L^∞ ESTIMATES FOR THE JKO SCHEME
IN PARABOLIC-ELLIPTIC KELLER-SEGEL SYSTEMS

JOSÉ-ANTONIO CARRILLO, FILIPPO SANTAMBROGIO

ABSTRACT. We prove L^∞ estimates on the densities that are obtained via the JKO scheme for a general form of a parabolic-elliptic Keller-Segel type system, with arbitrary diffusion, arbitrary mass, and in arbitrary dimension. Of course, such an estimate blows up in finite time, a time proportional to the inverse of the initial L^∞ norm. This estimate can be used to prove short-time well-posedness for a number of equations of this form regardless of the mass of the initial data. The time of existence of the constructed solutions coincides with the maximal time of existence of Lagrangian solutions without the diffusive term by characteristic methods.

1. INTRODUCTION

We consider in this work a general version of the so-called parabolic-elliptic Keller-Segel system, i.e.

(1.1)
$$
\begin{cases}
\partial_t \rho + \chi \nabla \cdot (\rho \nabla u) - \nabla \cdot (\rho \nabla f'(\rho)) = 0 & \text{in } [0, T] \times \Omega, \\
-\Delta u = \rho & \text{in } \Omega \text{ for every } t \in [0, T], \\
u = 0 & \text{on } \partial\Omega \text{ for every } t \in [0, T], \\
\rho(\nabla u - \nabla f'(\rho)) \cdot \mathbf{n} = 0 & \text{on } [0, T] \times \partial\Omega, \\
\rho(0, x) = \rho_0(x) & \text{in } x \in \Omega.
\end{cases}
$$

This classical system [27, 25, 24, 15] models the evolution of a population ρ of bacteria, which diffuse and are advected by a drift, being attracted by the high values of a chemoattractant substance whose concentration is given by u. The term $\nabla \cdot (\rho \nabla f'(\rho))$ is a linear diffusion $\Delta\rho$ whenever $f(s) = s \log s$, or acts as a nonlinear diffusion $\Delta\Psi(\rho)$ with diffusion coefficient $\Psi'(s) = s f''(s)$ in general. The chemoactractant density u is ruled by the distribution of the bacteria itself since they produce themselves the chemical/protein to which they are attracted to as a way of cell signalling. The initial density ρ_0 is supposed to be a probability density, which is always possible up to a suitable scaling. The constant $\chi > 0$ and the convex and superlinear function $f : \mathbb{R}_+ \to \mathbb{R}$ are parameters of the model. Note that the fact that the values $u(t, \cdot)$ only depend on the values $\rho(t, \cdot)$ at the same time is a strong modeling assumption, which corresponds to the fact that adjusement in the distribution of the chemoattractant u occur at a time scale much faster than the movement of the bacteria. This gives rise to the above parabolic-elliptic system, while different models involving $\partial_t u$ would be possible, and have been studied in, for instance [16, 8] and the references therein.

We concentrate, for simplicity, on the case where Ω is a bounded convex set, and we impose Dirichlet boundary conditions on u, together with no-flux boundary conditions on ρ through $\partial\Omega$. The results of the paper can easily be adapted to the case where the Dirichlet conditions on u are replaced by Neumann conditions by taking $-\Delta u = \rho - c$ (for $c = 1/|\Omega|$) or $-\Delta u + u = \rho$, with $\nabla u \cdot \mathbf{n} = 0$ on $\partial\Omega$. Also the whole space case $\Omega = \mathbb{R}^d$ can be treated. Both of these variants will be addressed at the end of Section 2 but, for simplicity of the exposition, the main theorems will be stated in the Dirichlet case in bounded domains.

The PDE system (1.1) is known to be the gradient-flow of the energy

$$
J(\rho) := \int f(\rho(x)) \mathrm{d}x - \frac{\chi}{2} \int |\nabla u(x)|^2 \mathrm{d}x,
$$

where u depends on ρ through $-\Delta u = \rho$ with $u = 0$ on $\partial\Omega$, with respect to the Wasserstein metric W_2. For the notion of gradient flows and of gradient flows for this metric, we refer to [37, 1, 35].

This means, among other things, that a way to produce a solution to (1.1) is to fix a time step $\tau > 0$ and then iteratively solve the minimization problem

$$\rho_{k+1} \in \text{argmin}\left\{J(\rho) + \frac{W_2^2(\rho, \rho_k)}{\tau} \ : \ \rho \in \mathcal{P}(\Omega)\right\},$$

where the minimum is taken over all probability densities $\rho \in \mathcal{P}(\Omega)$, which must be absolutely continuous because of the superlinear term $\int f(\rho)$. One defines a W_2-continuous curve in $\mathcal{P}(\Omega)$ setting $\rho^\tau(t) = \rho_k$ for $t = k\tau$ and interpolating using Wasserstein geodesics in the intervals $(k\tau, (k + 1)\tau)$. Then, it is possible to prove that the limit as $\tau \to 0$ provides a solution of (1.1). This discrete-in-time scheme to provide a solution is known under the name of Jordan-Kinderlehrer-Otto scheme. It was first introduced in [26] for the linear Fokker-Planck equation, and then subsequently used for the porous medium equation [32].

Yet, in this very precise case some difficulties arise due to the minus sign in front of the term $\int |\nabla u|^2$. This energy, in terms of ρ, corresponds to a squared norm in the dual of $H_0^1(\Omega)$ and it is clear that semicontinuity and bounds from below may fail when it is taken with a negative sign. On the other hand, it is an energy of order -1, in the sense that it acts as an L^2 norm of antiderivatives of ρ, and one can expect that the term $\int f(\rho)$ can somehow compensate it. This turns out to depend on f, on the dimension, and on χ. Indeed, in dimension $d = 2$, in the case of linear diffusion, i.e. $f(t) = t \log t$, it is proven that the above minimization problem has a solution, and the JKO scheme converges, as soon as $\chi < 8\pi$, see [11]. In the critical case $\chi = 8\pi$ solutions exist and they can also be constructed by variants of the JKO scheme, see [10, 7]. In higher dimension $d > 2$, this critical mass phenomena exists corresponding to other critical energies, of the form $f(t) = t^m$ with $m = 2(d - 1)/d$, see [9].

In this paper we will consider a slightly modified JKO scheme, adding a constraint, depending on τ and disappearing as $\tau \to 0$, on the L^∞ norm of ρ in the above minimization problem. We then prove that a solution at each time step exists, and that it satisfies an L^∞ bound that makes the constraint unbinding, and that we can iterate. Unfortunately, this estimate will explode in finite time T, with $T = (\chi\|\rho_0\|_{L^\infty})^{-1}$. Yet, the good news is the fact that this bound will be completely independent of the choice of f and of χ. Essentially, as the reader can easily guess, we are providing L^∞ bounds that hold for the system without diffusion, as in [33, 5], which is compatible with an estimate exploding in time. The interest of our work mainly resides in its discrete-in-time nature, namely the fact that the estimate is proven along the JKO scheme. Let us finally mention that L^∞ bounds at the level of the JKO scheme were recently shown for fractional porous medium equations [28] with totally different goals and strategy.

