note that for t small enough, $\boldsymbol{r}_t$ is a C^1 diffeomorphism and $\det D\boldsymbol{r}_t > 0$ since $\boldsymbol{\xi} \in C_c^\infty(\mathbb{R}^d; \mathbb{R}^d)$. By the change of variables $y = \boldsymbol{r}_t(x)$, we have

$$\begin{aligned}\int_{\mathbb{R}^d} |v_n(\boldsymbol{r}_t(x)) - v(\boldsymbol{r}_t(x))|^p \, dx &= \int_{\mathbb{R}^d} |v_n(y) - v(y)|^p \det(D\boldsymbol{r}_t^{-1}(y)) \, dy \\ &\le \sup_{(y,t)\in\mathbb{R}^d\times[0,\delta]} \Big(\det(D\boldsymbol{r}_t^{-1}(y))\Big) \|v_n - v\|_{L^p(\mathbb{R}^d)}^p. \quad \square\end{aligned}$$

Lemma 5.5 *(Gâteaux derivatives). If $(u, v) \in \mathcal{D}(|\partial_1\phi_m|)$ and $v \in W^{2,2}(\Omega)$, then $\nabla u^m \in L^1(\Omega)$ and for any $\boldsymbol{\xi} \in C_c^\infty(\mathbb{R}^d; \mathbb{R}^d)$ the function $t \mapsto \phi_m((\boldsymbol{id} + t\boldsymbol{\xi})_\# u, v)$ is differentiable at $t = 0$ with*

$$\begin{cases} \dfrac{d}{dt}\Big[\phi_m((\boldsymbol{id} + t\boldsymbol{\xi})_\# u, v)\Big]\Big|_{t=0} = \displaystyle\int_\Omega \langle \nabla u^m - \chi u \nabla v, \boldsymbol{\xi}\rangle \, dx, \\ \displaystyle\int_\Omega \frac{|\nabla u^m - \chi u \nabla v|^2}{u} \, dx \le |\partial_1\phi_m|^2(u, v). \end{cases} \tag{5.6}$$

On the other hand, if $(u, v) \in \mathcal{D}(|\partial_2\phi_m|)$ and $u \in L^2(\Omega)$, then $\Delta v \in L^2(\Omega)$ and for any $\eta \in C_c^\infty(\Omega)$ the function $t \mapsto \phi_m(u, v + t\eta)$ is differentiable at $t = 0$ with

$$\begin{cases} \dfrac{d}{dt}\big[\phi_m(u, v+t\eta)\big]\Big|_{t=0} = \dfrac{\chi}{\alpha}\displaystyle\int_\Omega (-\Delta v + \gamma v - \alpha u)\eta\,dx, \\ \dfrac{\chi}{\alpha\varepsilon}\displaystyle\int_\Omega |\Delta v - \gamma v + \alpha u|^2\,dx \le |\partial_2\phi_m|^2(u,v). \end{cases} \tag{5.7}$$

Remark 5.1. The inequality in (5.6) means

$$\int_\Omega |g_1|^2 u\,dx \le |\partial_1\phi_m|^2(u,v)$$

with

$$g_1 = \frac{\nabla u^m - \chi u \nabla v}{u} \text{ on the set } \{x \in \Omega : u > 0\}.$$

Proof of Lemma 5.5. Let us suppose $(u,v) \in \mathcal{D}(|\partial_1\phi_m|)$ and $v \in W^{2,2}(\Omega)$. Define

$$\mathcal{F}[u] := \frac{1}{m-1}\int_\Omega u^m\,dx, \quad I(u,v) := -\chi\int_\Omega uv\,dx.$$

Let $\boldsymbol{\xi} \in C_c^\infty(\mathbb{R}^d;\mathbb{R}^d)$ and $u_t := (\boldsymbol{id} + t\boldsymbol{\xi})_\# u$. By the definition of the push-forward and the change of variables, one easily obtains the relation:

$$u_t(x + t\boldsymbol{\xi}(x))\det(\boldsymbol{id} + tD\boldsymbol{\xi}(x)) = u(x).$$

Using this relation and taking into account that $(\det(\boldsymbol{id} + tD\boldsymbol{\xi}(x)) - 1)/t$ uniformly converges to $\operatorname{div}\boldsymbol{\xi}$ in Ω, we have

$$\begin{aligned} \lim_{t\to 0}\frac{\mathcal{F}[u_t] - \mathcal{F}[u]}{t} &= \lim_{t\to 0}\frac{1}{(m-1)t}\int_\Omega \left(\frac{1}{(\det(\boldsymbol{id} + tD\boldsymbol{\xi}(x)))^{m-1}} - 1\right)u^m(x)\,dx \\ &= -\int_\Omega u^m(x)\operatorname{div}\boldsymbol{\xi}\,dx. \end{aligned}$$

Next we calculate

$$\lim_{t\to 0}\frac{I(u_t,v) - I(u,v)}{t}.$$

Recall that $v \in W^{2,2}(\Omega) \hookrightarrow W^{2,2}(\mathbb{R}^d)$ by the extension operator (see for instance [40, Thm. 4.26]). Let $(v_n)_{n\in\mathbb{N}} \subset C_c^\infty(\mathbb{R}^d)$ be a sequence converging to v in $W^{2,2}(\mathbb{R}^d)$. We define the following two functions:

$$I(t) := -\chi\int_\Omega v(x + t\boldsymbol{\xi}(x))u(x)\,dx, \quad I_n(t) := -\chi\int_\Omega v_n(x + t\boldsymbol{\xi}(x))u(x)\,dx.$$

Please cite this article in press as: Y. Mimura, The variational formulation of the fully parabolic Keller–Segel system with degenerate diffusion, J. Differential Equations (2017), http://dx.doi.org/10.1016/j.jde.2017.03.020

By Lemma 5.4, $I_n(t)$ converges to $I(t)$ uniformly in $[0,\delta]$ and $I_n'(t)$ converges to

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

Here, note that $\tilde{I}'(t)$ is not defined if (u,v) is only assumed to be in the effective domain of ϕ_m, that is, $u \in L^m(\Omega)$ and $v \in H^1(\Omega)$ and this is the reason for requiring more regularity. Therefore, $I(t)$ is differentiable in $(0,\delta)$ and it holds that

$$\lim_{t\to 0} \frac{I(u_t,v) - I(u,v)}{t} = I'(0) = -\chi \int_\Omega \langle \nabla v(x), \boldsymbol{\xi}\rangle u(x)\,dx.$$

On the other hand, since $u_t = (\boldsymbol{id} + t\boldsymbol{\xi})_\# u$, Brenier's theorem (Proposition 2.5) gives

$$d_1^2(u,u_t) := \inf_{\{\boldsymbol{t}: u_t = \boldsymbol{t}_\# u\}} \int_{\mathbb{R}^d} |x - \boldsymbol{t}(x)|^2 u\,dx \le t^2 \int_{\mathbb{R}^d} |\boldsymbol{\xi}|^2 u dx,$$

and hence

$$\limsup_{t\to 0} \frac{d_1(u,u_t)}{t} \le \|\boldsymbol{\xi}\|_{L^2(u)}.$$

Combining all the relations above and recalling Definition 5.1, we obtain

$$\begin{aligned}
\int_\Omega u^m \operatorname{div}\boldsymbol{\xi} + \chi\langle \nabla v, \boldsymbol{\xi}\rangle u\,dx &= \lim_{t\to 0} \frac{\phi_m(u,v) - \phi_m(u_t,v)}{t} \\
&\le \limsup_{t\to 0} \frac{(\phi_m(u,v) - \phi_m(u_t,v))^+}{d_1(u,u_t)} \frac{d_1(u,u_t)}{t} \\
&\le |\partial_1 \phi_m|(u,v)\|\boldsymbol{\xi}\|_{L^2(u)} \\
&\le |\partial_1 \phi_m|(u,v)\|\boldsymbol{\xi}\|_{L^\infty}.
\end{aligned}$$

Since $u \in (L^1(\Omega) \cap L^m(\Omega)) \subset L^{\frac{2d}{d+2}}(\Omega)$ and $v \in W^{2,2}(\Omega)$, the Sobolev embedding theorem ensures that $u|\nabla v| \in L^1(\Omega)$. This and the above estimate imply

$$\int_\Omega u^m \operatorname{div}\boldsymbol{\xi}\,dx \le \left(|\partial_1\phi_m|(u,v) + \chi\|u\nabla v\|_{L^1}\right)\|\boldsymbol{\xi}\|_{L^\infty}.$$

Hence, by the Riesz representation theorem, there exists an $\mathbb{R}^d$-valued measure $\boldsymbol{\mu}^* = (\mu_1^*, \mu_2^*, \ldots, \mu_d^*)$ such that

$$\int_\Omega u^m \operatorname{div}\boldsymbol{\xi}\,dx = -\sum_{j=1}^d \int_\Omega \xi_j\,d\mu_j^*, \tag{5.8}$$

