ENTROPY ON THE PATH SPACE AND APPLICATION TO SINGULAR DIFFUSIONS AND MEAN-FIELD MODELS.

PATRICK CATTIAUX ♠

♠ Université de Toulouse

Abstract. In this paper we introduce a (partly) new approach for the study of McKean-Vlasov equations, including singular interactions. This approach is based on the relative entropy on the path space in the spirit of our previous works together with C. Léonard. It is also used in a recent work of D. Lacker.
We show how it can be used to derive existence and uniqueness for some singular diffusions, in particular linear mean field stochastic particle systems and non linear SDE of McKean-Vlasov type, including $\mathbf{L}^p - \mathbf{L}^q$ models, the 2D vortex model associated to the 2D Navier-Stokes equation, sub-Coulombic interactions models or the Patlak-Keller-Segel model.
We also show the convergence and propagation of chaos as the number of particles grows to infinity. This is obtained at the process level, not only at the Liouville equation (marginals flow) level.
The paper thus contains new proofs and extensions of known results, as well as new results. The main results are given at the end of the Introduction.

Key words : singular McKean-Vlasov equations, diffusion processes, Kullback information, propagation of chaos.

MSC 2010 : 35Q92, 60J60, 60K35.

1. Introduction and main results.

McKean-Vlasov non linear partial differential equations have been extensively used for modeling the dynamics of complex systems in physical and life sciences in particular. They are a macroscopic description of collective behaviour of particle systems for which the density of particles $\bar{\rho}(t)$ solves

$$\partial_t \bar{\rho}_t(x) = \Delta_x \, \bar{\rho}_t(x) + \nabla_x.((b + K * \bar{\rho}_t)\bar{\rho}_t)(x) \, . \tag{1.1}$$

Here $x \in \mathbb{R}^d$, b is some drift term describing a self-interaction and K describes interaction between particles. It is not difficult to see that (1.1) preserves positivity and mass, so that we may assume that $\bar{\rho}_0$ is a density of probability, i.e. $\bar{\rho}_0 \geq 0$ and $\int \bar{\rho}_t dx = \int \rho_0 dx = 1$.

Of course we have chosen here to only look at "diffusive" particles of Ito type and one can replace the Laplace operator by more general local or non local operators. Notice that up to a linear time change, one can include in this framework the case where the diffusion part is given by $\sigma^2 \, \Delta_x$.

Date: April 12, 2024.

In order to understand the microscopic behaviour of the system one has to introduce a "linear" stochastic version of (1.1), describing the individual behaviour of each particle, namely, for $i = 1, ..., N$, the system of Stochastic Differential Equations (SDE)

$$dX_t^{i,N} = \sqrt{2}\, dB_t^{i,N} \, - \, b(X_t^{i,N})dt \, - \, \frac{1}{N} \sum_{j=1}^{N} K(X_t^{i,N} - X_t^{j,N})\, dt\,, \tag{1.2}$$

where $B_\cdot^{i,N}$ is a collection of N i.i.d. standard Brownian motions on $\mathbb{R}^d$. If one assumes that the initial random vector $X_0^{\cdot,N}$ is a finite exchangeable sequence, it is easy to see that the same is true for $X_t^{\cdot,N}$ for all $t \geq 0$. Assuming that the variables $X_0^{\cdot,N}$ are i.i.d. with common distribution $\mu_0(dx)$, one can thus expect that the empirical measure $\nu_t^N = \frac{1}{N}\sum_{i=1}^N \delta_{X_t^{i,N}}$ which is a random variable taking its values in the set of probability measures, satisfies some "law of large numbers" i.e. converges in some sense (weakly or strongly) and hopefully that the limit is a non-random probability measure μ_t. This convergence implies that for all $k \in \mathbb{N}$, the distribution of $(X_t^{1,N}, ..., X_t^{k,N})$ weakly converges to the tensor product $\mu_t^{\otimes k}$. This property is called since its introduction by Kac, the *propagation of chaos* property, at the level of the marginals.

Since the interaction term in the drift of (1.2) writes as $K * \nu_t^N$ one can then expect that $X_\cdot^{1,N}$ converges in some sense to a *non-linear diffusion process* solution of the so called *non linear SDE*,

$$\begin{aligned} d\bar{X}_t &= \sqrt{2}\, dB_t \, - \, b(\bar{X}_t)\, dt \, - \, (K * \bar{\rho}_t)(\bar{X}_t)\, dt\,, \\ \mu_t = \bar{\rho}_t(x)\, dx &= \mathcal{L}(\bar{X}_t)\,, \end{aligned} \tag{1.3}$$

whose marginals flow satisfies (1.1).

The previous program was partly or fully completed in several situations during the last decades. The first basic result is described in [110] section 1 and reads as follows

Theorem 1.1. *Assume that b and K are global Lipschitz and bounded. Assume in addition that the distribution of the initial condition $X_0^{\cdot,N}$ is given by $\mu_0^{\otimes N}$.*

Then there exists a strongly unique solution for both (1.2) and (1.3), where for $t > 0$ the distributions of $X_t^{\cdot,N}$ and $\bar{X}_t$ admit a density $\rho_t^{\cdot,N}$ and $\bar{\rho}_t$. Furthermore $\bar{\rho}_t$ given by (1.3) solves (1.1).

In addition, building solutions with the same Brownian motion (synchronous coupling), it holds for any $i \geq 1$ and $T > 0$,

$$\sup_N \sqrt{N}\, \mathbb{E}[\sup_{t \leq T} |X_t^{i,N} - X_t|] = C(T) < +\infty\,,$$

and propagation of chaos holds true.

In particular, if $Q^{k,N}$, and $\bar{Q}$ denote respectively the distribution on $C^0([0,T], (\mathbb{R}^d)^N$ of the processes $(X_\cdot^{1,N}, ..., X_\cdot^{k,N})$ and $\bar{X}_\cdot$ on a finite time interval $[0,T]$, the Wasserstein distance $W_1(Q^{k,N}, \bar{Q}^{\otimes k})$ goes to 0 as N goes to infinity with a rate $\sqrt{N}$.

The last sentence implies the propagation of chaos at the level of the paths of the processes (we shall say *process level* for short), which is of course stronger than propagation of chaos for the marginals.

The existence of a density for the given laws is an application of elliptic regularity for the particle system and a little bit of additional work for the non-linear S.D.E. Notice that the densities ρ_t^N of the marginals flow of $Q^{N,N}$ solve a Fokker-Planck equation, often called the *Liouville equation* in the PDE literature

$$\partial_t \rho_t^N(x) = \sum_i \Delta_{x^i} \rho_t^N(x) + \sum_i \nabla_{x^i}.((b(x^i) + \frac{1}{N} \sum_{j \neq i} K(x^i - x^j))\rho_t^N)(x)\,, \tag{1.4}$$

and process level propagation of chaos implies the convergence of $\rho_t^{k,N}$ to $\bar{\rho}_t^{\otimes k}$, that is propagation of chaos at *marginals level*.

Since this time, a lot of works have been devoted to extend partly or fully Theorem 1.1 to more general situations. The first immediate extension consists in replacing the independence assumption $\mu_0^N = \mu_0^{\otimes N}$ by μ_0^N is μ_0 chaotic i.e. is exchangeable and satisfies for all fixed k that the distribution $\mu_0^{(1,...,k),N}$ of $(X_0^{1,N}, ..., X_0^{k,N})$ weakly converges to $\mu_0^{\otimes N}$ as N goes to infinity.

Many of these works are using the original scheme of proof: prove existence and uniqueness for (1.2), then for (1.3) using a fixed point theorem and finally prove propagation of chaos by using a synchronous coupling between the linear and the non-linear SDE's.

An alternate approach replacing the fixed point theorem consists in proving tightness for the empirical measures using moment bounds, to show that all limit points are solving the martingale problem associated with (1.3), show uniqueness for the latter and deduce propagation of chaos from this uniqueness and Proposition 2.2 in [110]. This alternate approach is described (in an even more general framework) in the survey [99] where the W_1 distance is replaced by the W_2 distance.

We also refer to the more recent surveys [12, 77], and especially to the very recent massive one [45, 44]. One would find in the bibliography of [77] and [109] a number of examples and references of mean-field diffusive models we will not add in this paper in order to save place.

The extensions of Theorem 1.1 we are interested in, are in particular the ones where b or K are no more bounded, nor global Lipschitz.

For non bounded examples let us mention aggregation models like the granular media equation (see [22, 23] for an analytic approach, [8, 7] for a probabilistic one based on a fixed point argument, [33] for a (more general) probabilistic one using the alternate approach and look at the references therein) or some stochastic FitzHugh-Nagumo equations for which we refer to [50] and the references therein.

In the present paper we shall focus on singular kernels.

A particularly large family of interesting kernels is given for $d \geq 2$ by

$$K(x) = \chi \frac{x}{|x|^{s+2}} \tag{1.5}$$

for some s, i.e. deriving from a Riesz potential. The constant χ determines whether the model is *attractive* for $\chi > 0$ or *repulsive* for $\chi < 0$. This terminology is motivated by the following property : looking at the squared distance between two particles $|X_t^{i,N} - X_t^{j,N}|^2$, the interaction part $-\langle K(X_t^{i,N} - X_t^{j,N}), X_t^{i,N} - X_t^{j,N}\rangle$ of the dynamics is negative for $\chi > 0$

(resp. positive for $\chi < 0$) when particles are close, inducing that this effect tends to make the distance decaying (attraction) or growing (repulsion) in each case.

It is worth noticing that even the existence of a solution for the particle system is far to be clear, so that all the program has to be performed.

For $d = 2$, $\chi > 0$, $b = 0$ and $s = 0$ correspond to the celebrated *parabolic-elliptic* (Patlak)-Keller-Segel model introduced in [80, 81] in order to model some chemotaxis phenomena observed in some bacteria populations like the Dictyostelium discoideum for instance. A detailed description of the most important results on this model is given in subsection 6. Let us simply indicate here that existence and uniqueness of a *free energy* solution to (1.1) is known under mild assumptions on the initial condition (see [10, 54]) provided $\chi \leq 4$, existence and uniqueness for the particle system is studied in [63, 38, 64] for $\chi < 4$, propagation of chaos at the level of marginals is studied in [17] but replacing the euclidean space $\mathbb{R}^d$ by the torus $\mathbb{T}^d$, partly studied in [112] in the whole space at the level of marginals and in [63] at the process level but for $\chi < 1$. Notice that [17] contains quantitative results on propagation of chaos, while the results in [112, 63] are only qualitative. The Keller-Segel model is presumably one of the most difficult to study due to the explosion of solutions for $\chi > 4$ and to various phenomena, like for instance the existence of collisions between particles with positive probability for the stochastic process.

A η relaxed Keller-Segel model corresponding to $s = -\eta$ for $\eta > 0$ is studied in [68]. The main difficulties of the classical one disappear, in particular there are no collisions, and one can use the strategy in [62] for the Biot-Savart kernel we shall introduce below, in order to perform the whole program and obtain propagation of chaos at the marginals level.

Still for $d = 2$, the 2D-vortex model corresponds to $b = 0$ and to the Biot-Savart kernel

$$K(x) = \chi \left(\frac{x_2}{|x|^2}, \frac{-x_1}{|x|^2} \right) := \chi \frac{x^\perp}{|x|^2}$$

for some constant χ. It is linked to the $2D$ Navier-Stokes equation. One can see the initial [104, 105] and the more generals [98, 62]. The key point for existence and (strong) uniqueness is thus that there are no collisions in finite time, i.e. that for all $T > 0$, all $i \neq j$,

$$\inf_{0<t\leq T} |X_t^{i,N} - X_t^{j,N}| \neq 0$$

almost surely as soon as the same holds at time 0. Once this is shown, the alternate strategy of proof is used in [62] to get a qualitative analogue of Theorem 1.1 replacing convergence in Wasserstein distance by convergence in distribution at the level process. For an analogue 3D model we refer to [59].

Quantitative results at the marginals level have been recently obtained in [69] on the torus and in [55] on the whole euclidean space. The results [69] are also uniform in time.

In higher dimension $d \geq 3$, the repulsive case is studied in [109] for *sub-Coulombic* potentials, i.e. $0 < s < d - 2$ and in the very recent [47] on the torus, the latter also containing some results in dimension 2 for the periodized Keller-Segel model studied in [17]. All these works only consider the Liouville equation, hence the marginals level.

Another particularly interesting case is the one of the Dyson Brownian or Ornstein-Uhlenbeck motion, corresponding to $d = 1$, $\chi < 0$ (repulsive situation), $b = 0$ or $b(x) = -ax$ for some

$a > 0$ and $K(x) = 1/x$ which is connected to the spectrum of random matrices, studied in the 90's in [46, 108, 41, 43] and very recently in [16, 70].

The notion of propagation of chaos has been extended in several directions and a very general theory is contained in [74]. In particular one may reinforce the weak convergence of $\mu_t^{(1,\dots,k),N}$ to $\mu_0^{\otimes N}$ by adding the convergence of some entropies (see [62] definition 2.3 for this notion). This is sometimes called entropic propagation of chaos in the PDE literature.

Recall that for two probability measures μ and ν on the same Polish space the relative entropy of ν w.r.t. μ (also called Kullback information or Kullback-Leibler divergence) is defined as

$$H(\nu|\mu) = \int \ln\left(\frac{d\nu}{d\mu}\right) d\nu \;=\; \sup_{||f||_\infty<+\infty} \left(\int f\, d\nu \;-\; \ln \int e^f\, d\mu\right), \tag{1.6}$$

this quantity being infinite if ν is not absolutely continuous w.r.t. μ. If Kullback information is not a distance it controls the total variation distance d_{TV} thanks to the celebrated Pinsker's inequality

$$\sup_{||f||_\infty\le 1} \left|\int f(d\nu - d\mu)\right| := d_{TV}(\nu,\mu) \;\le\; \sqrt{2\, H(\nu|\mu)}\,. \tag{1.7}$$

Entropic propagation of chaos is proved for the vortex and the relaxed Keller-Segel models.

Recently other approaches have been proposed for the study of singular mean-field models, in particular in order to obtain quantitative and/or uniform in time propagation of chaos, i.e. results like $\sup_{t\ge 0} d(\mu_t^{1,N}, \mu_t) \le \theta(N)$ for some distance d and some $\theta(N)$ going to 0 as N growths to infinity.

The first one, initiated in [76] and extended in [78], consists in studying the time evolution of $H(\rho_t^{(1,\dots,N),N}|\bar\rho_t^{\otimes N})$ where $\bar\rho_t$ is the solution of (1.3). The Biot-Savart kernel is also studied (as an example in a larger class) in [69] combining the Jabin-Wang approach and coupling methods introduced in [53]. In these papers the dynamics is built on the torus $\mathbb{T}^{\otimes N}$, in order to ensure some boundedness properties.

The second one, explained in [109], is based on the study of the time evolution of the *modulated energy* introduced by Serfaty. Convergence for the modulated energy implies weak convergence ([109] Remark 1.5).

Both approaches are combined in [17] in order to cover more singular kernels including the Patlak-Keller-Segel case for $d = 2$,

A new entropic approach was initiated by D. Lacker in [89] and [90], using probabilistic tools and some hierarchy principle for the entropies (BBGKY hierarchy) at the process level. The results in the previous references as well as in the last opus [91] of the method are written with very general assumptions that are difficult to check on explicit singular examples (as the authors themselves are claiming). In [75, 72] the authors combine Lacker's approach and Krylov estimates as e.g. in [85], that are used to check the assumptions made by Lacker. Propagation of chaos for singular kernels like the ones we have described, is not shown and seems very hard to obtain by using this approach.

Though singular, the kernels K we have discussed, satisfy $K \in \mathbb{L}^\alpha(\mathbb{R}^d)$ for some $\alpha > 0$. A general theory for S.D.E.

$$dX_t = \sqrt{2}\, dB_t + g(t, X_t)dt$$

with a drift $g \in \mathbb{L}^p([0,T], \mathbb{L}^q(\mathbb{R}^n))$ has been developed by Krylov and co-authors. Since we are mainly interested here in homogeneous processes, we shall essentially consider the case $p = +\infty$. Existence and uniqueness for such processes is obtained for $q > n$ (see e.g. [85]). Some results were obtained in the critical case $q = n$ (see the discussion in Remark 3.9). The method has been applied to interacting particles, with $q > d$ in e.g. [73, 113]. This limitation however is too strong for the most interesting singular kernels we want to look at, except the η relaxed Keller-Segel model. One can also look at [83] for another approach based on De Giorgi's method.

The main idea of the present paper is to look at the Kullback information (relative entropy) on the path space, i.e. for probability measures that are the laws of diffusion processes rather than looking at their time marginals. Except this use, our approach does not have many similarities with the one of Lacker (except in section 7 where we will recall and get tiny improvements of the results in [90]). We shall push further this idea including the proof of existence and uniqueness for the particle system and for the non linear diffusion process, and not only the propagation of chaos result.

This paper is thus some kind of propaganda for relative entropy on the path space. This notion offers an unified approach for many of the problems we have introduced, not for all.

Relative entropy on the path space has already been used in several other contexts, for instance the study of the celebrated Schrödinger problem (see e.g. [94]), stochastic analysis (see e.g. [92, 93, 95, 1, 30]), or large time behaviour and functional inequalities (see e.g. [60, 67, 27, 32, 31, 28, 115]).

An overview of the contents of the paper.

In the next section 2 we will recall first (subsection 2.1) the basics on relative entropy on path spaces and the associated Girsanov transforms, including the use of time reversal in order to connect relative entropy and Fisher information (Proposition 2.2 and Corollary 2.3). We also recall the main results of [34] for the construction of conservative diffusions with a given flow of marginals of finite energy. The stationary situation in connection with Dirichlet forms theory is discussed in subsection 2.2. We then relate absolute continuity of the law of the particle system (w.r.t. Wiener measure) to the absence of collisions (Theorem 2.14). Finally we apply these results to the construction of some singular particle systems (η relaxed Keller-Segel, sub-Coulombic or Dyson). If the results are mainly known, the method of construction is new.

In section 3 we first recall the useful Gagliardo-Nirenberg type inequalities (Lemma 3.1) and their application to generalized Hardy-Littlewood-Sobolev inequalities (Lemma 3.2). We then recall the standard way to relate Fisher information and entropy dissipation for the marginals, and compare with the results of the previous section. The next subsection 3.2 contains a new approach and new results about existence and uniqueness of drifted Brownian motions in $\mathbb{R}^m$ with a possibly singular drift $g \in \mathbb{L}^p(\mathbb{R}^m)$ for $p \geq m$ ($p > 2$ if $m = 2$), see Theorem 3.8. This Theorem completes the huge literature on the topic (see Remark 3.9) and is new in the critical case $p = m$. We then apply the method of construction to particle systems (Theorem 3.10) and rapidly discuss similar construction using the possible absence of collisions. We refer to remark 3.9 for a more complete discussion on the literature.

Section 4 is devoted to the study of the non linear S.D.E for general (not too degenerate) interaction kernels, including the aforementioned $\mathbb{L}^p$ case.

In Sections 5 and section 6, we study in details the four examples we are mainly interested in: the η relaxed keller-Segel model which is the prototype of the $\mathbb{L}^p$ case, the sub-Coulombic model which is the prototype of repulsive models, the $2D$ vortex model, possibly adding a confining potential, which is "neutral" (i.e. nor attractive, nor repulsive) and finally the Keller-Segel model with or without confining potential, which is the prototype of attractive models. In each case we discuss existence, uniqueness and absolute continuity of the law of the particle system, existence and uniqueness for the non-linear S.D.E. The results for the sub-Coulombic model and the $2D$ vortex model complete the existing one in the literature, in particular by estimating the relative entropy of the law of the particle system w.r.t. a well chosen product measure (see Theorem 5.1, Theorem 5.5 and Theorem 5.7) as well as the non-linear S.D.E (Theorem 5.4 and Proposition 5.9).

Section 6 is entirely devoted to the Keller-Segel model. We recall the existing results for the P.D.E. (1.1) and for the particle system. In particular we recall that collisions occur with a positive probability, so that the law of the particle system is not absolutely continuous w.r.t. the Wiener measure (or any product measure). We then spend some time to prove a result suggested in [17] without any proof, namely that the marginals flow of the particle system is an entropy solution of the associated Liouville equation (see Theorem 6.17) at least when adding a confining potential. The proof is delicate and technical. We also see how existence and some uniqueness for the non-linear S.D.E. is an immediate consequence of the results of section 2 and the known analytic results for the Keller-Segel P.D.E. We also extend the existing uniqueness results on the non linear PDE.

Section 7 is a little bit different. We mainly obtain tiny improvements of the results by Lacker, and then apply them to singular models with cut-off. These results are related to several papers in the literature.

Sections 8 and 9 are devoted to the proof of convergence and propagation of chaos for the four singular models we are interested in. Based on the material of the first six sections, we build on the usual strategy: proving tightness, thanks to some bounds on relative entropy (or on moments), identify all possible limits as solutions of the non linear S.D.E., prove or use some uniqueness result for the latter. We also derive a slightly original way to obtain asymptotic independence of k particles (i.e. propagation of chaos). Let us give the main results

In all what follows Q^N (resp. $Q^{k,N}$) denotes the law of the particle system defined in (1.2) (resp. the law of the first k particles, for $k \leq N$). Existence and uniqueness of the solution of (1.2) is in general not as simple, and is part of the Theorems below.

We start with the $\mathbb{L}^p$ case.

Theorem 1.2. *(see Theorem 8.8.) Assume that $K\mathbf{1}_{|K|>A} \in \mathbb{L}^p(\mathbb{R}^d)$ for some $p \geq d$ if $d \geq 3$ or $p > 2$ if $d = 2$ and some $A > 0$ and that the additional drift b is continuous and bounded. Also assume that the initial condition $\mu_0^N = \rho^N\, dx$ is chaotic so that $\mu_0^{k,N} \to (\bar{\rho}_0\, dx^1)^{\otimes k}$ and satisfies $H(\mu_0^N|\gamma_0^{\otimes N}) \leq CN$.*

Then Q^N is chaotic and for each k, $Q^{k,N}$ weakly converges to $(\bar{Q})^{\otimes k}$ where $\bar{Q}$ is the unique solution of the non linear SDE (1.3) with initial condition $\bar{\rho}_0\,dx$.
In the case $d=2$, one can improve the result, defining $K_A = K\mathbf{1}_{|K|>A}$ and assuming,

$$\int |K_A|^2\,|\ln(|K_A|)|\,|\ln|^{\beta}(|\ln|(|K_A|))\,dx \;<\; +\infty$$

for some $\beta \in \mathbb{R}$.

This result is new in the critical case $p=d$ and for $d=2$. The proof works for the all range $p \geq d$ and does not use $\mathbb{L}^q$ bounds for the Girsanov density as previous ones for $p>d$ (the simpler proof of this kind being the one in [113]). At this point we may also mention that our method of proof extends to diffusion processes with a non-constant or even a degenerate diffusion matrix (like kinetic models), while the former approaches based on Hasminski's result (see Remark 3.9) do not. This will be done in a forthcoming work.

The $2D$ vortex model is studied in details in [62], and despite intermediate new results, we do not obtain new results, except if we add some confinement potential (though even in this case one can presumably adapt the proof of [62])

Theorem 1.3. *(see Theorem 8.9.)* ***The 2D vortex case with confinement.*** *Let $d=2$, $K(x) = \chi\,\frac{x^{\perp}}{|x|^2}$.*
Assume that the initial condition $\mu_0^N = \rho_0^N\,dx$ is chaotic so that $\mu_0^{k,N} \to (\bar{\rho}_0\,dx^1)^{\otimes k}$, satisfies $H(\mu_0^N|\gamma_0^{\otimes N}) \leq CN$ and finally satisfies $\int |x|^2\,\bar{\rho}_0(x)dx < +\infty$. Also assume that the additional drift b is a confining potential as in Theorem 5.7.
Then Q^N is chaotic and for each k, $Q^{k,N}$ weakly converges to $(\bar{Q})^{\otimes k}$ where $\bar{Q}$ is the unique solution of the non linear SDE (1.3) with initial condition $\bar{\rho}_0\,dx$ satisfying $\int_0^T I(\bar{Q}\circ\bar{\omega}_t^{-1})dt < +\infty$.

For the sub-Coulombic model we obtain

Theorem 1.4. *(see Theorem 8.12.) Consider the sub-Coulombic case: $d \geq 3$, $\chi < 0$, , and*

$$K(x) = \chi\,\frac{x}{|x|^{s+2}}\,\mathbf{1}_{x\neq 0} \;\; \textit{for} \;\; 0 < s \leq d-2\,,$$

also assume that the additional drift b is bounded and Lipschitz. Assume that the initial condition $\mu_0^N = \rho_0^N\,dx$ is chaotic so that $\mu_0^{k,N} \to (\bar{\rho}_0\,dx^1)^{\otimes k}$, and satisfies $H(\mu_0^N|\gamma_0^{\otimes N}) \leq CN$. Finally assume that for some $q > d/d-s-2$, $\bar{\rho}_0 \in \mathbb{L}^q(\mathbb{R}^d)$.

Then Q^N is chaotic and for each k, $Q^{k,N}$ weakly converges to $(\bar{Q})^{\otimes k}$ where $\bar{Q}$ is the unique solution of the non linear SDE (1.3) with initial condition $\bar{\rho}_0\,dx$ satisfying $\int_0^T ||\bar{Q}\circ\omega_t^{-1}||_{\mathbb{L}^q}\,dt < +\infty$.

At the qualitative level this result extends [109] in various directions. First it holds at the process level and not only at the marginals flow level. Second, the conditions on $\bar{\rho}_0$ are more general, since only the case $\bar{\rho}_0$ bounded is studied in [109].

Finally for the Keller-Segel model we get

Theorem 1.5. *(see Theorem 9.8) Consider the Keller-Segel case $d = 2$, $0 < \chi < 4$ and*

$$K(x) = \chi \, \frac{x}{|x|^2} \, \mathbf{1}_{x \neq 0}$$

with or without a confining potential U such that $\int e^{-U} dx < +\infty$, i.e. an additional drift $b = \nabla U$ and assuming that b is Lipschitz.
Assume that the initial condition $\mu_0^N = \rho_0^N dx$ is chaotic so that $\mu_0^{k,N} \to (\bar{\rho}_0 \, dx^1)^{\otimes k}$, satisfies $\int \rho_0^N \, |\ln \rho_0^N| dx < CN$ and $\int |x^1|^2 \, \rho_0^{1,N} \, dx^1 \, < +\infty$.
Then Q^N is chaotic and for each k, $Q^{k,N}$ weakly converges to $(\bar{Q})^{\otimes k}$ where $\bar{Q}$ is the unique solution of the non linear SDE (1.3) with initial condition $\bar{\rho}_0 dx$ whose marginals flow is the (unique) free energy solution of the non linear P.D.E. (1.1).
Actually this non linear S.D.E. admits only one solution such that its marginals flow is bounded in entropy uniformly in time.

In [17] Theorem 1.1 a similar result is obtained for the Keller-Segel model on the torus, with in addition a quantitative rate of convergence, but at the level of the Liouville equation, for entropy solutions of the latter. Compactness of the state space is essential in the proof. Theorem 1.5 is thus the first one giving the convergence (not only tightness) and identifying the limit. The latter point is delicate and requires to use Orlicz spaces and the extension of Theorem 1.2 to this framework.

To conclude this long introduction, let us say that we decided to study at the same time all these models in order to compare their behaviour and to exhibit what are the differences (attractive vs repulsive or neutral for instance). It turns out that Theorem 1.2 is also useful in some of the other cases. As we briefly said before, the approach developed in the present work, can be adapted to much more general models. This will be done elsewhere. Of course the previous approach only furnishes qualitative results, while part of the mainstream, is now concerned with quantitative results. It is not hard to see that obtaining quantitative results using our approach essentially depends on our ability to get better controls in N for the relative entropy of the full particle system. This is exactly what is done in recent quantitative papers, at the Liouville equation (marginals flow) level. It should be interesting to relax Lacker's transportation assumption in this direction.

2. Entropy on the path space, Dirichlet forms and applications.

2.1. Kullback information on the path space.

On $\Omega_T^d = C([0,T], \mathbb{R}^d)$ equipped with its usual filtration, we introduce the probability measure P which is the law (up to time T) of a reversible diffusion process $Y_\cdot = (Y_\cdot^1, ..., Y_\cdot^d)$, the $Y_\cdot^i$'s being independent copies of the 1-dimensional process, satisfying

$$y_t = y_0 + \sqrt{2} \, B_t \, - \int_0^t V'(y_s) \, ds$$

where y_0 is a random variable with density $p_0(y) = Z_V^{-1} \, e^{-V(y)}$. We will assume that existence and strong uniqueness hold for this S.D.E. It is well known that $\gamma_0(dy) = \prod_{i=1}^d \, p_0(y^i) \, dy :=$

$Z_V^{-d}\, e^{-\bar V(y)}\, dy$ is a reversible measure for this process so that if we define on Ω_T^d the time reversal operator

$$R(\omega) = \bar\omega \quad \text{where} \quad \bar\omega_t = \omega_{T-t} \tag{2.1}$$

it holds $P \circ R^{-1} = P$.

The two main examples we have in mind are the Ornstein-Uhlenbeck process corresponding to $V_2(y) = y^2/2$ and a gaussian reversible distribution, and $V_1(y) = |y|$ or V_1 a smooth (C_b^2) non-negative function that coincides with $|y|$ for $|y| \geq 1$. The main advantage of the choice of V_1 is that V_1' and V_1'' are bounded. Other choices of such V's should also be interesting.