Since uniqueness results for the Keller-Segel system exist under L^∞ assumptions [18, 29] for certain nonlinearities, see [2, 20, 30] for related results, then this L^∞ bound allows for existence (and uniqueness with the additional assumptions in [18, 29]) results in short time. We prove these existence results in Section 3, and that they are of interest because they are very general in terms of the diffusion nonlinearity. They provide short-time existence even above critical mass. In short, we show in this work that solutions of the Keller-Segel model exists, and they are unique under additional assumptions on the nonlinearities, at least for the time of existence and uniqueness of Lagrangian bounded solutions of the model without diffusion.

2. L^∞-ESTIMATES

To make our strategy precise, we will consider the following minimization problem, for fixed $\tau > 0$, and for $g \in L^\infty(\Omega)$:

$$(2.1) \qquad \min\left\{J(\rho) + \frac{W_2^2(\rho, g)}{\tau} \ : \ \rho \in \mathcal{P}(\Omega), \rho \le M := \frac{1}{\chi\tau}\right\}$$

We will prove the following theorem

Theorem 1. *Problem* (2.1) *admits at least one solution; moreover, if Ω is convex,* $\inf_{t>0} tf''(t) > 0$ *and* $\log g \in C^{0,\alpha}(\Omega)$, *then for every $\lambda > 1$ there exists a constant $c_0 = c_0(\lambda, \chi, d)$ such that, if $\tau\|g\|_{L^\infty} \leq c_0$, then any solution ρ to this problem satisfies*

$$\|\rho\|_{L^\infty} \leq \|g\|_{L^\infty} \frac{1}{1 - \lambda\tau\chi\|g\|_{L^\infty}},$$

or, equivalently

$$(2.2) \qquad \|\rho\|_{L^\infty}^{-1} \geq \|g\|_{L^\infty}^{-1} - \lambda\tau\chi.$$

Removing the assumptions $\inf_{t>0} tf''(t) > 0$ *and* $\log g \in C^{0,\alpha}(\Omega)$, *the same L^∞ estimate is true for at least one particular solution of* (2.1).

Proof. First, we prove the existence of an optimizer ρ. We take a minimizing sequence $\rho_n \in \mathcal{P}(\Omega) \cap L^\infty(\Omega)$ and thus, we can extract a weakly converging subsequence $\rho_n \rightharpoonup \rho$ (we have weak-* convergence in L^∞ and in the space of measures, and weak in all $L^p(\Omega)$, $1 \leq p < \infty$). All the terms, except the potential energy, in the functional we minimize are classically known to be lower semicontinous for the weak convergence of probability measures. Since the potential energy term is the H^1 norm of u_n, solution of $-\Delta u_n = \rho_n$, modulo its sign, we make use of the L^∞ bound ρ_n to deduce that $\rho_n \to \rho$ in H^{-1} (using the compact injection of L^2 into H^{-1}), and hence the corresponding u_n strongly converges in H^1 to the solution of $-\Delta u = \rho$, which makes this term continuous. This proves that the limit measure ρ is a minimizer.

We now consider the optimality conditions in this minimization problem and we use the assumptions on Ω, f and g. First, by the same arguments in [34, Lemma 8.6], we observe that the optimal ρ has a density which is strictly positive a.e. (the assumption on f'' implies that f behaves at $t = 0$ at least as the entropy $f(t) = t \log t$, and in particular $f'(0) = -\infty$). Hence, the Kantorovich potential ϕ in the quadratic transport (for the cost $c(x, y) = |x - y|^2/2$) from ρ to g is unique, see [34, Proposition 7.18]. Then, the optimality conditions for this constrained problem read as follows: there exists a constant c and a continuous function p such that

$$(2.3) \qquad f'(\rho) + p - \chi u + \frac{\phi}{\tau} = c, \quad p \geq 0, \ \rho \leq M, \ p(M - \rho) = 0.$$

The proof of this fact is quite classical in optimization under density constraints, see [31]: we first note that the first variation of the functional that we minimize is $h := f'(\rho) - \chi u + \frac{\phi}{\tau}$ where ϕ is the Kantorovich potential such that the map $x - \nabla\phi$ gives the optimal map from ρ to g, for this derivation we refer to [34, Section 7.2.1, 7.2.2 and 7.2.3]. Then we observe that the optimality of ρ implies

$$\int h\tilde{\rho} \geq \int h\rho$$

for every other admissible $0 \leq \tilde{\rho} \leq M$, which means that ρ is equal to M on a sublevel set $\{h < c\}$, equal to 0 on $\{h > c\}$, and between 0 and M on the level set $\{h = c\}$. The result is then obtained by taking $p := (c - h)_+$, see for instance [4, Theorem 4] or [31, Lemma 3.2].

From these conditions, we deduce some regularity for ρ. Indeed, distinguishing the cases $\rho = M$ and $\rho < M$, we get

$$f'(\rho) = \min\left\{f'(M), c + \chi u - \frac{\phi}{\tau}\right\}.$$

Since $\rho \in L^\infty(\Omega)$, we have $u \in W^{2,p}(\Omega)$ for large p, hence u is Lipschitz continuous, and ϕ is also Lipschitz continuous. Hence, the same is true for $f'(\rho)$ and, using the lower bound on f'', for ρ itself. Moreover, we also obtain $\inf \rho > 0$. Of course this regularity depends on τ, but it allows us to perform some computations.

We next want to estimate the L^∞-norm of an optimal ρ. Using $\log g \in C^{0,\alpha}(\Omega)$, we are in a case where both g and ρ are $C^{0,\alpha}(\Omega)$ and bounded from below, and Caffarelli's theory (see [12, 14, 13, 21]) implies $\phi \in C^{2,\alpha}(\Omega)$. Note that from $\rho \in C^{0,\alpha}(\Omega)$ one also obtains $u \in C^{2,\alpha}(\Omega)$.

Now, take a point $x_0 \in \Omega$ which is a minimum point for $\phi - \tau\chi u$. Such a point exists since Ω is compact and these functions are continuous. Yet, from (2.3), one can see that this point maximizes $f'(\rho) + p$, and hence ρ, so that we have $\|\rho\|_{L^\infty} = \rho(x_0)$. First, let us prove $x_0 \notin \partial\Omega$. Indeed, if it were on the boundary, we would have $(\nabla\phi(x_0) - \tau\chi\nabla u(x_0)) \cdot \mathbf{n} \leq 0$. Yet, the optimal transport map T from ρ to g has the form $T(x) = x - \nabla\phi(x)$, and maps Ω onto Ω. Hence $(T(x_0) - x_0) \cdot \mathbf{n} \leq 0$, but this means $\nabla\phi(x_0) \cdot \mathbf{n} \geq 0$. Moroever, $u = 0$ on $\partial\Omega$ and $-\Delta u = \rho \geq 0$ in Ω. The strong maximum principle provides $\nabla u(x_0) \cdot \mathbf{n} < 0$ on every point of the boundary. Hence it is not possible to have $x_0 \in \partial\Omega$. Once we know that the minimizer of $\phi - \tau\chi u$ lies in the interior of Ω, we deduce $\Delta\phi(x_0) - \tau\chi\Delta u(x_0) \geq 0$, but this means $\Delta\phi(x_0) \geq -\tau\chi\rho(x_0)$. We now use the Monge-Ampère equation connecting ϕ, T, ρ and g:

$$\rho(x_0) = g(T(x_0)) \det(I - D^2\phi(x_0)).$$

Since $I - D^2\phi(x_0)$ is a positive symmetric matrix, the arithmetic-geometric inequality provides the well-known inequality $\det(I - D^2\phi(x_0))^{1/d} \leq 1 - \frac{\Delta\phi(x_0)}{d}$, hence

$$\rho(x_0) \leq g(T(x_0))\left(1 - \frac{\Delta\phi(x_0)}{d}\right)^d \leq \|g\|_{L^\infty}\left(1 + \tau\chi\frac{\rho(x_0)}{d}\right)^d.$$

