Intermediate Asymptotics for Critical and Supercritical Aggregation Equations and Patlak-Keller-Segel models

Jacob Bedrossian *

Abstract

We examine the long-term asymptotic behavior of dissipating solutions to aggregation equations and Patlak-Keller-Segel models with degenerate power-law and linear diffusion. The purpose of this work is to identify when solutions decay to the self-similar spreading solutions of the homogeneous diffusion equations. Entropy dissipation methods provide a natural solution to this question and make it possible to derive quantitative convergence rates in L^1. The estimated rate relates the nonlinearity of the diffusion and the decay of the interaction kernel at infinity. For supercritical problems with kernels in $W^{1,1}(\mathbb{R}^d)$, we obtain the optimal convergence rates associated with the diffusion equations.

1 Introduction

The most widely studied mathematical models of nonlocal aggregation phenomena are the Patlak-Keller-Segel (PKS) models, originally introduced to study the chemotaxis of microorganisms [38, 27, 24, 23]. Similar models are also used to study the formation of herds and flocks in ecological systems [11, 45, 35, 22]. A common theme is the competition between the tendency for organisms to diffuse, e.g. under Brownian motion or to avoid over-crowding, and for organisms to aggregate into groups through nonlocal self-attraction. The parabolic-elliptic PKS models are a subclass of the general aggregation-diffusion equations

$$u_t + \nabla \cdot (u \nabla \mathcal{K} * u) = \Delta A(u). \tag{1}$$

The local and global existence and uniqueness of models such as (1) is well studied (see for instance [5, 6, 8, 10, 42, 43, 44, 18]). However, less is known about the long-term qualitative behavior of solutions. In this work, we are interested in examining the asymptotic profiles of dissipating solutions to (1) in the special case

$$\begin{cases} u_t + \nabla \cdot (u \nabla \mathcal{K} * u) = \Delta u^m, \quad m \geq 1, \\ u(0,x) = u_0(x) \in L^1_+(\mathbb{R}^d; (1+|x|^2)dx) \cap L^\infty(\mathbb{R}^d), \end{cases} \tag{2}$$

where $L^1_+(\mathbb{R}^d;\mu) := \left\{ f \in L^1(\mathbb{R}^d;\mu) : f \geq 0 \right\}$. In particular, we are interested in determining when solutions to (2) converge in $L^1(\mathbb{R}^d)$ as $t \to \infty$ to the self-similar spreading solutions of the diffusion equation

$$u_t = \Delta u^m. \tag{3}$$

All dissipating solutions are weak* converging to zero as $t \to \infty$, but this kind of result implies that for $1 << t < \infty$, the dissipating solutions all look more or less like self-similar solutions of (3). For this reason, these results are often referred to as *intermediate asymptotics*.

* *jacob.bedrossian@math.ucla.edu*, University of California-Los Angeles, Department of Mathematics

Supercritical problems are those in which the aggregation is dominant at high concentrations, *subcritical* problems are those in which the diffusion dominates at high concentrations, and *critical* problems are those in which the effects are in approximate balance. It is known that supercritical problems exhibit finite time blow up for solutions of arbitrarily small mass and subcritical problems have global solutions [42, 43, 5, 8]. The critical case is more interesting; data with small mass exists globally, whereas finite time blow up is possible for large mass [8, 5, 10, 43]. In this work, we will refer to the case $m < 2 - 2/d$ as supercritical and $m = 2 - 2/d$ as critical. This is in contrast to the definition used in [5], where the critical diffusion exponent was taken to depend on the singularity of the kernel. Here, achieving such a precise balance is not the primary interest. In the sense of [5, 8, 43, 42], $m = 2 - 2/d$ is the critical exponent for the Newtonian potential, which is the most singular kernel known to have unique, local-in-time solutions [5].

As strong nonlinearities vanish quickly near zero, scaling heuristics suggest that the nonlocal aggregation term should become irrelevant for small data in the critical and supercritical regime. We use entropy dissipation methods [16, 46, 17, 14, 13, 7] to obtain several intermediate asymptotics results which show this to be true, and that solutions of (2) converge to self-similar solutions of (3). Entropy dissipation methods are well-suited for proving the convergence to equilibrium states of nonlinear Fokker-Plank-type equations for arbitrary data [16, 14]. Through a change of variables employed below, this also provides convergence to self-similarity of nonlinear homogeneous diffusion equations [17]. In contrast to these works, we employ such methods to prove a *small data* result, treating the nonlocal aggregation term as a perturbation. The change of variables will also clarify how the long-range behavior of the kernel affects the problem. The key step is to derive a sufficiently strong decay estimate to allow one to invoke entropy dissipation methods that normally only apply to the homogeneous diffusion equations.

The first result, Theorem 1, covers the case $\mathcal{K} \in W^{1,1}(\mathbb{R}^d)$. Here, the nonlocal term can be considered to have a finite characteristic length-scale which becomes vanishingly small relative to the length-scale of the solution as it dissipates. A result similar to Theorem 1 for L^p, $1 < p < \infty$, was proved for the special case of the Bessel potential in [32, 33] with the compactness method of [26] (see also [47]). In contrast to methods based on compactness, the entropy dissipation methods obtain quantitative convergence rates in L^1, which by interpolation against the decay estimates, provides convergence in all L^p, $1 \leq p < \infty$. For supercritical problems, the convergence rate is shown to be the same as the optimal rates for (3) [16, 46, 17, 14, 47].

In general, if the kernel does not have critical scaling at large length-scales, the long-range effects should still become irrelevant as the solution dissipates. That is, we should expect results similar to the $\mathcal{K} \in W^{1,1}(\mathbb{R}^d)$ case to hold, except when $m = 2 - 2/d$ and $\nabla\mathcal{K} \sim |x|^{1-d}$ as $|x| \to \infty$. Indeed, when $\mathcal{K}$ is the Newtonian potential, there exists at least one self-similar spreading solution to (2) when $m = 2 - 2/d$ [10, 8, 9, 12]. In the presence of linear diffusion, these are additionally known to be the global attractors [10, 9]. Theorem 2 below extends Theorem 1 to the general case of $\mathcal{K} \notin W^{1,1}(\mathbb{R}^d)$, where the decay of $\mathcal{K}$ is characterized by $\gamma \in [d-1, d]$ such that $|\nabla\mathcal{K}(x)| = \mathcal{O}(|x|^{-\gamma})$ as $|x| \to \infty$. We show that if $\gamma > d - 1$, then dissipating solutions converge to the self-similar spreading solutions of (3). However, in contrast to Theorem 1, the long-range effects appear to degrade the convergence rate and Theorem 2 provides a quantitative estimate of this effect in terms of m and γ. It is not known whether the rates obtained in Theorem 2 are sharp. When $\gamma = d - 1$, the kernel behaves like the Newtonian potential on large length-scales, and the result is no longer expected to hold if $m = 2-2/d$. Indeed, we expect solutions to converge to the self-similar solutions of (2) constructed in [10, 8]. However, Theorem 2 asserts that in supercritical cases, self-similar solutions to (3) again govern the intermediate asymptotics. Thus, Theorem 2 provides intermediate asymptotics for Patlak-Keller-Segel models with linear diffusion in dimensions $d \geq 3$.