Remark 2.1. The choice of a product measure P is arbitrary. All what follows remains true if $\bar V$ is a general potential such that the $e^{-\bar V} dx$ is a bounded measure and P is the reversible diffusion process associated to this potential. The only point is that the dimension dependence is not as explicit (see for instance Lemma 2.5 below). ♢

Let Q be another probability distribution on Ω_T^d with finite relative entropy w.r.t. P, i.e. such that

$$H(Q|P) := \int \ln\left(\frac{dQ}{dP}\right) dQ \, < \, +\infty\,.$$

Denote $\eta_0 = Q \circ \omega_0^{-1}$. First, since relative entropy is non-increasing by pushforward, it holds $H(\eta_0|\gamma_0) < +\infty$. If ρ_0 denotes the density of probability of η_0 w.r.t. the Lebesgue measure, the latter writes

$$H(\eta_0|\gamma_0) = \int (\ln \rho_0(x) + \bar V(x) + d \ln Z)\, \rho_0(dx) \, < \, +\infty. \tag{2.2}$$

Since relative entropy is non-negative we immediately deduce that for any density of probability ρ,

$$\int \rho(x) \ln \rho(x)\, dx \geq - d \ln Z_V - \int \bar V(x)\, \rho(x)\, dx\,, \tag{2.3}$$

so that in particular, as mentioned in [68] lemma 2.3, one can bound from below the entropy $h(\rho)$ by the opposite of some moments of ρ.
Remark that if ρ_0 is exchangeable, we get

$$\int \rho_0(x) \ln \rho_0(x)\, dx \geq - d \ln Z_V - \, d \int_{\mathbb{R}} V(u)\, \rho_0^1(u)\, du\,. \tag{2.4}$$

Notice that in [68] the entropy is normalized (by $2/d$ in our framework) explaining why the bound in lemma 2.3 is dimension free.

Second, Girsanov transform theory ensures that there exists a $\mathbb{R}^d$ previsible process $\beta_.$ (we will call the drift of Q) such that dQ/dP is given by the exponential (Doleans-Dade) martingale associated to the process $\beta_.$, i.e.

$$\frac{dQ}{dP}|\mathcal{F}_T = (d\eta_0/d\gamma_0)\, \exp\left(\int_0^T \langle \beta_s\,,\, \sqrt{2}\, dB_s\rangle \, - \, \int_0^T |\beta_s|^2\, ds\right) \tag{2.5}$$

where $B_s = (1\sqrt{2})(\omega_s - \int_0^s \nabla \bar V(\omega_u)\, du)$ is a Brownian motion. In addition it holds

$$H(Q|P) \, = \, H(\eta_0|\gamma_0) + \int \int_0^T |\beta_t|^2\, dt\, dQ\,. \tag{2.6}$$

More generally, if Q is absolutely continuous w.r.t. P with drift $\beta_\cdot$ furnished by the Girsanov transform theory, a necessary and sufficient condition for $H(Q|P)$ to be finite is that the right hand side of (2.6) is finite, and in this case $H(Q|P)$ is given by (2.6).

All this is developed in [34] (the correction [35] is not useful in the present framework), in connection with the construction of conservative diffusion processes with finite energy, we shall revisit later.

The key observation is that relative entropy is invariant under time reversal, so that here

$$H(Q \circ R^{-1}|P) = H(Q|P)\,. \tag{2.7}$$

This elementary observation was first made in [58, 57] in order to study time reversal on the Wiener space, then developed in the unpublished [39] and pushed forward recently in [29]. We refer to the latter for all the results we shall now describe. Readers who are not familiar with stochastic forward and backward derivatives should have a look at section 3 in [39].

Since $Q \circ R^{-1}$ has finite relative entropy we can use the same arguments as before and obtain some "backward" drift $\bar\beta_\cdot$. One then obtains (see [29] Theorem 4.9 or [39] Corollary 3.15)

Proposition 2.2. *Assume that $H(Q|P) < +\infty$. Then*

(1) *For all $t \in [0,T]$, the time marginal law $Q \circ \omega_t^{-1} = \eta_t$ is absolutely continuous w.r.t. the Lebesgue measure, and its density ρ_t satisfies*

$$\int (\ln \rho_t(x) + \bar V(x))\, \rho_t(x)\, dx \;<\; +\infty.$$

(2) *One can find two previsible processes $\beta_\cdot$ and $\bar\beta_\cdot$ such that,*

$$H(Q|P) \;=\; H(\eta_0|\gamma_0) + \int \int_0^T |\beta_t|^2\, dt\, dQ \;=\; H(\eta_T|\gamma_0) + \int \int_0^T |\bar\beta_{T-t}|^2\, dt\, dQ \circ R^{-1}$$

(3) *For all $f \in C_c^\infty(\mathbb{R}^d)$, any vector $u \in \mathbb{R}^d$ and almost any $t > 0$ it holds*

$$-\int \langle u, \nabla f(\omega_t)\rangle\, dQ \;=\; \int \big(\langle u, (\beta_t + \bar\beta_{T-t} \circ R) - \nabla \bar V(\omega_t)\rangle\big)\; f(\omega_t)\, dQ\,. \tag{2.8}$$

The final equation (2.8) is called the *duality equation*. The accurate reader will check that the previous result is written for "Markov" probability measures Q in [29] (but in the general form in [39]). One can of course take conditional expectations w.r.t. $\sigma(\omega_t)$ and obtain "Markov type" drifts. It is precisely the aim of section 3 in [34] to prove that the obtained probability measure has the same time marginals as the initial one. In any case, taking these conditional expectations we get the following corollary (see [29] (4.12) or [39] lemma 4.9)

Corollary 2.3. *In the situation of proposition 2.2, the density ρ_t satisfies the following duality equation*

$$\nabla_x \rho_t = \rho_t\; \big[\big(\mathbb{E}^Q(\beta_t|\omega_t = x) + \mathbb{E}^Q(\bar\beta_{T-t} \circ R|\omega_t = x)\big) - \nabla \bar V(x)\big]$$

for almost all $t \in (0,T)$ in the sense of Schwartz distributions on $\mathbb{R}^d$. As a consequence, the Fisher information

$$I(\rho_t) := \int \rho_t\, |\nabla \ln(\rho_t)|^2\, dx\,,$$

satisfies, for all $T>0$, all $\lambda>0$,

$$2(1+\lambda)H(\rho_T dx|\gamma_0)+\int_0^T I(\rho_t)\,dt \le 4(1+\lambda)\int\int_0^T |\beta_t|^2\,dt\,dQ+$$
$$+\left(1+\frac{1}{\lambda}\right)\int_0^T\int |\nabla\bar V(x)|^2\rho_t(x)\,dx\,dt+(1+\lambda)\int(\ln\rho_0(x)+\bar V(x))\,\rho_0(x)\,dx\,.$$

Remark 2.4. Recall that in the definition of $I(\rho)$, $|\nabla\rho|^2/\rho=0$ by convention on the set $\rho=0$. ♢

The proof of the final bound is an immediate application of $(a+b)^2\le(1+\lambda)a^2+(1+(1/\lambda)b^2$, of Cauchy-Schwarz inequality for the conditional expectations and of (2) in proposition 2.2. Since relative entropy $H(\eta_T|\gamma_0)\ge 0$ we thus get a bound for the Fischer information..
Notice that the same (2) in proposition 2.2 also furnishes

$$\int(\ln\rho_T(x)+\bar V(x))\,\rho_T(x)\,dx\le\int\int_0^T|\beta_t|^2\,dt\,dQ+\int(\ln\rho_0(x)+\bar V(x))\,\rho_0(x)\,dx\,. \tag{2.9}$$

These two bounds can replace (with different constants) equation (3.8) in [68]. They also will be related to (6.11) and the notion of entropy solution in [17].

In some situations one can go further. For instance a simple application of Ito's formula furnishes the next lemma.

Lemma 2.5. *In the situation of proposition 2.2, assume that V is smooth and that, $|V'|$ and $|V''|$ are bounded by A. Then*

$$\int\bar V(x)\,\rho_T(x)\,dx\le\int\bar V(x)\,\rho_0(x)\,dx+A\sqrt{dT}\left(\int\int_0^T|\beta_t|^2\,dt\,dQ\right)^{\frac12}+AdT\,.$$

As a consequence

$$\begin{aligned}\int\ln\rho_T\,\rho_T\,dx &\ge -(d\ln Z_V+\int\bar V(x)\,\rho_0(x)\,dx+A\sqrt{d}\left(\int\int_0^T|\beta_t|^2\,dt\,dQ\right)^{\frac12}+dAT)\\ \int\ln\rho_T\,\rho_T\,dx &\le \int\int_0^T|\beta_t|^2\,dt\,dQ+\int(\ln\rho_0(x)+\bar V(x))\,\rho_0(x)\,dx\,,\end{aligned} \tag{2.10}$$

and

$$\int_0^T I(\rho_t)\,dt\le 4(1+\lambda)\int\int_0^T|\beta_t|^2\,dt\,dQ+(1+(1/\lambda))\,d\,A^2\,T+(1+\lambda)\int(\ln\rho_0(x)+\bar V(x))\,\rho_0(x)\,dx\,. \tag{2.11}$$

The main result of [34] is a converse construction. Let $t\mapsto\nu_t$ be a flow of probability measures on $\mathbb{R}^d$ and $(t,x)\mapsto b(t,x)$ some measurable function.

Definition 2.6. We will say that $\nu_\cdot$ is an admissible flow if

(i) $\nu_\cdot$ satisfies the (b,C_b^∞) weak forward equation i.e. for all $f\in C_b^{1,\infty}(\mathbb{R}^+\times\mathbb{R}^d)$ and all $0\le s<t\le T$,

$$\int f(t,x)d\nu_t-\int f(s,x)d\nu_s=\int_s^t\int(\partial_u f+\Delta_x f+\langle 2b(u,x)-\nabla\bar V(x),\nabla_x f\rangle)\,d\nu_u\,du\,.$$

(ii) b is of finite ν-energy, i.e. for all $T \geq t > 0$,

$$\int_0^t \int |b(s,x)|^2 \nu_s(dx) ds \, < \, +\infty \, .$$

(iii) $H(\nu_0|\gamma_0) < +\infty$.

Theorem (4.18) in [34] then says

Theorem 2.7. *Let ν be an admissible flow. Define $\beta_s = b(s,\omega_s)\,\mathbf{1}_{s<\tau}$ where $\tau = \sup_n \tau_n$ and $\tau_n = \inf\{t > 0, \int_0^t |b|^2(u,\omega_u)\,du \geq n\}$.*
Then Q defined by (2.5) is a probability measure satisfying $H(Q|P) < +\infty$ and is a weak solution (up to time T) of the stochastic differential equation (SDE)

$$dX_t = \sqrt{2}\,dB_t \, + \, (2b(t,X_t) - \nabla \bar{V}(X_t))dt$$

with initial distribution ν_0. In addition $Q \circ \omega_t^{-1} = \nu_t$.

Remark 2.8. The proof in [34] uses stochastic calculus. In [36, 37] the same result (and more general ones) is obtained by using Sanov theorem, i.e. a large deviation approach. $\diamondsuit$

Actually the construction in [34] is more general. We may replace P by any measure with initial measure ν_0 which is no more reversible nor invariant, provided P is the unique solution of the associated martingale problem. Theorem 2.7 then becomes

Theorem 2.9. *Let P be the unique solution of the martingale problem associated to the generator $L_P = \Delta + 2\,g\,\nabla$, with initial measure θ_0. Let ν be an admissible flow in the following sense*

(i) *$\nu_\cdot$ satisfies the (b, C_b^∞) weak forward equation i.e. for all $f \in C_b^{1,\infty}(\mathbb{R}^+ \times \mathbb{R}^d)$ and all $0 \leq s < t \leq T$,*

$$\int f(t,x) d\nu_t - \int f(s,x) d\nu_s = \int_s^t \int \left(\partial_u f + \Delta_x f + \langle 2b(u,x), \nabla_x f\rangle\right) d\nu_u \, du \, .$$

(ii) *$b - g$ is of finite ν-energy, i.e. for all $T \geq t > 0$,*

$$\int_0^t \int |(b-g)(s,x)|^2 \nu_s(dx) ds \, < \, +\infty \, .$$

(iii) $H(\nu_0|\theta_0) < +\infty$.

Then Q defined by (2.5) and $\beta_s = (b-g)(s,\omega_s)$ is a probability measure satisfying

$$H(Q|P) = H(\nu_0|\theta_0) + \int_0^T \int |(b-g)(s,x)|^2 \nu_s(dx) ds \, < \, +\infty \, .$$

Q is a weak solution (up to time T) of the stochastic differential equation (SDE)

$$dX_t = \sqrt{2}\,dB_t \, + \, 2b(t,X_t)dt$$

with initial distribution ν_0. In addition $Q \circ \omega_t^{-1} = \nu_t$ and $\nu_t(dx) = \rho_t(x)\,dx$ for some flow of probability densities.

One can also use similar time reversal arguments, but this time the duality equation involves the marginals flow of P and is thus more delicate (see [29, 38]), except when P is time reversible, in which case what has previously been done extends immediately. We shall describe an application of this more general statement.

Remark 2.10. Let P be as in Theorem 2.9, with a time homogeneous drift g and assume that $P \circ R^{-1} = P$. Let $Q = \frac{d\nu_0}{d\theta_0} P$. Q is thus the solution of

$$dY_t = \sqrt{2}\, dB_t + 2g(Y_t)dt \quad ; \quad Q \circ \omega_0^{-1} = \nu_0 \,.$$

In addition

$$H(Q|P) = H(\nu_0|\theta_0)$$

is assumed to be finite. We thus deduce the existence of a time reversed drift $\hat{\beta}_s$ such that

$$H(\nu_0|\theta_0) = H(Q|P) = H(Q \circ R^{-1}|P) = H(\nu_T|\theta_0) + \int \int_0^T |\hat{\beta}_{T-t}|^2 dt dQ \circ R^{-1}$$

and satisfying (if ρ_t denotes the density of ν_t),

$$\nabla_x \rho_t = \rho_t \left[\mathbb{E}^Q (\hat{\beta}_{T-t} \circ R | \omega_t = x) + 2g(x) \right] \tag{2.12}$$

so that

$$H(\nu_0|\theta_0) \geq H(\nu_T|\theta_0) + \int_0^T \int |\nabla \ln(\rho_t) - 2g|^2 \, \rho_t \, dx \, dt \,. \tag{2.13}$$

When g derives from a potential i.e. $2g = -\nabla U$ (this is natural since P is reversible), one recognizes the *free energy* functional $F(\rho) = \int \rho(\ln(\rho) + U)dx$, used in several studies of non linear P.D.E., in particular, (2.12) can be rewritten

$$F(\rho_0) \geq F(\rho_T) + \int_0^T \int |\nabla \ln(\rho_t) + \nabla U|^2 \, \rho_t \, dx \, dt$$

which is similar to (1.2) in [10] concerned with the Keller-Segel model. In other words, if it exists an unique reversible solution to the martingale problem associated to L_P, the marginals flow, starting from a finite free energy condition, is naturally a *free energy solution* of the associated Liouville equation. We shall see that this property is true in many cases.

In the previous situation we may obtain additional informations on ρ_t. Indeed if $d\theta_0 = \theta_0(x)\, dx$, using reversibility we obtain

$$\rho_t(x) = \mathbb{E}_P \left[\frac{\rho_0}{\theta_0}(Y_t) | Y_0 = x \right] \theta_0(x) \,.$$

In particular

$$||\rho_t/\theta_0||_{\mathbb{L}^q} \leq ||\rho_0/\theta_0||_{\mathbb{L}^q}$$

for all $1 \leq q \leq +\infty$ and

$$\inf(\rho_t/\theta_0) \geq \inf(\rho_0/\theta_0) \,.$$

$\diamondsuit$

Remark 2.11. It is interesting to notice the difference between finite relative entropy and absolute continuity. It is indeed well known that if a stochastic process $Y_.$ is such that $Y_. - \int_0^. h(Y_s)ds$ is a Brownian motion, then the law Q of $Y_.$ on a finite time interval $[0, T]$ is absolutely continuous w.r.t. the Wiener measure if and only if $\int_0^T |h(Y_s)|^2 ds < +\infty$ almost surely. One can of course replace the Wiener measure by any P which is equivalent, as the one we are considering in this section. Passage to finite entropy is thus simply assuming that this quantity has finite expectation. $\diamondsuit$

Remark 2.12. Let us say a word about weak uniqueness, since this aspect is not explicitly stated in [34]. The two previous Theorems 2.7 and 2.9 are saying that a solution Q is built via (2.5). This measure is actually the Föllmer measure associated to the exponential martingale (see [34] 1.13, 1.14 and 1.15 for this notion and references). In particular, contained in the statement of these theorems, is the following fact: $\tau_n \to T$, Q almost surely as n goes to infinity. Now if Q' denotes a solution of the SDE in the theorems, Girsanov theory says that Q' and Q coincide up to the stopping times τ_n. In particular $\tau_n \to T$, Q' almost surely too and $Q = Q'$. ♢

Remark 2.13. The construction of a diffusion with given marginals flow is an old question. We discovered it in Carlen's paper [18] which was motivated by Nelson's stochastic mechanics. The method used in [18] is purely analytic. In our papers [34, 36, 37] we have shown that this construction can be done (and extended) using relative entropy on the path space. Note that the forward and backward derivatives used in in [29] are actually an extended version of the ones introduced by Nelson. It seems that this question did not attract interest (except in some works on mathematical finance) during the following twenty years. A new existence result appeared in [114] where the following is shown:

if a flow $\eta_.$ of probability measures satisfies the $(b, C_b^{1,\infty})$ weak forward equation and the integrability condition $\int \int_0^T |b(u,x)|\, \eta_u(dx)\, du < +\infty$, one can build a solution Q to the martingale problem such that $\eta_.$ is exactly the marginals flow of Q.

This is Theorem 2.5 in [114], called "superposition principle" therein, and later used in some works. It is worth to notice that, even if a less demanding $\mathbb{L}^1$ condition is required (instead of the energy $\mathbb{L}^2$ integrability), weak uniqueness is not contained in the superposition principle, contrary to the entropic case (Remark 2.12). As explained in Remark 2.11, absolute continuity is not ensured for the process built via the previous principle. ♢

2.2. Some useful elements of Dirichlet forms theory.

As we said the use of Dirichlet forms is (at least at a "theoretical" level, as said in [63]) very useful in our context. We shall here recall some useful results obtained in the 90th's. We refer to [97, 66, 30] and the references in the last mentioned paper for all the material described below.

Let $\nu(dx) = \rho(x)\, dx$ be a σ-finite, non necessarily bounded, positive measure. Consider the symmetric bilinear form

$$\mathcal{E}_\rho(f,g) = \int \langle \nabla f, \nabla g \rangle \, d\nu$$

defined on $C_c^\infty(\mathbb{R}^d)$. If

$$\rho^{\frac{1}{2}} \in H^1_{loc}(\mathbb{R}^d) \quad \text{i.e. for any } R, \quad \int_{|y|\le R} |\nabla \ln(\rho)|^2 \, d\nu < +\infty\,, \tag{2.14}$$

then the form $\mathcal{E}_\rho$ is closable and its minimal extension (still denoted by $\mathcal{E}_\rho$ on its domain) is a Dirichlet form.

Assume that ν is bounded (hence a probability measure after renormalization) and $\rho^{\frac{1}{2}} \in H^1(dx)$. (2.14) thus simply becomes the finite energy condition in Theorem 2.9 for the stationary flow $\rho_t = \rho$. We recall below the main results proved almost thirty years ago.

(1) If $\rho \neq 0$ dx almost everywhere, then it is also the maximal closed markovian extension, i.e. Markov uniqueness holds.
(2) Under the previous assumption, $\mathcal{E}_\rho$ is regular and local, so that there exists a ν-symmetric diffusion Q_x associated to this form.
(3) Q_ν is a weak solution of

$$dY_t = \sqrt{2}\, dB_t + \nabla(\ln \rho)(Y_t)dt$$

with initial distribution equal to ν. Furthermore $(Q_x)_{x\in\mathbb{R}^d}$ is a solution of the same SDE starting from x for quasi-every x.
(4) Q_ν is absolutely continuous w.r.t. the Wiener measure with initial distribution ν denoted by W_ν, and the density is given by (2.5) with $\beta = \nabla\rho/\rho$, hence for all $T > 0$,

$$H(Q_\nu|W_\nu) < +\infty \text{ in restriction to } [0,T].$$

For ν quasi every x, Q_x is absolutely continuous w.r.t the Wiener measure starting from x with the same Girsanov density.
As a consequence, for ν almost all x, $H(Q_x|W_x) < +\infty$ and

$$H(Q_\nu|W_\nu) = \int H(Q_x|W_x)\, \nu(dx)\,.$$

An interesting point, already proved in [100], is that, if ρ is a density of probability, the *nodal set* i.e. $\mathcal{N} = \{\rho = 0\}$ is not attained. According to section 4 in [30], one can "desintegrate" Q_ρ in order to get the density of Q_x w.r.t. the Wiener measure starting from x, for all x outside some explicit polar set, and the process does not hit the nodal set for those x's, Q_x almost surely.

If $\rho(x)dx$ is not bounded, some properties are preserved, in particular the last one, provided the process is conservative. An easy way to see it is to consider a family ρ_k of densities such that $\rho_k = \rho$ on the ball $B(0,k)$ and $\rho_K^{\frac{1}{2}} \in H^1(dx)$. This is easy to build. The corresponding process coincides with the initial one up to the exit time $\bar\tau_k$ of this ball, so that, if σ denotes the hitting time of $\mathcal{N}$, we have

$$Q_x(\bar\tau_k \wedge \sigma < +\infty) = Q_x^k(\bar\tau_k \wedge \sigma < +\infty) = Q_x^k(\bar\tau_k < +\infty) = Q_x(\bar\tau_k < +\infty) \to_{k\to+\infty} 0$$

since the process is conservative.
Notice that part of these results are the *stationary* version of Theorem 2.7 and are shown using large deviations results in Theorem 5.5 of [37].

2.3. **Absolute continuity and collisions.**

It is very natural to ask about a generalization of relative entropy, namely, is it possible to describe the absolute continuity of Q w.r.t. P in terms of the drift β_s. The Girsanov formula (2.5) tells us that in any case one has to consider $\int_0^T |\beta_s|^2 ds$. This presumably (but we have no proof of this) implies that a weaker integrability condition for dQ/dP, replacing $x \ln x$ by a smaller Young function cannot be described in terms of the sole β_s.
Nevertheless, if we consider Q^N the law solution of (1.2), if it exists, one may ask about the consequences of its absolute continuity w.r.t. P. Actually absolute continuity is simply connected to the absence of collisions. We shall state a general result in this direction.

Theorem 2.14. *Consider the system (1.2) for $N \geq 2$. Assume that $b = b_1 + b_2$ where b_1 is bounded, b_2 is local Lipschitz and the solution of $dY_t = \sqrt{2}\, dB_t - b_2(Y_t)dt$ is conservative, i.e. does not explode in finite time.*
Assume in addition that, for any $\varepsilon > 0$, K is local Lipschitz and bounded in the set

$$C_\varepsilon^c = \{(x^1, ..., x^N) \in (R^d)^N \; ; \; \inf_{i,j} |x^i - x^j| \geq \varepsilon\} \, .$$

Let $C_0 = \{(x^1, ..., x^N) \in (R^d)^N \; ; \; x^i = x^j$ for at least one pair $i \neq j\}$ be the collision set. By convention $K(0) = 0$.
Let Q be a solution of (1.2) such that $Q(X_0^N \in C_0) = 0$, assumed to exist. Then there is an equivalence between:

- *for all $T > 0$, Q is absolutely continuous w.r.t. the Wiener measure (actually the distribution of $\sqrt{2}$ times a dN-dimensional Brownian motion) in restriction to $[0, T]$,*
- *the hitting time of C_0 is Q almost surely infinite, i.e. the solution of (1.2) has no collisions.*

Proof. We may first consider the case $b_1 = 0$, and then get the result via a Girsanov transform. The first part is then a simple consequence of the fact that for $d \geq 1$, a $2d$-dimensional Brownian motion starting away from 0 never hits the origin. The same will hold for the distribution of any pair of particles under Q if Q is absolutely continuous. Conversely, one can build the solution starting from some $x \notin C_0$ denoted by Q_x up to the exit time τ_ε of $C_\varepsilon^c \cap \{|x| \leq \varepsilon^{-1}\}$ thanks to our regularity assumptions. This sequence of stopping times goes to infinity according to our assumptions on Q. One can thus build a strong solution up to infinity. In addition in restriction to $T \wedge \tau_\varepsilon$, Q_x is given by the Girsanov density

$$Z_T^\varepsilon = \exp\left(- \int_0^{T \wedge \tau_\varepsilon} \langle \beta_s \sqrt{2} d\omega_s \rangle \; - \; \frac{1}{2} \int_0^{T \wedge \tau_\varepsilon} |\beta_s|^2 ds \right)$$

where $\beta_s^j = b(\omega_s^j) + \sum_k K(\omega_s^j - \omega_s^k)$ and the reference measure is P_x. Let A be $\mathcal{F}_T$ measurable and such that $P_x(A) = 0$. Then $P_x(A \cap \{\tau_\varepsilon \geq T\}) = 0$ so that $Q_x(A \cap \{\tau_\varepsilon \geq T\}) = 0$. It remains to use the fact that on $\mathcal{F}_{T \wedge \tau_\varepsilon}$, $Q = \int Q_x \; Q \circ \omega_0^{-1}(dx)$, so that $Q(A \cap \{\tau_\varepsilon \geq T\}) = 0$ and since $\{\tau_\varepsilon \geq T\}$ is non-decreasing (ε is supposed to decrease to 0) and satisfies $\lim_\varepsilon Q(\tau_\varepsilon \geq T) = 1$, we get $Q(A) = 0$. □

2.4. **Applications to some singular models.**

In this subsection we will prove a general existence theorem for the particle system (1.2). In order to understand the assumptions required for this goal, we shall first look at some example.

2.4.1. ***The relaxed Keller-Segel model.*** Pick $\eta > 0$ and consider

$$\rho_\eta(x) = \exp\left(- \frac{\chi}{N\eta} \sum_{1 \leq i < j \leq N} |x^i - x^j|^\eta \right) \quad , \quad x = (x^1, ..., x^N) \in (\mathbb{R}^2)^N \, .$$

It is easily seen that $\rho_\eta^{\frac{1}{2}} \in H^1_{loc}(dx)$ but $\nu(dx) = \rho(x)dx$ is not bounded.

In order to apply the Dirichlet forms approach, one can add a confining potential i.e. consider

$$\rho_\eta^M(x) = \exp\left(-\frac{\chi}{N\eta}\sum_{1\le i<j\le N}|x^i-x^j|^\eta\right)\exp(-V_M(x)) \quad , \quad x=(x^1,...,x^N)\in(\mathbb{R}^2)^N\,,$$

where V_M is smooth, nonnegative, vanishes on $B(0,M)$ and coincides with $|x|^2$ outside $B(0,M+1)$. This time $\rho_\eta^M dx$ is a bounded measure and $\sqrt{\rho_\eta^M}\in H^1(dx)$. We may thus apply what precedes and get the existence and uniqueness of a weak solution of

$$dX_t^{i,N,M} = \sqrt{2}\,dB_t^{i,N} - \frac{\chi}{N}\sum_{j=1}^N \frac{X_t^{i,N,M}-X_t^{j,N,M}}{|X_t^{i,N,M}-X_t^{j,N,M}|^{2-\eta}}\,dt - \nabla_i V_M(X_t^{N,M})\,dt \quad , \quad i=1,...,N$$

with initial distribution any probability measure absolutely continuous w.r.t. ρ_η^M, hence w.r.t. dx. Actually one can also consider δ_x for quasi every x. Notice that arguing as in [38] subsection 3.1, one can prove that the sets of zero capacity are exactly the polar sets for the process. If the initial measure is absolutely continuous w.r.t. Lebesgue's measure, we deduce from what we said in subsection 2.2 that Q is absolutely continuous w.r.t. the Wiener measure, and thanks to Theorem 2.14, that there are no collisions.

If we start from a non collision point x, we may build a solution up to the exit time τ of a small ball centered at x that does not intersect C_0. The law of $X_{\varepsilon\wedge\tau}^{N,K}$ is supported by a non-polar set, so that we can build a solution starting from it. It remains to concatenate both laws to build a solution starting from x. Hence we get a solution starting from any non collision point x, and this solution never hits C_0, so that it is a strong solution as the authors of [68] are saying.

Since we have strong solutions, we may build a family of solutions for $K\in\mathbb{N}^*$ with the same Brownian motion and define $X_t^N = X_t^{N,M}\,\mathbf{1}_{T_{M-1}<t\le T_M}$ where T_M is the exit time of $B(0,M)$ for the process $X_\cdot^{N,M}$, hence for $X_\cdot^N$. It remains to show that $\lim_{M\to+\infty}T_M = T_\infty = +\infty$ almost surely, i.e. that the law Q_x of $X_\cdot^N$ is conservative.

The proof is standard. First, $\lim_{M\to+\infty}|X_{T_M}^N| = +\infty$ almost surely. Second,

$$\sum_i\left\langle x^i, \sum_j \frac{x^i-x^j}{|x^i-x^j|^{2-\eta}}\right\rangle = \sum_{i,j}|x^i-x^j|^\eta \ge 0\,.$$

Hence applying Ito's formula, and provided $\chi\ge 0$, we get

$$|X_{t\wedge T_M}^N|^2 \le |X_0^N|^2 + 2\sqrt{2}\int_0^t \mathbf{1}_{s\le T_M}\sum_i\langle X_s^{N,i}, dB_s^{N,i}\rangle + 2dN(t\wedge T_M) \tag{2.15}$$

so that $\mathbb{E}_x(|X_{t\wedge T_M}^N|^2)\le |x|^2+2dNt$ implying $M^2\mathbb{P}(T_M\le t)\le |x|^2+2dNt$ for all M, and finally $T_\infty=+\infty$ almost surely.