Setting $X := \frac{\tau\chi}{d}\|\rho\|_{L^\infty} = \frac{\tau\chi}{d}\rho(x_0)$ and $Y = \frac{\tau\chi}{d}\|g\|_{L^\infty}$, we get

$$Y \geq \frac{X}{(1 + X)^d}.$$

Looking at the behavior of $x \mapsto x/(1 + x)^d$, we observe that this function is maximal at $x = 1/(d - 1)$. By choosing the constant c_0 in the statement small enough, the value of Y will be such that the level set $\{x : x/(1 + x)^d \leq Y\}$ will be composed of two intervals, one before the maximal point and one after. Yet, since we know $X \leq M\tau\chi/d = 1/d < 1/(d-1)$, the value of X can only be in the first interval. If c_0, and hence Y, is small enough, this interval is a small neighborhood of the origin, on which we can use the inequality

$$\frac{x}{(1 + x)^d} \geq \frac{x}{1 + \lambda dx},$$

which is true for $\lambda > 1$ and for small x just by comparing the second derivatives of these two functions at $x = 0$ (actually the values of the functions and of their first derivatives coincide at $x = 0$). Hence, we get

$$\frac{X}{1 + \lambda dX} \leq Y,$$

which is equivalent to

$$\frac{1}{X} + \lambda d \geq \frac{1}{Y}$$

and, substiting the values of X and Y, we obtain exactly (2.2).

The last part of the statement is obtained by a standard approximation procedure. We take a sequence of smooth and bounded from below densities g_n converging to g, and a sequence of functions f_n defined via $f_n(t) = f(t) + n^{-1}t\log t$, and take the corresponding minimizers ρ_n. Up to subsequences, the minimizers ρ_n converge to a minimizer ρ for the Problem (2.1) with g and f. This means that the same estimate is true for ρ, which is only one particular minimizer corresponding to g and f, and not the unique one. This explains why the same estimate is not guaranteed for all the minimizers. $\square$

The previous estimate can be iterated leading to an estimate on the evolution of the L^∞-norm for this particular JKO scheme with constraints.

Theorem 2. *Let $\rho_0 \in L^\infty$ be a given probability density. Take $\chi t_0 < \|\rho_0\|_{L^\infty}^{-1}$ and let $\lambda > 1$ and $\varepsilon_0 > 0$ be such that $\|\rho_0\|_{L^\infty}^{-1} - \chi\lambda t_0 > \varepsilon_0$. Fix $\tau < \varepsilon_0 c_0$. Then there exists a sequence ρ_k for $k = 1, \ldots, \lfloor t_0/\tau \rfloor$ obtained by iteratedly solving* (2.1)*:*

$$\rho_{k+1} \in \operatorname{argmin}\left\{ J(\rho) + \frac{W_2^2(\rho, \rho_k)}{\tau} \; : \; \rho \in \mathcal{P}(\Omega), \, \rho \le \frac{1}{\chi\tau} \right\}$$

satisfying the L^∞ estimate

(2.4) $$\|\rho_k\|_{L^\infty} \le \frac{\|\rho_0\|_{L^\infty}}{1 - \lambda k\tau\|\rho_0\|_{L^\infty}} \le \frac{1}{\varepsilon_0} \quad \text{or equivalently} \quad \|\rho_k\|_{L^\infty}^{-1} \ge \|\rho_0\|_{L^\infty}^{-1} - \lambda k\tau \ge \varepsilon_0 \, .$$

Proof. We use the previous results, and prove by induction on k that we may choose a minimizer ρ_{k+1} satisfying the desired estimate. Note that ρ_0 satisfies the desired estimate. Now suppose we have ρ_k satisfying the corresponding estimate, then Theorem 2.1 can be applied with $g = \rho_k$ since, by assumption, we have $\tau < \varepsilon_0 c_0$ and $\|\rho_k\|_{L^\infty} \le \varepsilon_0^{-1}$, which imply $\tau\|\rho_k\|_{L^\infty} < c_0$. This provides the existence of a minimizer ρ_{k+1} with

$$\|\rho_{k+1}\|_{L^\infty}^{-1} \ge \|\rho_k\|_{L^\infty}^{-1} - \lambda\tau\chi.$$

This, together with (2.4), which is true by induction assumption and which reads $\|\rho_k\|_{L^\infty}^{-1} \ge \|\rho_0\|_{L^\infty}^{-1} - \lambda k\tau\chi$, gives

$$\|\rho_{k+1}\|_{L^\infty}^{-1} \ge \|\rho_0\|_{L^\infty}^{-1} - \lambda(k+1)\tau\chi,$$

which is the claim. □

2.1. **Variants: besides the case of Dirichlet conditions on bounded domains.** We consider here easy variants of our result, which was stated on a bounded domain Ω with Dirichlet boundary conditions for the equation $-\Delta u = \rho$. Of course, the boundary conditions on the continuity equation are unchanged, they are always of no-flux type.

1. *Neumann or periodic boundary conditions.-* We can change the equation relating u to ρ into a PDE with Neumann boundary conditions, either considering

$$-\Delta u = \rho - c \quad \text{(with } c = 1/|\Omega|\text{)}$$

or $-\Delta u + u = \rho$, and imposing in both cases $\nabla u \cdot \mathbf{n} = 0$ on $\partial\Omega$. The estimates that we can obtain are exactly the same and the arguments are very similar. The only differences are:
   – When proving that we have $x_0 \notin \partial\Omega$, we need to suppose that Ω is strictly convex, which guarantees $\nabla\phi(x_0) \cdot \mathbf{n} > 0$ and allows to find a contradiction together with $\nabla u \cdot \mathbf{n} = 0$. This assumption can be enforced by approximation, and the result stays true anyway on arbitrary convex domains.
   – It is no more true that we have $-\Delta u(x_0) = \rho(x_0)$, but we have in both cases $-\Delta u(x_0) \le \rho(x_0)$. Indeed, in the first case we just use $c > 0$, and in the second case $u \ge 0$ by the maximum principle.

   The case of periodic boundary conditions, when Ω is the flat torus, is even simpler: the equation can be taken as $-\Delta u = \rho - c$ or $-\Delta u + u = \rho$, and we do not need to distinguish the case $x_0 \in \partial\Omega$. Of course, this requires to use regularity results for the optimal transport map on the torus, see [19].