Please cite this article in press as: Y. Mimura, The variational formulation of the fully parabolic Keller–Segel system with degenerate diffusion, J. Differential Equations (2017), http://dx.doi.org/10.1016/j.jde.2017.03.020

where $\boldsymbol{\xi} = (\xi_1, \xi_2, \dots, \xi_d)$. Thus the above estimate can be rewritten as

$$-\sum_{j=1}^{d} \int_{\Omega} \xi_j \, d\mu_j^* + \chi \int_{\Omega} \langle \nabla v, \boldsymbol{\xi} \rangle u \, dx \le |\partial_1 \phi_m|(u, v) \|\boldsymbol{\xi}\|_{L^2(u)}.$$

Again, by applying the Riesz representation theorem, there exists an $\mathbb{R}^d$-valued function $g_1 \in L^2_u(\Omega, \mathbb{R}^d)$ such that $\|g_1\|_{L^2(u)} \le |\partial_1 \phi_m|(u, v)$ and that

$$-\sum_{j=1}^{d} \int_{\Omega} \xi_j \, d\mu_j^* + \chi \int_{\Omega} \langle \nabla v, \boldsymbol{\xi} \rangle u \, dx = -\int_{\Omega} \langle g_1, \boldsymbol{\xi} \rangle u \, dx. \tag{5.9}$$

Combining (5.8) and (5.9) yields

$$\int_{\Omega} u^m \operatorname{div} \boldsymbol{\xi} \, dx = -\sum_{j=1}^{d} \int_{\Omega} \xi_j \, d\mu_j^* = \int_{\Omega} \langle -u g_1 - \chi u \nabla v, \boldsymbol{\xi} \rangle \, dx.$$

Here we note that $u g_1 + \chi u \nabla v \in L^1(\Omega)$. Indeed, since $\|g_1\|_{L^2(u)} \le |\partial_1 \phi_m|(u, v)$, one has

$$\|u g_1\|_{L^1} \le \|u\|_{L^1}^{1/2} \|g_1\|_{L^2(u)} \le |\partial_1 \phi_m|(u, v) < +\infty,$$

and we have already seen that $u|\nabla v| \in L^1(\Omega)$, which shows $u g_1 + \chi u \nabla v \in L^1(\Omega)$. Thus the above identity implies

$$u g_1 = \nabla u^m - \chi u \nabla v$$

in the sense of distributions, which gives (5.6).

Next we prove (5.7). Let $\eta \in C_c^\infty(\Omega)$ and $v_t := v + t\eta$. Then, by the definitions of d_2 and the metric slope $|\partial_2 \phi_m|$ (see Definition 5.1),

$$\begin{aligned}
\frac{\chi}{\alpha} \int_{\Omega} (\alpha u - \gamma v)\eta - \langle \nabla v, \nabla \eta \rangle \, dx &= \lim_{t \to 0} \frac{\phi_m(u, v) - \phi_m(u, v_t)}{t} \\
&\le \limsup_{t \to 0} \frac{(\phi_m(u, v) - \phi_m(u, v_t))^+}{d_2(v, v_t)} \frac{d_2(v, v_t)}{t} \\
&\le \sqrt{\frac{\varepsilon \chi}{\alpha}} |\partial_2 \phi_m|(u, v) \|\eta\|_{L^2}.
\end{aligned}$$

As we are assuming $(u, v) \in \mathcal{D}(|\partial_2 \phi_m|)$, we have $|\partial_2 \phi_m|(u, v) < \infty$. Furthermore, since $u \in L^2(\Omega)$, the above inequality yields

$$\int_{\Omega} \langle \nabla v, \nabla \eta \rangle \, dx \le C \|\eta\|_{L^2}$$

for some constant $C > 0$, which implies $\Delta v \in L^2(\Omega)$. Consequently, we can rewrite the above inequality as follows:

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

from which we obtain

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

The lemma is proved. $\square$

Lemma 5.6. *If $(u, v) \in \mathcal{D}(|\partial_1\phi_m|)$ and $v \in W^{2,2}(\Omega)$, then u belongs to $L^2(\Omega)$. Conversely, if $(u, v) \in \mathcal{D}(|\partial_2\phi_m|)$ and $u \in L^2(\Omega)$, then v belongs to $W^{2,2}(\Omega)$.*

Proof. By Lemma 5.5 and the Cauchy–Schwarz inequality, we have

$$\begin{aligned}\int_\Omega |\nabla u^m - \chi u\nabla v|\,dx &\le \left(\int_\Omega \frac{|\nabla u^m - \chi u\nabla v|^2}{u}\,dx\right)^{1/2}\left(\int_\Omega u(x)\,dx\right)^{1/2}\\ &\le |\partial_1\phi_m|(u, v).\end{aligned}$$

Since $v \in W^{2,2}(\Omega)$ ensures $u|\nabla v| \in L^1(\Omega)$, we have

$$\|\nabla u^m\|_{L^1} \le |\partial_1\phi_m|(u, v) + \chi\|u\nabla v\|_{L^1}. \tag{5.10}$$

On the other hand, by the interpolation inequality, for $\theta = \frac{d}{2}(2 - m) \in [0, 1]$ and $p = \frac{m}{\theta} > 1$, it holds that

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

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

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

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

$$\|u\|_{L^2}^p \le C_1\|u\|_{L^m}^{p-m}\left(\|u\|_{L^m}^m + |\partial_1\phi_m|(u, v) + \chi\|u\nabla v\|_{L^1}\right). \tag{5.11}$$

The second assertion follows from the estimate (5.7) and the L^2-estimate $\|v\|_{W^{2,2}} \le C\|\Delta v\|_{L^2}$ for $v \in W^{2,2}(\Omega) \cap H_0^1(\Omega)$. $\square$

Please cite this article in press as: Y. Mimura, The variational formulation of the fully parabolic Keller–Segel system with degenerate diffusion, J. Differential Equations (2017), http://dx.doi.org/10.1016/j.jde.2017.03.020

Corollary 5.7 *(Regularity of discrete solutions). If $u_0 \in L^2(\Omega)$ and $(u_0, v_0) \in \mathcal{D}(\phi_m)$, then it holds that*

$$\overline{u}_\tau(t) \in L^2(\Omega) \qquad \forall t > 0,$$
$$\overline{v}_\tau(t) \in W^{2,2}(\Omega) \quad \forall t > 0.$$

Proof. From Lemma 5.2-(i), $(u_0, v_\tau^1) \in \mathcal{D}(|\partial_2\phi_m|)$. From the assumption $u_0 \in L^2(\Omega)$ and Lemma 5.6, $v_\tau^1 \in W^{2,2}(\Omega)$. Again, by Lemma 5.2-(i) and Lemma 5.6, $(u_\tau^1, v_\tau^1) \in \mathcal{D}(|\partial_1\phi_m|)$ 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 $t > 0$. □

The following lemma is fundamental in the minimizing scheme with the Wasserstein distance. One can find the proof in the section 5 of [26].

Lemma 5.8 *(Gâteaux derivative of the Wasserstein distance). Let $\mu := u\mathscr{L}^d \in \mathscr{P}_2(\Omega)$ and $\mu_* := u_*\mathscr{L}^d \in \mathscr{P}_2(\Omega)$. Then for any $\boldsymbol{\xi} \in C_c^\infty(\mathbb{R}^d; \mathbb{R}^d)$, the function $t \mapsto d_1^2((\boldsymbol{id} + t\boldsymbol{\xi})_\# u, u_*)$ is differentiable at $t = 0$ and*

$$\frac{d}{dt} d_1^2((\boldsymbol{id} + t\boldsymbol{\xi})_\# u, u_*)\Big|_{t=0} = \int\limits_{\mathbb{R}^d \times \mathbb{R}^d} \langle x - y, \boldsymbol{\xi}(y)\rangle \, dp_*(x, y)$$

where p_* is an optimal transport map from u to u_*.

Lemma 5.9 *(Euler–Lagrange equations). Let $\{(u_\tau^k, v_\tau^k)\}_{k=0}^\infty$ be a solution of the variational scheme (3.7). Then for all $\varphi \in C_c^\infty(\mathbb{R}^d)$ and $\psi \in C_c^\infty(\Omega)$, the function (u_τ^k, v_τ^k) satisfies the following relations:*

$$\begin{cases} \displaystyle\int\limits_\Omega (u_\tau^k - u_\tau^{k-1})\varphi \, dx + \tau \int\limits_\Omega \langle \nabla (u_\tau^k)^m - \chi u_\tau^k \nabla v_\tau^k, \nabla\varphi\rangle dx \\ \qquad\qquad\qquad\qquad\qquad\qquad = O(d_W^2(u_\tau^k, u_\tau^{k-1})), \\ \displaystyle\varepsilon \int\limits_\Omega (v_\tau^k - v_\tau^{k-1})\psi \, dx - \tau \int\limits_\Omega (\Delta v_\tau^k - \gamma v_\tau^k + \alpha u_\tau^{k-1})\psi \, dx = 0. \end{cases} \tag{5.12}$$

Proof. Let $\boldsymbol{\xi} \in C_c^\infty(\Omega; \mathbb{R}^d)$ and define $U_t := (\boldsymbol{id} + t\boldsymbol{\xi})_\# u_\tau^k$. Since u_τ^k is a minimizer of

$$u \in \mathscr{P}_2(\Omega) \mapsto \left\{ \phi_m(u, v_\tau^k) + \frac{1}{2\tau} d_1^2(u, u_\tau^{k-1}) \right\},$$

we have

$$\phi_m(U_t, v_\tau^k) - \phi_m(u_\tau^k, v_\tau^k) + \frac{1}{2\tau}\left(d_1^2(U_t, u_\tau^{k-1}) - d_1^2(u_\tau^k, v_\tau^k) \right) \ge 0. \tag{5.13}$$