If $\chi<0$, since $0<\eta<2$, we may use $|x^i-x^j|^\eta\le 1+2|x^i|^2+2|x^j|^2$ so that

$$\sum_i\left\langle x^i, \sum_j \frac{x^i-x^j}{|x^i-x^j|^{2-\eta}}\right\rangle = \sum_{i,j}|x^i-x^j|^\eta \le N(N+2|x|^2)\,.$$

It follows

$$|X^N_{t\wedge T_M}|^2 = |X^N_0|^2 + 2\sqrt{2}\int_0^t \mathbf{1}_{s\le T_M}\sum_i \langle X^{N,i}_s, dB^{N,i}_s\rangle + 2dN(t\wedge T_M) \qquad (2.16)$$
$$+2|\chi|\int_0^t \mathbf{1}_{s\le T_M}\left(N+2|X^N_s|^2\right)ds$$

so that

$$\mathbb{E}_x(|X^N_{t\wedge T_M}|^2) \le |x|^2 + 2dNt + 2|\chi|\int_0^t \mathbb{E}_x(\mathbf{1}_{s\le T_M}\left(N+2|X^N_s|^2\right))\,ds\,.$$

According to Gronwall's lemma we thus deduce that for any fixed t,

$$\mathbb{E}_x(|X^N_{t\wedge T_M}|^2) \le C(t,N,|\chi|)$$

and we may conclude as before.

This answers the question of existence and uniqueness for the η-sub-critical model in [68], i.e.

$$dX^{i,N}_t = \sqrt{2}\,dB^{i,N}_t - \frac{\chi}{N}\sum_{j=1}^N \frac{X^{i,N}_t - X^{j,N}_t}{|X^{i,N}_t - X^{j,N}_t|^{2-\eta}}\,dt \quad , \quad i=1,...,N \qquad (2.17)$$

and shows at the same time that there are no collisions. Notice that the previous construction also furnishes a solution in the repulsive case $\chi<0$.

2.4.2. ***A general result.*** We can now introduce the required assumptions

Assumption 2.15. $b=b_1+b_2$ where b_1 is bounded, b_2 is local Lipschitz and the solution of $dY_t=\sqrt{2}\,dB_t - b_2(Y_t)dt$ is conservative, i.e. does not explode in finite time.
In addition, for any $\varepsilon>0$, K is local Lipschitz and bounded in the set

$$C^c_\varepsilon = \{(x^1,...,x^N)\in(R^d)^N\ ;\ \inf_{i,j}|x^i-x^j|\ge\varepsilon\}\,.$$

Finally $C_0=\{(x^1,...,x^N)\in(R^d)^N\ ;\ x^i=x^j$ for at least one pair $i\ne j\}$ denotes the collision set. By convention $K(0)=0$.

Assumption 2.16.
(1) There exists a potential U such that $K=\chi\,\nabla U$, and a potential V such that $b=\nabla V$.
(2) The measure $d\nu=\rho dx$ is bounded, where

$$\rho(x)=\exp\left(-\frac{\chi}{N}\sum_{1\le i<j\le N}U(x^i-x^j)\right)\exp\left(-V(x)\right) \quad , \quad x=(x^1,...,x^N)\in(\mathbb{R}^d)^N\,.$$

(3) $\rho^{\frac12}\in H^1(dx)$.

and a second one

Assumption 2.17.
(1) There exists a potential U such that $K=\nabla U$, and a potential V such that $b=\nabla V$.
(2) One can find a sequence $(V_M)_M$ of smooth, nonnegative potentials such that $V_M=0$ on the ball $B(0,M)$, and such that the potentials K and $U+V_M$ satisfy Assumption 2.16.

(3)

$$(a) \quad \sum_i \left\langle x^i, \sum_j K(x^i - x^j) \right\rangle \geq 0$$

or

$$(b) \quad \left| \sum_i \left\langle x^i, \sum_j K(x^i - x^j) \right\rangle \right| \leq c(N)(1 + |x|^2)$$

for some $c > 0$.

(4)

$$(a) \quad \langle x, \nabla V(x) \rangle \geq 0$$

or

$$(b) \quad |\langle x, \nabla V(x) \rangle| \leq c|x|^2$$

for some constant $c > 0$.

We then have

Theorem 2.18. *Assume that ν_0 is a probability measure on $(\mathbb{R}^d)^N$ such that $\nu_0(C_0) = 0$ and that Assumption 2.15 is fulfilled.*
If Assumption 2.16 is satisfied, there exists an unique weak solution to the system (1.2). The law Q_{ν_0} of this solution is absolutely continuous w.r.t. the Wiener measure with the same initial condition W_{ν_0} and there are no collisions. Hence if b_1 is local Lipschitz too, the solution is strong. In addition for all $T > 0$, in restriction to $[0, T]$, $H(Q_\nu | W_\nu) < +\infty$, for ν defined in Assumption 2.16.
If Assumption 2.17 is satisfied, the same conclusion, except the finiteness of the relative entropy, is true.

The proof is exactly the same as in the particular relaxed Keller-Segel case.

2.4.3. ***Some more examples.***

(1) Granular media. The granular media equation studied in [33] enters the framework of Theorem 2.18. Of course the direct proof given in [33] is much simpler.

(2) sub-Coulombic potentials in dimension $d \geq 3$. Consider

$$\rho(x) = \exp\left(\frac{\chi}{sN} \sum_{1 \leq i < j \leq N} |x^i - x^j|^{-s} \right) \quad , \quad x = (x^1, ..., x^N) \in (\mathbb{R}^d)^N ,$$

i.e. $U(y) = -|y|^{-s}$ for some $s > 0$. The measure $\rho(x)dx$ is bounded on compact sets (Radon measure), if and only if $\chi < 0$. Hence we can only consider this situation.

Since $K(y) = \frac{\chi y}{|y|^{s+2}}$ satisfies (3) (b) in Assumption 2.17, $\rho^{\frac{1}{2}} \in H^1_{loc}(dx)$ and one can easily build a sequence V_M of confining potentials (gaussian outside some large ball for instance). We can thus apply Theorem 2.18 and get the existence (and uniqueness) of a weak (resp. strong) solution to

$$dX_t^{i,N} = \sqrt{2}\, dB_t^{i,N} - \frac{\chi}{N} \sum_{j=1}^N \frac{X_t^{i,N} - X_t^{j,N}}{|X_t^{i,N} - X_t^{j,N}|^{s+2}} dt - b(X_t^{i,N}) dt \quad , \quad i = 1, ..., N \tag{2.18}$$

for any $\chi < 0$ and $s > 0$ and $b = \nabla V + b_1$ where V satisfies Assumption 2.17 (4) and b_1 is bounded (resp. Lipschitz and bounded) . In addition no collision occurs.

This result entails the one obtained in section 4 of [109] where the range of s is reduced to $0 < s < d-2$. Actually we do not really see why this restriction is useful in the proof of [109]. Extending this range to larger s's is natural. Indeed the more s is large, the more C_0 becomes repulsive (since $\chi < 0$), while in C_1^c (for $\varepsilon = 1$) the drift term becomes smaller. However we shall see later that the condition $s \leq d-2$ naturally appears for entropic controls.

By the way, it seems us that the proof of Lemma 4.2 in [109] is wrong because inequality (4.13) therein is not correct, $\leq$ has to be changed into $\geq$. In other words Assumption 2.17 (3) (a) is not satisfied, but the previous proof shows that the statement of Lemma 4.2 in [109] is correct.

(3) *The Dyson processes*. The Dyson Ornstein-Uhlenbeck is the process associated to the S.D.E.

$$dX_t^{i,N} = \sqrt{2}\, dB_t^{i,N} - \frac{\chi}{N} \sum_{j=1}^{N} \frac{1}{X_t^{i,N} - X_t^{j,N}}\, dt - \beta X_t^{i,N} dt \quad , \quad i = 1, ..., N \tag{2.19}$$

where each particle $X^{i,N} \in \mathbb{R}$. The associated reversible measure ν^N is thus given on $\mathbb{R}^N$ by the density

$$\rho(dx) = \prod_{i \neq j} |x^i - x^j|^{-\chi/N}\, e^{-\beta |x|^2}\, .$$

Assumption 2.16 is satisfied if and only if $\chi/N < -1$ so that we get existence and absence of collisions in this repulsive case. We thus recover the result in [41] (Theorem 3.1 and Proposition 4.1) except that we do not cover the equality case $\chi/N = -1$ for which the same existence result is known (also see [108]). In addition we know that the law of the process has finite relative entropy w.r.t. the law of the Ornstein-Uhlenbeck process. In the equality case we only know thanks to Theorem 2.14 that the law is absolutely continuous.

We shall come back later to the proofs. Notice that recently [70] the authors proposed another proof, for $\chi/N < -1$ (simply use a linear time change to compare the results) based on the existence of a Lyapunov function and Khasminski explosion test.

In what precedes we may replace $e^{-\beta|x|^2}$ by $e^{-\beta\psi(x)}$ for some smooth ψ such that $\psi(x) = 0$ if $|x| \leq M$, and $\psi(x) = |x|^2$ for $|x| \geq M+1$. If T_M denotes the first exit time from $\max_i |x^i| \leq M$, this modified Dyson Ornstein-Uhlenbeck process coincides with the Dyson Brownian motion obtained for $\beta = 0$, up to time T_M. It is thus enough to show that T_M goes to infinity almost surely for the latter. As for (2.16) it is enough to remark that

$$\mathbb{E}(\sum_i |X^i_{T\wedge T_M}|^2) \leq \mathbb{E}(\sum_i |X^i_0|^2) + 2T(N + |\chi|(N-1))$$

since $\sum_i x^i \sum_{j\neq i} \frac{1}{x^i - x^j} = N-1$.

One thus see that many existing results in the literature enter the entropic framework.

3. Approximations and singular diffusions.

In this section we shall introduce and use some analytic tools in order to study the existence of singular diffusions without calling upon the Dirichlet forms theory, but still using relative entropy.

3.1. **Some analytic bounds.**

We will collect in this subsection some useful bounds for the sequel.
The finiteness of Fisher information has some regularity consequences. The following is the extension to $\mathbb{R}^d$ of Lemma 3.2 in [62] written for $d = 2$:

Lemma 3.1. *Let ρ be a density of probability in $\mathbb{R}^d$, $d \geq 2$. Then denoting by $||.||_p$ the $\mathbb{L}^p(dx)$ norm, and $I(\rho)$ its Fisher information $I(\rho) = \int \rho\, |\nabla \ln \rho|^2 dx$, it holds*

(1) *for all $q \in [1, (d/d-1))$,*

$$||\nabla \rho||_q \leq a^{\frac{d(q-1)}{q}} \left(I(\rho)\right)^{\frac{qd+q-d}{2q}}, \tag{3.1}$$

with $a = \frac{d-1}{d-q} \frac{q}{d^{\frac{1}{2}}}$,

(2) *for all $p \in [1, (d/d-2))$,*

$$||\rho||_p \leq a^{\frac{d(p-1)}{p}} \left(I(\rho)\right)^{\frac{d(p-1)}{2p}}, \tag{3.2}$$

with $a = \frac{2p(d-1)}{p(d-2)+d} \frac{1}{d^{\frac{1}{2}}}$.

(3) *for $d \geq 3$ (3.2) (resp. (3.1)) is still true for $p = d/(d-2)$ (resp. $p = d/d-1$).*

Proof. The proof of the two first items is the same as the one in [62]. First for any $1 \leq q < 2$, Hölder's inequality furnishes

$$||\nabla \rho||_q \leq \left(I(\rho)\right)^{\frac{1}{2}} ||\rho||_{q/(2-q)}^{\frac{1}{2}}.$$

Next according to the Sobolev inequality (see e.g. [51] p-42),

$$||\rho||_{q*} \leq \frac{d-1}{d-q} \frac{q}{d^{\frac{1}{2}}} ||\nabla \rho||_q \quad \text{for} \quad q^* = \frac{qd}{d-q}.$$

We may thus use interpolation between $\mathbb{L}^{q/(2-q)}$, $\mathbb{L}^1$ and $\mathbb{L}^{q*}$, provided $q^* > q/(2-q)$ i.e. $q < d/d-1$. We obtain since ρ is a density of probability,

$$||\rho||_{q/(2-q)} \leq ||\rho||_{q^*}^{\frac{2d(q-1)}{qd+q-d}} \leq a^{\frac{2d(q-1)}{qd+q-d}} ||\nabla \rho||_q^{\frac{2d(q-1)}{qd+q-d}} \tag{3.3}$$

and finally (3.1). Plugging this inequality in (3.3) we get (3.2) if we denote $p = q/(2-q)$.
Notice that for $d = 2$ and $q = 2$, (3.1) is still true but $a = +\infty$.
For $d \geq 3$ first remark that (3.1) and (3.2) extend to the limiting $d/d-1$ and $d/d-2$, first for compactly supported ρ. Indeed take some compactly supported $g \geq 0$ and an increasing sequence p such that $p \to p' < +\infty$. On $g \leq 1$, g^p decreases to $g^{p'}$ and on $g > 1$ it increases, so that, since g is compactly supported $\int g^p dx \to \int g^{p'} dx$ and the latter is bounded above by any bound of the family $\int g^p dx$.

Next one can build a non-negative smooth function ψ on $\mathbb{R}^+$ such that $\psi(u) = 1$ for $u \le 1$, $\psi(u) = 0$ if $u \ge 2$ and $|\nabla\psi|^2/\psi$ is bounded by some constant A where by convention we define $|\nabla\psi|^2/\psi = 0$ on the set $\psi = 0$. Define $\psi_M(x) = \psi(|x|/M)$.
Defining $\rho_M(x) = z_M\, \rho(x)\, \psi_M(x)$ where z_M is a normalizing constant, we have as $M \to +\infty$, $z_M \to 1$ since $\rho \in \mathbb{L}^1$, ρ_M increases to ρ so that $\lim_M \int \rho_M^{p'} dx = \int \rho^{p'} dx$ and finally

$$I(\rho_M) = \int \frac{|\nabla\rho|^2}{\rho}\, \psi_M\, dx + \frac{1}{M^2} \int \frac{|\nabla\psi|^2}{\psi}(|x|/M)\, \rho(x)\, dx\,,$$

so that $I(\rho_M)$ goes to $I(\rho)$, and the proof is complete. □

An immediate consequence is the following extending [62] Lemma 3.3:

Lemma 3.2. *Let (Y_1, Y_2) be a random variable taking values in $\mathbb{R}^d \times \mathbb{R}^d$ with distribution $\rho(y_1, y_2)\, dy_1\, dy_2$. Then for any $\gamma \in (0, d)$ and any $\beta \in (\gamma/d, 2/d)$ ($\beta = 2/d$ is allowed if $d \ge 3$), there exists $C(\gamma, \beta, d)$ such that*

$$\mathbb{E}(|Y_1 - Y_2|^{-\gamma}) = \int_{\mathbb{R}^d\times\mathbb{R}^d} \frac{\rho(y_1, y_2)}{|y_1 - y_2|^\gamma}\, dy_1\, dy_2 \;\le\; C(\gamma, \beta, d)\, (1 + I^{\frac{d\beta}{2}}(\rho))\,.$$

Proof. The proof is exactly the same as in [62] with a modification due to dimension d instead of 2. One performs the change of variable $(y_1, y_2) \mapsto (y_1 - y_2, y_1 + y_2)$ so that the integral is with respect to the marginal distribution of $Y_1 - Y_2$ denoted by $\tilde\rho$. Hence

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} \frac{\rho(y_1, y_2)}{|y_1 - y_2|^\gamma}\, dy_1\, dy_2 = \int_{\mathbb{R}^d} \frac{\tilde\rho(z)}{|z|^\gamma}\, dz \le 1 + \int_{|z|\le 1} \frac{\tilde\rho(z)}{|z|^\gamma}\, dz\,.$$

One then applies Hölder inequality yielding, provided $1 > \beta > \gamma/d$,

$$\int_{|z|\le 1} \frac{\tilde\rho(z)}{|z|^\gamma}\, dz \le \left(\int_{|z|\le 1} |z|^{-\gamma/\beta} dz\right)^\beta ||\tilde\rho||_{1/(1-\beta)} \le C(\gamma, \beta, d)\, ||\tilde\rho||_{1/(1-\beta)}\,.$$

We use next (3.2), provided $1 > \beta > 2/d$, yielding $||\tilde\rho||_{1/(1-\beta)} \le C(\beta, d)\, I^{\frac{d\beta}{2}}(\tilde\rho)$. It remains to recall that the Fisher information of a marginal is less than the Fisher information of the initial probability measure (super-additivity, see remark 3.5 below), so that $I(\tilde\rho) \le I(\rho')$ where ρ' is the joint density of $(y_1 - y_2, y_1 + y_2)$. It is immediately seen that $I(\rho') \le c(d)\, I(\rho^{1,2})$ where $\rho^{1,2}$ is the joint density of (y_1, y_2) so that we get the result using again the super-additivity of the Fisher information. □

Remark 3.3. The previous proof furnishes an intermediate result that will be useful in the sequel, namely for $\theta < 1$ and β as before

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} \frac{\rho(y_1, y_2)}{|y_1 - y_2|^\gamma}\, \mathbf{1}_{|y_1 - y_2|\le\theta}\, dy_1\, dy_2 \;\le\; C(\gamma, \beta, d)\, \theta^{d-(\gamma/\beta)}\, I^{\frac{d\beta}{2}}(\rho)\,.$$

♢

Remark 3.4. Consider

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} h(y_1 - y_2)\, \rho(y_1, y_2)\, dy_1\, dy_2\,,$$

and assume that $h \in \mathbb{L}^\alpha(\mathbb{R}^d)$ for $\alpha \geq d/2$ if $d \geq 3$, $\alpha > 1$ if $d = 2$. Then the same proof shows that

$$\int_{\mathbb{R}^d \times \mathbb{R}^d} |h|(y_1 - y_2)\, \rho(y_1, y_2)\, dy_1\, dy_2 \leq C(\alpha, d)\, ||h||_\alpha\, I^{\frac{d}{2\alpha}}(\rho)\,.$$

♢

Remark 3.5. The super-additivity property of the Fisher information was first proved in [19] Theorem 3. It says that for any density of probability ρ on $\mathbb{R}^{m+n}$ whose marginals on $\mathbb{R}^m$ and $\mathbb{R}^n$ respectively are denoted by ρ_1 and ρ_2, it holds $I(\rho) \geq I(\rho_1) + I(\rho_2)$ with equality if and only if $\rho = \rho_1 \otimes \rho_2$ (i.e. marginals are independent). It is amusing to remark that another proof, in a more general framework, similarly using entropy on the path space, is contained in [27] Proposition 3.5. ♢

Another consequence strongly used in the P.D.E literature is the following result, similar to [68] proof of Proposition 3.1. Notice that, up to the constants, this result is exactly the same as what we directly obtained in lemma 2.5 (2.11). We give a proof for the sake of completeness and also to see what kind of analytic arguments are necessary for the derivation of this inequality.

Lemma 3.6. *Let $b(.,.)$ be some bounded smooth function. Let $t \mapsto \nu_t(dx) = \rho_t(x)dx$ be the flow of probability laws of the solution of*

$$dX_t = \sqrt{2}\, dB_t + b(X_t)dt \quad , \quad Law(X_0) = \rho_0(x)dx\,.$$

Assume in addition that for some $\alpha > 0$,

$$\int \ln \rho_0(x)\, \rho_0(x)\, dx \, < \, +\infty \qquad \text{and} \qquad \int \bar{V}_\alpha(x)\rho_0(dx) < +\infty$$

where $\bar{V}_\alpha(x) = \sum_{i=1}^d |x_i|^\alpha$. Then

$$\int \rho_T \, \ln \rho_T \, dx + \frac{1}{2} \left(\int_0^T I(\rho_t) dt \right) \leq \int \rho_0 \, \ln \rho_0 \, dx + 2 \left(\int_0^T \int b^2 \, \rho_t \, dx \, dt \right) . \tag{3.4}$$

Proof. Since b is smooth and bounded, using elliptic regularity (or Malliavin calculus) so is ρ_t for $t > 0$. ρ_t then satisfies the forward equation

$$\partial_t \rho_t = -\, div_x(2b\, \rho_t) + \Delta_x \rho_t\,.$$

Let $h(\rho) = \int \rho \ln(\rho)\, dx$. The first point is to show that $-\infty < h(\rho_t) < +\infty$ for all $t > 0$. Since b is bounded, the law Q of the solution of

$$dX_t = \sqrt{2}\, dB_t + b(X_t)dt \quad , \quad Law(X_0) = \rho_0(x)dx$$

is absolutely continuous (up to any finite time) w.r.t. the Wiener measure $\tilde{P}$ with the same initial distribution (recall that $\tilde{P}$ is the distribution of $\sqrt{2}B_.$). In addition dQ/dP belongs to all the $\mathbb{L}^p$ for $1 \leq p < +\infty$. With our assumption, for all t, $\int \bar{V}_\alpha(\omega_t)\, d\tilde{P} < +\infty$, so that the same is true for Q for any $\alpha' < \alpha$. According to (2.3), $h(\rho_t) > -\infty$, while $h(\rho_t) < +\infty$ since $H(Q|\tilde{P}) < +\infty$.

The goal is now to study the time evolution of $h(\rho_t)$ i.e. to compute

$$\begin{aligned}\partial_t h(\rho_t) &= \int \partial_t \rho_t \,(1+\ln \rho_t)\, dx \\ &= \int (-div_x(2b\rho_t) + \Delta_x \rho_t)\, \ln \rho_t \, dx\,.\end{aligned}$$

In order to justify differentiation under the integral one can for instance recall that the heat kernel associated to a smooth parabolic equation as here, is smooth and that all its derivatives have gaussian tails, i.e. are bounded by $C(m)t^{-d/2}\, e^{-\frac{|x-y|^2}{c(m)t}}$ where m denotes the order of the derivatives. It follows that for $t>0$,

$$\sup_{s\in[t/2,3t/2]} (-div_x(2b\rho_s) + \Delta_x \rho_s) \in \mathbb{L}^1(dx)$$

justifying the differentiation.
Next, a simple integration by parts furnishes

$$\partial_t h(\rho_t) = \int 2b\, \nabla_x \rho_t \, dx \; - \; I(\rho_t)\,.$$

Using the finite energy condition we deduce

$$\partial_t h(\rho_t) \le 2 \left(\int b^2\, \rho_t(x)\, dx \right)^{\frac{1}{2}} \; I^{\frac{1}{2}}(\rho_t) \; - \; I(\rho_t)$$

so that integrating w.r.t. dt and using Cauchy-Schwarz inequality we obtain

$$h(\rho_T) - h(\rho_0) \le 2 \left(\int_0^T \int b^2\, \rho_t \, dx\, dt \right)^{\frac{1}{2}} \left(\int_0^T I(\rho_t) dt \right)^{\frac{1}{2}} \; - \; \left(\int_0^T I(\rho_t) dt \right)\,,$$

from which we deduce the result. □

The natural question is to extend the previous lemma to any admissible flow. In the following remark we explain how to do and compare with the results obtained in lemma 2.5.

Remark 3.7. Consider an admissible flow $t \mapsto \nu_t$ as in theorem 2.9. If b is not smooth or bounded we may find a sequence of smooth and bounded b_n such that

$$\int_0^T \int |b-b_n|^2\, \rho_t(x)\, dx\, dt \; \to \; 0 \quad \text{ as } n \to +\infty\,.$$

(first approximate b in $\mathbb{L}^2([0,T]\times\mathbb{R}^d, \rho_t\, dxdt)$ by a continuous and compactly supported $\tilde{b}_n$, and then using a nice mollifier η_n approximate uniformly $\tilde{b}_n$ by $\tilde{b}_n * \eta_n$).
Denoting by Q_n (resp. Q that exists thanks to theorem 2.9) the associated probability measures on the path space with the same initial ν_0, it holds

$$H(Q|Q_n) = \int_0^T \int |b-b_n|^2\, \rho_t(x)\, dx\, dt$$

that goes to 0. Since $H(Q|Q_n)$ goes to 0 so do all time marginals ρ_t^n for $t>0$.
We first claim that $\liminf_n I(\rho_t^n) \ge I(\rho_t)$. Many proofs appear in the literature as for instance the general results in [106]. We give here a direct (idea of) proof. Assume that $\liminf_n I(\rho_t^n) < +\infty$ and choose a subsequence $\bar{\rho}_t^n$ such that $\lim_n I(\bar{\rho}_t^n) = \liminf_n I(\rho_t^n)$. According to lemma 3.1 we know that $\bar{\rho}_t^n$ is bounded in some $\mathbb{L}^p$. Since $\sqrt{\bar{\rho}_t^n}$ is also bounded

in $\mathbb{L}^2$, it is not too difficult to get a subsequence (still denoted by $\bar\rho_t^n$) converging to some ρ_t weakly in $\mathbb{L}^p$ and such that $\sqrt{\bar\rho_t^n}$ converges to $\sqrt{\rho_t}$ weakly in $\mathbb{L}^2$. It follows that $\nabla\sqrt{\bar\rho_t^n}$ converges to $\nabla\sqrt{\rho_t}$ in the set of (Schwartz) distributions, and since this sequence is bounded in $\mathbb{L}^2$, the limit also belongs to $\mathbb{L}^2$. It remains to use Fatou's lemma to control the time integral. For a complete rigorous proof in this spirit see Proposition 13.2 in [11].
Choose some V such that $|V'|$ and $|V''|$ are bounded by K. Taking a subsequence if necessary, we also have that $\int \bar V\rho_T dx \le \liminf_n \int \bar V\rho_T^n dx$. According to lemma 2.5 we know that

$$\left|\int \rho_T \ln\rho_T dx\right| \le dC(K,T)\left(1+\ln Z_V+\int_0^T\int |b-\nabla\bar V|^2 dxdt\right)+\int(\ln\rho_0+\bar V)\rho_0 dx\,.$$

It follows that, we may control both $\int_0^T I(\rho_t)dt$ and $h(\rho_T)$ by the $\liminf_n \int_0^T\int b_n^2\,\rho_t^n dxdt$. These results are of the same nature as the ones we directly obtained in lemma 2.5.

One can also use the euclidean log-Sobolev inequality

$$\int \rho_T\ln\rho_T dx \le \frac{d}{2}\,\ln\left(\frac{4I(\rho_T)}{d\pi e}\right) \tag{3.5}$$

which furnishes, for large Fisher information, a better bound. ♢

3.2. **Existence and uniqueness of entropic singular diffusions.**

In this subsection we shall use all what precedes to study existence and uniqueness for Ito processes with singular drift. Comparison with the existing literature will be made at the end of the subsection.
We thus want to study the S.D.E. in $\mathbb{R}^m$

$$X_t = X_0+\sqrt{2}\,B_t+\int_0^t 2\,g(X_s)ds\,. \tag{3.6}$$

The law of X_0 is denoted by $\nu_0(dx)=\rho_0(x)dx$. We may first use a cut-off $g^M = g\,\mathbf{1}_{|g|\le M}$ which is bounded. One can thus, using Girsanov theory, build a (weak) solution Q^M which is absolutely continuous w.r.t. W the Wiener measure, and more precisely the law of $X_0+\sqrt{2}B_t$, on the time interval $[0,T]$ for any $T>0$. We fix once for all such a T and keep the same notation Q^M and W. As in subsection 2.1, we choose a smooth V (defined on $\mathbb{R}$) such that $|V'|$ and $|V''|$ are bounded by A, define $\bar V(x)=\sum_{j=1}^m V(x^j)$ and define P as the distribution of

$$Y_t = Y_0+\sqrt{2}\,B_t-\int_0^t \nabla\bar V(Y_s)\,ds$$

where Y_0 is distributed according to the probability measure $\gamma_0(dy)=Z^{-1}\,e^{-\bar V(y)}\,dy$.
Assume that

$$H(\nu_0|\gamma_0)=\int(\ln\rho_0+\bar V+\ln Z)\rho_0 dx<+\infty\,.$$

We thus have

$$H(Q^M|P)=H(\nu_0|\gamma_0)+\int\int_0^T |g^M+\nabla\bar V|^2(\omega_t)\,dt\,dQ^M$$

(recall that ω is the generic element of the path space). We thus know that the law of X_t^M admits a density ρ_t^M (with respect to Lebesgue measure) so that,

$$\int\int_0^T |g^M + \nabla\bar{V}|^2(\omega_t)\,dt\,dQ^M = \int\int_0^T |g^M + \nabla\bar{V}|^2(x)\,\rho_t^M(x)\,dt\,dx$$

satisfies for all $\lambda > 0$

$$\int\int_0^T |g^M + \nabla\bar{V}|^2(\omega_t)\,dt\,dQ^M \le (1+\lambda)\int_0^T\int |g^M|^2\,\rho_t^M dxdt + \left(1+\frac{1}{\lambda}\right) m\,A^2\,T\,. \quad (3.7)$$

According to Corollary 2.3 we also know that $\int_0^T I(\rho_t^M)\,dt < +\infty$ so that we may use (3.2) for almost all $t \in]0,T]$.

In particular for $m \ge 3$ one may use lemma 3.1 and Hölder's inequality with $p = m/(m-2)$ whose dual exponent is $q = m/2$ yielding (see the value of a in Lemma 3.1)

$$\begin{aligned}\int |g^M|^2\,\rho_t^M dx &= \int |g^M\,\mathbf{1}_{|g^M|>A}|^2\,\rho_t^M dx + \int |g^M\,\mathbf{1}_{|g^M|\le A}|^2\,\rho_t^M dx\\ &\le ||\rho_t^M||_p\,|||g^M\,\mathbf{1}_{|g^M|>A}|^2||_q + A^2\\ &\le a^2\,I(\rho_t^M)\,||g^M\,\mathbf{1}_{|g^M|>A}||_m^2 + A^2\\ &\le \frac{(m-1)^2}{m\,(m-2)^2}\,I(\rho_t^M)\,||g\,\mathbf{1}_{|g|>A}||_m^2 + A^2\,. \end{aligned} \quad (3.8)$$

If we plug this inequality into (2.11) we thus obtain

$$\int_0^T I(\rho_t^M)\,dt \le 4(1+\lambda)^2\,\frac{(m-1)^2}{m\,(m-2)^2}\,\left(\int_0^T I(\rho_t^M)\,dt\right)\,||g\,\mathbf{1}_{|g|>A}||_m^2 + C(\bar{V},\rho_0,T,\lambda,m,||g\mathbf{1}_{|g|>A}||_m,A)\,. \quad (3.9)$$

Of course (3.9) is useful only if

$$4\,||g\,\mathbf{1}_{|g|>A}||_m^2\,\frac{(m-1)^2}{m\,(m-2)^2} < 1$$

in which case we get that $\int I(\rho_t^M)dt \le C$ where C does not depend on M, hence $\sup_M \int_0^T I(\rho_t^M)\,dt < +\infty$. If $g\,\mathbf{1}_{|g|>C} \in \mathbb{L}^m$ for some $C > 0$ we may always find $A > C$ such that the previous is satisfied.