In what follows, we denote $\|u\|_p := \|u\|_{L^p(\mathbb{R}^d)}$ where $L^p(\mathbb{R}^d) := L^p$ is the standard Lebesgue space. The standard characteristic function for some $S \subset \mathbb{R}^d$ is denoted $\mathbf{1}_S$ and we denote the ball $B_R(x_0) := \left\{x \in \mathbb{R}^d : |x - x_0| < R\right\}$. In formulas we use the notation $C(p,k,M,..)$ to denote a generic constant, which may be different from line to line or term to term in the same formula. In general, these constants will depend on more parameters than those listed, for instance those which are fixed by the problem, such as $\mathcal{K}$ and the dimension, but these dependencies are suppressed. We use the notation $f \lesssim_{p,k,...} g$ to denote $f \leq C(p,k,..)g$ where again, dependencies that are not relevant are suppressed.

1.1 Statement of Results

We need the following definition from [5], which we restate here.

Definition 1. We say a kernel $\mathcal{K}$ is *admissible* if $\mathcal{K} \in W^{1,1}_{loc}(\mathbb{R}^d)$ and the following holds:

(R) $\mathcal{K} \in C^3 \setminus \{0\}$.

(KN) $\mathcal{K}$ is radially symmetric, $\mathcal{K}(x) = k(|x|)$ and $k(|x|)$ is non-increasing.

(MN) $k''(r)$ and $k'(r)/r$ are monotone on $r \in (0,\delta)$ for some $\delta > 0$.

(BD) $\left|D^2\mathcal{K}(x)\right| \leq C\,|x|^{-d}$ for some $C > 0$.

The definition ensures that the kernel is radially symmetric, attractive, reasonably well-behaved at the origin and has second derivatives which define bounded distributions on L^p for $1 < p < \infty$. It is important to note that all admissible kernels satisfy $\nabla\mathcal{K} \in L^{\frac{d}{d-1},\infty}$, where $L^{p,\infty}$ denotes the weak-L^p space, making the Newtonian potential effectively the most singular of admissible kernels [5]. Provided $\mathcal{K}$ is admissible, (2) has a unique local-in-time weak solution with values in $L^1_+(\mathbb{R}^d;(1+|x|^2)dx) \cap L^\infty(\mathbb{R}^d)$ [5, 6, 10, 44, 4]. For a given initial condition $u_0(x) \in L^1_+(\mathbb{R}^d;(1+|x|^2)dx) \cap L^\infty(\mathbb{R}^d)$, the unique weak solution to (2) satisfies $u(t) \in C([0,T);L^1_+(\mathbb{R}^d;(1+|x|^2)dx)) \cap L^\infty([0,T)\times\mathbb{R}^d)$. Moreover, $u(t)$ is a solution to (2) in a sense which is slightly stronger than a distribution solution, but the distinction is not important here [5, 6]. Weak solutions conserve mass and we define $M = \|u_0\|_1 = \|u(t)\|_1$.

The self-similar solutions to the diffusion equation (3) are well-known, see for instance [47] and [17]. In the linear case $m = 1$, the self-similar solution is simply the heat kernel,

$$\mathcal{U}(t,x;M) = \frac{M}{(4\pi t)^{d/2}} e^{\frac{-|x|^2}{4t}}. \tag{4}$$

In the case of degenerate diffusion $m > 1$, the self-similar solution is given by the Barenblatt solution,

$$\mathcal{U}(t,x;M) = t^{-\beta d}\left(C_1 - \frac{(m-1)\beta}{2m}|x|^2 t^{-2\beta}\right)_+^{\frac{1}{m-1}}, \tag{5}$$

where C_1 is determined from the conservation of mass and

$$\beta = \frac{1}{d(m-1)+2}. \tag{6}$$

The entropy dissipation methods of [16, 46, 17, 14] were used to determine the optimal rate of convergence in $L^1(\mathbb{R}^d)$ to self-similarity. That is, any solution $u(t)$ of (3) satisfies

$$t^{d\beta\left(1-\frac{1}{p}\right)}\|u(t) - \mathcal{U}(t;M)\|_p \lesssim (1+t)^{-\frac{2\beta}{p}\min\left(\frac{1}{2},\frac{1}{m}\right)},\ \forall p,\ 1 \leq p < \infty.$$

This rate should be contrasted with the rates obtained in Theorems 1 and 2, where it is shown that kernels with finite length-scales do not have much effect on the rate, but strong nonlocal effects do.

Theorem 1 (Intermediate Asymptotics I: Finite Length-Scale)**.** *Let $d \geq 2$, $m \in [1, 2-2/d]$ and $\mathcal{K} \in W^{1,1}$ be admissible. Let $f \in L^1_+(\mathbb{R}^d; (1+|x|^2)dx) \cap L^\infty(\mathbb{R}^d)$. Then there exists an $\epsilon_0(\|f\|_1, \|f\|_{(2-m)d/2}) > 0$ such that for all $\epsilon < \epsilon_0$, if $u_0 = \epsilon f$ then the weak solution $u(t)$ to (2) is global and satisfies*

$$\|u(t)\|_\infty \lesssim (1+t)^{-d\beta}. \tag{7}$$

Moreover, if $m < 2 - 2/d$, then $u(t)$ satisfies

$$t^{d\beta\left(1-\frac{1}{p}\right)}\|u(t) - \mathcal{U}(t;M)\|_p \lesssim (1+t)^{-\frac{\beta}{p}}, \; \forall p,\, 1 \leq p < \infty, \tag{8}$$

and if $m = 2 - 2/d$, then for all $\delta > 0$, $u(t)$ satisfies

$$t^{d\beta\left(1-\frac{1}{p}\right)}\|u(t) - \mathcal{U}(t;M)\|_p \lesssim_\delta (1+t)^{-\frac{\beta}{p}(1-\delta)}, \;\; \forall p,\, 1 \leq p < \infty. \tag{9}$$

Here β is defined in (6) and $\mathcal{U}(x,t;M)$ is the self-similar solution to (3) with mass $M = \epsilon\|f\|_1$ given in (4) or (5).

Theorem 2 (Intermediate Asymptotics II: Infinite Length-Scales)**.** *Let $d \geq 2$ and $\mathcal{K}$ be admissible with $\nabla\mathcal{K}(x) = \mathcal{O}(|x|^{-\gamma})$ as $|x| \to \infty$ for some $\gamma \in [d-1, d]$. If $\gamma = d-1$ then suppose $m \in [1, 2-2/d)$ and otherwise we may take $m \in [1, 2-2/d]$. Let $f \in L^1_+(\mathbb{R}^d; (1+|x|^2)dx) \cap L^\infty(\mathbb{R}^d)$. Then there exists an $\epsilon_0(\|f\|_1, \|f\|_{(2-m)d/2}) > 0$ such that for all $\epsilon < \epsilon_0$, if $u_0 = \epsilon f$ then the weak solution $u(t)$ to (2) is global and satisfies*

$$\|u(t)\|_\infty \lesssim (1+t)^{-d\beta}. \tag{10}$$

Moreover, for all $\delta > 0$, $u(t)$ satisfies

$$t^{d\beta\left(1-\frac{1}{p}\right)}\|u(t) - \mathcal{U}(t;M)\|_p \lesssim_\delta (1+t)^{-\frac{\beta}{p}\min\left(1, 1+\gamma-\beta^{-1}-\delta\right)}, \;\; \forall p,\, 1 \leq p < \infty. \tag{11}$$

Here β and $\mathcal{U}(t,x;M)$ are as above.