Dividing both sides by $t > 0$ and passing to the limit as $t \downarrow 0$, we obtain

Please cite this article in press as: Y. Mimura, The variational formulation of the fully parabolic Keller–Segel system with degenerate diffusion, J. Differential Equations (2017), http://dx.doi.org/10.1016/j.jde.2017.03.020

$$\frac{d}{dt}\big[\phi_m(U_t, v_\tau^k)\big]\Big|_{t=0} + \frac{1}{2\tau}\left(\limsup_{t\downarrow 0}\frac{d_1^2(U_t, u_\tau^{k-1}) - d_1^2(u_\tau^k, v_\tau^k)}{t}\right) \geq 0.$$

Consequently, we deduce from (5.6) in Lemma 5.5 (with $u = u_\tau^k$, $v = v_\tau^k$) and Lemma 5.8 (with $u = u_\tau^k$, $u_* = u_\tau^{k-1}$) that

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} \langle y - x, \boldsymbol{\xi}(y)\rangle dp_k(x, y) + \tau \int_\Omega \langle \nabla(u_\tau^k)^m - \chi u_\tau^k \nabla v_\tau^k, \boldsymbol{\xi}\rangle dx \geq 0.$$

Similarly, dividing (5.13) by $t < 0$ and passing to the limit as $t \uparrow 0$, we obtain the opposite inequality, hence

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} \langle y - x, \boldsymbol{\xi}(y)\rangle dp_k(x, y) + \tau \int_\Omega \langle \nabla(u_\tau^k)^m - \chi u_\tau^k \nabla v_\tau^k, \boldsymbol{\xi}\rangle dx = 0 \tag{5.14}$$

for $\boldsymbol{\xi} \in C_c^\infty(\Omega; \mathbb{R}^d)$. By Brenier's theorem, there exists the optimal map $\boldsymbol{t}_k : \mathbb{R}^d \to \mathbb{R}^d$ such that

$$\iint_{\mathbb{R}^d\times\mathbb{R}^d} \langle y - x, \boldsymbol{\xi}(y)\rangle dp_k(x, y) = \int_{\mathbb{R}^d} \langle y - \boldsymbol{t}_k(y), \boldsymbol{\xi}(y)\rangle u_\tau^k(y)\, dy.$$

Therefore, the relation (5.14) leads to

$$\frac{(\boldsymbol{id} - \boldsymbol{t}_k)}{\tau} u_\tau^k + \nabla(u_\tau^k)^m - u_\tau^k \nabla v_\tau^k = 0 \quad \text{a.e. on } \Omega.$$

Taking the inner product of both sides with $\nabla\varphi$, $\varphi \in C_c^\infty(\mathbb{R}^d)$, and integrating over Ω, one sees that

$$\int_\Omega \langle x - \boldsymbol{t}_k(x), \nabla\varphi(x)\rangle u_\tau^k(x)\, dx + \tau \int_\Omega \langle \nabla(u_\tau^k)^m - \chi u_\tau^k \nabla v_\tau^k, \nabla\varphi\rangle\, dx = 0.$$

Here, by Taylor's expansion $\varphi(\boldsymbol{t}_k(x)) - \varphi(x) = \langle \nabla\varphi, \boldsymbol{t}_k - x\rangle + O(|\boldsymbol{t}_k(x) - x|^2)$,

$$\begin{aligned}
&\int_\Omega \langle x - \boldsymbol{t}_k(x), \nabla\varphi\rangle u_\tau^k(x)\, dx \\
&\qquad = \int_\Omega (\varphi(x) - \varphi(\boldsymbol{t}_k(x)) u_\tau^k(x)\, dx + \int_\Omega O(|x - \boldsymbol{t}_k(x)|^2) u_\tau^k\, dx \\
&\qquad\qquad = \int_\Omega (u_\tau^k - u_\tau^{k-1})\varphi\, dx + O(d_W^2(u_\tau^k, u_\tau^{k-1})).
\end{aligned}$$

Therefore, we obtain the first assertion of (5.12).

Considering the perturbation $v_\tau^k + t\psi$ and the minimality of v_τ^k, the second assertion of (5.12) easily follows from (5.7). □

Please cite this article in press as: Y. Mimura, The variational formulation of the fully parabolic Keller–Segel system with degenerate diffusion, J. Differential Equations (2017), http://dx.doi.org/10.1016/j.jde.2017.03.020

5.3. Convergence to a solution of (1.1)

In this subsection, we prove the convergence result of Theorem 3.3. What we will show is that the limit function (u, v) in (5.4) is a weak solution of (1.1). Furthermore, we prove that convergence in (5.4) takes place in a stronger topology. We start by establishing the following uniform estimate.

Lemma 5.10 *(Uniform estimates). For $p = 2m/(d(2-m))$ and any $T > 0$,*

$$\sup_{\tau>0} \int_0^T \|\overline{u}_\tau(t)\|_{L^2}^{2p}\, dt < +\infty \tag{5.15}$$

holds. In addition,

$$\sup_{\tau>0} \int_0^T \|\nabla \overline{u}_\tau^m\|_{L^1}^2\, dt < +\infty, \quad \sup_{\tau>0} \int_0^T \|\Delta \overline{v}_\tau\|_{L^2}^2\, dt < +\infty$$

hold.

Proof. From the estimate (5.11) and Corollary 5.7, we have

$$\|\overline{u}_\tau\|_{L^2}^p \le C_1 \|u\|_{L^m}^{p-m} \left(\|\overline{u}_\tau\|_{L^m}^m + |\partial_1 \phi_m|(\overline{u}_\tau, \overline{v}_\tau) + \chi \|\nabla \overline{v}_\tau\|_{L^2} \|\overline{u}_\tau\|_{L^2} \right),$$

where $p = 2m/(d(2-m)) \ge m$. Since $p > 1$, by Young's inequality,

$$C_1 \|u\|_{L^m}^{p-m} \chi \|\nabla \overline{v}_\tau\|_{L^2} \|\overline{u}_\tau\|_{L^2} \le \frac{1}{p} \|\overline{u}_\tau\|_{L^2}^p + \frac{p-1}{p} (C_1 \chi \|u\|_{L^m}^{p-m} \|\nabla \overline{v}_\tau\|_{L^2})^{p/(p-1)},$$

and we deduce that

$$\|\overline{u}_\tau\|_{L^2}^p \le \frac{C_1 p \|u\|_{L^m}^{p-m}}{p-1} \left(\|\overline{u}_\tau\|_{L^m}^m + |\partial_1 \phi_m|(\overline{u}_\tau, \overline{v}_\tau) \right) + (C_1 \chi \|u\|_{L^m}^{p-m} \|\nabla \overline{v}_\tau\|_{L^2})^{p/(p-1)}. \tag{5.16}$$

Since

$$\|\overline{u}_\tau\|_{L^m}, \quad \|\nabla \overline{v}_\tau\|_{L^2}, \quad \int_0^T |\partial_1 \phi_m|^2 (\overline{u}_\tau(t), \overline{v}_\tau(t))\, dt$$

are uniformly bounded from (3.8) and Lemma 5.2-(iv), integrating the square of (5.16) over $(0, T)$, one obtains (5.15). Consequently, considering the estimates (5.10), (5.7) and Lemma 5.2-(iv), the second and the third assertions hold. □

Please cite this article in press as: Y. Mimura, The variational formulation of the fully parabolic Keller–Segel system with degenerate diffusion, J. Differential Equations (2017), http://dx.doi.org/10.1016/j.jde.2017.03.020

Recall that there exists a sequence τ_n with $\tau_n \downarrow 0$ such that, for every $t > 0$, the sequence of discrete solutions $(\overline{u}_{\tau_n}(t), \overline{v}_{\tau_n}(t))$ converges to some $(u(t), v(t))$ weakly in $L^m(\Omega) \times H_0^1(\Omega)$ (Lemma 5.3).

Now, we define

$$\begin{aligned}\Phi_\tau(t) := |\partial_1 \phi_m|^2(\overline{u}_\tau(t), \overline{v}_\tau(t)) + |\partial_2 \phi_m|^2(\overline{u}_\tau(t), \overline{v}_\tau(t))\\ + \|\overline{u}_\tau(t)\|_{L^2}^{2p} + \|\nabla \overline{u}_\tau^m(t)\|_{L^1}^2 + \|\Delta \overline{v}_\tau(t)\|_{L^2}^2\\ + \|\overline{u}_\tau(t)\|_{L^m} + \|\nabla \overline{v}_\tau(t)\|_{L^2}\end{aligned}$$

and

$$S_\tau(L) := \{t > 0 \,|\, \Phi_\tau(t) > L\}. \tag{5.17}$$

Then from Proposition 3.1, Lemma 5.2-(iv) and Lemma 5.10, we have

$$\int_0^\infty \Phi_\tau(t)\,dt < +\infty.$$

Consequently, for any $\tau > 0$ we have

$$|S_\tau(L)| < \frac{1}{L}\int_0^\infty \Phi_\tau(t)\,dt \to 0 \quad (L \to \infty). \tag{5.18}$$

First, we show that the following pointwise convergence holds at $t_0 \in [0, \infty) \setminus \bigcup_{n=1}^\infty S_{\tau_n}(L)$.