One can also be less demanding and apply Hölder's inequality for some $q > m/2$, i.e. $p < m/(m-2)$ so that $m(p-1)/2p < 1$ i.e. $m(p-1)/2p = 1 - \varepsilon_p$ for some $\varepsilon_p > 0$. We thus obtain, this time for all $m \ge 2$,

$$\int I(\rho_t^M)dt \le C(m,p)\,||g\,\mathbf{1}_{|g|>A}||_{2q}^2\,\int (I(\rho_t^M))^{1-\varepsilon_p}dt + C(\bar{V},\rho_0,T,q,m,||g\,\mathbf{1}_{|g|>A}||_{2q},A)\,, \quad (3.10)$$

and again if $||g\,\mathbf{1}_{|g|>A}||_{2q} < +\infty$ that $\sup_M \int_0^T I(\rho_t^M)\,dt < +\infty$.

In both cases, plugging this bound first in (3.8) and then the new obtained inequality in (3.7), we deduce that $\sup_M H(Q^M|P) < +\infty$. This will yield the following result

Theorem 3.8. *Let $m \ge 2$. Assume that $H(\nu_0|\gamma_0) < +\infty$ and in addition that $g\,\mathbf{1}_{|g|>A} \in \mathbb{L}^p(\mathbb{R}^m)$ for some $p \ge m$ if $m \ge 3$ or $p > 2$ if $m = 2$.*

Then equation (3.6) has an unique weak solution (solution in law) Q, and this Q satisfies $H(Q|P) < +\infty$ on $[0,T]$ for all $T > 0$.

Proof. According to the discussion before the Theorem's statement, the sequence Q^M is such that $\sup_M H(Q^M|P) < +\infty$. It follows that one can find a subsequence which is weakly converging to some Q with $H(Q|P) \le \liminf_M H(Q^M|P) < +\infty$, since relative entropy is lower semi-continuous. Thanks to Dunford-Pettis theorem the convergence of dQ^M/dP to dQ/dP holds for the stronger $\sigma(\mathbb{L}^1, \mathbb{L}^\infty)$ topology.

Also remark that we can replace g^M by g in the left hand side of (3.8), so that it holds $\sup_M \int_0^T \int |g|^2 \rho_t^M dxdt < +\infty$, and $\lim_{M\to+\infty} \int_0^T \int |g-g^M|^2 \rho_t^M dxdt = 0$ since $|||g-g^M|||_{2q} \to 0$.

To prove that Q solves (3.6) it remains to show that it solves the corresponding martingale problem i.e. to show that for all $t > s$, all bounded H defined on $C^0([0,s])$, all smooth φ with compact support, it holds

$$\mathbb{E}^Q \left[H(\omega_{v\le s}) \left(\varphi(\omega_t) - \varphi(\omega_s) + \int_s^t (\Delta\varphi(\omega_u) + \langle \nabla\varphi(\omega_u), 2g(\omega_u)\rangle) du \right) \right] = 0\,. \tag{3.11}$$

Since the previous is true replacing Q and g by Q^M and g^M, it is enough to prove

$$\lim_M \mathbb{E}^{Q^M} \left[\left(\int_s^t \langle \nabla\varphi(\omega_u), g^M(\omega_u)\rangle du \right) H(\omega_{v\le s}) \right] = \mathbb{E}^Q \left[\left(\int_s^t \langle \nabla\varphi(\omega_u), g(\omega_u)\rangle du \right) H(\omega_{v\le s}) \right]. \tag{3.12}$$

Since $\lim_{M\to+\infty} \int_s^t \int |g-g^M|^2 \rho_u^M dxdu = 0$, thanks to (3.8),

$$\lim_M \mathbb{E}^{Q^M} \left[\left(\int_s^t \langle \nabla\varphi(\omega_u), (g^M - g)(\omega_u)\rangle du \right) H(\omega_{v\le s}) \right] = 0\,,$$

while

$$\lim_M \mathbb{E}^{Q^M} \left[\left(\int_s^t \langle \nabla\varphi(\omega_u), g(\omega_u)\rangle du \right) H(\omega_{v\le s}) \right] = \mathbb{E}^Q \left[\left(\int_s^t \langle \nabla\varphi(\omega_u), g(\omega_u)\rangle du \right) H(\omega_{v\le s}) \right]$$

follows from the $\sigma(\mathbb{L}^1, \mathbb{L}^\infty)$ convergence of Q^M to Q. Indeed we can first choose A large enough for $\sup_M \int_s^t \int |g-g^A|^2 \rho_u^M dxdu \le \varepsilon$ and $\int_s^t \int |g-g^A|^2 \rho_u dxdu \le \varepsilon$ according to (3.8), and then replacing g by g^A get the desired limit in M so that it only remains to let ε go to 0.

Finally we have to prove uniqueness. Define $T_k = \inf\{t \ge 0, \int_0^t |g + \nabla\bar{V}|^2(\omega_s)\, ds \ge k\}$. Since $H(Q|P) < +\infty$ (on $[0,T]$), Girsanov theory tells us that $T_k \wedge T \to T$, Q almost surely. If Q' is another solution of (3.6), Girsanov theory again tells us that it coincides with Q on $[0, T\wedge T_k]$, hence $T_k \wedge T \to T$, Q' almost surely too and $Q = Q'$ on $[0,T]$. □

Remark 3.9. *About the literature.*

Theorem 3.8 is mainly known, except in the critical case $p = m$. For weak solutions it has been first obtained in [2] for homogeneous drift, the method being then extended to time dependent drifts in [79]. A different approach was proposed by N. Krylov and co-authors. For a complete bibliography on the topic see [82]. The proof we have given here can be extended to several other situations: particle system, non linear SDE, more general processes than Ito processes, thanks to the versatility of the entropic approach. This will be done in the next subsection for particle systems, and in another work for the other cases.

We shall recall below, for later use, the arguments allowing to get a weak solution, following Krylov's ideas. The case $g \in \mathbb{L}^p(\mathbb{R}^m)$ for some $p > m$ is contained for instance in [85]

including time inhomogeneous drift, while the case $p = m$ is studied and partly solved in e.g. [86, 88, 87, 3]. In all these papers existence of a strong solution is the main goal.

Actually a standard result by Khasminskii recalled in [72] Proposition 3.1 (also see [85]) says that for a $\mathbb{R}^m$ valued standard Brownian motion $w_.$, any $g \in \mathbb{L}^q([0,T], \mathbb{L}^p(dx))$ with $\frac{m}{p} + \frac{2}{q} < 1$, any $x \in \mathbb{R}^m$ and any $\kappa > 0$ it holds

$$\mathbb{E}\left(\exp\left(\kappa \int_0^T g^2(s, x + w_s) ds\right)\right) \leq C(\kappa, T, p, q, ||g||_{\mathbb{L}^q([0,T],\mathbb{L}^p(dx))}) < +\infty\,. \tag{3.13}$$

Notice that this upper bound does not depend on x since the $\mathbb{L}^p$ norm of $g(s, x + .)$ is equal to the one of $g(s, .)$ thanks to the invariance by translation of Lebesque's measure. It immediately follows from Novikov's criterion, that the Girsanov density associated to g is a true martingale, the existence of a weak solution Q_x of (3.6) starting from x follows. In addition the Girsanov density belongs to all the $\mathbb{L}^k(W_x)$ where W is Wiener measure starting from x. Indeed

$$\begin{aligned}
\mathbb{E}_x\left(e^{k(\int_0^T g\sqrt{2}\, d\omega_s - \int_0^T |g|^2\, ds)}\right) &= \mathbb{E}_x\left(e^{(k\int_0^T g\sqrt{2}\, d\omega_s - 2k^2 \int_0^T |g|^2\, ds)}\, e^{(2k^2-k)\int_0^T |g|^2\, ds)}\right) \\
&\leq \mathbb{E}_x^{\frac{1}{2}}\left(e^{\int_0^T 2k\, g\sqrt{2}\, d\omega_s - \int_0^T 4k^2 |g|^2\, ds}\right) \mathbb{E}_x^{\frac{1}{2}}\left(e^{2(2k^2-k)\int_0^T |g|^2\, ds}\right)
\end{aligned}$$

thanks to Cauchy-Schwarz inequality. The first factor is equal to 1 according to Novikov's criterion, hence for $k \geq 1$,

$$\mathbb{E}_x\left(e^{k(\int_0^T g\sqrt{2}\, d\omega_s - \int_0^T |g|^2\, ds)}\right) \leq \mathbb{E}_x^{\frac{1}{2}}\left(e^{2(2k^2-k)\int_0^T |g|^2\, ds}\right)\,. \tag{3.14}$$

In what precedes $\mathbb{E}_x$ is the expectation w.r.t. the Wiener measure starting from x.

It is worth to notice that the previous quantity can be bounded independently of x by $C(T, k, p, q, ||g||_{\mathbb{L}^q([0,T],\mathbb{L}^p(dx))})$.

It follows the same exponential integrability with respect to Q_x since

$$\mathbb{E}_{Q_x}((dQ_x/dW_x)^k) = \mathbb{E}_{W_x}((dQ_x/dW_x)^{k+1})\,.$$

It immediately follows that (3.13) is still true if one replaces $w_.$ by $X_.$, so that the same argument shows that the inverse of the Girsanov density also belongs to all the $\mathbb{L}^k$ w.r.t W or Q (this is (ii) in Proposition 3.3 of [72]).

Assume for simplicity that $g(s, x) = g(x)$ and that $p > m$. Girsanov theory tells us that Q_x solves (3.6) and that the family $(Q_x)_{x\in\mathbb{R}^m}$ is strong Markov. As a consequence, for any initial distribution ν_0, $Q_{\nu_0} = \int Q_x\, \nu_0(dx)$ solves (3.6) with initial distribution ν_0. If ν_0 satisfies the hypotheses in Theorem 3.8, Q_{ν_0} coincides with the solution built in this Theorem.

Notice that it is absolutely continuous w.r.t. W_{ν_0}, and that its density satisfies the same $\mathbb{L}^k$ controls than Q_x. Also notice that since we can always add or subtract a bounded drift via another Girsanov transform (it does not change the k integrability), the required $\mathbb{L}^p$ integrability can be restricted to $g\,\mathbf{1}_{|g|>A}$ as for our result.

Denote by $\rho_t(x,.)$ the density w.r.t. Lebesgue's measure at time t of Q_x. Let h be a compactly supported and bounded function. Then for $1/u+1/v=1$,

$$\begin{aligned}\int h(y)\, e^{\frac{|x-y|^2}{4ut}}\, \rho_t(x,y)\, dy &= \int h(\omega_t)\, e^{\frac{|x-\omega_t|^2}{4ut}}\, \frac{dQ_x}{dW_x}(\omega)\, W_x(d\omega)\\ &\le \left(\int \left(\frac{dQ_x}{dW_x}\right)^v W_x(d\omega)\right)^{1/v} \left(\int h^u(\omega_t)\, e^{\frac{|x-\omega_t|^2}{4t}}\, W_x(d\omega)\right)^{1/u}\\ &\le C(T,u,p,||g||_p)\, \left(\int h^u(y)\,(4\pi t)^{-m/2}\, dy\right)^{1/u}\end{aligned}$$

so that if h is non-negative

$$\int h(y)\rho_t(x,y)\, dy \le C(T,u,p,||g||_p)\, t^{-m/2u}\, ||h||_u\,.$$

Taking the positive and the negative part of h the previous extends to any compactly supported h, showing that for any $1\le v<+\infty$ and any x

$$||\rho_t(x,.)||_v \le C(T,v,p,||g||_p)\, t^{-\frac{(v-1)m}{2v}}. \tag{3.15}$$

It immediately follows that the density $\rho_t=\int \rho_t(x,.)\,\nu_0(dx)$ satisfies the same property, thanks to Hölder inequality. Notice that, according to (3.15),

$$\rho_. \in \mathbb{L}^1((0,T],\mathbb{L}^v(\mathbb{R}^m)) \quad \text{provided } v<m/(m-2)\,. \tag{3.16}$$

Our result (despite the fact we are not looking for a strong solution) seems to be new for $p=m$, for which we (presumably) do no more have a Girsanov density in some $\mathbb{L}^k$ for some $k>1$ but only in the Orlicz space $\mathbb{L}^{x\ln x}$.

It is worth noticing that our proof immediately extends to the case of time inhomogeneous drift $g(t,.)$ provided $\sup_{t\in[0,T]}||g(t,.)\,\mathbf{1}_{|g|>A}||_p<+\infty$ for some (time independent) A. Actually using some Hölder inequality w.r.t. time, it extends to the standard $\mathbb{L}^p$-$\mathbb{L}^q$ case. This will be explained for the non linear SDE in the next section. ♢

Finally in the case $d=2$ the result will be improved in the final section.

3.3. **Application to particle systems.**

We come back to (1.2) where for simplicity we will first assume that $b=0$. We will not use Theorem 3.8 since $m=dN$ is too big. Instead we may look at the proof. It is easily seen that all we have to do is to bound

$$\int_0^T \int K^2(\omega_t^i-\omega_t^j)\, dQ^M\, dt$$

so that we may apply (3.8) with $m=2d$, replacing ρ_t^M by $\rho_{t,i,j}^M$ the Q^M law of (ω_t^i,ω_t^j) and $||g\,\mathbf{1}_{|g|>A}||_m$ by $||K\,\mathbf{1}_{|K|>A}||_{2\alpha}$ provided $K\,\mathbf{1}_{|K|>A}\in\mathbb{L}^{2\alpha}(\mathbb{R}^d)$ with $\alpha\ge d/2$ if $d\ge 3$, $\alpha>1$ if $d=2$ according to Remark 3.4.

Using again the super-additivity of Fisher information and the convexity inequality

$$\sum_i \left| \frac{1}{N} \sum_{j \neq i} K(x^i - x^j) \right|^2 \leq \frac{1}{N} \sum_i \sum_{j \neq i} K^2(x^i - x^j)$$

we deduce the analogue of (3.10)

$$I(\rho_t^M) \leq C(d,\alpha)\,(N-1)\,||K\,\mathbf{1}_{|K|>A}||_{2\alpha}^2\,(I(\rho_t^M))^{1-\varepsilon_\alpha} + N\,C(\bar{V},\rho_0,T,\alpha,d,||K\mathbf{1}_{|K|>A}||_{2\alpha},A)\,, \tag{3.17}$$

for α as before and some $\varepsilon_\alpha > 0$ except if $\alpha = d/2$. In the latter case we may again choose a larger A as for the proof of Theorem 3.8. The remaining part of the proof of Theorem 3.8 is unchanged so that we may state

Theorem 3.10. *Assume that $K\mathbf{1}_{|K|>A} \in \mathbb{L}^p(\mathbb{R}^d)$ for some $p \geq d$ for $d \geq 3$ and $p > 2$ for $d = 2$, and some $A > 0$. Assume in addition that $b = b_1 + b_2$ where b_1 is bounded and $b_2 = \nabla U$.*
Assume that there exists a smooth V defined on $\mathbb{R}$ such that $|V'|$ and $|V''|$ are bounded and $\int e^{-(\bar{V}+U)}dx < +\infty$ where as before $\bar{V}(x) = \sum_i V(x^i)$. Introduce the probability distribution $\gamma_0(dx) = Z^{-1}\,e^{-(\bar{V}+U)(x)}dx$ and define P as the symmetric diffusion process with reversible measure γ_0 which is assumed to exist.
Finally assume that $H(\nu_0|\gamma_0) < +\infty$. Then there exists an unique weak solution Q to (1.2) with initial law ν_0. In addition $H(Q|P) < +\infty$.

For the proof it is enough to remark that the drift of Q w.r.t. P is simply given by the interaction kernel K, $\nabla\bar{V}$ and b_1 since we have included ∇U in the the drift of P, and to recall Remark 2.1.

Remark 3.11. *Discussion and examples.*
Theorem 3.10 is also partially already known (except the absolute continuity of Q) and is contained (for $p > d$) in [73] Theorem 1.1 for instance (also see [71] Theorem 1.4). The proof in [73] uses deep analytical results on singular PDE's (as for Krylov's approach), while ours is more probabilistic and direct. Notice that if we assume in addition that K is local Lipschitz and bounded in C_ε^c for all $\varepsilon > 0$ (recall Assumption 2.15), we get a strong (pathwise) solution, as for Theorem 2.18.
However, in our benchmark of examples the only one to which we may apply Theorem 3.10 is the relaxed Keller-Segel model. In dimension $d \geq 3$ we should also consider $K_s(x) = \chi\,x/|x|^{s+1}$ for $s < 1$ (which is example 1 (i) in [73]). This is a little bit disappointing in comparison with Theorem 2.18 (except that the previous K_s does not satisfy the assumptions in Theorem 2.18), but proves that approximations by smooth models is not efficient for very singular models.
Of course one explanation is the following: the sign of χ has no importance in what we have done in this section, while we know that repulsive or attractive models can have very different behaviours, as the Coulombic type potentials clearly show. ♢

3.4. **Existence via the absence of collisions.**

Consider the particle system in $(\mathbb{R}^d)^N$,

$$dX_t^{i,N} = \sqrt{2}\, dB_t^{i,N} \; - \; b(X_t^{i,N})\, dt \; - \; \frac{\chi}{N} \sum_j \frac{X_t^{i,N} - X_t^{j,N}}{|X_t^{i,N} - X_t^{j,N}|^m}\, dt \tag{3.18}$$

for $m > 1$ and b is a-Lipschitz (recall that for $i = j$ the ratio is equal to 0 by convention). A non explosive pathwise solution exists up to the hitting time T_0 of the collision set C_0 defined in Theorem 2.14. It is tempting to try to directly show that T_0 is almost surely infinite.
If $m > 2$ introduce $H(x) = \sum_{i \neq j} \frac{1}{|x^i - x^j|^{m-2}}$. Applying Ito formula we have that for $T < T_0$,

$$\begin{aligned} H(X_T^N) - H(X_0^N) \;\; = \;\; & \frac{2\chi(m-2)}{N} \int_0^T \sum_i \left| \sum_j \frac{X_t^{i,N} - X_t^{j,N}}{|X_t^{i,N} - X_t^{j,N}|^m} \right|^2 dt \\ & -2(m-2)(d-m) \int_0^T \sum_{i \neq j} \frac{dt}{|X_t^{i,N} - X_t^{j,N}|^m} \\ & + (m-2) \int_0^T \sum_{i \neq j} \frac{\langle b(X_t^{i,N}) - b(X_t^{j,N}), X_t^{i,N} - X_t^{j,N} \rangle}{|X_t^{i,N} - X_t^{j,N}|^m}\, dt \\ & + M_T \end{aligned} \tag{3.19}$$

where $M_.$ is a local martingale, and a true martingale up to τ_ε the first hitting time of C_ε defined in Theorem 2.14.
If $\chi < 0$ (repulsive case) the first term in the right hand side is non-positive. If $m \leq d$ the second one is also non-positive. Since b is a-Lipschitz, a simple use of Gronwall's lemma shows that

$$\mathbb{E}[H(X_{T \wedge \tau_\varepsilon}^N)] \leq \mathbb{E}[H(X_0^N)]\, e^{\,(m-2)a\,T}\,.$$

Hence if $\chi < 0$, we deduce that $\tau_0 = +\infty$ almost surely for $X_0^N = x \notin C_0$, hence $\tau_0 = +\infty$ almost surely if the law of X_0^N is absolutely continuous.

The previous applies thus in the sub-Coulombic situation for any $s > 0$ as we already saw in subsubsection 2.4.3, furnishing another proof of strong existence for the particle system. This is close to the proof chosen in [109] in a more general context (the g therein generalizes the sub-Coulombic potential). Notice that the second line in (4.19) of [109] is non-positive and not non-negative as it is said therein.

Of course the method is more general and relies on the existence of a Lyapunov function H which is smooth outside the collision set, equals $+\infty$ on C_0, goes to infinity at infinity and satisfies

$$\mathcal{L}^N\, H \leq D^N \; + \; C^N\, H$$

for some positive constants C^N and D^N, where $\mathcal{L}^N$ is the generator of the process X^N. The existence of such a Lyapunov function is known to ensure non-explosion (Khasminskii test). The construction of such a Lyapunov function for the system (3.18) when adding a self-interaction is made in [70] Lemma 2.1. For $d = 1$ it is also made with $m = 2$ (Dyson processes) provided $|\chi|$ is large enough.

4. The non linear S.D.E.

We turn to the study of (1.3) and its applications.

First consider a function $L : \mathbb{R}^d \mapsto \mathbb{R}$ which is assumed to be in C_b^∞, i.e. C^∞, bounded with bounded derivatives of any order, and b be bounded and global Lipschitz.
We may define the system of S.D.E.'s

$$dY_t^{i,N} = \sqrt{2}\, dB_t^{i,N} \,-\, b(Y_t^{N,i})\, dt \,-\, \frac{1}{N}\sum_{j=1}^N L(Y_t^{i,N} - Y_t^{j,N})\, dt\,,$$

starting from an initial configuration Y_0^N whose law is assumed to be exchangeable. If for all fixed k the distribution of $(Y_0^{1,N}, ..., Y_0^{k,N})$ converges to $(\rho_0(y)dy)^{\otimes k}$ as N goes to infinity, it is well known that, for all $T > 0$, the distribution of the process $(Y_t^{1,N}, ..., Y_t^{k,N})_{t\in[0,T]}$ converges to the one of an i.i.d. k-uple of processes $(Y_t^1, ..., Y_t^k)_{t\in[0,T]}$ each one being a solution of the "non linear" S.D.E.

$$\begin{aligned} dY_t &= \sqrt{2}\, dB_t \,-\, b(Y_t)\, dt \,-\, (L * \rho_t)(Y_t)\, dt \\ Law(Y_t) &= \rho_t(y)\, dy \quad \forall t \in [0,T]\,. \end{aligned} \tag{4.1}$$

In addition this solution is unique. For all this see e.g. [110, 99].
As long as we only consider weak solutions, we may assume that $b = 0$ since it can be introduced via a standard Girsanov transformation on the initial P.
Notice that $L * \rho_t$ is a bounded (time dependent) drift, so that the law Q_L of $Y_.$ restricted to the time interval $[0,T]$, satisfies $H(Q_L|P) < +\infty$ where P is as in subsection 2.1 for a V with bounded first and second derivatives. Recall that the drift β_t is given by $\beta_t = \frac{1}{2}(\nabla\bar{V} - (L * \rho_t))(Y_t)$. According to what we have done in the previous sections we thus know some bounds on the entropy and the Fisher information of ρ_t.

Assume that $L \in \mathbb{L}^\alpha(\mathbb{R}^d)$ for some $\alpha \geq 1$. We will first get a bound for $H(Q_L|P)$ that only depends on α and d. To this end we need to get controls for

$$\int |L * \rho_t|^2\, \rho_t\, dx\,.$$

Since $H(Q_L|P) < +\infty$ we know that $I(\rho_t) < +\infty$ for almost all t so that $\rho_t \in \mathbb{L}^p(\mathbb{R}^d)$ for $p \leq d/d-2$ if $d \geq 3$ and $p < +\infty$ for $d = 2$.
As in the previous section introduce $L_A = L\,\mathbf{1}_{|L|\geq A}$. $L_A * \rho_t \in \mathbb{L}^r(\mathbb{R}^d)$ for $1 + \frac{1}{r} = \frac{1}{p} + \frac{1}{\alpha}$ and $||L_A * \rho_t||_r \leq ||L_A||_\alpha\, ||\rho_t||_p$. In addition

$$\int |(L - L_A) * \rho_t|^2\, \rho_t dx \leq A^2\,.$$

We have for another $p' \leq d/d-2$ (but $p' < +\infty$),

$$\int |L_A * \rho_t|^2\, \rho_t\, dx \;\leq\; ||\rho_t||_{p'}\, |||L_A * \rho_t|||^2_{2p'/(p'-1)}$$

hence for $r = 2p'/(p'-1)$ i.e. $\frac{1}{\alpha} = \frac{3}{2} - \frac{1}{2p'} - \frac{1}{p}$,

$$\int |L_A * \rho_t|^2\, \rho_t\, dx \;\leq\; ||\rho_t||_{p'}\, ||\rho_t||_p^2\, |||L_A|||^2_\alpha\,. \tag{4.2}$$

We thus obtain the analogue of (3.8),

$$\int |L * \rho_t|^2 \, \rho_t \, dx \; \leq C(d) \, |||L_A|||_\alpha^2 \, (I(\rho_t))^{d(\frac{3}{2}-\frac{1}{2p'}-\frac{1}{p})} \; + \; A^2 \, . \tag{4.3}$$

In the context of an a priori bound for $I(\rho_t)$, the previous is interesting only if $d(\frac{3}{2}-\frac{1}{2p'}-\frac{1}{p}) \leq 1$, yielding $\alpha \geq d$ for $d \geq 3$ or $\alpha > 2$ if $d = 2$. Indeed it can be used for $L = K^M$, where K^M is a sequence of regular (C_b^∞) kernels converging to K in $\mathbb{L}^\alpha$ norm. We will thus obtain the following analogue of Theorem 3.8 and Theorem 3.10

Theorem 4.1. *Let V be a smooth function such that $e^{-V} \in \mathbb{L}^1(\mathbb{R})$ and such that V' and V'' are bounded. Define $\bar{V}(x) = \sum_{i=1}^d V(x^i)$. Let ρ_0 be a density of probability such that $\int \rho_0(\ln \rho_0 + \bar{V})dx < +\infty$.*
Assume that b is bounded and global Lipschitz and that $K\mathbf{1}_{|K|>A} \in \mathbb{L}^\alpha(\mathbb{R}^d)$ for some $A > 0$. Then, there exists a (weak) solution of (1.3) i.e.

$$\begin{aligned} dX_t &= \sqrt{2}\, dB_t \; - \; b(X_t)\, dt \; - \; (K * \rho_t)(X_t)\, dt \, , \\ \nu_t = \rho_t(x)\, dx &= \mathcal{L}(X_t) \, , \end{aligned}$$

provided $\alpha > 2$ if $d = 2$ or $\alpha \geq d$ for $d \geq 3$.
In addition this solution satisfies $H(Q|P) < +\infty$.

Proof. The proof is very similar to the one of Theorem 3.8 and here we may choose for simplicity $p = 1$, $p' = \alpha/(\alpha - 2)$ in (4.3) so that the exponent of $I(\rho_t)$ is d/α. First Q^M converges to Q in the $\sigma(\mathbb{L}^1, \mathbb{L}^\infty)$ topology with $H(Q|P) < +\infty$. In particular ρ_t^M weakly converges to ρ_t which is the marginal distribution of Q at time t, and $\int_0^T I(\rho_t)dt < +\infty$ thanks to (2.6). Hence (4.3) is still true when we replace L by K or by $K - K^M$.
Since b is Lipschitz and bounded (hence does not play any annoying role), it remains to prove that

$$\begin{aligned} \lim_M \mathbb{E}^{Q^M} \left[\left(\int_s^t \langle \nabla \varphi(\omega_u), (K^M * \rho_u^M)(\omega_u) \rangle du \right) \, H(\omega_{v \leq s}) \right] = \\ = \mathbb{E}^Q \left[\left(\int_s^t \langle \nabla \varphi(\omega_u), (K * \rho_u)(\omega_u) \rangle du \right) \, H(\omega_{v \leq s}) \right] . \end{aligned}$$

Using (4.3) with K^M and ρ_t^M in place of L and ρ_t, it is easily seen, since $\int_s^t I^{d/\alpha}(\rho_u^M)\, du \leq c(s,t)$ where $c(s,t)$ does not depend on M (recall that $d/\alpha \leq 1$), that

$$\lim_M \mathbb{E}^{Q^M} \left[\left(\int_s^t \langle \nabla \varphi(\omega_u), ((K^M - K) * \rho_u^M)(\omega_u) \rangle du \right) \, H(\omega_{v \leq s}) \right] = 0$$

so that

$$\begin{aligned} \lim_M \mathbb{E}^{Q^M} \left[\left(\int_s^t \langle \nabla \varphi(\omega_u), (K^M * \rho_u^M)(\omega_u) \rangle du \right) \, H(\omega_{v \leq s}) \right] = \\ = \lim_M \, \mathbb{E}^{Q^M} \left[\left(\int_s^t \langle \nabla \varphi(\omega_u), (K * \rho_u)(\omega_u) \rangle du \right) \, H(\omega_{v \leq s}) \right] . \end{aligned}$$

To show that the latter is what we want we may first choose some B large enough so that for all M,

$$\mathbb{E}^{Q^M} \left[\left(\int_s^t \langle \nabla \varphi(\omega_u), ((K - K^B) * \rho_u)(\omega_u) \rangle du \right) \, H(\omega_{v \leq s}) \right] \leq \varepsilon$$

and also

$$\mathbb{E}^Q\left[\left(\int_s^t \langle\nabla\varphi(\omega_u),((K-K^B)*\rho_u)(\omega_u)\rangle du\right) H(\omega_{v\leq s})\right]\leq\varepsilon$$

which is possible again using (4.3), and then use the convergence of Q^M to Q in the $\sigma(\mathbb{L}^1,\mathbb{L}^\infty)$ topology to control

$$\mathbb{E}^{Q^M}\left[\left(\int_s^t \langle\nabla\varphi(\omega_u),(K^B*\rho_u)(\omega_u)\rangle du\right) H(\omega_{v\leq s})\right].$$

It remains to let ε go to 0. □

Remark 4.2. Once again, Theorem 4.1 covers the relaxed Keller-Segel case for which a proof appears in [68]. The interested reader can compare both proofs. As a general result it also entails Theorem 1.2 in [72] for $q_1=+\infty$ in equation (1.15) therein, but with a much weaker integrability condition for the initial measure in our result. Actually, our approach can be generalized without too much efforts to time dependent interaction kernels as in [72] and then improve on the results therein. ♢

Remark 4.3. If L is time dependent, we may similarly introduce L^t_A for a fixed A. (**??**) is unchanged so that, using H$'$older inequality w.r.t. time we have to control

$$\left(\int_0^T ||L^t_A||^{2s}_\alpha dt\right)^{\frac{1}{s}}\left(\int_0^T I(\rho_t)^{ds'(\frac{3}{2}-\frac{1}{2p'}-\frac{1}{p})}\,dt\right)^{\frac{1}{s'}}$$

for $\frac{1}{s}+\frac{1}{s'}=1$. As before what is required for this to be interesting is $\alpha\geq ds'$ ($>$ if $d=2$), yielding, provided $L\in\mathbb{L}^{2s}([0,T],\mathbb{L}^\alpha(\mathbb{R}^d))$,

$$\frac{d}{\alpha}+\frac{1}{s}\leq 1$$

i.e. generalizing the time inhomogeneous case up to the critical case (of equality). ♢

Corollary 4.4. *Under the assumptions of Theorem 4.1 the solution of (1.3) admits for all $t>0$ a density ρ_t w.r.t. Lebesgue measure. In addition*

(1) *If $\alpha>d$, $\rho_t\in\mathbb{L}^p(\mathbb{R}^d)$ for all $1\leq p<+\infty$.*
(2) *If $\alpha\geq d\geq 3$, $\rho_.\in\mathbb{L}^1([0,T],\mathbb{L}^p(\mathbb{R}^d))$ for all $1\leq p\leq d/d-2$.*

Proof. Let Q be a solution of (1.3). First since $\rho_t\in\mathbb{L}^1(\mathbb{R}^d)$, $\beta(t,.)=K*\rho_t\in\mathbb{L}^\alpha(\mathbb{R}^d)$. We thus may apply Theorem 3.8 and Remark 3.9 (since β is time inhomogeneous) ensuring that the *linear* SDE

$$dY_t=\sqrt{2}\,dB_t-\,b(Y_t)dt\,-\,\beta(t,Y_t)dt$$

has an unique solution whose marginals flow satisfies the desired integrability properties. □

Theorem 4.1 also furnishes via Ito formula an existence result for the non linear McKean-Vlasov P.D.E.