Remark 1. Note that $f \in L^1_+(\mathbb{R}^d; (1+|x|^2)dx) \cap L^\infty(\mathbb{R}^d)$ implies $f \log f \in L^1(\mathbb{R}^d)$ by Jensen's inequality for probability measures.

Remark 2. The works of [31] and [3] prove that in many subcritical cases, there exist nontrivial stationary solutions to (2), and thus nothing analogous to Theorems 1 and 2 holds. However, between this work and [31, 3], still not every case is covered. For instance, if $\mathcal{K} \in L^1(\mathbb{R}^d)$ and $2 - 2/d < m < 2$, stationary solutions are only known to exist for sufficiently large mass, and the behavior of smaller solutions is unknown. Moreover, convergence to these stationary solutions is only known in certain cases [28].

Remark 3. The convergence rate in (8) is optimal, as it matches that of the corresponding diffusion equation. Optimality is not known for (9) or (11), however we suspect that these rates are nearly optimal. Note that the convergence rate obtained in (11) reduces to (8) and (9) when $\gamma = d$. Moreover, if $\gamma = d - 1$, then the convergence rate goes to zero as $m \nearrow 2 - 2/d$.

Remark 4. It seems natural to expect convergence in the Euclidean Wasserstein distance, a result which is well-known for (3) and related models [37, 2, 13, 15]. This has been shown in certain cases of the critical PKS model decaying to the self-similar solution of (2) [12].

Remark 5. From the proof, it will be clear that the decay estimates (7) and (10) imply the intermediate asymptotics results with the stated convergence rates. Indeed, the smallness conditions are only required to obtain these estimates. This has a number of implications, which we outline in the following remarks.

Remark 6. For critical problems with $\gamma > d-1$, it is natural to conjecture that (9) or (10) holds for all solutions satisfying $M < M_c$, where M_c is the critical mass [20, 10, 8, 5]. Of course, it is sufficient to provide the decay estimate (7) for all such solutions.

Remark 7. It will be evident from the proofs that ϵ_0 depends on the constants in the Gagliardo-Nirenberg-Sobolev inequality and we do not make an effort to determine the 'optimal' ϵ_0. Moreover, we do not make an effort to prove that ϵ_0 only depends on the $L^{(2-m)d/2}$ norm, which is the controlling norm for supercritical problems [18, 42, 43, 5]. That is, if $\mathcal{K}$ were the Newtonian potential, then (2) has a scaling symmetry which leaves the $L^{(2-m)d/2}$ norm invariant. We remark that global existence and uniform boundedness of solutions to (2) can be shown to only depend on this norm in supercritical cases [18], but here the stronger decay estimates (7) and (10) are required.

Remark 8. The results of [5] suggest that if the kernel $\mathcal{K}$ is less singular than the Newtonian potential at the origin, the $L^{(2-m)d/2}$ norm could, in some cases, possibly be replaced by a weaker one.

Remark 9. We consider only the case of power-law diffusion, however, the estimates (7),(10) hold for (1) provided $A'(z) \geq cz^{m-1}$ for some $c > 0$. Therefore, it is likely possible to apply the methods of [7, 13] to this more general case under some additional structural assumptions.

1.2 Outline of Proof

The proof of Theorems 1 and 2 involves several steps. As mentioned above, we use the entropy dissipation methods of [16, 17, 14] and in particular, the time-dependent rescaling used in [17]. All of the computations will be formal, they can be made rigorous for weak solutions either with a suitable parabolic regularization and passing to the limit, as in for instance [6, 5, 14, 4], or presumably also lifting to strictly positive solutions, as is common in the study of the porous media equation [47].

Following [17], we define $\theta(\tau,\eta)$ such that

$$e^{-d\tau}\theta(\tau,\eta) = u(t,x), \tag{12}$$

with coordinates $e^{\tau}\eta = x$ and $\beta e^{\beta^{-1}\tau} - \beta = t$, where β is given by (6). In what follows we denote $\alpha := d\beta$. In these coordinates, if $u(t,x)$ solves (2) then $\theta(\tau,\eta)$ solves,

$$\partial_\tau \theta = \nabla\cdot(\eta\theta) + \Delta\theta^m - e^{(1-\alpha-\beta)\beta^{-1}\tau}\nabla\cdot\left(\theta(e^{d\tau}\nabla\mathcal{K}(e^{\tau}\cdot) * \theta)\right). \tag{13}$$

Moreover, $\mathcal{U}(t,x;M)$ is stationary in these coordinates, and will be denoted by $F_M(\eta)$. That is (see [17]),

$$\mathcal{U}(t,x;M) = \left(1+\frac{t}{\beta}\right)^{-d\beta} F_M\left(\left(1+\frac{t}{\beta}\right)^{-\beta} x\right) = F_M(\eta). \tag{14}$$

In fact, $F_M(\eta)$ is the unique non-negative solution with mass M to the (degenerate, if $m>1$) elliptic equation

$$0 = \nabla\cdot(\eta\theta) + \Delta\theta^m. \tag{15}$$

In what follows we will refer to F_M as the *ground state Barenblatt solution.* Clearly, ground state solutions are stationary solutions of the homogeneous Fokker-Plank equation

$$\partial_\tau\theta = \nabla\cdot(\eta\theta) + \Delta\theta^m. \tag{16}$$

Therefore, the asymptotic convergence to self-similar profiles of solutions to (3) is equivalent to the convergence to the stationary profiles of (16). This was the fundamental observation made in [17] and is the purpose of the rescaling (12).

A primary step to proving Theorems 1 and 2 is establishing that $\theta(\tau,\eta) \in L^\infty_{\tau,\eta}(\mathbb{R}^+ \times \mathbb{R}^d)$. Note that by the change of variables, this estimate is the decay estimates (7) and (10). This estimate is what allows us to treat the inhomogeneous non-local term in (13) as a vanishing perturbation of (16). The decay estimate $\|u(t)\|_\infty \lesssim t^{-d\beta}$, or equivalently, $\|\theta(\tau)\|_\infty \lesssim 1$, is easily obtained for (3) in the linear case and the classical Aronson-Bénilan estimate proves it in the case $m > 1$ [47]. Clearly, no such analogues are available for (13). However, a standard Alikakos iteration [1] argument can be applied to (13) to prove a uniform bound in the rescaled variables. This method is commonly used in Patlak-Keller-Segel for obtaining $L^\infty(\mathbb{R}^+ \times \mathbb{R}^d)$ bounds of (2) [25, 29, 5, 8, 42, 43, 44]. It is at this step we require the smallness in L^1 and, in potentially supercritical cases, of the critical norm $L^{(2-m)d/2}$. Blow up results in [44, 43, 5, 8] imply that some smallness conditions on these norms are necessary, and similar conditions are taken in [32, 33, 9, 43, 42].