Lemma 5.11 *(Pointwise convergence). Let* $(\overline{u}_{\tau_n}, \overline{v}_{\tau_n})$ *and* (u, v) *be as above. Assume that* $\sup_n \Phi_{\tau_n}(t_0) < \infty$. *Then*

$$\begin{cases} \lim_{n\to\infty} \int_\Omega \langle \nabla \overline{u}_{\tau_n}^m(t_0) - \chi \overline{u}_{\tau_n}(t_0) \nabla \overline{v}_{\tau_n}(t_0), \boldsymbol{\xi} \rangle\,dx \\ \qquad\qquad = \int_\Omega \langle \nabla u^m(t_0) - \chi u(t_0) \nabla v(t_0), \boldsymbol{\xi}\rangle\,dx, \\ \lim_{n\to\infty} \int_\Omega (\Delta \overline{v}_{\tau_n}(t_0) - \gamma \overline{v}_{\tau_n}(t_0) + \alpha \underline{u}_{\tau_n}(t_0))\psi\,dx \\ \qquad\qquad = \int_\Omega (\Delta v(t_0) - \gamma v(t_0) + \alpha u(t_0))\psi\,dx, \end{cases}$$

holds for all $\boldsymbol{\xi} \in C_c^\infty(\mathbb{R}^d; \mathbb{R}^d)$ *and* $\psi \in C_c^\infty(\Omega)$.

Please cite this article in press as: Y. Mimura, The variational formulation of the fully parabolic Keller–Segel system with degenerate diffusion, J. Differential Equations (2017), http://dx.doi.org/10.1016/j.jde.2017.03.020

Proof. By assumption we have

$$\sup_n |\partial_1 \phi_m|(\overline{u}_{\tau_n}(t_0), \overline{v}_{\tau_n}(t_0)) < +\infty, \quad \sup_n |\partial_2 \phi_m|(\underline{u}_{\tau_n}(t_0), \overline{v}_{\tau_n}(t_0)) < +\infty,$$
$$\sup_n \|\overline{u}_{\tau_n}(t_0)\|_{L^2} < +\infty, \ \sup_n \|\nabla \overline{u}^m_{\tau_n}(t_0)\|_{L^1} < +\infty, \ \sup_n \|\overline{v}_{\tau_n}(t_0)\|_{W^{2,2}} < +\infty.$$

We drop t_0 for notational simplicity. The boundedness of $\|\overline{u}_{\tau_n}\|_{L^m}$ and $\|\nabla \overline{u}^m_{\tau_n}\|_{L^1}$ imply that the sequence $(\overline{u}^m_{\tau_n})_n$ is bounded in $W^{1,1}(\Omega)$, hence in $\mathrm{BV}(\Omega)$, functions of bounded variation. By the compactness theorem for functions of bounded variation (see [41, Thm. 3.23]), a bounded sequence in $\mathrm{BV}(\Omega)$ has a subsequence that is weakly* convergent in $\mathrm{BV}(\Omega)$, thus strongly convergent in $L^1(\Omega)$. On the other hand, since $\overline{u}_{\tau_n} \rightharpoonup u$ weakly in $L^m(\Omega)$, we deduce that $\overline{u}^m_{\tau_n} \rightharpoonup u^m$ weakly* in $\mathrm{BV}(\Omega)$ without extracting a subsequence. That is,

$$\overline{u}^m_{\tau_n} \to u^m \quad \text{in } L^1(\Omega) \tag{5.19}$$

and there exists an $\mathbb{R}^d$-valued measure $\boldsymbol{\mu}^* = (\mu_1, \mu_2, \dots, \mu_d)$ such that

$$\int_\Omega \langle \nabla \overline{u}^m_{\tau_n}, \boldsymbol{\xi} \rangle \, dx \to \sum_{j=1}^d \int_\Omega \xi_j \, d\mu^*_j = -\int_\Omega u^m \operatorname{div} \boldsymbol{\xi} \, dx \tag{5.20}$$

for every $\boldsymbol{\xi} = (\xi_1, \xi_2, \dots, \xi_d) \in C^\infty_c(\mathbb{R}^d; \mathbb{R}^d)$.

On the other hand,

$$\begin{aligned} &\left| \int_\Omega \langle \nabla v, \boldsymbol{\xi} \rangle u \, dx - \int_\Omega \langle \nabla \overline{v}_{\tau_n}, \boldsymbol{\xi} \rangle \overline{u}_{\tau_n} \, dx \right| \\ &\le \left| \int_\Omega \langle \nabla v, \boldsymbol{\xi} \rangle (u - \overline{u}_{\tau_n}) \, dx \right| + \int_\Omega |\langle \nabla v - \nabla \overline{v}_{\tau_n}, \boldsymbol{\xi} \rangle| \overline{u}_{\tau_n} \, dx. \end{aligned} \tag{5.21}$$

By the inequality $\|v\|_{W^{2,2}} \le \liminf_{n\to\infty} \|\overline{v}_{\tau_n}\|_{W^{2,2}} < +\infty$ and the Sobolev embedding theorem, $|\langle \nabla v, \boldsymbol{\xi} \rangle| \in L^{\frac{2d}{d-2}}(\Omega)$, which also belongs to the dual space of $L^m(\Omega)$. Because of this, the first term of the right-hand side in (5.21) tends to 0 as $n \to \infty$. The second term of the right-hand side in (5.21) is estimated by

$$\|\xi\|_{L^\infty} \left(\sup_n \|\overline{u}_{\tau_n}\|_{L^2} \right) \|\nabla v - \nabla \overline{v}_{\tau_n}\|_{L^2}.$$

From the boundedness of $\|\overline{v}_{\tau_n}\|_{W^{2,2}}$ and Rellich's compactness theorem, we see that the third factor tends to 0 as $n \to \infty$. We thus have

$$\int_\Omega \langle \nabla \overline{v}_{\tau_n}, \boldsymbol{\xi} \rangle \overline{u}_{\tau_n} \, dx \to \int_\Omega \langle \nabla v, \boldsymbol{\xi} \rangle u \, dx.$$

Please cite this article in press as: Y. Mimura, The variational formulation of the fully parabolic Keller–Segel system with degenerate diffusion, J. Differential Equations (2017), http://dx.doi.org/10.1016/j.jde.2017.03.020

In view of this and (5.20), one obtains

$$\int_\Omega \langle \nabla \overline{u}_{\tau_n}^m - \chi \overline{u}_{\tau_n} \nabla \overline{v}_{\tau_n}, \boldsymbol{\xi} \rangle \, dx \to \sum_{j=1}^d \int_\Omega \xi_j \, d\mu_j^* - \chi \int_\Omega \langle \nabla v, \boldsymbol{\xi} \rangle u \, dx.$$

This and (5.6) imply

$$\begin{aligned} \sum_{j=1}^d \int_\Omega \xi_j \, d\mu_j^* - \chi \int_\Omega \langle \nabla v, \boldsymbol{\xi} \rangle u \, dx &\le \liminf_{n\to\infty} |\partial_1 \phi_m|(\overline{u}_{\tau_n}, \overline{v}_{\tau_n}) \|\boldsymbol{\xi}\|_{L^2(\overline{u}_{\tau_n})} \\ &= \liminf_{n\to\infty} |\partial_1 \phi_m|(\overline{u}_{\tau_n}, \overline{v}_{\tau_n}) \|\boldsymbol{\xi}\|_{L^2(u)}. \end{aligned}$$

By the Riesz representation theorem, there exists $g_1 \in L^2_u(\Omega; \mathbb{R}^d)$ such that

$$\|g_1\|_{L^2} \le \liminf_{n\to\infty} |\partial_1 \phi_m|(\overline{u}_{\tau_n}, \overline{v}_{\tau_n})$$

and

$$\sum_{j=1}^d \int_\Omega \xi_j \, d\mu_j^* - \chi \int_\Omega \langle \nabla v, \boldsymbol{\xi} \rangle u \, dx = \int_\Omega \langle g_1, \boldsymbol{\xi} \rangle u \, dx.$$

Therefore, we have

$$\sum_{j=1}^d \int_\Omega \xi_j \, d\mu_j^* = \int_\Omega \langle u g_1 + \chi u \nabla v, \boldsymbol{\xi} \rangle \, dx = -\int_\Omega u^m \operatorname{div} \boldsymbol{\xi} \, dx.$$