Corollary 4.5. *Under the assumptions of Theorem 4.1, equation (1.1) admits a weak solution, i.e. a flow of measures $\nu_t = \rho_t(x)dx$ such that for all $f \in C_b^{1,\infty}(\mathbb{R}^+ \times \mathbb{R}^d)$ and all $0 \le s < t \le T$,*

$$\int f(t,x)d\nu_t - \int f(s,x)d\nu_s = \int_s^t \int (\partial_u f + \Delta_x f + \langle (b + K * \rho_u), \nabla_x f\rangle \, d\nu_u \, du \,. \tag{4.4}$$

In addition this solution belongs to $\mathbb{L}^1([0,T], \mathbb{L}^p(\mathbb{R}^2))$ for any $p \in (1, d/d-2)$ and $p = d/d-2$ if $d \ge 3$, and is a finite energy solution, i.e. $\int_0^T I(\rho_t)dt < +\infty$ and $\int \rho_t \, |\ln(\rho_t)|dx < +\infty$.

A converse statement directly follows from Theorem 2.9. Indeed (4.4) is exactly the weak forward equation Theorem 2.9 with the non homogeneous drift $(b + K * \rho_t)$. We may thus state

Theorem 4.6. *Let V, b, ρ_0 be as in Theorem 4.1 and K be a measurable kernel. Let $t \mapsto \rho_t$ be a flow of probability densities, weak solution of (4.4) such that $\int_0^T \int |K * \rho_t|^2 \, \rho_t \, dx \, dt < +\infty$. Then there exists a weak solution Q of (1.3) and $H(Q|P) < +\infty$.*

Example 4.7. Let us describe two situations (in addition to the one in Theorem 4.1) where one can check the energy condition of Theorem 4.6

(1) Assume that $K \in \mathbb{L}^\alpha(\mathbb{R}^d)$ for some $\alpha \ge 1$. Choose $p = p' = \frac{3\alpha}{3\alpha-2}$, provided $\frac{3\alpha}{3\alpha-2} \le \frac{d}{d-2}$ i.e. $\alpha \ge d/3$. (4.3) furnishes

$$\int |K * \rho_t|^2 \, \rho_t \, dx \le C(d) \, ||K||_\alpha^2 \, (I(\rho_t))^{\frac{d}{\alpha}}$$

so that if $I(\rho_t) \in \mathbb{L}^{d/\alpha}([0,T])$ (in particular if $\sup_{t\in]0,T]} I(\rho_t) < +\infty$) we may apply Theorem 4.6.

(2) Let K and $p = \frac{3\alpha}{3\alpha-2}$ as in the previous example. If $||\rho_t|| \in \mathbb{L}^3([0,T], \mathbb{L}^{3\alpha/(3\alpha-2)}(\mathbb{R}^d))$ (in particular if $\sup_{t\in]0,T]} ||\rho_t||_{3\alpha/(3\alpha-2)} < +\infty$) we may directly apply (4.2) to show that Theorem 4.6 applies.

We shall discuss explicit examples satisfying in particular the second condition in the next sections.

For uniqueness we will have to discuss separately each example. However the next result covers some cases

Proposition 4.8. *Assume that the assumptions of Theorem 4.1 are fulfilled for $\alpha \ge d$ (> 2 if $d = 2$). Then there is only one weak solution of (1.3).*

Proof. Let Q be a solution of (1.3). First since $\rho_t \in \mathbb{L}^1(\mathbb{R}^d)$, $\beta(t,.) = K * \rho_t \in \mathbb{L}^\alpha(\mathbb{R}^d)$. We thus may apply Theorem 3.8 and the final comment of Remark 3.9 (since β is time inhomogeneous) ensuring that the *linear* SDE

$$dY_t = \sqrt{2}\, dB_t - b(Y_t)dt \, - \, \beta(t, Y_t)dt$$

has an unique solution. Since Q is a solution, it is the only one and $H(Q|P) < +\infty$. It follows that $\int_0^T I(\rho_t) \, dt < +\infty$.

Let Q' be another solution of (1.3), with marginals ρ'_t and same initial ρ_0. Denote by Q_t and Q'_t the restriction of Q and Q' to $[0,t]$. Applying Pinsker inequality, it holds

$$\begin{aligned}\sup_{u\le t}||\rho'_u-\rho_u||_1^2 &\le 2\sup_{u\le t}H(\rho_u|\rho'_u)\le 2\,H(Q_t|Q'_t)\\ &\le 2\int_0^t\int|K*(\rho_s-\rho'_s)|^2\,\rho_s\,dx\,ds\\ &\le 2\,||K\,\mathbf{1}_{|K|>A}||_\alpha^2\sup_{u\le t}||\rho'_u-\rho_u||_p^2\int_0^t||\rho_s||_{p'}\,ds+2\,A^2\,t\sup_{u\le t}||\rho'_u-\rho_u||_1^2\end{aligned}$$

for any p and p' in $[1,d/d-2]$ if $d\ge 3$, $[1,+\infty)$ if $d=2$, and satisfying $\frac{3}{2}=\frac{1}{\alpha}+\frac{1}{2p'}+\frac{1}{p}$ (recall (4.2)). Choose $p=1$ and $p'=\alpha/(\alpha-2)\le d/(d-2)$. We deduce from Lemma 3.1 (2) and (3) that

$$\int_0^t||\rho_s||_{p'}\,ds\le C(p')\left(T+\int_0^T I(\rho_s)ds\right)\le C'(p',T,H(Q|P))\,,$$

where $M\mapsto C'(p',T,M)$ is non-decreasing. First choose A such that

$$2\,||K\,\mathbf{1}_{|K|>A}||_\alpha^2\,C'(p',T,H(Q|P))\le\frac{1}{2}\,.$$

We thus have

$$\sup_{u\le t}||\rho'_u-\rho_u||_1^2\le 4A^2\,t\sup_{u\le t}||\rho'_u-\rho_u||_1^2$$

so that if t is small enough, satisfying $4A^2t<1$ we deduce $\rho_u=\rho'_u$ for all $u\le t$.
We may now iterate the procedure, i.e.

$$\begin{aligned}\sup_{u\le 2t}||\rho'_u-\rho_u||_1^2 &\le 2\int_0^{2t}\int|K*(\rho_s-\rho'_s)|^2\,\rho_s\,dx\,ds=2\int_t^{2t}\int|K*(\rho_s-\rho'_s)|^2\,\rho_s\,dx\,ds\\ &\le 2\,||K\,\mathbf{1}_{|K|>A}||_\alpha^2\sup_{u\le 2t}||\rho'_u-\rho_u||_1^2\int_0^{2t}||\rho_s||_{p'}\,ds+2\,A^2\,t\sup_{u\le 2t}||\rho'_u-\rho_u||_1^2\end{aligned}$$

and deduce as before $\rho_u=\rho'_u$ for all $u\le 2t$ and finally for $0\le u\le T$. It follows that Q and Q' are solutions of the same *linear* SDE as before, so that as we already said that $Q=Q'$. □

The previous result extends [73] Theorem 4.8 to the critical case $\alpha=d$ and to a larger class of initial conditions. One can also see [107].

Remark 4.9. The previous proof actually shows that, in full generality, if $K\mathbf{1}_{|K|>A}\in\mathbb{L}^\alpha(\mathbb{R}^d)$ for some $\alpha\ge 2$, uniqueness holds in the set of solutions such that $\int_0^T||\rho_s||_{p'}\,ds<+\infty$, for $p'\ge\alpha/(\alpha-2)$. ♢

In some cases one can give a direct proof of uniqueness

Proposition 4.10. *Consider (1.3), assume that b is L-Lipschitz and that K satisfies $|K(x)-K(y)|\le C\,|x-y|\,(\frac{1}{|x|^m}+\frac{1}{|y|^m})$ for some $0\le m<d$. Then there exists at most one solution of (1.3) which is pathwise unique in the set of Probability measures whose marginals flow satisfies $\rho_.\in\mathbb{L}^1([0,T],\mathbb{L}^p(\mathbb{R}^d))$ in particular if $\sup_{t\in(0,T]}||\rho_t||_p\le M$ for some $p>d/(d-m)$.*

The proof is similar to the proof of Theorem 1.7 in [68] p-18.

Remark 4.11. It is interesting to notice that if $K(x) = x/|x|^m$ for some $m > 0$, the previous inequality is satisfied with the same m and $C = m + 1$. Look at the first two lines in the proof of Lemma 2.5 in [68], and consider the case $|x| \leq |y|$. $\diamondsuit$

5. SDE and non linear SDE for some singular examples.

In this section we shall review the models we have quoted in the introduction, and give results of existence and uniqueness for both the SDE system of particles and the non linear SDE.

5.1. The η relaxed Keller-Segel model.

Recall that this model corresponds to $K_\eta(x) = \chi \frac{x}{|x|^{2-\eta}} \mathbf{1}_{x\neq 0}$ for $x \in \mathbb{R}^2$.

The existence of a strong solution for the particle system can be shown either by using Dirichlet forms as in subsubsection 2.4.1 or approximations i.e. Theorem 3.10 (and no collisions to get a strong and not only weak solution). The second proof shows that this solution is of finite entropy.

The existence of a solution for the non-linear SDE is a consequence of Theorem 4.1. This solution has finite entropy. The marginals flow of this solution thus satisfies $\int_0^T I(\rho_t)dt < +\infty$ so that, according to (3.2), $\rho_. \in \mathbb{L}^1([0,T], \mathbb{L}^p(\mathbb{R}^2))$ for any $1 \leq p < +\infty$. According to Remark 4.11 we may apply Proposition 4.10 to get strong uniqueness in the corresponding set of Probability measures.

Actually we have strong uniqueness without restrictions. Indeed if Q' is another solution of the non linear SDE, it is also a solution of the *linear* SDE with drift $b + K * \rho_t$. Hence according to Remark 3.9 (see (3.16)) its marginals flow $\rho'_. \in \mathbb{L}^1([0,T], \mathbb{L}^p(\mathbb{R}^2))$ for all $p < +\infty$, so that we may apply the previous uniqueness result.

The η relaxed Keller-Segel model is thus the prototype to which all we have done before applies. Notice that we may take $\chi > 0$ or $\chi < 0$, as we said before the entropic approach cannot separate repulsive and attractive situations.

5.2. The sub-Coulombic repulsive model.

We directly state a first result for the particle system

Theorem 5.1. *For $d \geq 3$ look at $K(x) = \chi \frac{x}{|x|^{s+2}} \mathbf{1}_{x\neq 0}$ for $d-2 \geq s > 0$ and $\chi < 0$, i.e. the repulsive situation in [109]. Assume in addition that b is bounded and a-Lipschitz.*

Finally assume the following property: for all $i \neq j$, the distribution $\tilde{\rho}_{0,i,j}$ of $(X_0^i - X_0^j)$ satisfies $\tilde{\rho}_{0,i,j} \in \mathbb{L}^q(\mathbb{R}^d)$ for some $q > d/(d-s)$.

Then the particle system (1.2) admits a unique strong solution whose distribution Q satisfies $H(Q|\frac{\rho_0}{\gamma_0} P) < +\infty$, where P is the product measure introduced in subsection 2.1 with $|V'|$ bounded.

Accordingly, if $\int \rho_0 \ln(\rho_0/\gamma_0)\, dx < +\infty$, the marginals flow ρ_t satisfies $\int_0^T I(\rho_t)dt < +\infty$ so that $\rho_. \in \mathbb{L}^1([0,T], \mathbb{L}^p(\mathbb{R}^d))$ for $p \in [1, Nd/Nd-2]$.

Notice that, as explained in the proof of Lemma 3.2 the additional property is satisfied as soon as $I(\rho_{0,i,j}) < +\infty$ where $\rho_{0,i,j}$ is the joint density of (X_0^i, X_0^j).

Proof. According to subsubsection 2.4.3 we know that the particle system admits a unique strong solution which is absolutely continuous with respect to the Wiener measure with the same initial distribution, i.e. there is no collision. If we add a confining potential the stationary solution has finite relative entropy.
Subsection 3.4 for $d-2 \geq s > 0$ says more. Recall that we have defined $H(x) = \sum_{i \neq j} \frac{1}{|x^i - x^j|^s}$. The assumption on ρ_0 ensures that $\mathbb{E}(H(X_0^N)) < +\infty$. (3.19) ($m = s+2$) shows in particular that

$$\begin{aligned}\mathbb{E}\left[\frac{2|\chi|s}{N}\int_0^T \sum_i \left|\sum_j \frac{X_t^{i,N} - X_t^{j,N}}{|X_t^{i,N} - X_t^{j,N}|^{s+2}}\right|^2 dt\right] &\leq \mathbb{E}(H(X_0^N)) \\ &\quad + s\,a\int_0^T \mathbb{E}(H(X_t^N))\,dt \\ &\leq \mathbb{E}(H(X_0^N))\,(1 + e^{saT})\end{aligned}$$

since according to Subsection 3.4, $\mathbb{E}(H(X_t^N)) \leq \mathbb{E}(H(X_0^N))\,e^{sat}$. Hence if $\mathbb{E}(H(X_0^N)) < +\infty$, the left hand side is also finite. If $b = 0$, up to the pre-factor this left hand side is nothing else but the relative entropy of the law Q of $X_\cdot^N$ w.r.t. the Wiener measure with the same initial condition. Adding a bounded b does not change the finiteness of relative entropy, so that $H(Q|\frac{\rho_0}{\gamma_0}P)$ is also finite, where P is the product measure introduced in subsection 2.1 with $|V'|$ bounded as in Lemma 2.5. Arguing as in the proof of Lemma 3.2, we immediately see that for $s \leq d-2$, $\mathbb{E}(H(X_0^N)) < +\infty$.
The others statements in the Theorem follow from the discussion on relative entropy and Lemma 3.1. □

Hence the particle system has an entropic solution with or without confinement. This has to be related to some comments in the introduction (subsection 1.4 of [109]) about the confinement force.

Remark 5.2. In the previous Theorem, assume that $\max_{i,j} ||\tilde{\rho}_{0,i,j}||_q = C_q$. Then $\mathbb{E}(H(X_0^N)) \leq C(s,a,T,C_q)\,N^2$. It follows that

$$H(Q|\frac{\rho_0}{\gamma_0}P) \leq C(s,a,T,|\chi|,C_q)\,N\,.$$

This will be interesting in the sequel. ♢

Clearly $K \notin \mathbb{L}^\alpha(\mathbb{R}^d)$ for any $\alpha \geq d$ so that the entropic approach of the non-linear SDE fails. Fortunately the non-linear PDE (4.4) has some good properties (for our purpose). Let us recall the following result shown in section 3 of [109] (we emphasize that this result assumes that there is no confining potential, i.e. $b = 0$, see the discussion in Remark 1.6 of [109])

Proposition 5.3. *Suppose $b = 0$, $0 < s < d-2$ and that ρ_0 is a probability density which is bounded. Then there exists a flow $t \mapsto \rho_t$ of probability densities, satisfying (4.4) and belonging to $\mathbb{C}^0([0,T], \mathbb{L}^\infty(\mathbb{R}^d))$. This flow belongs to $C^0([0,T], \mathbb{L}^\infty(\mathbb{R}^d))$ and satisfies $||\rho_t||_\infty \leq ||\rho_0||_\infty$. It is the unique solution of (4.4) in the set $\mathbb{C}^0([0,T], \mathbb{L}^\infty(\mathbb{R}^d))$.*

Here are the precise references in [109]. Remark 3.7 shows that the non-linear PDE is positivity preserving and Remark 3.4 that mass is preserved. Proposition 3.1 shows existence of a bounded solution on small enough time interval, Remark 3.4 shows that the $\mathbb{L}^\infty$ norm of ρ_t is non-increasing in time, so that one can conclude (see Remark 3.5) to the existence of a global (in time) bounded solution. Uniqueness is part of Proposition 3.1.

Notice that even if ρ_0 is bounded we do not know (at this stage) whether the density of the particle system at time t is bounded or not, but we can apply Theorem 5.1.

We can thus apply Theorem 4.6 since for A large enough, $K\,\mathbf{1}_{|K|>A} \in \mathbb{L}^1(\mathbb{R}^d)$ (this is true as soon as $s < d-1$) so that

$$\int_0^T \int |K * \rho_t|^2 \, \rho_t \, dx \, dt \;\leq\; 2T \, (\sup_{0\leq t\leq T} ||\rho_t||_\infty^2 \, ||K\,\mathbf{1}_{|K|>A}||_1^2 \;+\; A^2) \, .$$

We thus have obtained the existence of a weak solution to the non-linear SDE, which is of finite entropy. Of course the previous bound is obtained via

$$|K * \rho_t| \;\leq\; A \;+\; \sup_{0\leq t\leq T} ||\rho_t||_\infty \, ||K\,\mathbf{1}_{|K|>A}||_1 \tag{5.1}$$

showing that the "drift" $|K * \rho_t|$ is bounded. The existence of a *linear* diffusion process with this drift is thus standard via Girsanov theory. That this diffusion process has marginals ρ_t is more or less classical and it seems that calling upon the results of subsection 2.1, if presumably not necessary, is not as strange. Nevertheless this remark shows that dQ/dP belongs to all the $\mathbb{L}^p$'s for $1 \leq p < +\infty$ and not only L^{xlnx}. Finally one can apply the uniqueness part in Proposition 5.3.

Let us gather the results we have just described

Theorem 5.4. *For $d \geq 3$ look at $K(x) = \chi \, \frac{x}{|x|^{s+2}} \, \mathbf{1}_{x\neq 0}$ for $d-2 > s > 0$ and $\chi < 0$. Finally assume that ρ_0 is a bounded density of probability.*
Then there exists a solution $\bar{Q}$ of the non linear SDE (1.3). This solution is absolutely continuous w.r.t. $(\rho_0/\gamma_0)P$ with a density $Z_T \in \bigcap_{1\leq p<+\infty} \mathbb{L}^p((\rho_0/\gamma_0)P)$ (in particular with finite entropy) and such that the marginals flow satisfies $\sup_{0\leq t\leq T} ||\rho_t||_\infty < +\infty$.
It is the pathwise unique solution in the set of Probabilities on the path space such that the marginals flow satisfies $\sup_{0\leq t\leq T} ||\rho_t||_\infty < +\infty$.

Notice that the limitation to the repulsive case seems to be necessary for Proposition 5.3 too.

We shall see in section 8 another approach of existence. A natural question will thus be to get uniqueness in more general situations. To this end we may see how to apply Remark 4.9 or Proposition 4.10 and Remark 4.11. Since $K \in \mathbb{L}^\alpha(\mathbb{R}^d)$ for $\alpha < d/(1+s)$, in order to apply Remark 4.9 we need $d > 2(1+s)$ and $\rho_. \in \mathbb{L}^1([0,T], \mathbb{L}^{d/(d-2s-2)}(\mathbb{R}^d))$ which is of course worse than the $d/(d-s-2)$ obtained in Proposition 4.10.
We are nevertheless led to look at a priori q integrability of the marginals of solutions to the non linear SDE. This will be also done in section 8.

5.3. The $2D$ vortex model.

In dimension $d = 2$ we consider the Biot-Savart kernel $K(x) = \chi \frac{x^\perp}{|x|^2} \mathbf{1}_{x\neq 0}$ where for $x = (x_1, x_2)$ we define $x^\perp = (-x_2, x_1)$. The case of $\chi < 0$ is studied in [98], we will also consider the case $\chi > 0$. Notice that in [62] a more general situation with some random weights χ_j is also considered.

Look first at the particle system.

When $b = 0$, the natural candidate as invariant measure is given by the density

$$\rho_\infty(x) = \prod_{i<j} e^{-\frac{\chi}{N} \arctan((x_2^i - x_2^j)/(x_1^i - x_1^j))} \tag{5.2}$$

which is bounded (we may define $\arctan(0/0) = 0$ if necessary), but does not satisfy (2.14), so that we may not use the strategy of subsection 2.2. For $b \neq 0$ the same is true.

Nevertheless one can use the strategy in subsection 3.4 and find a Lyapunov function proving the absence of collisions. This is done in [111]. Actually, we may use the function

$$H(x) = -\sum_{i\neq j} \ln(|x^i - x^j|^2)$$

as in subsection 3.4.

Up to a multiplicative constant the first part of the drift is given by

$$\sum_{i,j,k} \left\langle \frac{X_t^{i,N} - X_t^{j,N}}{|X_t^{i,N} - X_t^{j,N}|^2}, \left(\frac{(X_t^{i,N} - X_t^{k,N})^\perp}{|X_t^{i,N} - X_t^{k,N}|^2} - \frac{(X_t^{j,N} - X_t^{k,N})^\perp}{|X_t^{j,N} - X_t^{k,N}|^2} \right) \right\rangle$$

i.e.

$$\sum_{i,j,k} \left\langle \frac{X_t^{i,N} - X_t^{j,N}}{|X_t^{i,N} - X_t^{j,N}|^2}, \frac{(X_t^{i,N} - X_t^{k,N})^\perp}{|X_t^{j,N} - X_t^{k,N}|^2} \right\rangle - \sum_{i,j,k} \left\langle \frac{X_t^{i,N} - X_t^{j,N}}{|X_t^{i,N} - X_t^{j,N}|^2}, \frac{(X_t^{j,N} - X_t^{k,N})^\perp}{|X_t^{j,N} - X_t^{k,N}|^2} \right\rangle .$$

Exchanging the role of i and j we see that both terms in the difference are opposite, so that this difference is twice the first term. Now exchanging the role of j and k we have

$$\sum_{i,j,k} \left\langle \frac{X_t^{i,N} - X_t^{j,N}}{|X_t^{i,N} - X_t^{j,N}|^2}, \frac{(X_t^{i,N} - X_t^{k,N})^\perp}{|X_t^{j,N} - X_t^{k,N}|^2} \right\rangle = \sum_{i,j,k} \left\langle \frac{X_t^{i,N} - X_t^{k,N}}{|X_t^{i,N} - X_t^{k,N}|^2}, \frac{(X_t^{i,N} - X_t^{j,N})^\perp}{|X_t^{j,N} - X_t^{k,N}|^2} \right\rangle$$

so that using that $\langle u, v^\perp \rangle = -\langle u^\perp, v \rangle$ the previous sum equals 0.

Finally, using that $\Delta_x \ln(|x|^2) = 0$ for $x \neq 0$ (3.19) becomes

$$H(X_u^N) - H(X_0^N) = 2 \int_0^u \sum_{i\neq j} \frac{\langle b(X_t^{i,N}) - b(X_t^{j,N}), X_t^{i,N} - X_t^{j,N} \rangle}{|X_t^{i,N} - X_t^{j,N}|^2} \, dt \, + \, M_u \tag{5.3}$$

for all stopping time u with $u < \tau_0$ (τ_ε denotes the hitting time of the set C_ε defined in Theorem 2.14), ensuring again that there are no collisions starting from a non collision x arguing as follows. Since H is not everywhere positive we have to introduce S_L the first time $H(X_\cdot^N)$ becomes less than $-L$, for a positive L. We thus have

$$\mathbb{E}[H(X_{T\wedge\tau_\varepsilon\wedge S_L}^N)] \leq \mathbb{E}[H(X_0^N)] + 2\,T\,a$$

so that if $X_0^N = x$ a non collision point, $\tau_0 \geq T \wedge S_L$ almost surely for all L. It remains to make L go to infinity.
Remark that the result does not depend on the sign of χ.
According to Theorem 2.14 the distribution Q^N of $X_\cdot^N$ on $C([0,T],(\mathbb{R}^2)^N)$ is thus absolutely continuous with respect to the corresponding Wiener measure or w.r.t P provided the distribution at time 0 is absolutely continuous w.r.t. γ_0.

Contrary to the sub-Coulombic case, we cannot use (5.3) for studying the relative entropy. One can however obtain interesting informations. Indeed, we may explicitly write the martingale part M_t. It holds

$$\begin{aligned} M_t &= -2\sqrt{2}\int_0^t \sum_{i\neq j} \frac{\langle X_s^{i,N}-X_s^{j,N}, dB_s^i - dB_s^j\rangle}{|X_s^{i,N}-X_s^{j,N}|^2} \\ &= -4\sqrt{2}\sum_i \int_0^t \left\langle \sum_{j\neq i} \frac{X_s^{i,N}-X_s^{j,N}}{|X_s^{i,N}-X_s^{j,N}|^2}, dB_s^i \right\rangle \end{aligned}$$

the second equality being obtained by exchanging the roles of i and j, so that

$$\mathbb{E}(M_T^2) = C\sum_i \int_0^T \left| \sum_{j\neq i} \frac{X_t^{i,N}-X_t^{j,N}}{|X_t^{i,N}-X_t^{j,N}|^2} \right|^2 dt\,.$$

Since $|U^\perp| = |U|$, once again, up to a constant and to the additional regular drift b we recognize $H(Q|\frac{\rho_0}{\gamma_0}P)$. More precisely

$$H(Q|\frac{\rho_0}{\gamma_0}P) \leq \frac{\chi^2}{32\,N^2}\,\mathbb{E}(M_T^2) + N\,||b||_\infty^2\,. \tag{5.4}$$

We are thus led to estimate $\mathbb{E}(H^2(X_u^N))$ for $u = 0, T$ since the expectation of the square of remaining term involving b in (5.3) is less than $4a^2T\,N^4$. We will use the rough estimate $|\sum_{i\neq j} a_{i,j}|^2 \leq N^2 \sum_{i\neq j} |a_{i,j}|^2$ and the elementary

$$\ln^2(v) \leq C + v + \frac{1}{v}$$

for some constant C and all $v > 0$. It allows us to reduce the problem to the study of $G_\alpha(x) = \sum_{i\neq j} |x^i - x^j|^\alpha$ for $\alpha = 1$ and $\alpha = -1$.
Using Ito's formula, the same manipulations as for H show that

$$\begin{aligned} G_\alpha(X_u^N) - G_\alpha(X_0^N) &= \alpha \int_0^u \sum_{i\neq j} \langle b(X_t^{i,N}) - b(X_t^{j,N}), X_t^{i,N} - X_t^{j,N}\rangle\, |X_t^{i,N} - X_t^{j,N}|^{\alpha-2}\, dt \\ &\quad + 2\,\alpha^2 \int_0^u \sum_{i\neq j} |X_t^{i,N} - X_t^{j,N}|^{\alpha-2}\, dt \;+\; M_{u,\alpha} \end{aligned}$$

for all $u < \tau_0$. We immediately deduce using Gronwall lemma that for $\alpha = 2$, and $t \leq T$,

$$\mathbb{E}(G_2(X_t^N)) \leq \mathbb{E}(G_2(X_0^N)) + C(a,T)\,N^2\,.$$

For any $0 < \alpha \leq 2$, using $|v|^\alpha \leq 1 + |v|^2|$, it follows

$$\mathbb{E}(G_\alpha(X_t^N)) \leq \mathbb{E}(G_2(X_0^N))) + C'(a,T)N^2\,.$$

We thus deduce

$$\begin{aligned}\int_0^T \sum_{i\neq j} \mathbb{E}(|X_t^{i,N} - X_t^{j,N}|^{\alpha-2})\, dt &\leq \frac{a}{2\alpha} \int_0^T G_\alpha(X_t^N)\, dt \,+\, (1/2\alpha^2)\, \mathbb{E}(G_\alpha(X_T^N)) \\ &\leq C(T,a,\alpha)\, (N^2 + \mathbb{E}(G_2(X_0^N)))\,. \end{aligned} \tag{5.5}$$

and finally

$$\int_0^T \mathbb{E}(H^2(X_t^N))\, dt \,\leq\, C(T)\, N^2 \left(N^2 \,+\, \int_0^T \mathbb{E}(G_1(X_t^N))\, dt \,+\, \int_0^T \sum_{i\neq j} \mathbb{E}(|X_t^{i,N} - X_t^{j,N}|^{-1})\, dt \right) \tag{5.6}$$

so that using (5.5) with $\alpha = 1$ and what precedes

$$\int_0^T \mathbb{E}(H^2(X_t^N))\, dt \,\leq\, C(T,a)\, N^2\, (N^2 + \mathbb{E}(G_2(X_0^N)))\,. \tag{5.7}$$

From (5.3) we deduce for all $t \leq T$,

$$\begin{aligned}\mathbb{E}(M_t^2) &\leq 3\,\mathbb{E}(H^2(X_t^N)) + 3\,\mathbb{E}(H^2(X_0^N)) + c\, a^2\, T^2\, N^4 \\ &\leq 3\,\mathbb{E}(H^2(X_t^N)) + C(a,T) N^2 (N^2 + \mathbb{E}(G_2(X_0^N))) + c\, a^2\, T^2\, N^4\,. \end{aligned}$$

so that integrating with respect to t it holds for all time T,

$$\int_0^{2T} \mathbb{E}(M_t^2)\, dt \,\leq\, C(a,T)\, N^2 (N^2 + \mathbb{E}(G_2(X_0^N))) \,+\, c\, a^2\, T^2\, N^4\,. \tag{5.8}$$

Using that for $T \leq t \leq 2T$, $\mathbb{E}(M_t^2) \geq \mathbb{E}(M_T^2)$, we deduce from (5.8) that

$$\mathbb{E}(M_T^2) \leq T^{-1}\, C(a,T)\, N^2 (N^2 + \mathbb{E}(G_2(X_0^N)))\,. \tag{5.9}$$

so that we have obtained the analogue of Theorem 5.1

Theorem 5.5. *For $d = 2$ look at $K(x) = \chi\, \frac{x^\perp}{|x|^2}\, \mathbf{1}_{x\neq 0}$ and assume in addition that b is bounded and a-Lipschitz.*
If $\int \sum_{i,j} |x^i - x^j|^2\, \rho_0(dx) < +\infty$, then the particle system (1.2) admits a unique strong solution whose distribution Q satisfies $H(Q|\frac{\rho_0}{\gamma_0} P) < +\infty$, where P is the product measure introduced in subsection 2.1 with $|V'|$ bounded.
Accordingly, if $\int \rho_0\, \ln(\rho_0/\gamma_0)\, dx < +\infty$, the marginals flow ρ_t satisfies $\int_0^T I(\rho_t) dt < +\infty$ so that $\rho_. \in \mathbb{L}^1([0,T], \mathbb{L}^p(\mathbb{R}^d))$ for $p \in [1, 2N/2(N-1)]$.