2. *The whole space case $\Omega = \mathbb{R}^d$.-* In this case define u as

$$u(x) = -U * \rho := -\int U(x - y)\rho(y)\mathrm{d}y$$

   where U is the fundamental solution of the Laplacian, i.e. $\Delta U = \delta_0$, given by $U(z) = \frac{1}{2\pi}\log(|z|)$ for $d = 2$, and $U(z) = c_d|z|^{2-d}$ for $d > 2$. In this way it is still true that we have $-\Delta u = \rho$, but we choose

a precise representation formula instead of fixing the boundary conditions. Instead of the Dirichlet energy $-\int |\nabla u|^2$, we use

$$H(\rho) := -\int u(x)\rho(x)\mathrm{d}x = \int \int \rho(x)\rho(y)U(x-y)\mathrm{d}x\mathrm{d}y.$$

We use the fact that, in the space $L^1 \cap L^\infty$, the energy $H(\rho)$ and the convolution $U * \rho$ satisfy the following estimates:

– if $d > 2$ then for each sequence ρ_n bounded in $L^1 \cap L^\infty$, $u_n := -U * \rho_n$ is bounded in L^∞ and hence $H(\rho_n)$ is bounded. This is a consequence of a simple splitting of the convolution in near and far field contributions.

– if $d = 2$ then for each sequence ρ_n bounded in $L^1 \cap L^\infty$, $U * \rho_n$ is bounded from below by a uniform constant, again by splitting in near and far field contributions in the convolution and discarding the positive part of the logarithmic kernel. This also implies that $H(\rho_n)$ is bounded from below.

Moreover, we need to assume that our initial bounded probability density g also has finite second moment. This is important to show the existence of minimizers, since one can take advantage of the $W_2^2(\rho, g)$ term to make appear the second moment of the minimizing sequence to help to control the negative part of the entropy in case $f(\rho) = \rho \log \rho$ or the negative part of the f_n approximations for degenerate diffusions (in particular, for simplicity, we suppose in the unbounded case that f satisfies $f(s) \geq C(s \log(Cs))_-$ for some constant C). We can use the Carleman inequality [6, Lemma 2.2], ot similar arguments (see [6, 7, 17, 36]) which allows to show that there is no mass escaping to infinity and that the weak limit of the minimizing sequence as measures is a probability density. In fact, the convergence happens also weakly in $L^1(\mathbb{R}^d)$ by Dunford-Pettis theorem.

Again, in the space $L^1 \cap L^\infty$, the energy $H(\rho)$ has good convergence properties:

– if $d > 2$ then for each sequence ρ_n bounded in $L^1 \cap L^\infty$ such that $\rho_n \rightharpoonup \rho$, in the sense of weak convergence in duality with bounded and continuous functions on $\mathbb{R}^d$, then $H(\rho_n) \to H(\rho)$.

– if $d = 2$ then for each sequence ρ_n bounded in $L^1 \cap L^\infty$ such that the second moments are uniformly bounded in n and $\rho_n \rightharpoonup \rho$ weakly in $L^1(\mathbb{R}^d)$, then $H(\rho_n) \to H(\rho)$, see [6, Lemma 3.1] for the proof.

Then, the proof goes on in the same way, but in the approximation we will suppose that g is a compactly supported density with uniformly convex and smooth support, smooth and bounded away from zero on its support (but discontinuous on the boundary of the support). This allows to apply the regularity theory for the Monge-Ampère equation. More precisely, we need that the optimal transport map between two $C^{0,\alpha}$ probability densities with a source density being strictly positive on $\mathbb{R}^d$ and a target density being bounded away from zero on a bounded smooth convex domain is $C^{1,\alpha}(\mathbb{R}^d)$, and the Kantorovich potential is $C^{2,\alpha}(\mathbb{R}^d)$. This fact, for an unbounded source domain, but keeping boundedness and convexity of the target, is not easy to find in the literature, but can be proven by an easy adaptation of the arguments in [23, Theorem 4.23, Remark 4.25]. By the way, the fact that the source domain is the whole space makes the proof even shorter. Moreover, in this case ϕ grows quadratically, and u is bounded from above, which guarantees that a minimizer for $\phi - \tau\chi u$ exists.

A final point to observe is that the results in [34, Chapter 7] about the first variation of the W_2^2 cost are only proven in [34] in the case of a compact domain. Yet, generalizing them to the case where g is compactly supported (and the comparison is done with measures $\tilde{\rho}$ which are compactly supported themselves) is an easy exercise. The rest of the proof is unchanged.

3. Local existence of a solution, for arbitrary non-linearity and L^∞ data

In this section we apply the previous estimates to prove the existence, local in time, of a weak solution of the parabolic-elliptic Keller-Segel system (1.1). For simplicity we will start from the case of linear diffusion, and then turn to more arbitrary functions f. The proof will only underline the general strategy and the peculiarity of each case.

Linear diffusion.- We consider the particular case

(3.1)
$$\begin{cases} \partial_t \rho + \chi \nabla \cdot (\rho \nabla u) - \Delta \rho = 0 & \text{in } [0, T] \times \Omega, \\ -\Delta u = \rho & \text{in } \Omega \text{ for every } t \in [0, T], \\ u = 0 & \text{on } \partial\Omega \text{ for every } t \in [0, T], \\ (\rho \nabla u - \nabla \rho) \cdot \mathbf{n} = 0 & \text{on } [0, T] \times \partial\Omega, \\ \rho(0, \cdot) = \rho_0 \end{cases}$$

Theorem 3. *Suppose Ω is a bounded convex domain in $\mathbb{R}^d$. For any $\rho_0 \in L^\infty(\Omega)$ and $T < (\chi \|\rho_0\|_{L^\infty})^{-1}$, the system (3.1) admits at least a weak solution on $[0, T) \times \Omega$, and this solution is bounded for all $0 \le t < T$.*

Proof. The construction follows the same scheme summarized in [22] and [35].

By using the modified JKO scheme of the previous section, for each $\tau > 0$ we are able to build a sequence $(\rho_k^\tau)_k$, iteratively solving (2.1), where we choose $f(\rho) = \rho \log \rho$. the sequence is bounded in $L^\infty(\Omega)$ independently of τ. We also define a sequence of velocities $\mathbf{v}_k^\tau = (id - \mathrm{T})/\tau$, taking as T the optimal transport from ρ_k^τ to ρ_{k-1}^τ. Now, we notice that the optimality conditions of (2.1), since we know that in the end the constraint $\rho \le M$ is not binding for τ small enough due to (2.4), may be re-written without the pressure p, thus getting

$$\mathbf{v}_k^\tau = \frac{\nabla \phi}{\tau} = -\frac{\nabla \rho_k^\tau}{\rho_k^\tau} + \nabla u_k^\tau,$$

where u_k^τ is the solution of $-\Delta u = \rho_k^\tau$ with Dirichlet boundary conditions on $\partial\Omega$.