Taking into account that $u g_1 + \chi u \nabla v \in L^1(\Omega)$, we have

$$u g_1 + \chi u \nabla v = \nabla u^m.$$

Combining all relations above, we obtain

$$\lim_{n\to\infty} \int_\Omega \langle \nabla \overline{u}_{\tau_n}^m - \chi \overline{u}_{\tau_n} \nabla \overline{v}_{\tau_n}, \boldsymbol{\xi} \rangle \, dx = \int_\Omega \langle \nabla u^m - \chi u \nabla v, \boldsymbol{\xi} \rangle \, dx$$

and

$$\left(\int_\Omega \frac{|\nabla u^m - \chi u \nabla v|^2}{u} \, dx \right)^{1/2} \le \liminf_{n\to\infty} |\partial_1 \phi_m|(\overline{u}_{\tau_n}, \overline{v}_{\tau_n}). \tag{5.22}$$

Please cite this article in press as: Y. Mimura, The variational formulation of the fully parabolic Keller–Segel system with degenerate diffusion, J. Differential Equations (2017), http://dx.doi.org/10.1016/j.jde.2017.03.020

Next, from the boundedness of $\|\underline{u}_{\tau_n}\|_{L^2}$ and $\|\overline{v}_{\tau_n}\|_{W^{2,2}}$, we deduce that

$$\Delta\overline{v}_{\tau_n} - \gamma\overline{v}_{\tau_n} + \alpha\underline{u}_{\tau_n} \rightharpoonup \Delta v - \gamma v + \alpha u \quad \text{weakly in } L^2(\Omega), \tag{5.23}$$

which implies the second assertion. □

Lemma 5.12 *(L^1-convergence). Let (τ_n) be the sequence in Lemma 5.11. Then*

$$\begin{cases} \lim_{n\to\infty} \int_a^b\int_\Omega \langle \nabla\overline{u}_{\tau_n}^m - \chi\overline{u}_{\tau_n}\nabla\overline{v}_{\tau_n}, \boldsymbol{\xi}\rangle\, dxdt = \int_a^b\int_\Omega \langle \nabla u^m - \chi u\nabla v, \boldsymbol{\xi}\rangle\, dxdt, \\ \lim_{n\to\infty} \int_a^b\int_\Omega (\Delta\overline{v}_{\tau_n} - \gamma\overline{v}_{\tau_n} + \alpha\overline{u}_{\tau_n})\psi\, dxdt = \int_a^b\int_\Omega (\Delta v - \gamma v + \alpha u)\psi\, dxdt, \end{cases}$$

for any $a, b \in [0,\infty)$ and for all $\boldsymbol{\xi} \in C^\infty(\mathbb{R}^d;\mathbb{R}^d)$ and $\psi \in C_c^\infty(\Omega)$.

Proof. Let

$$\rho_{\tau_n}(t) := \left| \int_\Omega \langle \nabla\overline{u}_{\tau_n}^m - \chi\overline{u}_{\tau_n}\nabla\overline{v}_{\tau_n}, \boldsymbol{\xi}\rangle\, dx - \int_\Omega \langle \nabla u^m - \chi u\nabla v, \boldsymbol{\xi}\rangle\, dx \right|$$

and let

$$\tilde{\rho}_{\tau_n,L}(t) := \begin{cases} \rho_{\tau_n}(t), & t \in [0,\infty) \setminus S_{\tau_n}(L), \\ 0, & t \in S_{\tau_n}(L), \end{cases}$$

where $S_{\tau_n}(L)$ is the set defined in (5.17).

For arbitrary $L > 0$ and $t_0 \in [0,\infty)$ if there exists a subsequence (τ_n') such that $t_0 \notin \bigcup_{n=1}^\infty S_{\tau_n'}(L)$, then we deduce from Lemma 5.11 that $\tilde{\rho}_{\tau_n',L}(t_0) \to 0$ $(n \to \infty)$. Consequently, for arbitrary $L > 0$ and $t \in [0,\infty)$, we obtain $\tilde{\rho}_{\tau_n,L}(t) \to 0$ $(n \to \infty)$. Therefore the Lebesgue dominated convergence theorem yields

$$\lim_{n\to\infty} \int_a^b \frac{\tilde{\rho}_{\tau_n,L}}{1+\tilde{\rho}_{\tau_n,L}}\, dt = 0$$

for any $a, b \in [0,\infty)$.

On the other hand, by (5.6) and the Cauchy–Schwarz inequality, together with the fact that $\int_\Omega \overline{u}_{\tau_n}\, dx = 1$, we have

$$\left(\int_\Omega |\nabla \overline{u}_{\tau_n}^m - \chi \overline{u}_{\tau_n} \nabla \overline{v}_{\tau_n}|\, dx\right)^2 \le \left(\int_\Omega \frac{|\nabla \overline{u}_{\tau_n}^m - \chi \overline{u}_{\tau_n} \nabla \overline{v}_{\tau_n}|^2}{\overline{u}_{\tau_n}}\, dx\right)\left(\int_\Omega \overline{u}_{\tau_n}\, dx\right) \tag{5.24}$$
$$\le |\partial_1 \phi_m|^2(\overline{u}_{\tau_n}, \overline{v}_{\tau_n}) \le \Phi_{\tau_n}.$$

Next, we prove

$$\left(\int_\Omega |\nabla u^m(t) - \chi u(t) \nabla v(t)|\, dx\right)^2 \le \liminf_{n\to\infty} \Phi_{\tau_n}(t) \tag{5.25}$$

for $\mathscr{L}^1$-a.e. $t \in [0,\infty)$. Since Fatou's lemma leads to

$$\int_0^\infty \liminf_{n\to\infty} \Phi_{\tau_n}(t)\, dt \le \liminf_{n\to\infty} \int_0^\infty \Phi_{\tau_n}(t)\, dt < +\infty,$$

we obtain

$$\liminf_{n\to\infty} \Phi_{\tau_n}(t) < +\infty \quad \text{for } \mathscr{L}^1\text{-a.e. } t \in [0,\infty).$$

Consequently, for $\mathscr{L}^1$-a.e. $t \in [0,\infty)$, there exists a subsequence (τ_n') depending on t such that

$$\lim_{n\to\infty} \Phi_{\tau_n'}(t) = \liminf_{n\to\infty} \Phi_{\tau_n}(t) < +\infty.$$

Then $(\overline{u}_{\tau_n'}(t), \overline{v}_{\tau_n'}(t))$ satisfies the assumption of Theorem 5.11. Therefore, by the same derivation as (5.22) we obtain

$$\left(\int_\Omega |\nabla u^m(t) - \chi u(t) \nabla v(t)|\, dx\right)^2 \le \liminf_{n\to\infty} |\partial_1 \phi_m|^2(\overline{u}_{\tau_n'}(t), \overline{v}_{\tau_n'}(t))$$
$$\le \lim_{n\to\infty} \Phi_{\tau_n'}(t) = \liminf_{n\to\infty} \Phi_{\tau_n}(t).$$

Combining (5.24) with (5.25) and using Fatou's lemma, we have

$$\int_a^b \rho_{\tau_n}^2\, dt \le 2\|\nabla\varphi\|_{L^\infty}^2 \int_a^b \left(\Phi_{\tau_n} + \liminf_{n\to\infty} \Phi_{\tau_n}\right) dt$$
$$\le 2\|\nabla\varphi\|_{L^\infty}^2 \left(\int_a^b \Phi_{\tau_n}\, dt + \liminf_{n\to\infty} \int_a^b \Phi_{\tau_n}\, dt\right) < +\infty,$$

hence we also obtain

Please cite this article in press as: Y. Mimura, The variational formulation of the fully parabolic Keller–Segel system with degenerate diffusion, J. Differential Equations (2017), http://dx.doi.org/10.1016/j.jde.2017.03.020

$$\sup_n \int_a^b \tilde{\rho}^2_{\tau_n,L}(t)\,dt < +\infty.$$

Therefore, for arbitrary $L > 0$, we have

$$\begin{aligned}
\int_a^b \tilde{\rho}_{\tau_n,L}\,dt &= \int_a^b \left(\frac{\tilde{\rho}_{\tau_n,L}}{1+\tilde{\rho}_{\tau_n,L}}\right)^{1/2} \left(\tilde{\rho}_{\tau_n,L}(1+\tilde{\rho}_{\tau_n,L})\right)^{1/2} dt \\
&\le \left(\int_a^b \frac{\tilde{\rho}_{\tau_n,L}}{1+\tilde{\rho}_{\tau_n,L}}\,dt\right)^{1/2} \left(\int_a^b \tilde{\rho}_{\tau_n,L} + \tilde{\rho}^2_{\tau_n,L}\,dt\right)^{1/2} \\
&\to 0 \quad (n \to \infty).
\end{aligned}$$

Consequently, Cauchy–Schwarz inequality leads to

$$\begin{aligned}
\int_a^b \rho_{\tau_n}\,dt &= \int_{S_{\tau_n}(L)} \rho_{\tau_n}\,dt + \int_{[a,b]\setminus S_{\tau_n}(L)} \rho_{\tau_n}\,dt \\
&\le \left(\sup_n |S_{\tau_n}(L)|\right)^{1/2} \left(\sup_n \int_a^b \rho^2_{\tau_n}\,dt\right)^{1/2} + \int_a^b \tilde{\rho}_{\tau_n,L}\,dt.
\end{aligned}$$

Here, the estimate (5.18) implies that the first term of the right-hand side becomes arbitrary small if $L > 0$ large enough. The second term of the right-hand side tends to 0 as $n \to \infty$. This implies the first assertion. The second assertion can be shown by a similar argument. □