Remark 5.6. The previous calculation shows that

$$H(Q|\frac{\rho_0}{\gamma_0} P) \leq C(a,T) N^2 \left(1 + \sup_{i,j} \int |x^i - x^j|^2\, \rho_0(dx) \right) + N\, ||b||_\infty^2\,.$$

This bound will be unfortunately insufficient for studying propagation of chaos. ♢

One can nevertheless obtain a somehow better bound for the free energy functional, assuming that the additional drift $b = \nabla \tilde{U}$ as in subsection 2.1, with $\tilde{U}(x) = \sum_j U(x^j)$ and $\int e^{-U} < +\infty$, so that $\theta_0 = Z^{-1}\, \rho_\infty\, e^{-\tilde{U}}\, dx$ is a reversible probability measure for the particle system. We will assume that U' is Lipschitz, but non necessarily bounded in order to allow a gaussian confinement.

We may then use Remark 2.10 and obtain

$$H(\rho_0\,dx|\theta_0) \geq H(\rho_T\,dx|\theta_0) + \int_0^T \int \sum_{i=1}^N \left|\nabla_i \ln(\rho_t) + b(x) + \frac{\chi}{2N}\sum_{j\neq i} \frac{(x^i-x^j)^\perp}{|x^i-x^j|^2}\right|^2 \rho_t\,dx\,dt\,.$$

Since

$$H(\rho_0\,dx|\theta_0) = \int \rho_0 \left(\ln(\rho_0) + \sum_i U(x^i) + \frac{\chi}{2N}\sum_{i\neq j} \arctan((x_2^i - x_2^j)/(x_1^i - x_1^j))\right) dx$$

we obtain that provided $\int \rho_0 |\ln(\rho_0)| dx = O(N)$ and $\max_j \int |U(x^j)|\,\rho_0\,dx < +\infty$, the free energy dissipation

$$\int_0^T \sum_i \int \left|\nabla_i \ln(\rho_t) + b(x) + \frac{\chi}{2N}\sum_{j\neq i} \frac{(x^i-x^j)^\perp}{|x^i-x^j|^2}\right|^2 \rho_t\,dx\,dt$$

is of order N.

According to the discussion in subsection 2.1, the marginals flow of the solution is (up to some constants) almost an entropy solution of the Liouville equation in the sense of [78] definition 2 (so that we partly recover Proposition 1 therein).

First using $|a+b|^2 \geq \frac{1}{2}|a|^2 - |b|^2$ we have

$$H(\rho_0\,dx|\theta_0) + \int_0^T \int |b|^2\,\rho_t dx\,dt \geq H(\rho_T\,dx|\theta_0) +$$

$$\frac{1}{2}\int_0^T \sum_i \int \left|\nabla_i \ln(\rho_t) + \frac{\chi}{2N}\sum_{j\neq i} \frac{(x^i-x^j)^\perp}{|x^i-x^j|^2}\right|^2 \rho_t\,dx\,dt$$

Expanding the squared norm in the right hand side, using a truncation $\psi(x)$ with compact support, the product term becomes

$$\frac{2\chi}{N}\int \left\langle \nabla\rho, \psi \sum_{i\neq j} \frac{(x^i-x^j)^\perp}{|x^i-x^j|^2}\right\rangle dx$$

one can integrate by parts. Since the divergence of the Biot-Savart kernel vanishes it remains to look at

$$\int \rho \left\langle \sum_{i\neq j} \frac{(x^i-x^j)^\perp}{|x^i-x^j|^2}, \nabla\psi\right\rangle dx$$

which goes to 0 by correctly choosing a sequence of ψ's, since

$$\int \left|\sum_{i\neq j} \frac{(x^i-x^j)^\perp}{|x^i-x^j|^2}\right| \rho\,dx < +\infty\,.$$

For the left hand side we remark that $|b|^2$ is at most quadratic. Using one more time the specificity of the kernel, it is immediately seen, since b is Lipschitz, that

$$\mathbb{E}(|X_t|^2) \leq \mathbb{E}(|X_0|^2) + c\left(t + \int_0^t \mathbb{E}(|X_s|^2)ds\right)$$

so that $\mathbb{E}(|X_t|^2) \leq c(T)\,\mathbb{E}(|X_0|^2)$ by using Gronwall.
Since $H(\rho_T dx|\theta_0) \geq 0$, we finally have obtained a better result

Theorem 5.7. *Assume that $\tilde{U}(x) = \sum_j U(x^j)$ is a confining potential such that ∇U is Lipschitz, and that $\int |x|^2 \rho_0(x)dx < +\infty$. Then there exists a constant $c(T)$ such that*

$$H(\rho_0\, dx|\theta_0) + c(T) \int |x|^2 \rho_0(x)dx \;\geq H(\rho_T\, dx|\theta_0) \;+\; \frac{1}{2}\int_0^T I(\rho_t)\, dt \;+\; \frac{1}{8}\, H(Q|\frac{\rho_0}{\gamma_0}P)\,.$$

Recall that here $\gamma_0(dx) = Z^{-1}\, e^{-\tilde{U}(x)}dx$ and P is the distribution of the γ_0 symmetric diffusion process $dY_t = dB_t - \nabla\tilde{U}(Y_t)dt$.
In particular if $H(\rho_0\, dx|\theta_0) = O(N)$ and $\int |x|^2\rho_0(x)dx = O(N)$, $H(Q|\frac{\rho_0}{\gamma_0}P) = O(N)$.

Remark 5.8. In the case without confinement, using the calculations in the proof of Lemma 3.6 and an ad-hoc regularization K_ε of K such that $div K_\varepsilon = 0$, one obtains

$$\int \rho_T\, \ln(\rho_T)\, dx + \int_0^T I(\rho_t)\, dt \;=\; \int \rho_0\, \ln(\rho_0)\, dx\,,$$

i.e. formula (5.7) in [62].
As said in [62] this equation is interesting provided one can get a lower bound for the Boltzmann entropy $\int \rho_T\, \ln(\rho_T)\, dx$ (which is no more necessarily non-negative). Such a (uniform) lower bound is obtained provided one can find an upper bound for $\int |x|^k\, d\rho_T$ (see [62] Lemma 3.1). That is why, the whole paper [62] is studying together polynomial moments, Boltzmann entropy and Fisher information. Notice that provided $\int |x|^k\, d\rho_0 = O(N)$ for some fixed $k > 0$, then, according to Proposition 5.1 in [62], as soon as $\int \rho_0\, |\ln(\rho_0)|dx = O(N)$, $\int_0^T I(\rho_t)\, dt = O(N)$. ♢

In the sequel we assume that $b = 0$. It is shown in [6] Theorem B, that for any initial density of probability ρ_0, the nonlinear (4.4) has a solution belonging to $C^0((0,T], \mathbb{L}^\infty(\mathbb{R}^2))$. If in addition ρ_0 is bounded, then $||\rho_t||_\infty \leq ||\rho_0||_\infty$. Assuming some more regularity on ρ_0 furnishes higher regularity for ρ_t see [6] Theorem A or [55] Lemma 2.2. In addition this solution is unique in the set of bounded functions.
Another existence result, using convergence of the particle system is shown in [62] Lemma 3.5 and Theorem 2.5 (one can also look at the former [98]).

In [62] section 7 another uniqueness result is shown in the set of solutions such that $\nabla\rho_\cdot \in \mathbb{L}^{2q/(3q-2)}([0,T], \mathbb{L}^q(\mathbb{R}^2))$ for all $1 \leq q < 2$. If $\int_0^T I(\rho_t)\, dt < +\infty$, the latter condition is satisfied according to 3.1 (1). This uniqueness result is self contained i.e. the proof does not use the particle approximation, contrary to the existence part.
We thus have the analogue of proposition 5.3

Proposition 5.9. *Assume that ρ_0 is a probability density such that $\int \rho_0\, (\ln(\rho_0) + |x|^2)dx < +\infty$. Then, when $b = 0$, if ρ_0 is bounded, there exists a flow $t \mapsto \rho_t$ of probability densities, satisfying (4.4) and belonging to $\mathbb{C}^0([0,T], \mathbb{L}^\infty(\mathbb{R}^d))$.*
Even if ρ_0 is not bounded, there exists at most one solution in the set of probability densities flows such that $\int_0^T I(\rho_t)\, dt < +\infty$.

Accordingly there exists a solution $\bar{Q}$ of the non linear SDE, satisfying the same properties as in Theorem 5.4 and at most one weak solution in the set of probability measures such that $\int_0^T I(Q \circ (X_t)^{-1})\, dt < +\infty$.

As we said, the proof of existence for this solution given in [62] is totally different, based on the relative compactness (tightness) of the law of $X^{1,N}_{\cdot}$. Notice that the former [61] also contains results in this direction. We shall look at this approach later.

Remark 5.10. We claim that the above result is still true if we add a confining potential as before. To show this, it is enough to extend Ben Artzi's proof following the presentation made in the proof of Theorem 2 in [69]. We will not give the details, so that this result is still a claim, and we shall see later another approach for uniqueness. ♢

5.4. **Dyson processes.**

To complete the picture, let us say a word on Dyson type models.

We gave an existence proof in subsubsection 2.4.3 for the particle system, recovering the results in [41, 70] for $\chi/N < -1$. The existence of the particle system was obtained in [41] Theorem 3.1 for $\chi < 0$ using the theory of multivalued SDE's developed by Cépa in his PHD Thesis. Transporting the system on the torus, the authors have shown in [42] that for $0 > \chi/N > -1$ collisions of two particles always occur, while for $\chi/N \leq -1$ they do not (remark that the equality case $\chi/N = -1$ is out of reach of our approach). In [43] they also have shown that no k-collisions occur for $k \geq 3$. Actually the proof in [38, 64] using Dirichlet forms presumably also works for the Dyson model, for $\chi < 0$.

The non linear SDE is not studied in [41, 70] where the limit for a large number of particles is studied for a varying $\chi = \chi_N \to +\infty$ after a linear time change (vanishing noise).

6. The $2D$ parabolic-elliptic (Patlak)-Keller-Segel attractive model.

The (Patlak) Keller-Segel system introduced in [80], is a tentative model to describe chemotaxis phenomenon, an attractive chemical phenomenon between organisms. Since we do not yet speak of this model, that contains new interesting features, we will give a more complete description of the situation.

In two dimensions, the classical 2-D parabolic-elliptic Keller-Segel model reduces to the single non linear P.D.E.,

$$\partial_t \rho_t(x) = \Delta_x\, \rho_t(x) + \chi\, \nabla_x.((K * \rho_t)\rho_t)(x) \tag{6.1}$$

with some initial ρ_0. It is not difficult to see that (6.1) preserves positivity and mass, so that we may assume that ρ_0 is a density of probability i.e. the model enters the framework of this work with $K(x) = \chi\, \frac{x}{|x|^2}$ defined on $\mathbb{R}^2$. K is the gradient of the harmonic kernel, i.e. $K(x) = \nabla\, \log(|x|)$. In order to compare our model to the usual formulation, the reader can think that the parameter χ is actually given by

$$\chi \,=\, \chi_0\, \frac{\alpha\, m}{2\pi\, D}$$

where χ_0 is the chemotactic sensitivity, α is the rate of production of chemoattractant by the cells, m is the total mass and D is the product of the diffusivities.

As usual, ρ is modeling a density of cells, and $c_t = K * \rho_t$ is (up to some constant) the concentration of chemo-attractant.

A very interesting property of such an equation is a blow-up phenomenon. The following is easily obtained by looking at the time evolution of the variance of a solution (see [10]):

Proposition 6.1. *If $\chi > 4$, the maximal time interval of existence of a classical solution of (6.1) is $[0, T^*)$ with*

$$T^* \leq \frac{1}{2\pi\,\chi\,(\chi-4)} \int |x|^2\,\rho_0(x)\,dx\,.$$

The existence of a solution, up to this explosion time, is more delicate. One classically says that $\chi < 4$ is the sub-critical case, $\chi = 4$ the critical case and $\chi > 4$ the super-critical one.

As in [10] one can consider a weak version of (6.1), i.e. looking for a continuous flow $s \mapsto \mu_s$ of probability measures on $\mathbb{R}^2$ satisfying for all smooth function φ and all $t > 0$

$$\begin{aligned}\int \varphi\,d\mu_t \;&=\; \int \varphi\,d\mu_0 + \int_0^t \int \Delta\varphi\,d\mu_s\,ds + \\ &\quad - \frac{\chi}{2}\int_0^t \int\int \langle K(x-y), \nabla\varphi(x) - \nabla\varphi(y)\rangle\,\mu_s(dx)\mu_s(dy)\,ds\,. \qquad (6.2)\end{aligned}$$

It is worth to remark that the integrand in the last term in (6.2), which is obtained thanks to the symmetry of K, is bounded, so that the integral is well defined for any probability distribution μ. The oddness of K is a key point for writing the last term in this convenient form.

We then have (see [10] Theorem 1.1 for $\chi < 4$ and [9] Theorem 1.3 for $\chi = 4$, also see [20])

Theorem 6.2. *For $\chi \leq 4$, assume that $\mu_0(dx) = \rho_0(x)\,dx$ is an initial probability distribution. If*

$$\int (|x|^2 + |\ln\rho_0|(x))\,\mu_0(dx) < +\infty\,,$$

then there exists a flow of probability densities $t \mapsto \rho_t(x)$ defined for all $t \geq 0$, such that $\mu_t(dx) = \rho_t(dx)dx$ is a solution of (6.2).

Actually the solution built in these papers satisfies additional regularity properties. First for any $T > 0$,

$$\sup_{t\in[0,T]} \int (|x|^2 + |\ln\rho_t|(x))\,\rho_t(x)\,dx \;< +\infty\,. \qquad (6.3)$$

Next, if we introduce the free energy of a density of probability ρ, defined by

$$F(\rho) \;=\; \int \rho\,\ln\rho\,dx \;+\; \frac{\chi}{2}\int c(\rho)\,\rho\,dx\,, \qquad (6.4)$$

where

$$c(\rho)(x) = \int \ln(|x-y)|)\,\rho(y)\,dy$$

then the solution in theorem 6.2 satisfies

$$F(\rho_t) \;+\; \int_0^t \int (|\nabla\ln\rho_s(x) + \chi\nabla c(\rho_s)(x)|^2)\,\rho_s(x)\,dx\,ds \;\leq\; F(\rho_0)\,. \qquad (6.5)$$

Such a solution is called a *free energy* solution. This notion is particularly relevant for uniqueness, for which we have the following (see [54] Theorem 1.3)

Theorem 6.3. *If $\chi \leq 4$, and ρ_0 satisfies the assumption in theorem 6.2, there exists at most one free energy solution of (6.2), i.e. satisfying (6.3) and (6.5).*

Remark 6.4. Actually [10] contains slightly better results. Indeed according to Lemma 2.11 and the proof of Lemma 2.12 therein, for all $t > 0$,

$$\int_0^t \int |\nabla \ln \rho_s(x)|^2 \, \rho_s(x) \, dx \, ds < +\infty$$

and

$$\int_0^t \int |\chi \nabla c(\rho_s)(x)|^2 \, \rho_s(x) \, dx \, ds < +\infty$$

too.

Actually this point is automatic, as shown in [54] Lemma 2.2. Denote by

$$D_t(F) = \int_0^t \int (|\nabla \ln \rho_s(x) + \chi \nabla c(\rho_s)(x)|^2) \, \rho_s(x) \, dx \, ds \, . \tag{6.6}$$

Using integration by parts and $\Delta c(\rho) = 2\pi \, \rho$, any solution of (6.1) satisfies

$$\int_0^t I(\rho_s) ds + \int_0^t \int |\chi \nabla c(\rho_s)(x)|^2 \, \rho_s(x) \, dx \, ds \leq \, D_t(F) + \, 4\pi \, \chi \int \int_0^t \rho_s^2(x) \, dx ds \, .$$

$\int_0^t I(\rho_s) ds < +\infty$ and $\int_0^t \int |\chi \nabla c(\rho_s)(x)|^2 \, \rho_s(x) \, dx \, ds < +\infty$ follow from Lemma 3.1 (2) and some clever manipulations, see the end of the proof of Lemma 2.2 in [54]. Conversely if these two quantities are finite, $D_t(F) < +\infty$ is immediate. Remark that this can be rewritten using

$$\int_0^t \int |\chi \nabla c(\rho_s)(x)|^2 \, \rho_s(x) \, dx \, ds = \int_0^t \int |K * \rho_s|^2 \, \rho_s \, dx \, ds < +\infty \, .$$

At least formally one has $F(\rho_t) + D_t F(\rho) = F(\rho_0)$, provided $D_t F(\rho) < +\infty$. However a rigorous proof requires some additional regularity on $\rho_.$ (see e.g. [10] Lemma 2.2). This regularity is shown in [54], in particular Lemmas 2.6, 2.7, 2.8. This regularity holds on time intervals $[t_0, T]$ but for $t_0 > 0$. An accurate reading shows that assuming (6.5) is not used in these proofs. One can thus deduce for $t \geq t_0 > 0$, $F(\rho_t) + D_t F(\rho) = F(\rho_{t_0}) + D_{t_0} F(\rho)$. Step 2 in the proof of Theorem 1.4 p. 1176 of [54] together with Lemma 2.9 therein (lower semi continuity of the free energy) furnishes that (6.5) is actually satisfies as soon as $D_t F < +\infty$. The price to pay is that equality becomes an inequality in (6.5) because of lower semi-continuity.

Hence uniqueness holds once, in addition to (6.3),

$$\int_0^t \int (|\nabla \ln \rho_s(x) + \chi \nabla c(\rho_s)(x)|^2) \, \rho_s(x) \, dx \, ds < +\infty \, .$$

This will be of particular interest in the sequel. Notice that according to the results in subsection 3.1, as said in Lemma 2.4 in [54],

$$\rho_. \in \mathbb{L}^{p/(p-1)}(]0, T], \mathbb{L}^p(\mathbb{R}^2)) \quad \text{for all } 1 < p < +\infty. \tag{6.7}$$

Stronger regularity is known since, according to [54] Lemma 2.8, for all $\varepsilon > 0$, $\rho_. \in C_b^\infty([\varepsilon, T] \times \mathbb{R}^2)$.

When ρ_0 is bounded, this result is shown in to Theorem 1.3 in [21], for bounded solutions. $\diamondsuit$

Remark 6.5. Another notion of solution called "mild" solution is discussed in [4, 116]. Existence for general initial data, and uniqueness in the set of such solutions is obtained therein. A mild solution is a solution that can be expressed via Duhamel formula for the heat semi-group. In particular an a priori $\mathbb{L}^{\frac{4}{3}}$ bound is required. It is shown in [54] section 3 that the (unique) free energy solution is a mild solution. ♢

The assumptions in theorem 6.2 can be relaxed, when $\chi < 4$, the following is proved in the recent [65] (Theorem 2 therein), where the moment condition (6.8) is shown

Theorem 6.6. *If $\chi < 4$, for any initial probability measure μ_0, there exists a flow of probability measures $t \mapsto \mu_t$ defined for all $t > 0$ and solution of (6.2) . Furthermore, for all $T > 0$, and all $2 \geq \gamma > \frac{\chi}{2}$,*

$$\int_0^T \int\int |x-y|^{\gamma-2}\, \mu_s(dx)\, \mu_s(dy) \;\leq\; C(\gamma,\chi)\,(1+T)\,. \tag{6.8}$$

Unfortunately, no uniqueness result is known in this more general context. Notice that if ρ_0 satisfies the assumptions in Theorem 6.2, according to Remark 6.4 and Lemma 3.2 the previous integrability (6.8) holds for all $0 < \gamma < 2$.

Remark 6.7. In the sequel we will also be interested by the model with an added drift $b(x) = \nabla\tilde{U}(x) = \sum_{j=1}^2 U'(x^j)$ for some nice U we will call a confinement potential and the corresponding drift an advection term. As for the $2D$ vortex model we will assume that U'' is bounded. When U is convex, $-\langle x, \nabla\tilde{U}\rangle \leq 0$ inducing a true confinement.

Despite our tentatives, we did not find in the jungle of the literature on Keller-Segel models, the analogue of the results recalled before for $\alpha = 0$, except one case: a change of coordinates in self similar variables naturally introduces a quadratic confinement potential (see e.g [54] p.1164 formula (1.16) and what precedes). One can thus transfer the previous results to this situation.

We nevertheless claim that all that we recalled before is still true if we add a confinement as we defined before. The only way we found to see it is to rewrite the proofs in [10] and [54]. We shall not give all the details here, only explain how to modify some notations, and why the confinement does not perturb the heart of the problem.

First, we may look at the second moment

$$\int |x|^2 \rho_T dx = \int |x|^2 \rho_0 dx + (4-\chi)T - 2\int_0^T \int \langle x, \nabla\tilde{U}\rangle\, \rho_t\, dx\, dt$$

where the last term is non-positive. It follows that, as for the usual Keller-Segel model, existence of a global solution is only possible for $\chi \leq 4$. In this case one also sees that the second moment is finite at time T if it is finite at time 0.

In order to mimic the proofs in [10] it is enough to modify the free energy functional by adding $\frac{1}{2}\int \tilde{U}\,\rho\, dx$ i.e. replace χc by $\chi c + \tilde{U}$. We follow the numbering in [10]. Lemma 2.3 is still true, up to this modification, and since $\int f\,\tilde{U}\, dx \geq 0$ for a non-negative f, adding this term in the Hardy-Littlewood-Sobolev inequality (2.6) does not modify the lower bound, and Lemma 2.5 is unchanged. The approximation method in 2.5.3 is modified, replacing χc^ε by $\theta^\varepsilon = \chi c^\varepsilon + \tilde{U}$. The key lemma 2.11 is also unchanged, once one remarks that Δc^ε is replaced by $\Delta c^\varepsilon + \Delta\tilde{U}$

with a bounded $\Delta\tilde{U}$, so that for instance in the proof of (iv) p.14, one simply has to add a constant in the right hand side of the first inequality. We stop here, but the reader can easily follow line by line [10] and see that, in the worst case, only constant additional terms appear and do not modify the main statements, the most important results being Lemma 2.12, proposition 3.3 or Lemma 3.4, where for instance the dissipation (or production) of the free energy is a simple consequence of the non positivity of its time derivative, as in Lemma 2.3.

One can do exactly the same with [54]. The modification of χc does not introduce any difference in the proofs. For instance, Step 4 on p. 1177 uses the self similar coordinates, introducing a quadratic term $|x|^2/2$ in formula (2.27). This term is changed into $|x|^2/2+h(x)$ for an at most quadratic h and all the estimates starting from (2.29) are still true. $\diamondsuit$

From the point of view of the non linear PDE the situation is close to the case of the sub-Coulombic potential. However for the particle system, the situation is very different. Indeed one cannot use nor subsection 2.4 with the invariant measure candidate

$$G_N(dx) = \prod_{1\le i<j\le N} |x^i - x^j|^{-\frac{\chi}{N}}\, dx\,,$$

nor subsection 3.3 since $K \notin \mathbb{L}^2(\mathbb{R}^2)$. Nevertheless, Dirichlet forms theory is a powerful tool to get the following existence and uniqueness Theorem obtained in [38] (also see Theorem 7 in [63] with a different approach)

Theorem 6.8. *Assume that $b = 0$ and $K(x) = \chi\, \frac{x}{|x|^2}\, \mathbf{1}_{x\neq 0}$ for some $\chi > 0$.*
Let $M^N = \{x \in (\mathbb{R}^2)^N$ there exists at most one pair $i \neq j$ such that $x^i = x^j\}$. Then for $N \geq 4$ and $\chi < 4\,\frac{N-2}{N-1}$, there exists a unique (in distribution) non explosive solution of (1.2) starting from any $x \in M^N$. Moreover the process is strong Markov, lives in M^N and admits a symmetric σ-finite (in M^N) invariant (actually reversible) measure given by

$$G_N(dx) = \prod_{1\le i<j\le N} |x^i - x^j|^{-\frac{\chi}{N}}\, dx\,.$$

In addition, for all $T > 0$,

$$\sum_{i\neq j} \int_0^T \mathbf{1}_{X_t^{i,N}=X_t^{j,N}}\, dt \;=\; 0 \;a.s. \tag{6.9}$$

We are using here the notation G_N which is the one used in the recent [17]. The latter property (6.9) is Lemma 3.4 in [38].
Actually the limitation to $\chi < 4\,\frac{N-2}{N-1}$ prevents for multiple 2-collisions or k-collisions with $k \geq 3$. For $\chi > 4$ it is also shown in [38] that the solution explodes.
The method of [38] was recently pushed further in [64] in order to bypass the limitation to $\chi < 4\,\frac{N-2}{N-1}$. This time k-collisions are allowed up to a certain size k depending on χ and a very precise description of these collisions is made. Let us give the result extending our previous Theorem (due to the normalizing coefficients our χ is 2θ in [64])

Theorem 6.9. *In the situation of Theorem 6.8 the existence of a non explosive solution starting from any non collision point is ensured as soon as $N \geq 2$ and $\chi < 4$.*

Uniqueness in the sense of Hunt processes is also shown. Notice that uniqueness starting from a collision point in M^N shown in [38] is using the specific nature of this set (at most a 2-collision) via an explicit computation made in [63] Lemmata 19 and 20.

Another crucial result is that 2-collisions are not only allowed but always occur with positive probability as shown in [63] Proposition 4 (with an initial exchangeable condition) or in [38] subsubsection 2.5.2

Proposition 6.10. *For any solution of (1.2) in the Keller-Segel setting, and any $t > 0$,*

$$\mathbb{P}(\exists s \in [0,t], \exists i \neq j \; ; \; X_s^{i,N} = X_s^{j,N}) > 0 \, .$$

In particular a solution Q to (1.2) cannot be absolutely continuous w.r.t. P or to the Wiener measure with the same initial condition.

Remark 6.11. It is interesting to see that Keller and Segel themselves proposed in [81] a *microscopic* description of their aggregation model, which is nothing else (in a modern formulation) than the particle system (1.2). ♢

Remark 6.12. In all what precedes one can add a linear confinement potential as in Remark 6.7 simply using a Girsanov transform, since the additional drift is bounded. ♢

In the recent [17], the authors have studied the Liouville equation (recall (1.4)) associated to (1.2) for the Keller-Segel model, i.e.

$$\partial_t \mu_t^N = \sum_{i=1}^N \nabla_{x^i} . \left(\mu_t^N \, \frac{\chi}{N} \sum_{j=1}^N \frac{x^i - x^j}{|x^i - x^j|^2} \right) + \Delta_x \, \mu_t^N \tag{6.10}$$

where as before $0 = 0/0$. The authors introduce in their definition 2.1 the notion of (relative) entropy solution, as follows:

Definition 6.13. A solution (in the sense of Schwartz distributions) μ_t^N of (6.10) is said to be an entropy solution if for all $T > 0$,

$$\int_0^T I(\mu_t^N | G_N) \, dt \; \leq \; H(\mu_0^N | G_N) - H(\mu_T^N | G_N) \tag{6.11}$$

where as usual $H(\mu|\nu)$ and $I(\mu|\nu)$ denote the relative entropy (Kullback-Leibler information) and the relative Fisher information, i.e.

$$I(\mu|\nu) = \int \left| \nabla \ln \left(\frac{d\mu}{d\nu} \right) \right|^2 \, d\mu \, .$$

Since G_N is not bounded one has to be careful with these definitions. It is immediate that both hand sides in (6.11) are unchanged if we replace G_N by $c\, G_N$ for any nonnegative constant. Remark that one can always choose an appropriate c for the relative entropy H to be nonnegative as in the usual probabilistic situation, but c depends on ρ.