Then, we build two interpolating curves in the space of measures:

- first we can define some piecewise constant curves, i.e. $\bar\rho_t^\tau := \rho_{k+1}^\tau$ for $t \in]k\tau, (k+1)\tau]$; associated with this curve we also define the velocities $\overline{\mathbf{v}_t^\tau} = \mathbf{v}_{k+1}^\tau$ for $t \in]k\tau, (k+1)\tau]$ and the momentum variable $\overline{E}^\tau = \bar\rho^\tau \overline{\mathbf{v}^\tau}$;
- then, we can also consider the densities $\widehat{\rho_t^\tau}$ that interpolate the discrete values $(\rho_k^\tau)_k$ along geodesics:

$$\widehat{\rho_t^\tau} = (id - (k\tau - t)\mathbf{v}_k^\tau)_\sharp \rho_k^\tau, \quad \text{for } t \in](k-1)\tau, k\tau[;$$

the velocities $\widehat{\mathbf{v}_t^\tau}$ are defined so that $(\widehat{\rho^\tau}, \widehat{\mathbf{v}^\tau})$ satisfy the continuity equation, taking

$$\widehat{\mathbf{v}_t^\tau} = \mathbf{v}_t^\tau \circ (id - (k\tau - t)\mathbf{v}_k^\tau)^{-1};$$

as before, we define: $\widehat{E}^\tau = \widehat{\rho^\tau}\widehat{\mathbf{v}^\tau}$.

After these definitions we look for a priori bounds on the curves and the velocities that we defined. Note that the optimality properties of the sequence $(\rho_k^\tau)_k$ give

$$\frac{W_2^2(\rho_k^\tau, \rho_{k-1}^\tau)}{\tau} \le J(\rho_{k-1}^\tau) - J(\rho_k^\tau).$$

Hence, for $N = \lfloor T/\tau \rfloor$, we get

(3.2)
$$\sum_{k=1}^N \tau \left(\frac{W_2(\rho_k^\tau, \rho_{k-1}^\tau)}{\tau} \right)^2 \le J(\rho_0) - J(\rho_N^\tau) \le C,$$

where the constant C is independent of τ, since $J(\rho_0) < +\infty$ and J is bounded from below on the set of measures with density bounded by a given constant. This is the discrete version of an L^2 estimate of the time derivative. As for $\widehat{\rho_t^\tau}$, it is an absolutely continuous curve in the Wasserstein space and its velocity on the time interval $[(k-1)\tau, k\tau]$ is given by the ratio $W_2(\rho_{k-1}^\tau, \rho_k^\tau)/\tau$. Hence, the L^2 norm of its velocity on $[0, T]$ is given by

$$(3.3) \qquad \int_0^T |(\widehat{\rho^\tau})'|^2(t)dt = \sum_k \frac{W_2^2(\rho_k^\tau, \rho_{k-1}^\tau)}{\tau},$$

and, thanks to (3.2), it admits a uniform bound independent of τ. In our case, thanks to results on the continuity equation and the Wasserstein metric, this metric derivative is also equal to $\|\widehat{\mathbf{v}_t^\tau}\|_{L^2(\widehat{\rho_t^\tau})}$. This gives compactness of the curves $\widehat{\rho^\tau}$, as well as an Hölder estimate on their variations in time (since $H^1 \subset C^{0,1/2}$). The characterization of the velocities $\overline{\mathbf{v}^\tau}$ and $\widehat{\mathbf{v}^\tau}$ allows to deduce bounds on these vector fields from the bounds on $W_2(\rho_{k-1}^\tau, \rho_k^\tau)/\tau$.

Considering all these facts, one obtains the following situation.

- The norm $\int \|\overline{\mathbf{v}_t^\tau}\|_{L^2(\overline{\rho_t^\tau})}^2 dt$ is τ-uniformly bounded.
- In particular, the above L^2 bound is valid in L^1 as well, which implies that $\overline{E}^\tau$ is bounded in the space of measures over $[0, T] \times \Omega$.
- The very same estimates are true for $\widehat{\mathbf{v}^\tau}$ and $\widehat{E}^\tau$.
- The curves $\widehat{\rho^\tau}$ are bounded in $H^1([0, T], \mathbb{W}_2(\Omega))$ and hence compact in $C^0([0, T], \mathbb{W}_2(\Omega))$.
- Up to a subsequence, one has $\widehat{\rho^\tau} \to \rho$, as $\tau \to 0$, uniformly according to the W_2 distance.
- From the estimate $W_2(\overline{\rho_t^\tau}, \widehat{\rho_t^\tau}) \leq C\tau^{1/2}$ one gets that ρ^τ converges to the same limit ρ in the same sense.
- If we denote by E a weak limit of $\widehat{E}^\tau$, since $(\widehat{\rho^\tau}, \widehat{E}^\tau)$ solves the continuity equation, by linearity, passing to the weak limit, also (ρ, E) solves the same equation.
- It is possible to prove, see [31, Section 3.2, Step 1] or [34, Chapter 8], that the weak limits of $\widehat{E}^\tau$ and $\overline{E}^\tau$ are the same.
- From the bounds in L^2 one gets that also the measure E is absolutely continuous w.r.t. ρ and has an L^2 density, so that we have for a.e. time t a measure E_t of the form $\rho_t v_t$.
- It is only left to prove that one has $E_t = -\Delta\rho_t + \nabla \cdot (\rho_t \nabla u_t)$, where u_t is the solution of $-\Delta u = \rho_t$ with Dirichlet boundary conditions on $\partial\Omega$.

In this case this last fact is easy to prove. Indeed, from $\bar{\rho}_t^\tau \rightharpoonup \rho_t$ one immediately deduces $-\Delta\bar{\rho}_t^\tau \to -\Delta\rho_t$ (in the sense of distributions). The L^∞ bound implies that the convergence $\bar{\rho}_t^\tau \rightharpoonup \rho_t$ is weak in L^p, and hence $u(\bar{\rho}_t^\tau) \rightharpoonup u(\rho_t)$ weakly in $W^{2,p}$. For $p > d$, this gives uniform convergence $\nabla u(\bar{\rho}_t^\tau) \to \nabla u(\rho_t)$ and guarantees that also the term $\nabla \cdot (\nabla u(\bar{\rho}_t^\tau)\bar{\rho}_t^\tau)$ converges in the sense of distribution (to $\nabla \cdot (\nabla u(\rho_t)\rho_t)$).   □

Remark 1. The previous proof works in the whole space case almost entirely. Similarly to the proof of the existence of minimizers in the previous section, in order to have a bound from below on J, one applies Carleman estimates, for instance [6, Lemma 2.2], to use second moments to estimate from below the entropy term $\int \rho \log \rho$. In order to do that, we can use part of the distance $W_2(\rho, \rho_0)$ as in [6, 7, 17] or use a sublinear power of the left-hand side of (3.2). For sharp lower bounds on the entropy in terms of moments, see [36]. Finally, passing to the limit in the equations only requires local convergences instead of global in the whole space, as this is enough for the equations to pass to the limit in the distributional sense.

Non-linear diffusion.- The problem becomes trickier in case of non-linear diffusions. However, a modification of the above proof, requiring the following lemma, inspired by the well-known Aubin-Lions compactness Lemma (see for instance [3]), allows to handle non-linear cases. The modification concerns the use of the Wasserstein distance and the simplification using the L^∞ estimates, see [31] for similar computations.