Once we have established $\theta(\tau,\eta) \in L^\infty_{\tau,\eta}(\mathbb{R}^+ \times \mathbb{R}^d)$, we prove that solutions to (13) converge to F_M and estimate the convergence rate in L^1. In fact, these are done together, as the quantitative estimate is direct and removes the need for compactness arguments. The primary step of the entropy dissipation method is an estimate of the decay of the entropy associated to (16). In the case $m = 1$, the entropy is given by,

$$H(\theta) = \int \theta \log \theta d\eta + \frac{1}{2}\int |\eta|^2 \theta d\eta, \tag{17}$$

and the entropy production functional by

$$I(\theta) = \int \theta \left|\nabla \log \theta + \eta\right|^2 d\eta. \tag{18}$$

In the nonlinear case $m > 1$, the corresponding quantities are,

$$H(\theta) = \frac{1}{m-1}\int \theta^m d\eta + \frac{1}{2}\int |\eta|^2 \theta d\eta, \tag{19}$$

and the entropy production functional,

$$I(\theta) = \int u \left|\frac{m}{m-1}\nabla u^{m-1} + \eta\right|^2 d\eta. \tag{20}$$

In the nonlinear case, these entropies were originally introduced for studying (16) in [36, 40]. Both (17) and (19) are displacement convex [34] and in fact, (16) is a gradient flow for (19) or (17) in the Euclidean Wasserstein distance [37, 2], and if $f(\tau,\eta)$ solves (16), then

$$\frac{d}{d\tau}H(f(\tau)) = -I(f(\tau)).$$

For a given mass M, (19) has a unique non-negative minimizer which is the ground state F_M. That is, if we define the relative entropy

$$H(\theta|F_M) = H(\theta) - H(F_M), \tag{21}$$

then $H(\theta|F_M) \geq 0$ with equality if and only if $\theta = F_M$ [17, 39]. In order to estimate a convergence rate, it is therefore sensible to measure how quickly $H(\theta|F_M) \to 0$. Following the method of [17], this is made possible by the following two crucial theorems. The first relates the entropy production functional (20) to the relative entropy (21). This represents a generalization of the Gross logarithmic inequality [21] (see also [39]).

Theorem 3 (Generalized Gross Logarithmic Sobolev Inequality [17, 14, 39, 21]). *Let $f \in L^1_+(\mathbb{R}^d)$ with $\|f\|_1 = M$ and let F_M be the ground state Barenblatt solution with mass M. Then,*

$$H(f|F_M) \leq \frac{1}{2} I(f). \tag{22}$$

For the Fokker-Plank equation (16), Theorem 3 implies $H(\theta(\tau)|F_M) \lesssim e^{-2\tau}$. The (generalized) Csiszar-Kullback inequality [19, 30] relates the relative entropy to the L^1 norm.

Theorem 4 (Csiszar-Kullback Inequality [14]). *Let $f \in L^1_+(\mathbb{R}^d)$ with $\|f\|_1 = M$ and let F_M be the ground state Barenblatt solution with mass M. Then,*

$$\|f - F_M\|_1 \lesssim H(f|F_M)^{\min\left(\frac{1}{2}, \frac{1}{m}\right)}. \tag{23}$$

Note that since we are interested in $1 \leq m \leq 2 - 2/d$, we will only apply the inequality with exponent $1/2$.

To prove Theorems 1 and 2, the purpose of proving $\theta(\tau, \eta) \in L^\infty_{\tau,\eta}(\mathbb{R}^+ \times \mathbb{R}^d)$ is to control the growth of $\|e^{d\tau} \nabla \mathcal{K}(e^\tau \cdot) * \theta\|_\infty$. Ultimately, this provides a bound essentially of the form,

$$\frac{d}{d\tau} H(\theta(\tau)) \leq -I(\theta(\tau)) + C(M, \|\theta\|_{L^\infty_{\tau,\eta}(\mathbb{R}^+ \times \mathbb{R}^d)}) e^{-\gamma\tau},$$

for some $\gamma > 0$ (in reality, it is not quite as clean). Theorem (3) then implies,

$$\frac{d}{d\tau} H(\theta(\tau)|F_M) \leq -2H(\theta(\tau)|F_M) + C(M, \|\theta\|_{L^\infty_{\tau,\eta}(\mathbb{R}^+ \times \mathbb{R}^d)}) e^{-\gamma\tau}.$$

Integrating this and applying Theorem 4 implies,

$$\|\theta - F_M\|_1 \lesssim e^{-\frac{\tau}{2}\min(2,\gamma)},$$

which after rescaling and interpolation against the decay estimates (7),(10), will prove Theorems 1 and 2.

2 Theorem 1

Proof. (Theorem 1: **Intermediate Asymptotics I**) Let $\overline{q} = (2-m)d/2$ and let η, τ and $\theta(\eta, \tau)$ be as defined in §1.2. As detailed above, we first establish that $\theta(\eta, \tau) \in L^\infty_{\tau,\eta}(\mathbb{R}^+ \times \mathbb{R}^d)$ using Alikakos iteration [1] (see also [25, 29, 5, 8, 43, 44, 42]). The first step is to prove the following lemma which allows control over finite L^p norms. In what follows we denote $\theta_0(\eta) := \theta(\eta, 0) = u(x, 0)$.

Lemma 1. *For all $\overline{q} \leq p < \infty$, there exists $C_{\overline{q}} = C_{\overline{q}}(p, M)$ and $C_M = C_M(p, \|\theta_0\|_{\overline{q}})$ such that if $\|\theta_0\|_{\overline{q}} < C_{\overline{q}}$ and $M < C_M$, then $\|\theta(\tau)\|_p \in L^\infty_\tau(\mathbb{R}^+)$.*

Proof. Define $\mathcal{I} = \int \theta^{m-1} \left|\nabla\theta^{p/2}\right|^2 dx$. We estimate the time evolution of $\|\theta\|_p$ using integration by parts, Hölder's inequality and Lemma 4 in the appendix,

$$\begin{aligned}
\frac{d}{d\tau}\|\theta\|_p^p &= -\frac{4mp}{(p+1)^2}\mathcal{I} + (p-1)e^{(1-\alpha-\beta)\beta^{-1}\tau}\int \theta^p \nabla \cdot (e^{d\tau}\nabla\mathcal{K}(e^\tau \cdot) * \theta) d\eta + d(p-1)\|\theta\|_p^p \\
&\leq -C(p)\mathcal{I} + C(p)e^{(1-\alpha-\beta)\beta^{-1}\tau}\|\theta\|_{p+1}^p \|\nabla(e^{d\tau}\nabla\mathcal{K}(e^\tau \cdot) * \theta)\|_{p+1} + C(p)\|\theta\|_p^p \\
&\leq -C(p)\mathcal{I} + C(p)e^{(1-\alpha)\beta^{-1}\tau}\|\theta\|_{p+1}^{p+1} + C(p)\|\theta\|_p^p.
\end{aligned}$$

We bound the second term using the using the homogeneous Gagliardo-Nirenberg-Sobolev inequality (Lemma 3 in appendix),

$$\|\theta\|_{p+1}^{p+1} \lesssim \|\theta\|_{\overline{q}}^{\alpha_2(p+1)} \mathcal{I}^{\alpha_1(p+1)/2},$$

where $\alpha_2 = 1 - \alpha_1(p+m-1)/2$ and

$$\alpha_1 = \frac{2d(\overline{q}-p-1)}{(p+1)\left(\overline{q}(d-2) - d(p+m-1)\right)}.$$

By the definition of $\overline{q}$ we have that,

$$\frac{\alpha_1(p+1)}{2} = \frac{d(\overline{q}-p-1)}{\overline{q}(d-2) - d(p+m-1)} = 1.$$