Proof of the convergence result of Theorem 3.3. Fix $a, b \ge 0$ arbitrarily. For any $\tau > 0$, there exist $\ell^a_\tau \in \mathbb{N}$ and $\ell^b_\tau \in \mathbb{N}$ such that

$$\begin{aligned}
&(\ell^a_\tau - 1)\tau < a \le \ell^a_\tau \tau, \quad \lim_{\tau \downarrow 0} \ell^a_\tau \tau = a, \\
&(\ell^b_\tau - 1)\tau < b \le \ell^b_\tau \tau, \quad \lim_{\tau \downarrow 0} \ell^b_\tau \tau = b.
\end{aligned} \tag{5.26}$$

Summing up (5.12) from $k = \ell^a_\tau$ to ℓ^b_τ, we obtain

$$\begin{cases}
\displaystyle\int_\Omega (\overline{u}_\tau(b) - \overline{u}_\tau(a))\varphi\,dx + \int_{\ell^a_\tau \tau}^{\ell^b_\tau \tau}\int_\Omega \langle \nabla \overline{u}^m_\tau - \chi \overline{u}_\tau \nabla \overline{v}_\tau, \nabla\varphi\rangle\,dxdt = R(\tau), \\
\displaystyle\varepsilon \int_\Omega (\overline{v}_\tau(b) - \overline{v}_\tau(a))\psi\,dx - \int_{\ell^a_\tau \tau}^{\ell^b_\tau \tau}\int_\Omega (\Delta \overline{v}_\tau - \gamma \overline{v}_\tau + \alpha \overline{u}_\tau)\psi\,dxdt = 0,
\end{cases}$$

where $R(\tau) = O(\sum_{k\in\mathbb{N}} d_1^2(u_\tau^k, u_\tau^{k-1})) = O(\tau)$. By Lemma 5.3, we can extract a sequence τ_n such that for every $t \geq 0$, the discrete solution $(\overline{u}_{\tau_n}(t), \overline{v}_{\tau_n}(t))$ converges to some function $(u(t), v(t))$ as $n \to \infty$ weakly in $L^m(\Omega) \times H_0^1(\Omega)$. Thus by Lemma 5.12, we see that this function (u, v) satisfies (1.5).

Finally, we check the regularity of the limit function (u, v). The property (i) in Definition 1.3 follows from (3.8). From Lemma 5.10, we see that $u \in L^p(0, T; L^2(\Omega))$ for $p = 2m/(d(2-m)) \geq m$ and every $T > 0$. Consequently, we have $u^m \in L^2(0, T; W^{1,1}(\Omega))$ and $v \in L^2(0, T; W^{2,2}(\Omega))$ by the same argument as in the proof of Lemma 5.10. This confirms the property (ii) in Definition 1.3. The proof of Lemma 5.3 implies that the limit function (u, v) is continuous with respect to the topology of $\mathscr{P}_2(\Omega) \times L^2(\Omega)$ endowed with the distance $d = \sqrt{d_1^2 + d_2^2}$. From this, the property (iii) in Definition 1.3 follows. It remains to check the property (iv). Revisiting the proof of Lemma 5.11 and Lemma 5.12, one sees that the two lemmas hold true for $\boldsymbol{\xi} \in C_c^\infty(\mathbb{R}^d \times [0, \infty); \mathbb{R}^d)$ and $\psi \in C_c^\infty(\mathbb{R}^d \times [0, \infty))$ instead of $\boldsymbol{\xi} \in C_c^\infty(\mathbb{R}^d; \mathbb{R}^d)$ and $\psi \in C_c^\infty(\mathbb{R}^d)$, respectively. Therefore, we have

$$\begin{aligned}
\int_a^b\!\!\int_\Omega \langle \nabla u^m - \chi u \nabla v, \boldsymbol{\xi}\rangle\, dx dt &= \lim_{n\to\infty} \int_a^b\!\!\int_\Omega \langle \nabla \overline{u}_{\tau_n}^m - \chi \overline{u}_{\tau_n} \nabla \overline{v}_{\tau_n}, \boldsymbol{\xi}\rangle\, dx dt \\
&\leq \liminf_{n\to\infty} \left\{ \left(\int_a^b\!\!\int_\Omega |\Delta \overline{v}_{\tau_n} - \gamma \overline{v}_{\tau_n} + \alpha \underline{u}_{\tau_n}|^2\, dx dt \right)^{1/2} \left(\int_a^b\!\!\int_\Omega |\boldsymbol{\xi}|^2 \overline{u}_{\tau_n}\, dx dt \right)^{1/2} \right\} \\
&= \liminf_{n\to\infty} \left(\int_a^b\!\!\int_\Omega |\Delta \overline{v}_{\tau_n} - \gamma \overline{v}_{\tau_n} + \alpha \underline{u}_{\tau_n}|^2\, dx dt \right)^{1/2} \left(\int_a^b\!\!\int_\Omega |\boldsymbol{\xi}|^2 u\, dx dt \right)^{1/2}.
\end{aligned}$$

Consequently, by duality and the inequality (5.6), we obtain

$$\begin{aligned}
\int_a^b\!\!\int_\Omega \frac{|\nabla u^m - \chi u \nabla v|^2}{u}\, dx dt &\leq \liminf_{n\to\infty} \int_a^b\!\!\int_\Omega \frac{|\overline{u}_{\tau_n} - \chi \overline{u}_{\tau_n} \nabla \overline{v}_{\tau_n}|^2}{\overline{u}_{\tau_n}}\, dx dt \\
&\leq \liminf_{n\to\infty} \int_a^b |\partial_1 \phi_m|^2 (\overline{u}_{\tau_n}, \overline{v}_{\tau_n})\, dt.
\end{aligned} \tag{5.27}$$

The right-hand side has a finite value by Lemma 5.2-(iv). Similarly,

$$\begin{aligned}
\int_a^b\!\!\int_\Omega |\Delta v - \gamma v + \alpha u|^2\, dx dt &\leq \liminf_{n\to\infty} \int_a^b\!\!\int_\Omega |\Delta \overline{v}_{\tau_n} - \gamma \overline{v}_{\tau_n} + \alpha \underline{u}_{\tau_n}|^2\, dx dt \\
&\leq \liminf_{n\to\infty} \int_a^b |\partial_2 \phi_m|^2 (\underline{u}_{\tau_n}, \overline{v}_{\tau_n})\, dt.
\end{aligned} \tag{5.28}$$

These prove the property (iv) in Definition 1.3. Therefore, the limit function (u, v) is a weak solution of (1.1) that exists globally in time. □

6. Energy dissipative inequality

In this final section, we prove Theorem 1.5. For that purpose, we need a certain continuity property of ϕ_m with respect to the weak topology and Rellich's theorem plays a key role in that argument. This is a reason why we are mainly focusing on the case where Ω is a bounded domain in $\mathbb{R}^d$.

Lemma 6.1 *(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*

$$\phi_m(u, v) \leq \liminf_{n\to\infty} \phi_m(u_n, v_n).$$

Proof. It suffices to show that $\int_\Omega uv\,dx$ is continuous with respect to the weak topology in $L^m(\Omega) \times H_0^1(\Omega)$, because it is well known that $\phi_m(u, v) + \chi \int_\Omega uv\,dx$ 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 $L^{m'}(\Omega)$, where $m' := \frac{m}{m-1} \leq \frac{2d-2}{d-2}$. Therefore, by Hölder's inequality we have

$$\left| \int_\Omega (u_n v_n - uv)\,dx \right| \leq \|u_n\|_{L^m} \|v_n - v\|_{L^{m'}} + \left| \int_\Omega (u_n - u)v\,dx \right| \to 0 \quad (n \to \infty). \qquad \square$$

Remark 6.1. Lemma 6.1 still holds in the case $\Omega = \mathbb{R}^d$, under the condition $\sup_n \int_{\mathbb{R}^d} |x|^2 u_n(x)\,dx < +\infty$.