Lemma 1. *Suppose that μ^τ is a sequence (indexed in τ) of time-dependent probability densities such that*

(1) *$\|\mu_t^\tau\|_{L^\infty}$ is bounded*
(2) *$W_2(\mu_t^\tau, \mu_s^\tau) \leq C\sqrt{\tau} + \int_t^s g^\tau(r)dr$, where g^τ is bounded in $L^2([0,T])$ and C is a constant.*
(3) *For some strictly increasing and convex function $K : [0,+\infty) \to [0,+\infty)$ such that the function $s \mapsto sK^{-1}(s)$ is strictly convex, we have $\|\nabla(K(\mu_t^\tau))\|_{L^2(\Omega)} \leq Cg^\tau(t)$, for a given constant C and the same function g^τ as above.*

Then, the sequence of μ^τ is compact in all the spaces L^p with $p < \infty$ and in particular μ^τ converges a.e. up to subsequences.

Before proving this lemma, we apply it to the case we are interested in.

Theorem 4. *Suppose that Ω is a bounded convex domain in $\mathbb{R}^d$. Suppose that $f : \mathbb{R} \to \mathbb{R}$ is a continuous function on $[0,+\infty[$ and C^2 on $]0,+\infty[$, with $f'' > 0$. Then, for any $\rho_0 \in L^\infty(\Omega)$ and $T < (\chi\|\rho_0\|_{L^\infty})^{-1}$, the system (1.1) admits at least a weak solution on $[0,T] \times \Omega$, and this solution is bounded.*

Proof. We follow the same proof as in Theorem 3. The only difficulty left is to pass to the limit the term $\nabla \cdot (\bar\rho_t^\tau \nabla f'(\bar\rho_t^\tau))$, as it has no more the linear form as it had before. Yet, this term can be written as $\Delta(\Psi(\bar\rho_t^\tau))$, where $\Psi(t) = tf'(t) - f(t)$. We know that we have weak convergence $\bar\rho_t^\tau \rightharpoonup \rho_t$, and we just need to turn it into a.e. convergence to prove the desired convergence. This can be done by means of Lemma 1, applied to $\mu^\tau = \bar\rho^\tau$. We need to check the assumptions of this lemma. Assumption (1) is guaranteed by our iterated L^∞ estimate, and we call M a constant such that $\bar\rho^\tau \leq M$ for all τ.

For assumptions (2) and (3), we define $g^\tau(t) := \|\overline{v_t^\tau}\|_{L^2(\bar\rho_t^\tau)} + C$, for a suitable large constant C. Then, the integrability of g^τ comes from (3.3), and we have

$$W_2(\bar\rho_t^\tau, \bar\rho_s^\tau) \leq W_2(\bar\rho_t^\tau, \widehat{\rho_t^\tau}) + W_2(\bar\rho_s^\tau, \widehat{\rho_s^\tau}) + W_2(\widehat{\rho_t^\tau}, \widehat{\rho_s^\tau}) \leq C\tau^{1/2} + \int_t^s g^\tau(r)dr.$$

In order to check the validity of assumption (3), we first choose, using the following Lemma 2, a function k such that its antiderivative K, with $K(0) = 0$, satisfies the requirements of assumption (3), and such that $k(s) \leq s^{1/2}f''(s)$ for all $s \in [0, M]$. This means that we have

$$\|\nabla K(\mu_t^\tau)\|_{L^2(\Omega)}^2 = \int_\Omega k(\bar\rho^\tau)^2 |\nabla\bar\rho^\tau|^2 \leq \int_\Omega \bar\rho^\tau |f''(\bar\rho^\tau)|^2 |\nabla\bar\rho^\tau|^2,$$

but this last quantity is the squared norm in $L^2(\bar\rho^\tau)$ of $\nabla(f'(\bar\rho^\tau))$. From the optimality conditions defining $\bar\rho^\tau$, we have

$$|\nabla f'(\bar\rho^\tau)| = \left|\nabla u + \frac{\nabla\phi}{\tau}\right| \leq C + \left|\frac{\nabla\phi}{\tau}\right|,$$

where the constant C is chosen so as to be larger than the L^∞ norm of ∇u (which is bounded by the L^∞ norm of $\bar\rho^\tau$). If we take the norms in $L^2(\bar\rho^\tau)$, we exactly get $\|\nabla K(\mu_t^\tau)\|_{L^2(\Omega)} \leq Cg^\tau(t)$. This concludes the proof. $\square$

Lemma 2. *Given any strictly positive and continuous function $k_0 : (0, M] \to (0, \infty)$, there exists a function $k : (0, M] \to (0, \infty)$ with $k \leq k_0$ such that*

- *k is non-decreasing and K, its antiderivative defined through $K(s) = \int_0^s k$, is strictly increasing and convex;*
- *the map $s \mapsto E(s) := sK^{-1}(s)$ is strictly convex.*

Proof. In order to obtain a non-decreasing function $k \leq k_0$ it is enough to consider a strictly decreasing sequence $\ell_n \to 0$ with $\ell_0 = M$ and define $k(s) = a_n := \inf_{[\ell_{n+1}, M]} h_0$ for all $s \in (\ell_{n+1}, \ell_n]$. This value is strictly positive since h_0 is continuous and strictly positive. Automatically, if k is strictly positive and non-decreasing, then K is strictly increasing and convex.

Then, we need to arrange this function so as to satisfy the second condition as well. For this we will choose a new function k, smaller than the one that we have just defined (i.e., smaller than a_n on each interval (ℓ_{n+1}, ℓ_n)), and locally Lipschitz. In this way K will be locally $C^{1,1}$ and it is not difficult to check that the strict convexity of E is satisfied whenever we have $KK'' < 2|K'|^2$. In order to be convinced of this, just notice that $E'(s) = K^{-1}(s) + s/K'(K^{-1}(s))$; we want E' to be strictly increasing, which is equivalent to $E' \circ K$ to be strictly increasing, hence we need to look at $s \mapsto s + K(s)/K'(s)$. If we differentiate it we obain $2 - KK''/|K'|^2$.

We will define k iteratively on each interval $(\ell_{n+1}, \ell_n]$ in the following way: define three sequences of numbers $b_n, p_n, I_n > 0$ such that

$$(3.4a) \qquad\qquad b_{n+1} = b_n(\ell_{n+1}/\ell_n)^{p_n},$$

$$(3.4b) \qquad\qquad b_n \le a_n,$$

$$(3.4c) \qquad\qquad I_n p_n (\ell_n/\ell_{n+1})^{p_n}/(b_n \ell_{n+1}) \le 1,$$

and

$$(3.4d) \qquad\qquad \ell_n b_n/(p_n + 1) + b_n \ell_{n+1}(\ell_{n+1}/\ell_n)^{p_n} \le I_{n-1}.$$

In order to do so, fix $b_0 = a_0, p_{-1} = 1$ and $I_{-1} = 2a_0 M$. Then, if (b_n, p_{n-1}, I_{n-1}) are given, we then choose p_n large enough so that, when b_{n+1} is defined using (3.4a), both (3.4b) and (3.4d) are satisfied. This defines p_n and b_{n+1}. We then choose I_n small enough to that (3.4c) is satisfied, and iterate.