We also estimate the second term using Lemma 3,

$$\|\theta\|_p^p \lesssim M^{\beta_2 p} \mathcal{I}^{\beta_1 p/2}, \tag{24}$$

where $\beta_2 = 1 - \beta_1 p/2$ and,

$$\frac{\beta_1 p}{2} = \frac{d(p-1)}{2 - d + d(p+m-1)} < 1,$$

by $1 - 2/d < m$. Then applying weighted Young's inequality for products,

$$\frac{d}{d\tau}\|\theta\|_p^p \leq \left(C_1(p) e^{(1-\alpha)\beta^{-1}\tau} \|\theta\|_{\overline{q}}^{\alpha_2(p+1)} - C_2(p)\right) \mathcal{I} + C_3(p) M^{\gamma(p)}, \tag{25}$$

for $\gamma(p) = 2\beta_2 p/(2-\beta_1 p) > 0$. If $m = 2 - 2/d$, then $\overline{q} = 1$ and $1 - \alpha = 0$, therefore by conservation of mass it is possible to choose M sufficiently small such that the first term in (25) is less than $-\delta\mathcal{I}$ for some $\delta > 0$. If $m < 2 - 2/d$, then $\overline{q} > 1$ and we must take advantage of $1 - \alpha < 0$. Note that (25) holds for $p = \overline{q}$; therefore since $1 - \alpha < 0$, a continuity argument establishes that for $\|\theta_0\|_{\overline{q}}$ and M sufficiently small,

$$\|\theta(\tau)\|_{\overline{q}}^{\overline{q}} \leq \|\theta_0\|_{\overline{q}}^{\overline{q}} + C_3(\overline{q}) M^{\gamma(\overline{q})} \tau.$$

Then by (25) for $p > \overline{q}$, if M and $\|\theta_0\|_{\overline{q}}$ additionally satisfy

$$C_1(p) e^{(1-\alpha)\beta^{-1}\tau} (C_3(\overline{q}) M^{\gamma(\overline{q})}\tau + \|\theta_0\|_{\overline{q}}^{\overline{q}})^{\alpha_2(p+1)/\overline{q}} - C_2(p) < -\delta,$$

for all $\tau > 0$, then the first term is less than $-\delta I$. By $1 - \alpha < 0$ we may always choose M and $\|\theta_0\|_{\overline{q}}$ such that this is possible. Therefore, whether $\overline{q} > 1$ or $\overline{q} = 1$, for small initial data in the suitable sense, we have

$$\frac{d}{d\tau}\|\theta\|_p^p \leq -\delta\mathcal{I} + C(M,p).$$

Using (24) and Young's inequality for products, we have a lower bound on $\mathcal{I}$,

$$\|\theta\|_p^p - C(M) \leq \mathcal{I}.$$

This proves,

$$\frac{d}{d\tau}\|\theta\|_p^p \leq -\delta\|\theta\|_p^p + C(M,p),$$

which immediately concludes the lemma with $\|\theta\|_p^p \leq \max(\|\theta_0\|_p^p, C(M,p)\delta^{-1})$. □

Alikakos iteration [1] is a standard method for using a result such as Lemma 1 to imply a result of the following form.

Lemma 2. *There exists $C_{\overline{q}} = C_{\overline{q}}(M)$ and $C_M = C_M(\|\theta_0\|_{\overline{q}})$ such that if $\|\theta_0\|_{\overline{q}} < C_{\overline{q}}$ and $M < C_M$, then $\|\theta(\tau)\|_\infty \in L^\infty_\tau(\mathbb{R}^+)$.*

Proof. Standard iteration implies $\|\theta(\tau)\|_\infty \in L^\infty_\tau(\mathbb{R}^+)$, provided

$$\vec{v} := e^{(1-\alpha-\beta)\beta^{-1}\tau} e^{d\tau} \nabla\mathcal{K}(e^\tau \cdot) * \theta \in L^\infty_{\tau,\eta}(\mathbb{R}^+ \times \mathbb{R}^d).$$

See [25, 5, 29, 42, 43]. For instance, an iteration lemma due to Kowalczyk [29] may be extended easily to the case $\mathbb{R}^d$, $d \geq 2$ and to include the $\nabla \cdot (\eta\theta)$ term in (13) [5].

Fix $p > d$. Then by Lemma 1, for sufficiently small M and $\|\theta_0\|_{\overline{q}}$, $\|\theta(\tau)\|_p \in L^\infty_\tau(\mathbb{R}^+)$. Therefore by Lemma 4 in the appendix,

$$\|\nabla\vec{v}\|_p = \|e^{(1-\alpha-\beta)\beta^{-1}\tau}\nabla\left(e^{d\tau}\nabla\mathcal{K}(e^\tau \cdot) * \theta\right)\|_p \lesssim e^{(1-\alpha)\beta^{-1}\tau}\|\theta\|_p \lesssim e^{(1-\alpha)\beta^{-1}\tau}.$$

Moreover, by $\nabla\mathcal{K} \in L^1(\mathbb{R}^d)$,

$$\|\vec{v}\|_p \leq e^{(1-\alpha-\beta)\beta^{-1}\tau}\|\theta\|_p \lesssim e^{(1-\alpha-\beta)\beta^{-1}\tau}.$$

Since $1 - \alpha \leq 0$, Morrey's inequality implies $\vec{v} \in L^\infty_{\tau,\eta}(\mathbb{R}^+ \times \mathbb{R}^d)$ and the lemma follows. □

By Lemma 2 and the definition of τ,

$$\|u(t)\|_{L^\infty_x(\mathbb{R}^d)} = e^{-d\tau}\|\theta\|_{L^\infty_\eta(\mathbb{R}^d)} \lesssim (1+t)^{-d\beta},$$

establishing (7).

Now that the requisite decay estimate has been established, we proceed by estimating the decay of the relative entropy (21). By Young's inequality, $\nabla\mathcal{K} \in L^1(\mathbb{R}^d)$ and Lemma 2,

$$\|e^{d\tau}\nabla\mathcal{K}(e^\tau \cdot) * \theta\|_\infty \leq \|\nabla\mathcal{K}\|_1\|\theta\|_\infty \lesssim 1. \tag{26}$$

We first settle the case $m > 1$. By a standard computation, (26) and Cauchy-Schwarz, for all $\delta > 0$,

$$\begin{aligned}\frac{d}{d\tau}H(\theta(\tau)|F_M) &= -I(\theta) + e^{(1-\alpha-\beta)\beta^{-1}\tau}\int \nabla\left(\frac{1}{m-1}\theta^m + \frac{1}{2}|\eta|^2\right)\cdot \theta e^{d\tau}\nabla\mathcal{K}(e^\tau \cdot) * \theta d\eta \\ &\leq -I(\theta) + e^{(1-\alpha-\beta)\beta^{-1}\tau}I(\theta)^{1/2}\left(\int \theta\left|e^{d\tau}\nabla\mathcal{K}(e^\tau \cdot) * \theta\right|^2 d\eta\right)^{1/2} \\ &\leq (1 - e^{-2\delta\tau})I(\theta) + Ce^{(2-2\alpha-2\beta)\beta^{-1}\tau + 2\delta\tau}.\end{aligned}$$

Let $\gamma(\delta) := (2\alpha + 2\beta - 2)\beta^{-1} - 2\delta > 0$. By Theorem 3 we therefore have,

$$\frac{d}{d\tau}H(\theta(\tau)|F_M) \leq -2I(1 - e^{-2\delta\tau})H(\theta|F_M) + Ce^{-\gamma\tau}. \tag{27}$$

Solving the differential inequality (27) implies,

$$H(\theta|F_M) \lesssim e^{-\tau\min(2,\gamma(\delta))}.$$

Now by Theorem 4,

$$\|\theta(\tau) - F_M\|_1 \lesssim e^{-\frac{\tau}{2}\min(2,\gamma(\delta))}.$$

Re-writing in terms of x and t and using (14),

$$\|u(t) - \mathcal{U}(t;M)\|_1 \lesssim (1+t)^{-\frac{\beta}{2}\tau \min(2,\gamma(\delta))}.$$

If $m < 2-2/d$, it can be verified that $\delta > 0$ may always be chosen small enough such that $2 < \gamma(\delta)$. If instead $m = 2-2/d$, then $2d+2-2\beta^{-1} = 2$. This establishes (9) in the case $p = 1$. Interpolation against (7) completes the proof.