Lemma 6.2 *(Continuity of ϕ_m). Let $(\overline{u}_{\tau_n}, \overline{v}_{\tau_n})$ be a sequence of discrete solutions converging to a weak solution (u, v) of (1.1). Then, for $\mathscr{L}^1$-a.e. $t \in [0, \infty)$ there exists a subsequence $(\overline{u}_{\tau_n'}, \overline{v}_{\tau_n'})$ such that*

$$\lim_{n\to\infty} \phi_m(\overline{u}_{\tau_n'}(t), \overline{v}_{\tau_n'}(t)) = \phi_m(u(t), v(t)).$$

Proof. The lemma is clear when $t = 0$. By Fatou's lemma, we have

$$\int_0^\infty \liminf_{n\to\infty} \Phi_{\tau_n}(t)\,dt \leq \liminf_{n\to\infty} \int_0^\infty \Phi_{\tau_n}(t)\,dt < +\infty.$$

Hence

$$\liminf_{n\to\infty} \Phi_{\tau_n}(t) < +\infty \quad \mathscr{L}^1\text{-a.e. } t \in [0, \infty).$$

Consequently, for $\mathscr{L}^1$-a.e. $t_0 \in [0, \infty)$, there exists a subsequence (τ_n') such that

Please cite this article in press as: Y. Mimura, The variational formulation of the fully parabolic Keller–Segel system with degenerate diffusion, J. Differential Equations (2017), http://dx.doi.org/10.1016/j.jde.2017.03.020

$$\lim_{n\to\infty} \Phi_{\tau_n'}(t_0) = \liminf_{n\to\infty} \Phi_{\tau_n}(t_0) < +\infty.$$

Since $(u_{\tau_n'}, v_{\tau_n'})$ and t_0 satisfy the assumption of Theorem 5.11, from (5.19), it follows that

$$\overline{u}_{\tau_n'}(t_0) \to u(t_0) \text{ in } L^m(\Omega)$$

On the other hand, we deduce from he boundedness of $\|\overline{v}_{\tau_n'}(t_0)\|_{W^{2,2}}$ and Rellich's compactness theorem that $\overline{v}_{\tau_n'}(t_0) \to v(t_0)$ strongly in $H_0^1(\Omega)$. Consequently,

$$\lim_{n\to\infty} \phi_m(\overline{u}_{\tau_n'}(t_0), \overline{v}_{\tau_n'}(t)_0) = \phi_m(u(t_0), v(t_0)). \quad \square$$

The gradient energy estimate listed in Lemma 5.2-(iv) leads to the following rough energy estimate:

$$\begin{aligned} &2(\phi_m(u(a), v(a)) - \phi_m(u(b), v(b))) \\ &\qquad \geq \int_a^b\!\!\int_\Omega \frac{|\nabla u^m - \chi u \nabla v|^2}{u}\, dxdt + \frac{\chi}{\alpha\varepsilon} \int_a^b\!\!\int_\Omega |\Delta v - \gamma v + \alpha u|^2\, dxdt \end{aligned} \tag{6.1}$$

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

In fact, letting $\tau = \tau_n'$ and passing to the limit as $n \to \infty$ in Lemma 5.2-(iv), and considering the inequalities (5.27) and (5.28) with Lemmas 6.1 and 6.2, we obtain (6.1).

We now improve the estimate (6.1) by using the Moreau–Yosida Approximation.

Proof of Theorem 1.5. By applying the derivative of Moreaux–Yosida approximation [28, Thm. 3.1.4] to our minimizing problem (3.7), it holds for $k = 1, 2, \dots$ that

$$\begin{aligned} \frac{d_2^2(v_\tau^k, v_\tau^{k-1})}{2\tau^2} + \int_0^\tau \frac{d_2^2(V_\sigma^k, v_\tau^{k-1})}{2\sigma^2}\, d\sigma &= \phi_m(u_\tau^{k-1}, v_\tau^{k-1}) - \phi_m(u_\tau^{k-1}, v_\tau^k), \\ \frac{d_1^2(u_\tau^k, u_\tau^{k-1})}{2\tau^2} + \int_0^\tau \frac{d_1^2(U_\sigma^k, u_\tau^{k-1})}{2\sigma^2}\, d\sigma &= \phi_m(u_\tau^{k-1}, v_\tau^k) - \phi_m(u_\tau^k, v_\tau^k), \end{aligned} \tag{6.2}$$

where

$$V_\sigma^k \in \operatorname*{argmin}_{H_0^1(\Omega)} \left\{ \phi_m(u_\tau^{k-1}, v) + \frac{1}{2\sigma} d_2^2(v, v_\tau^{k-1}) \right\},$$

$$U_\sigma^k \in \operatorname*{argmin}_{\mathscr{P}_2(\Omega)} \left\{ \phi_m(u, v_\tau^k) + \frac{1}{2\sigma} d_1^2(u, u_\tau^{k-1}) \right\}.$$

Note that for any $\sigma \in (0, \tau)$ and $k \in \mathbb{N}$, V_σ^k and U_σ^k are uniquely defined by an argument similar to the proof of Proposition 3.1. Moreover, by the slope estimates (5.2) and (5.3), it follows that

Please cite this article in press as: Y. Mimura, The variational formulation of the fully parabolic Keller–Segel system with degenerate diffusion, J. Differential Equations (2017), http://dx.doi.org/10.1016/j.jde.2017.03.020

$$\begin{cases} |\partial_2\phi_m|(u_\tau^{k-1}, V_\sigma^k) \le \dfrac{d_2(V_\sigma^k, v_\tau^{k-1})}{\sigma}, \\ |\partial_1\phi_m|(U_\sigma^k, v_\tau^k) \le \dfrac{d_1(U_\sigma^k, u_\tau^{k-1})}{\sigma}. \end{cases} \tag{6.3}$$

We now define the *De Giorgi variational interpolations* by

$$\begin{cases} \tilde{U}_\tau(t) := U_\sigma^k & \text{if } t = (k-1)\tau + \sigma, \\ \tilde{V}_\tau(t) := V_\sigma^k & \text{if } t = (k-1)\tau + \sigma. \end{cases}$$

Note that $\tilde{U}_\tau$ and $\tilde{V}_\tau$ have the same limits as $\overline{u}_\tau$ and $\overline{v}_\tau$, respectively. Indeed,

$$\begin{aligned} \phi_m(U_\sigma^k, v_\tau^k) + \frac{1}{2\sigma} d_1^2(U_\sigma^k, u_\tau^{k-1}) &\le \phi_m(u_\tau^k, v_\tau^k) + \frac{1}{2\sigma} d_1^2(u_\tau^k, u_\tau^{k-1}) \\ &= \left\{ \phi_m(u_\tau^k, v_\tau^k) + \frac{1}{2\tau} d_1^2(u_\tau^k, u_\tau^{k-1}) \right\} + \frac{1}{2}\left(\frac{1}{\sigma} - \frac{1}{\tau}\right) d_1^2(u_\tau^k, u_\tau^{k-1}) \\ &\le \phi_m(U_\sigma^k, u_\tau^k) + \frac{1}{2\tau} d_1^2(U_\sigma^k, u_\tau^{k-1}) + \frac{1}{2}\left(\frac{1}{\sigma} - \frac{1}{\tau}\right) d_1^2(u_\tau^k, u_\tau^{k-1}) \end{aligned}$$

from which we obtain $d_1(U_\sigma^k, u_\tau^{k-1}) \le d_1(u_\tau^k, u_\tau^{k-1})$. Hence for $t \in ((k-1)\tau, k\tau]$,

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

This implies that $\tilde{U}_\tau$ has the same limit as $\overline{u}_\tau$. By a similar argument, $\tilde{V}_\tau$ has the same limit as $\overline{v}_\tau$. Thus

$$(\tilde{U}_{\tau_n}(t), \tilde{V}_{\tau_n}(t)) \rightharpoonup (u(t), v(t)) \ \text{ weakly in } L^m(\Omega) \times H_0^1(\Omega) \text{ for all } t > 0.$$

Considering (6.2), (6.3) and Lemma 5.2-(i), we deduce that

$$\begin{aligned} &\phi_m(\underline{u}_\tau(a), \underline{v}_\tau(a)) - \phi_m(\overline{u}_\tau(b), \overline{v}_\tau(b)) \\ &\ge \frac{1}{2} \int_{(\ell_\tau^a - 1)\tau}^{\ell_\tau^b \tau} |\partial_1\phi_m|^2(\overline{u}_\tau, \overline{v}_\tau) + |\partial_2\phi_m|^2(\underline{u}_\tau, \overline{v}_\tau)\, dt \\ &\quad + \frac{1}{2} \int_{(\ell_\tau^a - 1)\tau}^{\ell_\tau^b \tau} |\partial_1\phi_m|^2(\tilde{U}_\tau, \overline{v}_\tau) + |\partial_2\phi_m|^2(\underline{u}_\tau, \tilde{V}_\tau)\, dt \end{aligned} \tag{6.4}$$

$$\geq \frac{1}{2}\int_a^b |\partial_1\phi_m|^2(\overline{u}_\tau, \overline{v}_\tau) + |\partial_2\phi_m|^2(\underline{u}_\tau, \overline{v}_\tau)\,dt$$

$$+\frac{1}{2}\int_a^b |\partial_1\phi_m|^2(\tilde{U}_\tau, \overline{v}_\tau) + |\partial_2\phi_m|^2(\underline{u}_\tau, \tilde{V}_\tau)\,dt,$$

where ℓ_τ^a and ℓ_τ^b are the integers defined in (5.26). By Lemma 5.6, we see that $\tilde{U}_\tau(t) \in L^2(\Omega)$ and $\tilde{V}_\tau(t) \in W^{2,2}(\Omega)$ for any $t > 0$. Therefore, Lemma 5.10 and holds true if we replace $\overline{u}_\tau$ and $\overline{v}_\tau$ by $\tilde{U}_\tau$ and $\tilde{V}_\tau$, respectively. Hence revisiting the proof of Lemmas 5.11 and 5.12, we see that (5.27) remains true if $\overline{u}_{\tau_n}$ on the right-hand side is replaced by $\tilde{U}_{\tau_n}$, and (5.28) remains true if $\overline{v}_{\tau_n}$ is replaced by $\tilde{V}_{\tau_n}$. Consequently,

$$\begin{aligned}
\int_a^b\int_\Omega \frac{|\nabla u^m - \chi u \nabla v|^2}{u}\,dxdt &\leq \liminf_{n\to\infty}\int_a^b |\partial_1\phi_m|^2(\tilde{U}_{\tau_n}, \overline{v}_{\tau_n})\,dt, \\
\frac{\chi}{\alpha\varepsilon}\int_a^b\int_\Omega |\Delta v - \gamma v + \alpha u|^2\,dxdt &\leq \liminf_{n\to\infty}\int_a^b |\partial_2\phi_m|^2(\underline{u}_{\tau_n}, \tilde{V}_{\tau_n})\,dt.
\end{aligned} \tag{6.5}$$

Thus, setting $\tau = \tau_n'$ in (6.4) and passing to the limit as $n \to \infty$, and applying Lemmas 6.1 and 6.2, we obtain the energy dissipative inequality (1.6). $\square$

Acknowledgments

The author would like to thank Professor Hiroshi Matano for his valuable advise and continued encouragement, and also Professor Filippo Santambrogio for stimulating discussions. The author would also like to thank the anonymous referees for numerous valuable suggestions and comments.