Once these numbers $(b_n, p_n, I_n)_n$ are fixed, we define $k(s) = b_n(s/\ell_n)^{p_n} \le b_n \le a_n$ for $s \in (\ell_{n+1}, \ell_n]$. In this way for every n we have $k(\ell_n) = b_n$. Moreover, the function h defined in this way is strictly increasing, satisfies $k \le k_0$, it is locally Lipschitz continuous, continuity at the points ℓ_n is guaranteed by (3.4a), and we have, for $s \in (\ell_{n+1}, \ell_n]$

$$K(s) = \int_0^s k(r)dr \le b_{n+2}\ell_{n+2} + \int_{\ell_{n+2}}^{\ell_{n+1}} k(r)dr + \int_{\ell_{n+1}}^s k(r)dr$$

$$\le b_{n+2}\ell_{n+2} + \int_0^{\ell_{n+1}} b_{n+1}(r/\ell_{n+1})^{p_{n+1}}dr + \int_0^s b_n(r/\ell_n)^{p_n}dr$$

$$= b_{n+1}(\ell_{n+2}/\ell_{n+1})^{p_{n+1}}\ell_{n+2} + \frac{\ell_{n+1}b_{n+1}}{p_{n+1}+1} + \frac{b_n s^{p_n+1}}{(p_n+1)(\ell_n)^{p_n}} \le I_n + \frac{b_n s^{p_n+1}}{(p_n+1)(\ell_n)^{p_n}},$$

where we used (3.4a) to pass from the second to the third line and (3.4d) for the last inequality. Now, in order to verify that we have $KK'' < 2|K'|^2$, we just need to check that we have, for $s \in (\ell_{n+1}, \ell_n]$,

$$\left(I_n + \frac{1}{p_n+1}b_n \frac{s^{p_n+1}}{\ell_n^{p_n}}\right) p_n b_n \frac{s^{p_n-1}}{\ell_n^{p_n}}\left(b_n \frac{s^{p_n}}{\ell_n^{p_n}}\right)^{-2} < 2,$$

where we used the exact expression of $K' = k$ and $K'' = k'$ on such an interval. Yet, the left hand side may be re-written as

$$\frac{p_n}{p_n+1} + I_n p_n \frac{\ell_n^{p_n}}{b_n s^{p_n+1}} < 2,$$

due to (3.4c).   □

We now prove Lemma 1.

Proof. First of all we assume, up to subsequences, that the sequence μ^τ converges to a function μ, $K(\mu^\tau)$ to a function A and $\mu^\tau K(\mu^\tau)$ to a function B, all the convergences being weak-* in $L^\infty([0, T] \times \Omega)$. We now fix

two instants of time $a < b \in [0, T]$ and consider

$$\fint_a^b \int_\Omega \mu^\tau K(\mu^\tau) \, dx dt \to \fint_a^b \int_\Omega B \, dx dt.$$

Yet, we can also write

$$\int_\Omega \mu_t^\tau K(\mu_t^\tau) dx = \int_\Omega K(\mu_t^\tau) \mu_a^\tau dx + \int_\Omega K(\mu_t^\tau)(\mu_t^\tau - \mu_a^\tau) dx.$$

We use the H^1 bound on $K(\mu^\tau)$ to write

$$\left| \int_\Omega K(\mu^\tau)(\mu_t^\tau - \mu_a^\tau) dx \right| \le g^\tau(t) \|\mu_t^\tau - \mu_a^\tau\|_{X'},$$

where X is the Hilbert space of zero-mean H^1 functions endowed with the L^2 norm of the gradient, and X' is its dual. It is well-known, see [30, 31] and [34, Section 5.5], that we have the interpolation estimate

$$\|\mu - \nu\|_{X'} \le \max\{\|\mu\|_{L^\infty}, \|\nu\|_{L^\infty}\}^{1/2} W_2(\mu, \nu).$$

Defining $G : [a, b] \to \mathbb{R}$ through $G(a) = 0$ and $G' = g^\tau$, and using $W_2(\mu_t^\tau, \mu_a^\tau) \le C \sqrt{\tau} + G(t)$, we get

$$\left| \fint_a^b \int_\Omega \mu^\tau K(\mu^\tau) dx dt - \fint_a^b \int_\Omega K(\mu_t^\tau) \mu_a^\tau dx dt \right| \le \frac{C \sqrt{\tau}}{b - a} G(b) + \frac{1}{b - a} G^2(b).$$

We use $G^2(b) = |\int_a^b g^\tau(r) dr|^2 \le (b-a) \int_a^b g^\tau(r)^2 dr$ and we define σ^τ to be the measure on $[0, T]$ with density $(g^\tau)^2$. Suppose (up to subsequences) that σ^τ weakly converges to a measure σ. If we pass to the limit this inequality as $\tau \to 0$ we get

$$\left| \fint_a^b \int_\Omega B \, dx dt - \fint_a^b \int_\Omega \mu A \, dx dt \right| \le \sigma([a, b]).$$

Note that in the above limit we could pass to the limit the term

$$\fint_a^b \int_\Omega K(\mu_t^\tau) \mu_a^\tau dx dt = \int_\Omega \mu_a^\tau \left(\fint_a^b K(\mu_t^\tau) dt \right)$$

since the function $x \mapsto \fint_a^b K(\mu_t^\tau(x)) dt$ is bounded in H^1 for fixed $a < b$. This means that it converges strongly and can be multiplied times the weak convergence of μ_a^τ, which converges to μ_a thanks to the continuity bound on $t \mapsto \mu_t^\tau$.

Now, for every a which is a Lebesgue point for $t \mapsto (A, B) \in L^2(\Omega)^2$ and which is not an atom of σ, we get

$$\int_\Omega B(x, a) dx = \int_\Omega \mu_a A(x, a) dx \le \int_\Omega K^{-1}(A) A \, dx,$$

where we use the fact that $\mu^\tau = K^{-1}(K(\mu^\tau))$, and the function $s \mapsto K^{-1}(s)$ is concave on $\mathbb{R}_+$ (since K is convex and increasing). This guarantees that the weak limit μ of μ^τ must be smaller or equal than K^{-1} of the weak limit of $K(\mu^\tau)$.

Then, we use the function $E(s) = sK^{-1}(s)$, and we have a sequence, $v^\tau := E(\mu^\tau)$, of L^∞ functions on $[0, T] \times \Omega$ such that $v^\tau \rightharpoonup v$ and $\limsup_\tau \int E(v^\tau) \le \int E(v)$ for a strictly convex function E. Because of strict convexity, we can write the inequality

$$E(s) \ge E(s_0) + E'(s_0)(s - s_0) + \omega(|s - s_0|),$$

where ω plays the role of a modulus of strict convexity and is a continuous function with $\omega(0) = 0$ and $\omega(s) > 0$ for $s > 0$. Applying this inequality to $s = v^\tau$ and $s_0 = v$, we obtain

$$\limsup_\tau \int \omega(|v^\tau - v|) \leq \limsup_\tau \int E(v^\tau) - \int E(v) - \int E(v)(v^\tau - v) \leq 0,$$

where we used the weak convergence $v^\tau \rightharpoonup v$ to handle the very last term. This implies, up to subsequences, that $\omega(|v^\tau - v|) \to 0$ a.e. and hence that $v^\tau \to v$ a.e. In our case, composing with E^{-1}, this provides a.e. convergence for μ^τ to μ and, if we also use the L^∞ bound, L^p convergence for every p. □

References

[1] L. Ambrosio, N. Gigli and G. Savaré, *Gradient Flows in Metric Spaces and in the Space of Probability Measures*, Lectures in Mathematics, Birkhäuser, 2005.