We now settle the case $m = 1$. The time evolution of the relative entropy is similar to above. By (26) and Cauchy-Schwarz, for all $\delta > 0$,

$$\begin{aligned}
\frac{d}{d\tau} H(\theta(\tau)|F_M) &= -I(\theta) + e^{(1-\alpha-\beta)\beta^{-1}\tau} \int \nabla \left(\log\theta + \frac{1}{2}|\eta|^2\right) \cdot \theta e^{d\tau} \nabla\mathcal{K}(e^\tau \cdot) * \theta d\eta \\
&\leq -I(\theta) + e^{(1-\alpha-\beta)\beta^{-1}\tau} I(\theta)^{1/2} \left(\int \theta \left|e^{d\tau}\nabla\mathcal{K}(e^\tau\cdot) * \theta\right|^2 d\eta\right)^{1/2} \\
&\leq (1-e^{-\delta\tau}) I(\theta) + Ce^{(2-2\alpha-2\beta)\beta^{-1}\tau + \delta\tau}.
\end{aligned}$$

The rest of the proof follows similarly to the case $m > 1$ using Theorems 3 and 4. This concludes the proof of Theorem 1. □

3 Theorem 2

The proof of Theorem 2 is a technical refinement of Theorem 1.

Proof. (Theorem 2: **Intermediate Asymptotics II**) In order to properly extend Theorem 1 we must estimate the quantities $\|e^{d\tau}\nabla\mathcal{K}(e^\tau\cdot) * \theta\|_p$ appearing in (26) and the proof of Lemma 2. However, $\nabla\mathcal{K} \notin L^1(\mathbb{R}^d)$ and Young's inequality is not sufficient; indeed, $\|e^{d\tau}\nabla\mathcal{K}(e^\tau\cdot)*\theta\|_p$ will not be bounded uniformly in time. We separately estimate the growth of the quantities $\|\lambda^d\nabla\mathcal{K}(\lambda\cdot)\mathbf{1}_{B_1(0)}\|_1$ and $\|\lambda^d\nabla\mathcal{K}(\lambda\cdot)\mathbf{1}_{\mathbb{R}^d\setminus B_1(0)}\|_p$ as $\lambda \to \infty$. Using $|\nabla\mathcal{K}(x)| \lesssim |x|^{-\gamma}$ for sufficiently large $|x|$, if $\gamma < d$, then for large λ,

$$\begin{aligned}
\int \lambda^d |\nabla\mathcal{K}(\lambda y)| \mathbf{1}_{B_1(0)}(|y|) dy &= \int_{|y|\leq\lambda} |\nabla\mathcal{K}(y)|\, dy \\
&= \int_{S^{d-1}} \int_0^\lambda |\nabla\mathcal{K}(\rho\omega)|\, r\rho^{d-1} d\rho d\omega \\
&\lesssim 1 + \lambda^{d-\gamma}.
\end{aligned} \tag{28}$$

Similarly, if $\gamma = d$, then for large λ,

$$\int \lambda^d |\nabla\mathcal{K}(\lambda y)| \mathbf{1}_{B_1(0)}(|y|) dy \lesssim 1 + \log\lambda. \tag{29}$$

On the other hand, for $\infty > q > d/(d-1)$ and λ sufficiently large, since $\gamma \geq d-1$,

$$\begin{aligned}
\int \lambda^{qd} |\nabla\mathcal{K}(\lambda y)|^q \mathbf{1}_{\mathbb{R}^d\setminus B_1(0)}(|y|) dy &= \int_{|y|\geq\lambda} \lambda^{qd-d} |\nabla\mathcal{K}(y)|^q\, dy \\
&= \lambda^{qd-d} \int_{S^{d-1}} \int_\lambda^\infty |\nabla\mathcal{K}(\rho\omega)|^q \rho^{d-1} d\rho d\omega \\
&\lesssim \lambda^{q(d-\gamma)}.
\end{aligned} \tag{30}$$

Similarly,

$$\sup_{|x|\geq 1}\left|\lambda^d \nabla\mathcal{K}(\lambda x)\right| \lesssim 1+\lambda^{d-\gamma}. \tag{31}$$

The proof of Lemma 1 extends to cover Theorem 2 since Lemma 4 holds by admissibility of $\mathcal{K}$. Lemma 2 extends provided we can bound $\vec{v} := e^{(1-\alpha-\beta)\beta^{-1}\tau}e^{d\tau}\nabla\mathcal{K}(e^\tau\cdot)*\theta$ in $L^\infty_\eta(\mathbb{R}^d)$ uniformly in time. Indeed, fix $p > d$. Then for M and $\|\theta_0\|_{\overline{q}}$ sufficiently small, we have by Lemma 1, $\|\theta(\tau)\|_p \in L^\infty_\tau(\mathbb{R}^+)$. By Lemma 4,

$$\|\nabla\vec{v}\|_p \lesssim e^{(1-\alpha)\beta^{-1}\tau}\|\theta\|_p \lesssim e^{(1-\alpha)\beta^{-1}\tau}.$$

If $\gamma < d$ then, by (28), (30) and Young's inequality, for some $d/(d-1) < q \leq p$,

$$\begin{aligned}\|\vec{v}\|_q &\leq e^{(1-\alpha-\beta)\beta^{-1}\tau}\left(\|e^{d\tau}\nabla\mathcal{K}(e^\tau\cdot)\mathbf{1}_{B_1(0)}*\theta\|_q + \|e^{d\tau}\nabla\mathcal{K}(e^\tau\cdot)\mathbf{1}_{\mathbb{R}^d\setminus B_1(0)}*\theta\|_q\right)\\ &\lesssim e^{(1-\alpha-\beta)\beta^{-1}\tau}\left(e^{(d-\gamma)\tau}\|\theta\|_q + e^{(d-\gamma)\tau}M\right).\\ &\lesssim e^{(1-\beta-\gamma\beta)\beta^{-1}\tau}.\end{aligned}$$

Similarly if $\gamma = d$, then (29) and (30) imply,

$$\|\vec{v}\|_q \lesssim e^{(1-\beta-\gamma\beta)\beta^{-1}\tau} + \tau e^{(1-\beta-\gamma\beta)\beta^{-1}\tau}.$$

Since $1-\beta-\gamma\beta \leq 0$ and $1-\alpha\leq 0$, by Morrey's inequality we may conclude $\vec{v}\in L^\infty_{\tau,\eta}(\mathbb{R}^+\times\mathbb{R}^d)$, thus allowing Lemma 2 to apply under the hypotheses of Theorem 2. Re-writing in terms of x and t, this implies (10).