Actually the authors of [17] replace $\mathbb{R}^2$ by the 2 dimensional torus Π, and consider periodic solutions after periodizing the potential $\ln(|x|^2)$. The meaning and definition of the stochastic system (1.2) has to be reformulated. The existence of an entropy solution of (6.10) is shown in [17] Proposition 4.1.

Notice that, formally in $(R^2)^N$

$$\int_0^T I(\mu_t^N|G_N)\,dt = \int_0^T \int \sum_i \left(\left| \nabla_i \ln \rho_t^N + \frac{\chi}{N} \sum_{j\neq i} \frac{x^i - x^j}{|x^i - x^j|^2} \right|^2 \right) \rho_t^N(x)\,dx\,dt \tag{6.12}$$

so that (6.11) is similar to (6.5) in the context of the particle system. Recall that [17] works in a Π^N the N-tensor product of the $2D$ torus, so that the previous equality has to be modified.

Another point is concerned with a sentence at the top of p.4 of [17] i.e. "Of course any strong solution to (1.2) (in the probabilistic sense) would also yield an entropy solution to (6.10)." Since the notion of *strong solution in the probabilistic sense* is not defined in [17], the meaning of this sentence is only speculating. However, the proof of the existence of an entropy solution given in the Appendix of [17], is based on a ε regularization different from, but similar to the one used in [63] for proving the existence of a solution of the particle system. Since we know the (weak) uniqueness of a solution at the particles level, it indicates that, provided one can extend this result to our situation, $\rho_\cdot^{N,U}$ will be an entropy solution. Actually some key uniforms (in ε) estimates in the [17] proof are obtained via a large deviation estimate (Proposition 2.1 therein) inspired by [78]. If replacing the Lebesgue measure which is bounded on the torus by $e^{-\tilde{U}}dx$ on the whole space is presumably what has to be done to extend the argument in [17], we confess that we were not able to understand all the steps of the proof in sections 2 and 3 of [17].

We shall thus give a proof of a precise statement. The (potential) reader in a hurry can skip what follows up to the statement of Theorem 6.17.

Instead of working on the torus we will add a confining smooth potential U defined on $\mathbb{R}$, as for the 2D vortex model.

Next, we replace G_N by the measure $G_N^U(dx) = e^{-\tilde{U}(x)}\,G_N(dx)$. The existence and uniqueness of a G_N^U symmetric diffusion process

$$dX_t^{i,N,U} = \sqrt{2}\,dB_t^{i,N} - \frac{1}{M}\nabla\tilde{U}(X_t^{i,N,U}) - \frac{\chi}{N}\sum_{j\neq i} \frac{X_t^{i,N,U} - X_t^{j,N,U}}{|X_t^{i,N,U} - X_t^{j,N,U}|^2}\,dt \tag{6.13}$$

can be shown exactly as for Theorem 6.8 in the same range $\chi < 4(N-2)/(N-1)$. If $|\nabla\tilde{U}|$ is bounded, one can also use a Girsanov transformation.

Our goal is to use Remark 2.10. The main difficulty is that if N is large, G_N^U is *not* a bounded measure. To overcome this difficulty, first consider the approximate model $Q^{N,U,\varepsilon}$ where we replace $|x-y|^2$ by $\varepsilon + |x-y|^2$. This time the associated

$$G_N^{U,\varepsilon} = \prod_{1\le i<j\le N} (\varepsilon + |x^i - x^j|^2)^{-\frac{\chi}{2N}}\,e^{-\tilde{U}(x)}\,dx$$

is bounded with a normalization constant denoted by Z_ε. Applying (2.13) we thus have

$$H(\rho_0^N\,dx|Z_\varepsilon^{-1}\,G_N^{U,\varepsilon}) - H(\rho_T^{N,U,\varepsilon}dx|Z_\varepsilon^{-1}\,G_N^{U,\varepsilon}) \ge \tag{6.14}$$

$$\ge \int_0^T \int \left(\sum_i \left| \nabla_i \ln \rho_t^{N,U,\varepsilon} + \nabla\tilde{U} + \frac{\chi}{N}\sum_{j\neq i} \frac{x^i - x^j}{\varepsilon + |x^i - x^j|^2} \right|^2 \right) \rho_t^{N,U,\varepsilon}(x)\,dx\,dt\,.$$

The key is that the normalization constant, that already disappeared in the right hand side thanks to the gradient, also disappears in the left hand side which is equal to

$$\int \ln(\rho_0^N/G_N^{U,\varepsilon})\,\rho_0^N\,dx - \int \ln(\rho_T^{N,U,\varepsilon}/G_N^{U,\varepsilon})\,\rho_T^{N,U,\varepsilon}\,dx\,.$$

It turns out that the methodology of [63] precisely consists in showing that $Q^{N,U,\varepsilon}$ weakly converges to $Q^{N,U}$ on $[0,S_l]$ where S_l denotes the first time where the minimal sum of the distances between three particles becomes less than $1/l$ for large l i.e.

$$D_l^N = \{x \in (\mathbb{R}^2)^{\otimes N};\ \text{for all distinct indices } i,j,k\ ,\ |x^i - x^j| + |x^j - x^k| + |x^k - x^i| > 1/l\} \tag{6.15}$$

(see the proof of Lemma 14 in [63]) using the same approximation in ε we are using. It is not difficult to see that the same is true for the confined system.

Introduce now

$$A_l^N = \{x \in (\mathbb{R}^2)^{\otimes N};\ \min_{i\neq j} |x^i - x^j| > 1/l\}\,. \tag{6.16}$$

Notice that on this set $\rho_t^{N,U,\varepsilon}$ and $\rho_t^{N,U}$ are smooth thanks to ellipticity results (probabilists can call upon Malliavin calculus).

The accurate reader will remark that [63] always assumes that the initial condition is exchangeable. If many arguments in [63] can be modified in order to bypass this assumption (see e.g the proof of Lemma 6.14 below), some of them are really using the exchangeable setting. Notice that if $\rho_0^N = \kappa_0^N\,\nu_0^N$ where ν_0^N is exchangeable (for instance $\nu_0^N = \eta_0^{\otimes N}$ for some probability η_0 on $\mathbb{R}^2$) and κ_0^N is continuous and bounded, the above weak convergence is still true.

Actually we can say much more, namely that

$$||\rho_t^{N,U,\varepsilon}||_{C_b^j(A_l^N\cap B(0,R))} \leq C(j,V,l)\,t^{-\alpha(j)} \tag{6.17}$$

for any $j \in \mathbb{N}^*$ and some $\alpha(j) > 0$ where $B(0,R)$ is the euclidean ball of radius R. What is important here is that the previous bound is uniform in ε, because all the coefficients and their derivatives are bounded on $A_l^N \cap B(0,R)$ uniformly in ε. A similar bound is true for $\rho_t^{N,U}$. Once again, aficionados of Malliavin calculus will find a proof in [24] Theorem 1.15 and its proof, and the above statement in [25] Theorem 1.5.(i), in a more general hypoelliptic framework.

It is then standard to show that $\rho_t^{N,U,\varepsilon}$ and $\nabla\rho_t^{N,U,\varepsilon}$ weakly converge to $\rho_t^{N,U}$ and $\nabla\rho_t^{N,U}$ in $\mathbb{L}^2(A_l^N \cap B(0,R))$ so that one easily gets, for $s > 0$, both

$$\int_s^T \int_{A_l^N\cap B(0,R)} \left(\left|\frac{\chi}{N}\sum_{j\neq i}\frac{x^i - x^j}{\varepsilon + |x^i - x^j|^2}\right|^2\right)\rho_t^{N,U,\varepsilon}(x)\,dx\,dt \to$$

$$\to \int_s^T \int_{A_l^N\cap B(0,R)} \left(\left|\frac{\chi}{N}\sum_{j\neq i}\frac{x^i - x^j}{|x^i - x^j|^2}\right|^2\right)\rho_t^{N,U}(x)\,dx\,dt\,,$$

and

$$\int_s^T \int_{A_l^N\cap B(0,R)} \left\langle \nabla_i\rho_t^{N,U,\varepsilon}, \frac{\chi}{N}\sum_{j\neq i}\frac{x^i - x^j}{\varepsilon + |x^i - x^j|^2}\right\rangle dx\,dt \to$$

$$\to \int_s^T \int_{A_l^N \cap B(0,R)} \left\langle \nabla_i \rho_t^{N,U}, \frac{\chi}{N} \sum_{j\neq i} \frac{x^i - x^j}{|x^i - x^j|^2} \right\rangle dx\, dt\,,$$

as $\varepsilon \to 0$.

In addition, at least for l and R large enough,

$$\sup_\varepsilon \int_s^T \int_{A_l^N \cap B(0,R)} |\nabla \ln(\rho_t^{N,U,\varepsilon})|^2 \rho_t^{N,U,\varepsilon} dxdt < +\infty\,.$$

This bound follows from the fact that

$$\inf_\varepsilon \inf_{s\le t\le T, y\in A_l^N \cap B(0,R)} \rho_t^{N,U,\varepsilon} \;=\; c(l,s,T,R,U,N) \;>0\,. \tag{6.18}$$

Here is a proof with a probabilistic flavour. First, since the initial measure has a density, one may find $l_0 > 0$ and R_0 such that $\rho_0^N(A_{l_0}^N \cap B(0,R_0)) > 0$. If $l > 2l_0$ and $R > 2R_0$ we may consider the processes killed when they exit $A_{2l_0}^N \cap B(0,2R_0)$. Since their laws are equivalent to the one of a similarly killed Brownian motion (with variance $2t$), with densities and inverse densities bounded in all the $\mathbb{L}^p$'s uniformly in ε, all these laws are equivalent with densities and inverse densities bounded in all the $\mathbb{L}^p$'s uniformly in ε. The uniform in ε lower bound (6.18) follows from a similar uniform lower bound for the densities of the killed processes, which itself follows from the previous argument once a strictly positive lower bound is obtained for a given ε_0. The latter is standard see e.g [26] p.612-613.

Consider random variables $Z_t^{N,U,\varepsilon}$ supported by $A_l^N \cap B(0,R)$ with densities proportional to $\rho_t^{N,U,\varepsilon}$ in this set. It is easily seen that they converge in distribution to a random variable $Z_t^{N,U}$ with density proportional to $\rho_t^{N,U}$ and that the normalizing constants in ε also converge to the normalizing constant for the limiting distribution. According to Proposition 13.2 in [11] again and to Fatou's lemma, we deduce that

$$\liminf_\varepsilon \int_s^T \int_{A_l^N \cap B(0,R)} |\nabla \ln(\rho_t^{N,U,\varepsilon})|^2 \rho_t^{N,U,\varepsilon} dxdt \ge$$

$$\ge \int_s^T \int_{A_l^N \cap B(0,R)} |\nabla \ln(\rho_t^{N,U})|^2 \rho_t^{N,U} dxdt\,.$$

We have thus shown

$$\begin{aligned} \liminf_\varepsilon \int_s^T \int_{A_l^N \cap B(0,R)} \left(\sum_i \left| \nabla_i \ln \rho_t^{N,U,\varepsilon} + \frac{\chi}{N} \sum_{j\neq i} \frac{x^i - x^j}{\varepsilon + |x^i - x^j|^2} \right|^2 \right) \rho_t^{N,U,\varepsilon}(x)\, dx\, dt \ge \\ \ge \int_s^T \int_{A_l^N \cap B(0,R)} \left(\sum_i \left| \nabla_i \ln \rho_t^{N,U} + \frac{\chi}{N} \sum_{j\neq i} \frac{x^i - x^j}{|x^i - x^j|^2} \right|^2 \right) \rho_t^{N,U}(x)\, dx\, dt\,, \end{aligned} \tag{6.19}$$

and finally

$$\liminf_\varepsilon \int_0^T \int \left(\sum_i \left| \nabla_i \ln \rho_t^{N,U,\varepsilon} + \frac{\chi}{N} \sum_{j\neq i} \frac{x^i - x^j}{\varepsilon + |x^i - x^j|^2} \right|^2 \right) \rho_t^{N,U,\varepsilon}(x)\, dx\, dt \ge$$

$$\liminf_\varepsilon \int_s^T \int_{A_l^N\cap B(0,R)} \left(\sum_i \left| \nabla_i \ln \rho_t^{N,U,\varepsilon} + \frac{\chi}{N} \sum_{j\neq i} \frac{x^i - x^j}{\varepsilon + |x^i - x^j|^2} \right|^2 \right) \rho_t^{N,U,\varepsilon}(x)\, dx\, dt \geq$$

$$\geq \int_s^T \int_{A_l^N\cap B(0,R)} \left(\sum_i \left| \nabla_i \ln \rho_t^{N,U} + \frac{\chi}{N} \sum_{j\neq i} \frac{x^i - x^j}{|x^i - x^j|^2} \right|^2 \right) \rho_t^{N,U}(x)\, dx\, dt\,.$$

Now we have to take care of the potential $\nabla\tilde{U}$ in (6.14). We will argue as for the $2D$ vortex case, and map this term in the left hand side, using $|a+b|^2 \geq \frac{1}{2}|a|^2 - |b|^2$.

It remains to take the increasing limit as l and R go to infinity, since $A_l^N \cap B(0,R)$ grows to the whole space, and then as s goes to 0. We first have the existence of the density $\rho_t^{N,U}$ which is almost everywhere defined as well as $\nabla\rho_t^{N,U}$. Next we obtain

$$\liminf_\varepsilon \left(\int \ln(\rho_0^N/G_N^{U,\varepsilon})\, \rho_0^N dx + \int_0^T \int |\nabla\tilde{U}|^2\, \rho_t^{N,U,\varepsilon} dx\, dt \right) \geq \liminf_\varepsilon \int \ln(\rho_T^{N,U,\varepsilon}/G_N^{U,\varepsilon})\, \rho_T^{N,U,\varepsilon} dx + \tag{6.20}$$

$$+ \frac{1}{2} \int_0^T \int \left(\sum_i \left| \nabla_i \ln \rho_t^{N,U} + \frac{\chi}{N} \sum_{j\neq i} \frac{x^i - x^j}{|x^i - x^j|^2} \right|^2 \right) \rho_t^{N,U}(x)\, dx\, dt\,.$$

It remains to look at the left hand side and the first term on the right hand side of (6.20). Notice that the first term can be rewritten

$$H(\rho_0^N dx | e^{-\tilde{U}} dx) + \int \rho_0^N \frac{\chi}{2N} \sum_{i<j} \ln(\varepsilon + |x^i - x^j|^2)\, dx$$

and similarly replacing ρ_0^N by $\rho_T^{N,U,\varepsilon}$. Also note that we may always assume that $\tilde{U}$ is normalized so that $e^{-\tilde{U}}$ is a density of probability. We thus have to look at five terms

(i) First, $H(\rho_0^N dx | e^{-\tilde{U}} dx)$ does not depend on ε and will be assumed to be finite.

(ii) In second place, for $\int \rho_0^N \frac{\chi}{2N} \sum_{i<j} \ln(\varepsilon + |x^i - x^j|^2)\, dx$ we will assume that

$$\int \rho_0^N \sum_{i\neq j} |\ln(|x^i - x^j|)|\, dx < +\infty$$

so that we may use Lebesgue's theorem and get that the sum of the first two terms goes to $\int \rho_0^N \ln(G_N^U) dx$ as $\varepsilon \to 0$.

(iii) The third term

$$\liminf_\varepsilon H(\rho_T^{N,U,\varepsilon} dx | e^{-\tilde{U}} dx) \geq H(\rho_T^{N,U} | e^{-\tilde{U}})$$

can be shown using lower semi-continuity in the weak topology (it is important here to deal with probability measures) and a similar restriction to A_l^N first. A more direct proof will be given below.

(iv) The term

$$\int \rho_T^{N,U,\varepsilon} \frac{\chi}{2N} \sum_{i<j} \ln(\varepsilon + |x^i - x^j|^2)\, dx$$

which is not as simple.

(v) The energy $\int_0^T \int |\nabla\tilde{U}|^2\, \rho_t^{N,U,\varepsilon} dx\, dt$.

To get the result and for later purpose, we shall use the following Lemma

Lemma 6.14. *Let $X^N_\cdot$ be the particle system of the Keller-Segel model introduced in Theorem 6.8 and Theorem 6.9 with or without an additional the confinement potential. For all $0 \leq \varepsilon \leq 1$ (1 is arbitrary and we emphasize that $\varepsilon = 0$ is allowed), it holds*

(1) *For all $t > 0$, $\sum_{j=1}^N \mathbb{E}(|X_t^{j,N,\varepsilon}|^2) \leq c(T,a)(\sum_{j=1}^N \mathbb{E}(|X_0^{j,N}|^2) + NT)$, where a is the Lipschitz constant of ∇U.*

(2) *For all $0 < \gamma < 2$,*

$$\int_0^T \sum_{i\neq j} \mathbb{E}\left(\frac{1}{|X_t^{j,N,\varepsilon} - X_t^{i,N,\varepsilon}|^\gamma}\right) dt \leq CN\,(\sum_{j=1}^N \mathbb{E}(|X_0^{j,N}|^2) + N)\,,$$

where C is a constant that only depends on T, a (the Lipschitz constant of ∇U), γ and χ.

This Lemma is partly contained in [112] Lemma 16 and Proposition 10 in the case $U = 0$ and with $\gamma < 2 - (\chi/2)$ for (2). The proofs in [112] are written under the assumption that the initial distribution is exchangeable and only concern the process $X^N_\cdot$, not the approximations. We shall give here a complete proof for extending the range of allowed γ's and taking care of the dependence in ε. The Lemma also has an important consequence

Corollary 6.15. *$Q^{N,U,\varepsilon}$ weakly converges to $Q^{N,U}$ as $\varepsilon \to 0$.*

The Corollary extends Theorem 6 in [63] obtained for $\chi < 1$. The key is that since (2) holds for $\gamma > 1$ the family $Q^{N,U,\varepsilon}$ is tight. A proof of this fact will be given in subsection 8.1 in the framework of large N instead of small ε. One can also directly follow the proof of Theorem 6 in [63] (a small improvement is contained in our derivation in subsection 8.1). It is then standard to show (as in [63]) that any weak limit solves the martingale problem with $\varepsilon = 0$. One can also use the weak convergence up to time S_l we have already recalled, so that any weak limit coincides with a solution of (1.2) built in [63]. The result follows from the uniqueness of the latter.
We turn to the proof of the Lemma.

Proof. of the lemma. Proof of (1). Using Ito's formula and the usual care (localizing first with stopping times), one obtains

$$\begin{aligned}\sum_j |X_t^{j,N,\varepsilon}|^2 &= \sum_j |X_0^{j,N}|^2\, M_t + 4Nt - \sum_j \int_0^t 2\langle X_s^{j,N,\varepsilon}, \nabla U(X_s^{j,N,\varepsilon})\rangle\, ds \\ &\quad - \frac{2\chi}{N}\sum_{i\neq j}\int_0^t \frac{\langle X_s^{j,N,\varepsilon}, X_s^{j,N,\varepsilon} - X_s^{i,N,\varepsilon}\rangle}{\varepsilon + |X_s^{j,N,\varepsilon} - X_s^{i,N,\varepsilon}|^2}\, ds\,,\end{aligned}$$

where $M_\cdot$ is a local martingale. Exchanging again i and j in the final term, we see that it is equal to

$$\frac{1}{2}\sum_{i\neq j} \frac{\langle X_s^{j,N,\varepsilon} - X_s^{i,N,\varepsilon}, X_s^{j,N,\varepsilon} - X_s^{i,N,\varepsilon}\rangle}{\varepsilon + |X_s^{j,N,\varepsilon} - X_s^{i,N,\varepsilon}|^2} \leq N(N-1)/2$$

and in addition is non-negative, so that finally, using again localization for some stopping time in order to control the martingale term

$$\sum_j \mathbb{E}(|X_t^{j,N,\varepsilon}|^2) \leq \sum_j \mathbb{E}(|X_0^{j,N}|^2) + CNT + a \int_0^t \sum_j \mathbb{E}(|X_s^{j,N,\varepsilon}|^2)\, ds. \tag{6.21}$$

Here we have used $|\langle x, \nabla \tilde{V}(x)\rangle| \leq c + a|x|^2$ for some constant c. We may conclude thanks to Gronwall's lemma. When U is convex, $\langle \nabla \tilde{U}(y), y\rangle \geq 0$ and the bound does not depend on U.

For the proof of (2) we consider the function

$$G_{\alpha,\eta,c}(x) = \sum_{i\neq j} \frac{(\eta + |x^i - x^j|^2)^\alpha}{c + (\eta + |x^i - x^j|^2)^\alpha}$$

for $\eta > 0$, $c > 0$ and $0 < \alpha < 1$. Again we may use Ito's formula to get

$$G_{\alpha,\eta,c}(X_t^{N,\varepsilon}) = G_{\alpha,\eta,c}(X_0^{N,\varepsilon}) \,+\, M_t \,-\, S_t^1 \,-\, S_t^2 \,+\, S_t^3$$

where $M_\cdot$ is a martingale. We will describe the three other terms. Denote

$$f(u) = \frac{(\eta + |u|^2)^\alpha}{c + (\eta + |u|^2)^\alpha}$$

for $u \in \mathbb{R}^2$. Then

$$\nabla f(u) = 2c\alpha\, \frac{(\eta + |u|^2)^{\alpha-1}}{(c + (\eta + |u|^2)^\alpha)^2}\, u$$

and

$$\begin{aligned}\Delta f(u) &= \frac{4c\alpha\,(\eta + |u|^2)^{\alpha-1}}{(c + (\eta + |u|^2)^\alpha)^2} + \frac{4c\alpha(\alpha-1)|u|^2(\eta + |u|^2)^{\alpha-2}}{(c + (\eta + |u|^2)^\alpha)^2} - \frac{8c\alpha^2|u|^2\,(\eta + |u|^2)^{2\alpha-2}}{(c + (\eta + |u|^2)^\alpha)^3} \\ &= \frac{4c\alpha\,(\eta + |u|^2)^{\alpha-2}\,(c\eta + c\alpha|u|^2 + (\eta + |u|^2)^\alpha(\eta - \alpha|u|^2)}{(c + (\eta + |u|^2)^\alpha)^3}\,.\end{aligned}$$

We thus have

$$S_t^1 = \int_0^t \sum_{i\neq j} \nabla_i f(X_s^{i,N,\varepsilon} - X_s^{j,N,\varepsilon})\, \langle X_s^{i,N,\varepsilon} - X_s^{j,N,\varepsilon}, \nabla U(X_s^{i,N,\varepsilon}) - \nabla U(X_s^{j,N,\varepsilon})\rangle\, ds$$

which satisfies

$$|S_t^1| \leq C(\alpha, a, c) \sum_{i\neq j} \int_0^t (\eta + |X_s^{i,N,\varepsilon} - X_s^{j,N,\varepsilon}|^2)^\alpha\, ds$$

since ∇U is a-Lipschitz, where $C(\alpha, a, c)$ does not depend on ε.
Similarly

$$S_t^2 = \frac{2c\alpha\,\chi}{N} \int_0^t \sum_{i\neq j} \frac{(\eta + |X_s^{i,N,\varepsilon} - X_s^{j,N,\varepsilon}|^2)^{\alpha-1}}{(c + (\eta + |X_s^{i,N,\varepsilon} - X_s^{j,N,\varepsilon}|^2)^\alpha)^2} \left(\sum_k (A_{i,k} - A_{j,k})\right) ds$$

where

$$A_{i,k} = \frac{\langle X_s^{i,N,\varepsilon} - X_s^{j,N,\varepsilon}, X_s^{i,N,\varepsilon} - X_s^{k,N,\varepsilon}\rangle}{\varepsilon + |X_s^{i,N,\varepsilon} - X_s^{k,N,\varepsilon}|^2}\,.$$

Denote

$$u = X_s^{i,N,\varepsilon} - X_s^{j,N,\varepsilon}\,,\; v = X_s^{j,N,\varepsilon} - X_s^{k,N,\varepsilon}\,,\; w = X_s^{k,N,\varepsilon} - X_s^{i,N,\varepsilon}\,.$$

Each term of the sum under the integral is written as

$$-\,(\eta+|u|^2)^{\alpha-1}\,(c+(\eta+|u|^2)^\alpha)^{-2}\,\left\langle u, \left(\frac{w}{\varepsilon+|w|^2}\,+\,\frac{v}{\varepsilon+|v|^2}\right)\right\rangle$$

so that exchanging the role of the indices one gets

$$\begin{aligned} S_t^2 \;=\; & -\frac{2\alpha\chi}{3N}\int_0^t \sum_{i,j,k} \\ & (\eta+|u|^2)^{\alpha-1}\,(c+(\eta+|u|^2)^\alpha)^{-2}\,\left\langle u, \left(\frac{w}{\varepsilon+|w|^2}\,+\,\frac{v}{\varepsilon+|v|^2}\right)\right\rangle \\ & +(\eta+|v|^2)^{\alpha-1}\,(c+(\eta+|v|^2)^\alpha)^{-2}\,\left\langle v, \left(\frac{u}{\varepsilon+|u|^2}\,+\,\frac{w}{\varepsilon+|w|^2}\right)\right\rangle \\ & +(\eta+|w|^2)^{\alpha-1}\,(c+(\eta+|w|^2)^\alpha)^{-2}\,\left\langle w, \left(\frac{v}{\varepsilon+|v|^2}\,+\,\frac{u}{\varepsilon+|u|^2}\right)\right\rangle \\ & ds\,. \end{aligned}$$

If $0<\alpha<1$, $r\mapsto(\eta+r^2)^{\alpha-1}\,(c+(\eta+r^2)^\alpha)^{-2}$ is non increasing on $\mathbb{R}^+$, as well as $r\mapsto(\varepsilon+r)^{-2}$. According to Lemma 9 in [112], the sum of the three terms above is larger than

$$\begin{aligned} & -\quad (\eta+|u|^2)^{\alpha-1}\,(c+(\eta+|u|^2)^\alpha)^{-2}\,\frac{|u|^2}{\varepsilon+|u|^2} \\ & -\quad (\eta+|v|^2)^{\alpha-1}\,(c+(\eta+|v|^2)^\alpha)^{-2}\,\frac{|v|^2}{\varepsilon+|v|^2} \\ & -\quad (\eta+|w|^2)^{\alpha-1}\,(c+(\eta+|w|^2)^\alpha)^{-2}\,\frac{|w|^2}{\varepsilon+|w|^2} \end{aligned}$$

so that summing up over (i,j,k) (recall that in the sum we only consider the cases where the three indices are different), we finally get, since $\alpha>0$,

$$\begin{aligned} S_t^2 \;&\le\; \frac{2\alpha\,\chi\,(N-2)}{N}\int_0^t \sum_{i\neq j} \frac{(\eta+|X_s^{i,N,\varepsilon}-X_s^{j,N,\varepsilon}|^2)^{\alpha-1}}{(c+(\eta+|X_s^{i,N,\varepsilon}-X_s^{j,N,\varepsilon}|^2)^\alpha)^2}\,\frac{|X_s^{i,N,\varepsilon}-X_s^{j,N,\varepsilon}|^2}{(\varepsilon+|X_s^{i,N,\varepsilon}-X_s^{j,N,\varepsilon}|^2)}\,ds \\ &\le\; 2\alpha\,\chi\int_0^t \sum_{i\neq j} \frac{(\eta+|X_s^{i,N,\varepsilon}-X_s^{j,N,\varepsilon}|^2)^{\alpha-1}}{(c+(\eta+|X_s^{i,N,\varepsilon}-X_s^{j,N,\varepsilon}|^2)^\alpha)^2}\,ds\,. \end{aligned} \tag{6.22}$$

Finally S_t^3 is given by Ito's compensation term (second order term), i.e.

$$S_t^3 = \int_0^t \sum_{i\neq j} \Delta f(X_s^{i,N,\varepsilon}-X_s^{j,N,\varepsilon})\,ds\,. \tag{6.23}$$

It remains to look at $S_t^3-S_t^2$. It holds

$$S_t^3-S_t^2 \;\ge\; 2\alpha\int_0^t \sum_{i\neq j} \frac{(\eta+|X_s^{i,N,\varepsilon}-X_s^{j,N,\varepsilon}|^2)^{\alpha-2}}{(c+(\eta+|X_s^{i,N,\varepsilon}-X_s^{j,N,\varepsilon}|^2)^\alpha)^3}\,H(X_s^{i,N,\varepsilon}-X_s^{j,N,\varepsilon})\,ds\,,$$

where

$$\begin{aligned} H(u) &= 2c(c\eta + c\alpha|u|^2 + (\eta+|u|^2)^\alpha(\eta - \alpha|u|^2)) - \chi(\eta+|u|^2)(c + (\eta+|u|^2)^\alpha) \\ &= c\eta(2c-\chi) + 2c^2\alpha|u|^2 + (\eta+|u|^2)^\alpha(\eta(2c-\chi) - |u|^2(2c\alpha+\chi)) \\ &\geq |u|^2\left(2c^2\alpha - (\eta+|u|^2)^\alpha(2c\alpha+\chi)\right) \end{aligned}$$

as soon as $2c \geq \chi$ in particular if $c > 2$. For $|u| \geq 1$, one immediately sees that $|H(u)| \leq C$ where C only depends on c as soon as $0 \leq \eta < 1$ (recall that $0 < \alpha < 1$ and $\chi < 4$). Accordingly, the part of $S_t^3 - S_t^2$ corresponding to $|X_s^{i,N,\varepsilon} - X_s^{j,N,\varepsilon}| \geq 1$ is bounded by a constant. The previous minoration is thus only interesting for $|u| \leq 1$ in which case

$$H(u) \geq (2c\alpha(c-2) - \chi)|u|^2 \geq |u|^2$$

provided $c \geq 4$ and

$$c \geq \frac{1+\chi}{4\,\alpha}\,.$$

Gathering all what precedes and taking expectations, we deduce that for such a c,

$$\int_0^t \sum_{i\neq j} \mathbb{E}((\eta + |X_s^{i,N,\varepsilon} - X_s^{j,N,\varepsilon}|^2)^{\alpha-2}\,|X_s^{i,N,\varepsilon} - X_s^{j,N,\varepsilon}|^2)\,ds \leq C$$

where C only depends on t, a and c (itself depending on α), and linearly depends on $\mathbb{E}(G_{\alpha,\eta}(X_u^{N,\varepsilon}))$ for $0 \leq u \leq t$. For the latter we may use the first part of the Lemma as well as

$$|x-y|^{2\alpha} \leq 1 + |x-y|^2 \leq 1 + 2|x|^2 + 2|y|^2$$

so that for $\eta \leq 1$, C does not depend on η. It remains to let η go to 0 in order to obtain the integrability of

$$\mathbb{E}(|X_s^{i,N,\varepsilon} - X_s^{j,N,\varepsilon}|^2)^{\alpha-1}$$

as expected with $\gamma = 2 - \alpha$. □

Remark 6.16. The bounding constant C depends on c^3 i.e. on α^{-3}, hence explodes as α goes to 0, i.e. $\gamma \to 2$, in accordance with the fact that for $\gamma = 2$ the integral in (2) is infinite. This part of the Lemma is much stronger than the similar bounds in [63] or [112] (where $\gamma < 2 - \chi/2$, though its proof is inspired by the one in [112]. It will be crucial in the sequel. ♢

Once the Lemma is proved it should easily furnish the proof of the desired convergence. First, (iii) is a consequence of the Corollary. Second, remark that for all $0 \leq \varepsilon$,

$$\ln(\varepsilon + |x^i - x^j|) \geq \ln(|x^i - x^j|)\,.$$

For $|x^i - x^j| \geq 1$ one can use a continuous approximation of $\mathbf{1}_{|u|\geq 1}$, Lemma 6.14 (1) and Lebesgue theorem, since $\ln(u) \leq u$ for $u \geq 1$. For $|x^i - x^j| \leq 1$, it holds

$$|\ln(|x^i - x^j|)| \leq c(\gamma)\,|x^i - x^j|^{-\gamma}$$

for all $\gamma > 0$ and according to Lemma 6.14 (2) we may find some $\gamma_0 > 0$ such that the latter is integrable, uniformly w.r.t. ε.