[2] L. Ambrosio and S. Serfaty, A gradient flow approach to an evolution problem arising in superconductivity, *Communications on Pure and Applied Mathematics* 61, 1495-1539, 2008.

[3] J.-P. Aubin, Un théorème de compacité, *C. R. Acad. Sci. Paris* 256, 5042-5044, 1963.

[4] D. Balagué, J. A. Carrillo, T. Laurent and G. Raoul, Dimensionality of local minimizers of the interaction energy, *Arch. Ration. Mech. Anal.* 209, 1055-1088, 2013.

[5] A. L. Bertozzi, T. Laurent and F. Leger, Aggregation and spreading via the Newtonian potential: the dynamics of patch solutions, *Mathematical Models and Methods in Applied Sciences* 22, 1140005, 2012.

[6] A. Blanchet, V. Calvez and J.-A. Carrillo, Convergence of the mass-transport steepest descent scheme for the subcritical Patlak-Keller-Segel model, *SIAM Journal on Numerical Analysis* 46, 691-721, 2008.

[7] A. Blanchet, E.A. Carlen and J.-A. Carrillo, Functional inequalities, thick tails and asymptotics for the critical mass Patlak-Keller-Segel model, *Journal of Functional Analysis* 262, 2142-2230, 2012.

[8] A. Blanchet, J.-A. Carrillo, D. Kinderlehrer, M. Kowalczyk, P. Laurençot and S. Lisini, A Hybrid Variational Principle for the Keller-Segel System In $\mathbb{R}^2$, *ESAIM M2AN* 49, 1553-1576, 2015.

[9] A. Blanchet, J.-A. Carrillo and P. Laurençot, Critical mass for a Patlak-Keller-Segel model with degenerate diffusion in higher dimensions, *Calculus of Variations and Partial Differential Equations* 35, 133-168, 2009.

[10] A. Blanchet, J.-A. Carrillo and N. Masmoudi, Infinite time aggregation for the critical Patlak-Keller-Segel model in $\mathbb{R}^2$, *Communications on Pure and Applied Mathematics* 61, 1449-1481, 2008.

[11] A. Blanchet, J Dolbeault and B. Perthame, Two-dimensional Keller-Segel model: optimal critical mass and qualitative properties of the solutions, *Electronic Journal of Differential Equations* 44, 1-33, 2006.

[12] L. Caffarelli, A localization property of viscosity solutions to the Monge-Ampère equation and their strict convexity, *Ann. of Math.* 131, 129-134, 1990.

[13] L. Caffarelli, Interior $W^{2,p}$ estimates for solutions of the Monge-Ampère equation. *Ann. of Math.* 131, 135-150, 1990.

[14] L. Caffarelli, Some regularity properties of solutions of Monge Ampère equation, *Comm. Pure Appl. Math.* 44, 965-969, 1991.

[15] V. Calvez and J. A. Carrillo, Volume effects in the Keller-Segel model: energy estimates preventing blow-up, *J. Math. Pures Appl.* 86, 155-175, 2006.

[16] V. Calvez and L. Corrias, The parabolic-parabolic Keller-Segel model in $\mathbb{R}^2$, *Commun. Math. Sci.* 6, 417-447, 2008.

[17] J. A. Carrillo, M. DiFrancesco, A. Figalli, T. Laurent and D. Slepčev, Global-in-time weak measure solutions and finite-time aggregation for nonlocal interaction equations, *Duke Math. J.* 156, 229-271, 2011.

[18] J. A. Carrillo, S. Lisini and E. Mainini, Uniqueness for Keller-Segel-type chemotaxis models, *Discrete Contin. Dyn. Syst.-Series A* 34, 1319-1338, 2014.

[19] D. Cordero-Erausquin, Sur le transport de mesures périodiques, *C. R. Acad. Sci. Paris* Sér. I Math., vol. 329, no. 3, pp. 199-202, 1999

[20] K. Craig, Nonconvex gradient flow in the Wasserstein metric and applications to constrained nonlocal interactions, *Proc. Lond. Math. Soc.* 114, 60-102, 2017.

[21] G. De Philippis and A. Figalli, The Monge–Ampère equation and its link to optimal transportation, *Bulletin of the American Mathematical Society* 51, 527-580, 2014.

[22] S. Di Marino, B. Maury and F. Santambrogio, Measure sweeping processes, *Journal of Convex Analysis* 23, 567-601, 2016.

[23] A. Figalli, *The Monge-Ampère Equation and Its Applications*, Zurich Lectures in Advanced Mathematics, European Mathematical Society (EMS), Zürich, 2017.

[24] T. Hillen and K. J. Painter, A user's guide to PDE models for chemotaxis, *Journal of Mathematical Biology* 58, 183-217, 2009.

[25] W. Jäger and S. Luckhaus, On explosions of solutions to a system of partial differential equations modelling chemotaxis, *Trans. Amer. Math. Soc.* 329, 819-824, 1992.

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

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

[28] S. Lisini, E. Mainini and A. Segatti, A gradient flow approach to the porous medium equation with fractional pressure, Preprint arXiv:1606.06787.

[29] J.-G. Liu and J. Wang, A note on L^∞-bound and uniqueness to a degenerate Keller-Segel model, *Acta Appl. Math.* 142, 173–188, 2016.

[30] G. Loeper, Uniqueness of the solution to the Vlasov-Poisson system with bounded density, *J. Math. Pures Appl.* 86, 68-79, 2006.

[31] B. Maury, A. Roudneff-Chupin and F. Santambrogio, A macroscopic crowd motion model of gradient flow type, *Mathematical Models and Methods in Applied Sciences* 20, 1787-1821, 2010.

[32] F. Otto, The geometry of dissipative evolution equations: The porous medium equation, *Comm. Partial Differential Equations* 26, 101-174, 2001.

[33] F. Poupaud, Diagonal defect measures, adhesion dynamics and Euler equation, *Methods Appl. Anal.* 9, 2002, 533–561.

[34] F. Santambrogio, *Optimal Transport for Applied Mathematicians*, Progress in Nonlinear Differential Equations and Their Applications 87, Birkhäuser Basel 2015.

[35] F. Santambrogio, {Euclidean, Metric, and Wasserstein} Gradient Flows: an overview, *Bulletin of Mathematical Sciences* 7, 87-154, 2017.

[36] F. Santambrogio, Dealing with moment measures via entropy and optimal transport, *J. Functional Analysis* 271, 418-436, 2016.

[37] C. Villani, *Optimal Transport. Old and New*, Grundlehren der Mathematischen Wissenschaften 338, Springer-Verlag, Berlin, 2009.

J.-A. C. Department of Mathematics, Imperial College London, South Kensington Campus, SW7 2AZ London UK,
Email: carrillo@imperial.ac.uk,

F. S. Laboratoire de Mathématiques d'Orsay, Univ. Paris-Sud, CNRS, Université Paris-Saclay, 91405 Orsay Cedex, France,
Email: filippo.santambrogio@math.u-psud.fr.