To complete the proof of Theorem 2, we estimate the decay of the relative entropy (21). The proof of Theorem 1 used the estimate (26). Here we use (31) and (28) to imply, if $\gamma < d$,

$$\begin{aligned}\|e^{d\tau}\nabla\mathcal{K}(e^\tau\cdot)*\theta\|_\infty &\leq \left(\|e^{d\tau}\nabla\mathcal{K}(e^\tau\cdot)\mathbf{1}_{B_1(0)}*\theta\|_\infty + \|e^{d\tau}\nabla\mathcal{K}(e^\tau\cdot)\mathbf{1}_{\mathbb{R}^d\setminus B_1(0)}*\theta\|_\infty\right)\\ &\lesssim e^{(d-\gamma)\tau}(\|\theta\|_\infty + M) \lesssim e^{(d-\gamma)\tau}.\end{aligned} \tag{32}$$

Similarly, if $\gamma = d$ then, for all $\delta > 0$, for τ sufficiently large,

$$\|e^{d\tau}\nabla\mathcal{K}(e^\tau\cdot)*\theta\|_\infty \lesssim \tau \leq e^{\delta\tau}.$$

The growth of (32) in time is the source of the degraded convergence rate observed in (11). Indeed, computing the decay of the relative entropy (with linear or nonlinear diffusion) as above with (32),

$$\begin{aligned}\frac{d}{d\tau}H(\theta(\tau)|F_M) &= \leq -I(\theta) + e^{(1-\alpha-\beta)\beta^{-1}\tau}I(\theta)^{1/2}\left(\int\theta\left|e^{d\tau}\nabla\mathcal{K}(e^\tau\cdot)*\theta\right|^2 d\eta\right)^{1/2}\\ &\leq (1-e^{-2\delta\tau})I(\theta) + Ce^{(2(1-\alpha-\beta)\beta^{-1}+2(d-\gamma)+2\delta)\tau}.\end{aligned}$$

As before, Theorems 3 and 4 imply,

$$\|\theta(\tau)-F_M\|_1 \lesssim e^{-\tau\min\left(1,1+\gamma-\beta^{-1}-\delta\right)}.$$

Re-writing in terms of x and t and interpolating against (10) completes the proof. The corresponding argument follows also for $\gamma = d$, absorbing the mild growth of $\|e^{d\tau}\mathcal{K}(e^\tau\cdot)*\theta\|_\infty$ into the δ already introduced. □

4 Acknowledgments

The author would like to thank Andrea Bertozzi, Thomas Laurent and Nancy Rodríguez for helpful discussions and guidance, and to Inwon Kim for helpful discussions as well suggesting the problem. This work was in part supported by NSF grant DMS-0907931.

5 Appendix

Lemma 3 (Homogeneous Gagliardo-Nirenberg-Sobolev)**.** *Let $d \geq 2$ and $f : \mathbb{R}^d \to \mathbb{R}$ satisfy $f \in L^p \cap L^q$ and $\nabla f^k \in L^r$. Moreover let $1 \leq p \leq rk \leq dk$, $k < q < rkd/(d-r)$ and*

$$\frac{1}{r} - \frac{k}{q} - \frac{s}{d} < 0. \tag{33}$$

Then there exists a constant C_{GNS} which depends on s, p, q, r, d such that

$$\|f\|_{L^q} \leq C_{GNS} \|f\|_{L^p}^{\alpha_2} \| \, |\nabla|^s \, f^k \|_{L^r}^{\alpha_1}, \tag{34}$$

where $0 < \alpha_i$ satisfy

$$1 = \alpha_1 k + \alpha_2, \tag{35}$$

and

$$\frac{1}{q} - \frac{1}{p} = \alpha_1 (\frac{-s}{d} + \frac{1}{r} - \frac{k}{p}). \tag{36}$$

The following lemma verifies that the distributions defined by the second derivatives of admissible kernels behave as expected under mass-invariant scalings.

Lemma 4. *Let $\mathcal{K}$ be admissible. Then $\forall \, p$, $1 < p < \infty$, $u \in L^p$ and $t > 0$, we have*

$$\|\nabla \left(t^d \nabla \mathcal{K}(t\cdot) * u \right) \|_p \lesssim_p t \|u\|_p. \tag{37}$$

Proof. We take the second derivative in the sense of distributions. Let $\phi \in C_c^\infty$, then by the dominated convergence theorem,

$$\begin{aligned}
\int t^d \partial_{x_i} \mathcal{K}(tx) \partial_{x_j} \phi(x) dx &= \lim_{\epsilon \to 0} \int_{|x| \geq \epsilon} t^d \partial_{x_i} \mathcal{K}(tx) \partial_{x_j} \phi(x) dx \\
&= -t \lim_{\epsilon \to 0} \int_{|x| = \epsilon} t^{d-1} \partial_{x_y} \mathcal{K}(tx) \frac{x_j}{|x|} \phi(x) dS - t \mathrm{PV} \int t^d \partial_{x_i, x_j} \mathcal{K}(tx) \phi(x) dx.
\end{aligned}$$

By $\nabla \mathcal{K} \in L^{d/(d-1), \infty}$, we have $\nabla \mathcal{K} = \mathcal{O}(|x|^{1-d})$ as $x \to 0$. Therefore for ϵ sufficiently small, there exists $C > 0$ such that,

$$\begin{aligned}
\left| t \int_{|x| = \epsilon} t^{d-1} \partial_{x_i} \mathcal{K}(tx) \frac{x_j}{|x|} \phi(x) dS \right| &\leq C t \int_{|x| = \epsilon} |x|^{1-d} \left| \phi(x) \right| dS \\
&= C t \int_{|x| = 1} |\epsilon x|^{1-d} \left| \phi(\epsilon x) \right| \epsilon^{d-1} dS = C t \left| \phi(0) \right|.
\end{aligned}$$

The admissibility conditions **(R)**,**(BD)** and **(KN)** are sufficient to apply the Calderón-Zygmund inequality [Theorem 2.2 [41]], which implies that the principal value integral in the second term is a bounded linear operator on L^p for all $1 < p < \infty$. The operator norms, which are the implicit constants in (37), only depend on the bound in **(BD)** and on the smoothness condition

$$\int_{|x|>2|y|} |K(x-y) - K(x)|\, dx \leq B.$$

Both of these conditions are clearly left invariant under the rescaling in (37) and this concludes the proof. $\square$

References

[1] N.D. Alikakos. L^p bounds of solutions to reaction-diffusion equations. *Comm. Part. Diff. Eqn.*, 4:827–868, 1979.

[2] L.A. Ambrosio, N. Gigli, and G. Savaŕe. *Gradient flows in metric spaces and in the space of probability measures.* Lectures in Mathematics, Birkhäuser, 2005.

[3] J. Bedrossian. Global minimizers for free energies of subcritical aggregation equations with degenerate diffusion. *Preprint,* arXiv:1009.5370, 2010.

[4] J. Bedrossian and N. Rodríguez. Local and global well-posedness for aggregation equations and Patlak-Keller-Segel models in $\mathbb{R}^d$. *in preparation*, 2010.

[5] J. Bedrossian, N. Rodríguez, and A.L. Bertozzi. Local and global well-posedness for aggregation equations and Patlak-Keller-Segel models with degenerate diffusion. *Preprint,* arXiv:1009.2674, 2010.

[6] A.L. Bertozzi and D. Slepcev. Existence and uniqueness of solutions to an aggregation equation with degenerate diffusion. *To appear in Comm. Pure. Appl. Anal.*, 2010.

[7] P. Biler, J. Dolbeault, and M.J. Esteban. Intermediate asymptotics in L^1 for general nonlinear diffusion equations. *Appl. Math. Let.*, 15:101–107, 2002.