References

[1] E.F. Keller, L.A. Segel, Initiation of slime mold aggregation viewed as in instability, J. Theoret. Biol. 26 (3) (1970) 399–415.
[2] S. Ishida, T. Yokota, Global existence of weak solutions to quasilinear degenerate Keller–Segel systems of parabolic–parabolic type, J. Differential Equations 252 (2) (2012) 1421–1440.
[3] R. Kowalczyk, Z. Szymańska, On the global existence of solutions to an aggregation model, J. Math. Anal. Appl. 343 (1) (2008) 379–398.
[4] Y. Sugiyama, Global existence in sub-critical cases and finite time blow-up in super-critical cases to degenerate Keller–Segel systems, Differential Integral Equations 19 (8) (2006) 841–876.
[5] Y. Sugiyama, Time global existence and asymptotic behavior of solutions to degenerate quasi-linear parabolic systems of chemotaxis, Differential Integral Equations 20 (2) (2007) 133–180.
[6] Y. Sugiyama, H. Kunii, Global existence and decay properties for a degenerate Keller–Segel model with a power factor in drift term, J. Differential Equations 227 (1) (2006) 333–364.
[7] T. Senba, T. Suzuki, A quasi-linear parabolic system of chemotaxis, Abstr. Appl. Anal. (2006) 23061.
[8] L. Corrias, B. Perthame, H. Zaag, Global solutions of some chemotaxis and angiogenesis systems in high space dimensions, Milan J. Math. 72 (2004) 1–28.

Please cite this article in press as: Y. Mimura, The variational formulation of the fully parabolic Keller–Segel system with degenerate diffusion, J. Differential Equations (2017), http://dx.doi.org/10.1016/j.jde.2017.03.020

[9] S. Ishida, T. Yokota, Global existence of weak solutions to quasilinear degenerate Keller–Segel systems of parabolic–parabolic type with small data, J. Differential Equations 252 (3) (2012) 2469–2491.

[10] L. Chen, J.-G. Liu, J. Wang, Multidimensional degenerate Keller–Segel system with critical diffusion exponent $2n/(n+2)$, SIAM J. Math. Anal. 44 (2) (2012) 1077–1102.

[11] Y. Sugiyama, Application of the best constant of the Sobolev inequality to degenerate Keller–Segel models, Adv. Differential Equations 12 (2) (2007) 121–144.

[12] T. Cieślak, C. Stinner, Finite-time blowup and global-in-time unbounded solutions to a parabolic–parabolic quasilinear Keller–Segel system in higher dimensions, J. Differential Equations 252 (10) (2012) 5832–5851.

[13] J. Bedrossian, N. Rodríguez, A.L. Bertozzi, Local and global well-posedness for aggregation equations and Patlak–Keller–Segel models with degenerate diffusion, Nonlinearity 24 (6) (2011) 1683–1714.

[14] A. Blanchet, J.A. Carrillo, P. Laurençot, Critical mass for a Patlak–Keller–Segel model with degenerate diffusion in higher dimensions, Calc. Var. Partial Differential Equations 35 (2) (2009) 133–168.

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

[16] T. Suzuki, R. Takahashi, Degenerate parabolic equation with critical exponent derived from the kinetic theory. II. Blowup threshold, Differential Integral Equations 22 (11–12) (2009) 1153–1172.

[17] H. Gajewski, K. Zacharias, Global behaviour of a reaction–diffusion system modelling chemotaxis, Math. Nachr. 195 (1998) 77–114.

[18] T. Nagai, T. Senba, K. Yoshida, Application of the Trudinger–Moser inequality to a parabolic system of chemotaxis, Funkcial. Ekvac. 40 (3) (1997) 411–433.

[19] V. Calvez, L. Corrias, The parabolic–parabolic Keller–Segel model in $\mathbb{R}^2$, Commun. Math. Sci. 6 (2) (2008) 417–447.

[20] M.A. Herrero, J.J.L. Velázquez, Chemotactic collapse for the Keller–Segel model, J. Math. Biol. 35 (2) (1996) 177–194.

[21] D. Horstmann, G. Wang, Blow-up in a chemotaxis model without symmetry assumptions, European J. Appl. Math. 12 (2) (2001) 159–177.

[22] A. Blanchet, V. Calvez, J.A. Carrillo, Convergence of the mass-transport steepest descent scheme for the subcritical Patlak–Keller–Segel model, SIAM J. Numer. Anal. 46 (2) (2008) 691–721.

[23] A. Blanchet, E.A. Carlen, J.A. Carrillo, Functional inequalities, thick tails and asymptotics for the critical mass Patlak–Keller–Segel model, J. Funct. Anal. 262 (5) (2012) 2142–2230.

[24] F. Otto, Doubly degenerate diffusion equations as steepest descent, 1996.

[25] F. Otto, The geometry of dissipative evolution equations: the porous medium equation, Comm. Partial Differential Equations 26 (1–2) (2001) 101–174.

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

[27] M. Agueh, Existence of solutions to degenerate parabolic equations via the Monge–Kantorovich theory, Adv. Differential Equations 10 (3) (2005) 309–360.

[28] L. Ambrosio, N. Gigli, G. Savaré, Gradient Flows in Metric Spaces and in the Space of Probability Measures, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 2005.

[29] S. Lisini, Nonlinear diffusion equations with variable coefficients as gradient flows in Wasserstein spaces, ESAIM Control Optim. Calc. Var. 15 (3) (2009) 712–740.

[30] D. Matthes, R.J. McCann, G. Savaré, A family of nonlinear fourth order equations of gradient flow type, Comm. Partial Differential Equations 34 (10–12) (2009) 1352–1397.

[31] C. Villani, Optimal Transport, Grundlehren der Mathematischen Wissenschaften (Fundamental Principles of Mathematical Sciences), vol. 338, Springer-Verlag, Berlin, 2009, old and new.

[32] C. Villani, Topics in Optimal Transportation, Graduate Studies in Mathematics, vol. 58, American Mathematical Society, Providence, RI, 2003.

[33] H. Ohtsuka, T. Senba, T. Suzuki, Blowup in infinite time in the simplified system of chemotaxis, Adv. Math. Sci. Appl. 17 (2) (2007) 445–472.

[34] P. Biler, D. Hilhorst, T. Nadzieja, Existence and nonexistence of solutions for a model of gravitational interaction of particles. II, Colloq. Math. 67 (2) (1994) 297–308.

[35] P. Biler, T. Nadzieja, Existence and nonexistence of solutions for a model of gravitational interaction of particles. I, Colloq. Math. 66 (2) (1994) 319–334.

[36] N.I. Kavallaris, P. Souplet, Grow-up rate and refined asymptotics for a two-dimensional Patlak–Keller–Segel model in a disk, SIAM J. Math. Anal. 40 (5) (2008–2009) 1852–1881.

[37] Y. Sugiyama, 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 (2) (2011) 319–337.

Please cite this article in press as: Y. Mimura, The variational formulation of the fully parabolic Keller–Segel system with degenerate diffusion, J. Differential Equations (2017), http://dx.doi.org/10.1016/j.jde.2017.03.020

[38] L. Ambrosio, G. Savaré, Gradient flows of probability measures, in: Handbook of Differential Equations: Evolutionary Equations, vol. III, in: Handb. Differ. Equ., Elsevier/North-Holland, Amsterdam, 2007, pp. 1–136.

[39] C. Dellacherie, P.-A. Meyer, Probabilities and Potential, North-Holland Mathematics Studies, vol. 29, North-Holland Publishing Co., Amsterdam, 1978.

[40] R.A. Adams, Sobolev Spaces, Pure and Applied Mathematics, vol. 65, Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York, London, 1975.

[41] L. Ambrosio, N. Fusco, D. Pallara, Functions of Bounded Variation and Free Discontinuity Problems, Oxford Mathematical Monographs, The Clarendon Press, Oxford University Press, New York, 2000.

Please cite this article in press as: Y. Mimura, The variational formulation of the fully parabolic Keller–Segel system with degenerate diffusion, J. Differential Equations (2017), http://dx.doi.org/10.1016/j.jde.2017.03.020