[8] A. Blanchet, J. Carrillo, and P. Laurencot. Critical mass for a Patlak-Keller-Segel model with degenerate diffusion in higher dimensions. *Calc. Var.*, 35:133–168, 2009.

[9] A. Blanchet, J. Dolbeault, M. Escobedo, and J. Fernández. Asymptotic behavior for small mass in the two-dimensional parabolic-elliptic Keller-Segel model. *J. Math. Anal. Appl.*, 361:533–542, 2010.

[10] A. Blanchet, J. Dolbeault, and B. Perthame. Two-dimensional Keller-Segel model: Optimal critical mass and qualitative properties of the solutions. *E. J. Diff. Eqn*, 2006(44):1–33, 2006.

[11] Silvia Boi, Vincenzo Capasso, and Daniela Morale. Modeling the aggregative behavior of ants of the species *polyergus rufescens*. *Nonlinear Anal. Real World Appl.*, 1(1):163–176, 2000. Spatial heterogeneity in ecological models (Alcalá de Henares, 1998).

[12] V. Calvez and J.A. Carrillo. Refined asymptotics for the subcritical Keller-Segel system and related functional inequalities. *Preprint,* arXiv:1007.2837, 2010.

[13] J.A. Carrillo, M. Di Francesco, and G. Toscani. Intermediate asymptotics beyond homoegeneity and self-similarity: long time behavior for $u_t = \delta\phi(u)$. *Arch. Rat. Mech. Anal.*, 180:127–149, 2006.

[14] J.A. Carrillo, A. Jüngel, P.A. Markowich, G. Toscani, and A. Unterreiter. Entropy dissipation methods for degenerate parabolic problems and generalized Sobolev inequalities. *Montash. Math.*, 133:1–82, 2001.

[15] J.A. Carrillo, R.J. McCann, and C. Villani. Contractions in the 2-Wasserstein length space and thermalization of granular media. *Arch. Rat. Mech. Anal.*, 179:217–263, 2006.

[16] J.A. Carrillo and G. Toscani. Exponential convergence toward equilibrium for homogeneous Fokker-Plank-type equations. *Math. Meth. Appl. Sci*, 21:1269–1286, 1998.

[17] J.A. Carrillo and G. Toscani. Asymptotic L^1 decay of solutions of the porous medium equation to self-similarity. *Ind. Univ. Math. J.*, 49, 2000.

[18] L. Corrias, B. Perthame, and H. Zaag. Global solutions of some chemotaxis and angiogenesis systems in high space dimensions. *Milan J. Math.*, 72:1–28, 2004.

[19] I. Csiszar. Information-type measures of difference of probability distributions and indirect observation. *Stud. Sci. Math. Hung.*, 2:299–318, 1967.

[20] J. Dolbeault and B. Perthame. Optimal critical mass in the two dimensional Keller-Segel model in $\mathbb{R}^2$. *C. R. Acad. Sci. Paris*, 339:611–616, 2004.

[21] L. Gross. Logrithmic Sobolev inequalities. *Amer. J. of Math.*, 97:1061–1083, 1975.

[22] E. M. Gurtin and R.C McCamy. On the diffusion of biological populations. *Math. Biosci.*, 33:35–47, 1977.

[23] T. Hillen and K. J. Painter. A user's guide to PDE models for chemotaxis. *J. Math. Biol.*, 58(1-2):183–217, 2009.

[24] D. Hortsmann. From 1970 until present: the Keller-Segel model in chemotaxis and its consequences I. *Jahresberichte der DMV*, 105:103–165, 2003. No. 3.

[25] W. Jager and S. Luckhaus. On explosions of solutions to a system of partial differntial equations modelling chemotaxis. *Trans. Amer. Math. Soc.*, 329(2), 1992.

[26] S. Kamin and J.L. Vázquez. Fundamental solutions and asymptotic behavior for the p-Laplacian equation. *Ref. Mat. Iberoamericana*, 4:34–45, 1988.

[27] E. F. Keller and L.A. Segel. Model for chemotaxis. *J. Theor. Biol.*, 30:225–234, 1971.

[28] I. Kim and Y. Yao. *personal communication.*

[29] R. Kowalczyk. Preventing blow-up in a chemotaxis model. *J. Math. Anal. Appl.*, 305:566–588, 2005.

[30] S. Kullback. A lower bound for discrimination information in terms of variation. *IEEE Trans. Inf.*, 4:126–127, 1967.

[31] P.L. Lions. The concentration-compactness principle in calculus of variations. the locally compact case, part 1. *Ann. Inst. Henri. Poincare, Anal. non lin*, 1(2):109–145, 1984.

[32] S. Luckhaus and Y. Sugiyama. Large time behavior of solutions in super-critical case to degenerate Keller-Segel systems. *Math. Model. Numer. Anal.*, 40:597–621, 2006.

[33] S. Luckhaus and Y. Sugiyama. Asymptotic profile with optimal convergence rate for a parabolic equation of chemotaxis in super-critical cases. *Indiana Univ. Math. J.*, 56(3):1279–1297, 2007.

[34] R.J. McCann. A convexity principle for interacting gases. *Adv. Math.*, 128:153–179, 1997.

[35] Paul A. Milewski and Xu Yang. A simple model for biological aggregation with asymmetric sensing. *Comm. Math. Sci.*, 6(2):397–416, 2008.

[36] W.J. Newman. A Lyapunov functional for the evolution to the porous medium equation to self-similarity. I. *J. Math. Phys*, 25:3120–3123, 1984.

[37] F. Otto. The geometry of dissipative evolution equations: the porous medium equation. *Comm. Part. Diff. Eqn.*, 26(1):101–174, 2001.

[38] Clifford S. Patlak. Random walk with persistence and external bias. *Bull. Math. Biophys.*, 15:311–338, 1953.

[39] M. Del Pino and J. Dolbeault. Best constants for Galiardo-Nirenberg inequalities and applications to nonlinear diffusions. *J. Math. Pures. Appl.*, 81:847–875, 2002.

[40] J. Ralston. A Lyapunov functional for the evolution to the porous medium equation to self-similarity. II. *J. Math. Phys*, 25:3124–3127, 1984.

[41] E. Stein. *Singualar Integrals and Differentiability Properties of Functions*. Princeton University Press, 1970.

[42] Y. Sugiyama. Global existence in sub-critical cases and finite time blow-up in super-critical cases to degenerate Keller-Segel systems. *Diff. Int. Eqns.*, 19(8):841–876, 2006.

[43] Y. Sugiyama. Application of the best constant of the Sobolev inequality to degenerate Keller-Segel models. *Adv. Diff. Eqns.*, 12(2):121–144, 2007.

[44] Y. Sugiyama. The global existence and asymptotic behavior of solutions to degenerate to quasi-linear parabolic systems of chemotaxis. *Diff. Int. Eqns.*, 20(2):133–180, 2007.

[45] Chad M. Topaz, Andrea L. Bertozzi, and Mark A. Lewis. A nonlocal continuum model for biological aggregation. *Bull. Math. Biol.*, 68(7):1601–1623, 2006.

[46] G. Toscani. Entropy production and the rate of convergence to equillibrium for the Fokker-Plank equation. *Quart. Appl. Math.*, 57:521–541, 1999.

[47] J.L. Vázquez. *The Porous Medium Equations*. Clarendon Press, Oxford, 2007.