Unfortunately, this integrability property holds w.r.t. space and time so that we cannot conclude at integrability for a given fixed T in (iv). We thus have first to come back to equation (6.14) integrating w.r.t. T up to some new time S. It thus holds

$$S\left(H(\rho_0^N dx|e^{-\tilde U})+\frac{\chi}{N}\int\rho_0^N\sum_{i<j}|\ln(\varepsilon+|x^i-x^j|^2)|dx\right)+ \tag{6.24}$$

$$+\frac{\chi}{2N}\int_0^S\int\rho_T^{N,U,\varepsilon}\sum_{i<j}|\ln(\varepsilon+|x^i-x^j|^2)|dxdT\ \geq\ \int_0^S H(\rho_T^{N,U,\varepsilon}dx|e^{-\tilde U})\,dT+$$

$$+\int_0^S\int_0^T\int\left(\sum_i\left|\nabla_i\ln\rho_t^{N,U,\varepsilon}+\nabla\tilde U\ +\frac{\chi}{N}\sum_{j\neq i}\frac{x^i-x^j}{\varepsilon+|x^i-x^j|^2}\right|^2\right)\rho_t^{N,U,\varepsilon}(x)\,dx\,dt\,dT\,.$$

Using the positivity of the square, hence the fact that the integral w.r.t. T is non-decreasing, we can bound the final term from below by

$$\frac{S}{2}\int_0^{S/2}\int\left(\sum_i\left|\nabla_i\ln\rho_t^{N,U,\varepsilon}+\nabla\tilde U\ +\frac{\chi}{N}\sum_{j\neq i}\frac{x^i-x^j}{\varepsilon+|x^i-x^j|^2}\right|^2\right)\rho_t^{N,U,\varepsilon}(x)\,dx\,dt\,.$$

Since the left hand side of (6.24) is bounded, independently of ε, we may thus conclude as before, using lower semi-continuity and Fatou's lemma for the first term on the right hand side.

We have thus shown the following result

Theorem 6.17. *When adding a confining potential such that ∇U is Lipschitz, the marginals flow $\rho_\cdot^N$ of the unique solution of Theorem 6.8 is almost an entropy solution in the sense of (6.11) of the Liouville equation (6.10) as soon as the initial condition satisfies*

$$\int\rho_0^{N,U}\,|\ln\rho_0^{N,U}|\,dx<+\infty\,,\quad \sup_{i\neq j}\int\rho_0^{N,U}\,|\ln(|x^i-x^j|)|\,dx<+\infty\quad \textit{and}\quad \int|x|^2\,\rho_0^{N,U}\,dx<+\infty\,,$$

and $\rho_0^{N,U}=\kappa_0^N\,\nu_0^N$ where ν_0^N is exchangeable (for instance $\nu_0^N=\eta_0^{\otimes N}$ for some probability η_0 on $\mathbb{R}^2$) and κ_0^N is continuous and bounded.

Almost means that the following inequality is true, for some constant A depending on N,T,U,

$$A+2\left(H(\rho_0^{N,U}dx|e^{-\tilde U})+\frac{\chi}{N}\int\rho_0^{N,U}\sum_{i<j}|\ln(|x^i-x^j|)|dx\right)\geq\frac{1}{T}\int_0^T H(\rho_t^{N,U}dx|e^{-\tilde U})\,dt+$$

$$+\frac{1}{4}\int_0^T\int\left(\sum_i\left|\nabla_i\ln\rho_t^{N,U}+\frac{\chi}{N}\sum_{j\neq i}\frac{x^i-x^j}{|x^i-x^j|^2}\right|^2\right)\rho_t^{N,U}(x)\,dx\,dt$$

This partially confirms the [17] prediction when a confining potential is added.

Remark 6.18. Recall that, contrary to the solution of the non linear PDE, the density ρ_t^N, which exists, cannot satisfy $\int_0^T I(\rho_t^N)dt<+\infty$. Indeed, otherwise, the interaction drift would be square integrable, according to the previous Theorem, and Theorem 2.7 would tell us that the law of the particle system has finite relative entropy, hence is absolutely

continuous w.r.t. a probability P equivalent to the Wiener measure. This is impossible since we know that 2-collisions occur with a strictly positive probability.

This shows that some (mysterious) cancellations have to appear in (6.12). As for the non linear PDE one should, formally, develop

$$\int \sum_i \left| \nabla_i \ln \rho_t^{N,U} + \frac{\chi}{N} \sum_{j\neq i} \frac{x^i - x^j}{|x^i - x^j|^2} \right|^2 \rho_t^{N,U} dx = I(\rho_t^{N,U})$$

$$+ \int \sum_i \left| \frac{\chi}{N} \sum_{j\neq i} \frac{x^i - x^j}{|x^i - x^j|^2} \right|^2 \rho_t^{N,U} dx + \int \sum_i \left\langle \rho_t^{N,U}, \frac{2\chi}{N} \sum_{j\neq i} \frac{x^i - x^j}{|x^i - x^j|^2} \right\rangle dx$$

so that integrating by parts the scalar product

$$\begin{aligned} \int \sum_i \left\langle \nabla_i \rho_t^{N,U}, \frac{2\chi}{N} \sum_{j\neq i} \frac{x^i - x^j}{|x^i - x^j|^2} \right\rangle dx &= - \int \sum_i \rho_t^{N,U} \frac{4\pi\chi}{N} \sum_{j\neq i} \delta_{x^i = x^j}(dx) \\ &= - \frac{4\pi\chi}{N} \sum_{i\neq j} \int \rho_t^{N,U}(x_{i,j}) d\tilde{x}_{i,j} \end{aligned}$$

where $\tilde{x}_{i,j} = (x^1, ..., x^{j-1}, x^{j+1}, ..., x^N)$ and $x_{i,j} = (x^1, ..., x^{j-1}, x^i, x^{j+1}, ..., x^N)$. Since the squared terms are infinite, the latter has also to be infinite. ♢

Remark 6.19. The assumption $\int |\ln(|x^i - x^j|)| \rho_0^{N,U} dx < +\infty$ is satisfied, according to the proof of Lemma 3.2, as soon as $\tilde{\rho}_0^{i,j,N,U} := \rho_0^{N,U} \circ (x^i, x^j)^{-1} \in \mathbb{L}^q(\mathbb{R}^2)$ for some $q > 1$. Actually since $z \mapsto |\ln(|z|)|$ belongs to the Orlicz space $\mathbb{L}_\Phi(B(0,1))$, where $B(0,1)$ is the unit ball of $\mathbb{R}^2$ and $\Phi(u) = e^{|u|} - 1$, Orlicz-Hölder inequality ensures that $\int |\ln(|x^i - x^j|)| \rho_0^{N,U} dx < +\infty$ as soon as $\int \rho_0^{N,U} |\ln(\rho_0^{N,U})| dx < +\infty$, i.e. the second assumption on $\rho_0^{N,U}$ is satisfied as soon as the first one is satisfied.

Similarly, it is easy to see that, if ν and $h\nu$ are probability measures,

$$\int h \, |\ln h| d\nu \leq \int h \ln h \, d\nu + 2e^{-1} .$$

It follows that for a density of probability ρ, and U as before, assuming that $e^{-\tilde{U}}$ is normalized as a density of probability,

$$\int \rho \, |\ln \rho| dx \leq \int \rho \, |\ln(\rho/e^{-\tilde{U}})| dx + \int |\tilde{U}| \rho dx \leq \int \rho \ln \rho \, dx + 2 \int |\tilde{U}| \rho dx \, + \, 2e^{-1} .$$

It means that the first assumption on $\rho_0^{N,U}$ in Theorem 6.17 is satisfied as soon as the third one is satisfied and $\int \rho_0^{N,U} \ln \rho_0^{N,U} dx < +\infty$.

Since U has at most a quadratic growth, the same argument, together with the inequality in Theorem and Lemma 6.14 (1), shows that

$$\rho_\cdot^{N,U} \, |\ln \rho_\cdot^{N,U}| \in \mathbb{L}^1([0,T] \times ((\mathbb{R})^2)^{\otimes N}) .$$

♢

Finally let us look at the $2D$ Keller-Segel non-linear SDE.

Actually one can directly use theorem 6.2 and its proof in [10], since as recalled in remark 6.4, the drift $b_t = K * \rho_t = -\chi \nabla c$ satisfies the finite energy condition

$$\int_0^T \int |b_t|^2 \rho_t \, dx \, dt < +\infty .$$

We may thus use theorem 4.6 in order to get

Theorem 6.20. *Assume that* $\int (\ln \rho_0(x) + |x|^2) \rho_0(x) dx < +\infty$ *and* $0 < \chi < 4$. *If* $t \mapsto \rho_t$ *is the solution of (6.2) built in theorem 6.2 for* $K(x) = \frac{\chi x}{|x|^2}$, *then there exists a (weak) solution of*

$$dX_t = \sqrt{2} dB_t - (K * \rho_t)(X_t) \, dt \quad \text{with} \quad \mathcal{L}(X_t) = \rho_t(x) \, dx .$$

In addition the law Q *of this solution restricted to* $[0, T]$ *satisfies* $H(Q|P) < +\infty$ *where* P *denotes the law of a standard* 2*-D stationary Ornstein-Uhlenbeck process, i.e. with invariant probability measure the centered gaussian distribution with covariance matrix* $a\, Id$ *for any* $a > 0$ *and* $T > 0$.

Finally this solution is unique in the set of solutions such that their time marginals satisfy (6.3) and

$$\int_0^T \int |\nabla \ln \rho_t + K * \rho_t|^2(x) \, \rho_t(x) dx dt < +\infty .$$

For uniqueness it is enough to use Theorem 6.3. Accordingly, since for two solutions Q_1 and Q_2, their marginals flow satisfy (6.2), these marginals flow are the same denoted by ρ_t. Q^1 and Q^2 are thus both solutions of the linear S.D.E.

$$dY_t = \sqrt{2} \, dB_t - (K * \rho_t)(Y_t) \, dt$$

with finite relative entropy w.r.t. P, hence are given by the same Girsanov density.

Remark 6.21. As we already discussed, the same holds true for the Keller-Segel model with confinement. In the special case of $U(v) = |v|^2$ this result is a simple consequence of the correspondence with the so called "self similar" coordinates. Indeed if we define

$$g_t(x) = \frac{1}{1+2t} \, \rho_{\ln(\sqrt{1+2t})} \left(\frac{x}{\sqrt{1+2t}} \right) ,$$

as soon as ρ is a solution to the Keller-Segel P.D.E., g is a solution to the confined Keller-Segel P.D.E. with confinement given by the quadratic U above. ♢

Thanks to what we have done before, uniqueness in Theorem 6.20 can be improved. Indeed $K \mathbf{1}_{K>A} \in \mathbb{L}^p(\mathbb{R}^2)$ for any $p < 2$ so that if $\sup_{t \in [0,T]} ||\rho_t||_q < +\infty$ for some $q > 1$,

$$\sup_{t \in [0,T]} ||K \mathbf{1}_{K>A} * \rho_t||_r < +\infty \quad \text{for some } r > 2 .$$

We may thus apply Theorem 3.8 and get that the *linear* SDE

$$dY_t = \sqrt{2} \, dB_t - b(Y_t) dt - (K * \rho_t)(Y_t) dt$$

has a unique weak solution $\tilde{Q}$ and $H(\tilde{Q}|P) < +\infty$. If Q solves the non linear SDE (1.3), it is a weak solution of the previous linear one. Accordingly $Q = \tilde{Q}$ satisfies $H(Q|P) < +\infty$ and we have shown

Corollary 6.22. *Under the assumptions of Theorem 6.20, there exists at most one solution of the non-linear SDE such that its marginals flow satisfies* $\sup_{t\in[0,T]} ||\rho_t||_q < +\infty$ *for some* $q > 1$.

Remark 6.23. The previous result is (almost) more general than the uniqueness result in Theorem 6.20 (even if we need this uniqueness result in the proof of the Corollary), according to the discussion in Remark 6.4. ♢

Remark 6.24. One can also be interested by the *repulsive* Keller-Segel model with $\chi < 0$. This process is studied in [13] under the name *dynamics of a planar Coulomb gas* with a gaussian confinement potential, generalizing to dimension $d = 2$ the Ornstein-Uhlenbeck Dyson process. Existence and uniqueness for the particle system can be shown by using a Lyapunov function for proving the absence of collisions. Non linear SDE and propagation of chaos is mentioned in [13] subsubsection 1.4.6 but does not seem to have received a complete treatment since that time. ♢

7. Large number of particles: convergence in relative entropy and chaos.

We come to the large number of particles problem. In this section we denote by Q^N (resp. $Q^{k,N}$ and $\bar{Q}^{\otimes j}$) the law of the solution of the particle system (1.2) (resp. the law of $(X^{1,N}_\cdot, ..., X^{k,N}_\cdot)$ in the system (1.2) and the j tensor product of the law of the solution of the non linear S.D.E. (1.3)), when they are well defined. For simplicity we will sometimes denote by η^N and $\bar{\eta}^j$ the full drifts, including interactions and self-interaction, of Q^N and $\bar{Q}^{\otimes j}$, i.e. for instance

$$\eta^{i,N}(t) = b(X^{i,N}_t) + \frac{1}{N}\sum_j K(X^{i,N}_t - X^{j,N}_t)\,.$$

In the whole section, the initial distribution μ_0^N *is assumed to be exchangeable* and to converge, in a sense and with a rate to define, to some $\mu_0^{\otimes N}$.

We also introduce an additional definition

Definition 7.1. Let $\bar{Q}$ a solution of the non linear SDE (1.3). Denote by $\bar{\rho}_\cdot$ its marginals flow. The *linearized* McKean-Vlasov equation associated to $\bar{Q}$ is the linear (time inhomogeneous) SDE

$$dY_t = \sqrt{2}\,dW_t - b(Y_t)dt - (K * \bar{\rho}_t)(Y_t)dt\,.$$

As explained in the introduction, it is expected that $Q^{k,N}$ (which is also exchangeable) converges to $\bar{Q}^{\otimes k}$ as N goes to infinity for any fixed k, in some sense (total variation, Wasserstein distance, relative entropy) implying weak convergence. Notice that convergence in relative entropy is stronger than the entropic convergence described in the introduction and shown in [62, 68]. We shall first recall some existing results (and methods) concerned with the relative entropy convergence and then apply them to the situations we are interested in.

Exchangeability and the variational equivalent definition of relative entropy in (1.6) immediately show that for $k \le N-1$,

$$H(Q^{k,N}|\bar{Q}^{\otimes k}) \le \frac{1}{\lfloor N/k\rfloor}\,H(Q^N|\bar{Q}^{\otimes N}) \le \frac{k}{N-k}\,H(Q^N|\bar{Q}^{\otimes N})\,. \tag{7.1}$$

It is thus tempting to try to get some upper bound for $H(Q^N|\bar{Q}^{\otimes N})$, ideally some bound that does not depend on N. It is the approach developed some times ago in [5] where the authors are using a description of Q^N as a Gibbs measure on the path space, provided η^N is of gradient type (for a similar description in a non-mean field framework one can look at [40]). A uniform in N bound is obtained in Theorem 1 of [5] and applied to the particle system in their Theorem 3. The gibbsian description requires η^N to be a bounded smooth gradient.

As remarked by Lacker (this remark is part of the "old" folklore, and was already made by Föllmer, but seemingly not published), looking at the "reversed" relative entropy $H(\bar{Q}^{\otimes N}|Q^N)$ is interesting because, with less stringent assumptions than before, an uniform in N upper bound is true. However, the "projected" $H(\bar{Q}^{\otimes k}|Q^{k,N})$ does no more satisfy (7.1) so that the previous bound, if interesting, is far to be enough.

Notice that in our singular framework, looking at the reversed relative entropy has another interest. When Q^N allows collisions (as for the Keller-Segel model), $\bar{Q}^{\otimes N}$ does not allow collisions (due to the independence of its components). In particular Q^N is not absolutely continuous w.r.t. $\bar{Q}^{\otimes N}$ and the relative entropy is thus infinite, while $\bar{Q}^{\otimes N}$ has a chance to be absolutely continuous w.r.t. Q^N.

7.1. **Lacker's approach.**

Recently Lacker proposed in [90] a new method in order to show the convergence in relative entropy (and entropic chaos at the same time). We shall briefly recall his approach, in order to introduce some slight modifications.

Assume that both (1.2) and (1.3) have a weak solution Q^N and $\bar{Q}$. The key in Lacker's proof is that $Q^{k,N}$ is still an Ito process whose drift is obtained by conditioning w.r.t. $(X^{1,N}_{.\le t},...,X^{k,N}_{.\le t})$, i.e is given by the non markovian drift

$$\begin{aligned}
\hat{b}^{i,k}(t,X^{1,N}_{.\le t},...,X^{k,N}_{.\le t}) &= b(X^{i,N}_t)+\frac{1}{N}\sum_{i\neq j,j=1,...,k}K(X^{i,N}_t-X^{j,N}_t)+\\
&\quad+\mathbb{E}\left[\frac{1}{N}\sum_{j=k+1,...,N}K(X^{i,N}_t-X^{j,N}_t)\Big|X^{1,N}_{.\le t},...,X^{k,N}_{.\le t}\right]\\
&= b(X^{i,N}_t)+\frac{1}{N}\sum_{i\neq j,j=1,...,k}K(X^{i,N}_t-X^{j,N}_t)+\\
&\quad+\frac{N-k}{N}E\left[K(X^{i,N}_t-X^{k+1,N}_t)\Big|X^{1,N}_{.\le t},...,X^{k,N}_{.\le t}\right], \qquad (7.2)
\end{aligned}$$

since almost surely, for all $j\ge k+1$,

$$E\left[K(X^{i,N}_t-X^{k+1,N}_t)\Big|X^{1,N}_{.\le t},...,X^{k,N}_{.\le t}\right]=E\left[K(X^{i,N}_t-X^{j,N}_t)\Big|X^{1,N}_{.\le t},...,X^{k,N}_{.\le t}\right] \qquad (7.3)$$

the latter equality being a consequence of exchangeability. Of course here $|X^{i,N}_{.\le t}$ denotes the conditioning w.r.t. the whole path of $X^{i,N}_{.}$ up to time t.

Introduce the $\bar{Q}^{\otimes k}$ exponential local martingale

$$Z_t^{k,N} = \frac{d\mu_0^{k,N}}{d\mu_0^{\otimes k}} \, \exp\left(\int_0^t \langle \beta_s^{k,N}, \sqrt{2}\, dB_s \rangle - \int_0^t |\beta_s^{k,N}|^2 \, ds \right)$$

where

$$\beta_s^{i,k,N} = \hat{b}^{i,k}(s, \omega^{\leq k,N}) \, - \, b(\omega_s^{i,N}) \, - \, (K * \bar{\rho}_s)(\omega_s^{i,N})$$

$\omega^{\leq k,N}$ being the generic element in $C^0([0,t],(\mathbb{R}^d)^k)$ and $\sqrt{2}\, B_.$ being the $\bar{Q}^{\otimes k}$ dk dimensional Brownian motion (with variance $2t$)

$$\omega_.^{k,N} - \omega_0^{k,N} + \int_0^. (b(\omega_s^{k,N}) + (K * \bar{\rho}_s)(\omega_s^{k,N}))ds \, .$$

If we assume that the linearized McKean-Vlasov associated to $\bar{Q}$ has a unique weak solution, and that

$$I_T^N = \int_0^T M_t^N \, dt < +\infty \, ; \;\; \text{where} \;\; M_t^N = \mathbb{E}[|K(X_t^{1,N} - X_t^{2,N}) \, - \, (K * \bar{\rho}_t)(X_t^{1,N})|^2] \, , \quad (7.4)$$

(where thanks to exchangeability, we may replace the indices $(1,2)$ by any pair (i,j) with $i \neq j$), then the entropic Girsanov theory introduced in subsection 2.1 and detailed in this more general context in [92, 34], tells us that $Q^N = Z_T^{N,N} \, \bar{Q}^{\otimes N}$ and $H(Q^N|\bar{Q}^{\otimes N}) < +\infty$. In particular Q^N is the unique weak solution of (1.2).
Contraction of relative entropy w.r.t. measurable mapping says that for all $k \leq N$,

$$H(Q^{k,N}|\bar{Q}^{\otimes k}) \leq H(Q^N|\bar{Q}^{\otimes N}) < +\infty \, .$$

Also remark that for any $k = 1, ..., N$, all $i \leq k$ and $j > k$,

$$\mathbb{E}\left[|E[K(X_t^{i,N} - X_t^{j,N})|X_{.\leq t}^{1,N}, ..., X_{.\leq t}^{k,N}] - (K * \bar{\rho}_t)(X_t^{i,N})|^2 \right] \, \leq \, M_t^N \, .$$

Define

$$H_t^{k,N} = H(Q_{|\mathcal{F}_t}^{k,N}|\bar{Q}_{|\mathcal{F}_t}^{\otimes k}) \, .$$

Thanks to exchangeability, it holds

$$\frac{d}{dt} H_t^{k,N} \leq \frac{k(k-1)^2}{2N^2} \, M_t^N + \frac{k}{2} \, \mathbb{E}\left[|\mathbb{E}[K(X_t^{1,N} - X_t^{N,N})|X_{.\leq t}^{1,N}, ..., X_{.\leq t}^{k,N}] - (K * \bar{\rho}_t)(X_t^{1,N})|^2 \right] . \quad (7.5)$$

Lacker's idea is to get some *hierarchy* for these H_t^k, assuming that,

$$\begin{aligned} &|\mathbb{E}[K(X_t^{1,N} - X_t^{N,N})|X_{.\leq t}^{1,N}, ..., X_{.\leq t}^{k,N}] - (K * \bar{\rho}_t)(X_t^{1,N})|^2 \\ &\quad \leq \, \gamma(K) \, H([Q_t^{k+1,N}]_{|X_{.\leq t}^{1,N}, ..., X_{.\leq t}^{k,N}} | \bar{Q}_t) \end{aligned} \quad (7.6)$$

where $[Q_t^{k+1,N}]_{|X_{.\leq t}^{1,N}, ..., X_{.\leq t}^{k,N}}$ is the conditional distribution of $X_{.\leq t}^{N,N}$ knowing $X_{.\leq t}^{1,N}, ..., X_{.\leq t}^{k,N}$. Notice that both hand side are random, depending in particular on $X_t^{1,N}$, so that (7.6) has

to hold almost surely for all value of $X_t^{1,N}$. It is thus sufficient (and much more tractable) to replace (7.6) by the stronger (still random)

$$\sup_{s\mapsto x_s} |\mathbb{E}[K(x_t - X_t^{N,N})|X_{.\le t}^{1,N},...,X_{.\le t}^{k,N}] - (K*\bar\rho_t)(x_t)|^2 \\ \le \gamma(K)\, H([Q_t^{k+1,N}]_{|X_{.\le t}^{1,N},...,X_{.\le t}^{k,N}}|\bar Q_t) \tag{7.7}$$

which will be called the Lacker transportation assumption, (LTA) for short.
If (7.6) is satisfied, then

$$\frac{d}{dt}\, H_t^{k,N} \le \frac{k(k-1)^2}{2N^2}\, M_t^N + \frac{k}{2}\,\gamma(K)\,(H_t^{k+1,N} - H_t^{k,N})\,, \tag{7.8}$$

since

$$\mathbb{E}(H([Q_t^{k+1,N}]_{|X_{.\le t}^{1,N},...,X_{.\le t}^{k,N}}|\bar Q_t)) = H_t^{k+1,N} - H_t^{k,N}\,.$$

The constants are slightly different due to the $\sqrt{2}\,dW_t$ in our equations. Gronwall's lemma yields

$$\begin{aligned} H_t^{k,N} &\le e^{-(\gamma(K)/2)kt}\, H_0^{k,N} + \frac{k(k-1)^2}{2N^2}\int_0^t e^{-(\gamma(K)/2)k(t-s)}\, M_s^N\, ds + \\ &+ \frac{1}{2}\,\gamma(K)\,k\int_0^t e^{-(\gamma(K)/2)k(t-s)}\, H_s^{k+1,N}\, ds \end{aligned}$$

so that

$$H_t^{k,N} \le e^{-(\gamma(K)/2)kt}\, H_0^{k,N} + \frac{k(k-1)^2}{2N^2}\, I_T^N + \frac{1}{2}\,\gamma(K)\,k\int_0^t e^{-(\gamma(K)/2)k(t-s)}\, H_s^{k+1,N}\, ds\,, \tag{7.9}$$

where we used a crude upper bound for the second term in the sum and $I_t^N \le I_T^N$.
Iterating the previous we get

$$H_T^{k,N} \le \sum_{l=k}^{N-1} B_k^l(T)\, H_0^{l,N} + \sum_{l=k}^{N-1} \frac{k(l-1)^2 I_T^N}{2N^2}\, A_{k+1}^l(T) + A_k^{N-1}(T)\, H_T^{N,N}$$

where

$$A_j^l(T) = \left(\prod_{i=j}^{l} i(\gamma(K)/2)\right)\int_0^T\int_0^{t_{j+1}} ... \int_0^{t_l} e^{-\sum_{i=j}^{l} i(\gamma(K)/2)(t_i - t_{i+1})}\, dt_{l+1}...dt_{j+1}$$

and

$$B_j^l(T) = \left(\prod_{i=j}^{l-1} i(\gamma(K)/2)\right)\int_0^T\int_0^{t_{j+1}} ... \int_0^{t_{l-1}} e^{-l(\gamma(K)/2)t_l - \sum_{i=j}^{l-1} i(\gamma(K)/2)(t_i - t_{i+1})}\, dt_l...dt_{j+1}$$

if $N \ge l \ge j$, $A_{k+1}^k(T) = 1$ and $B_k^k(T) = e^{-k(\gamma(K)/2)T}$.
Compared with (4.18) in [90] we have to shift the subscript of the $A's$ due to the expression in (7.9).

If we assume that $H(Q_0^{l,N}|\bar{Q}_0^{\otimes l}) \le C_0\, l^2\, \varepsilon_N$, and after noticing that $H_T^{N,N} \le H_0^{N,N} + \frac{N}{4}\, I_T^N$, one obtains, thanks to Lemma 4.8 in [90], the analogue of (4.24) in [90],

$$\begin{aligned} H_T^{k,N} &\le C_0\, \varepsilon_N\, 2k^2\, e^{\gamma(K)T} + \frac{I_T^N (k+3)^3}{6N^2}\, e^{3\gamma(K)T/2} + \frac{k^3\, I_T^N}{2N^2} \\ &+ (C_0 N^2 \varepsilon_N + \frac{1}{4} N I_T^N)\, \exp\left(-2N(e^{-\gamma(K)T/2} - \frac{k}{N})_+^2\right). \end{aligned} \tag{7.10}$$

Since we are not interested in optimizing the dependence in k (i.e. replace k^3 by k^2 as in [90]), we may stop here and state the following version of [90] Theorem 2.2

Theorem 7.2. ***Lacker's theorem***
Assume that both (1.2) and (1.3) have a weak solution Q^N and $\bar{Q}$. Also assume that the linearized McKean-Vlasov equation associated to $\bar{Q}$ of Definition 7.1 has a unique weak solution. Assume in addition that (7.4) and (7.6) are satisfied.
Then if for all $k \le N$, $H(\mu_0^{k,N}|\mu_0^{\otimes k}) \le C_0\, k^2\, \varepsilon_N$ for some $C_0 < +\infty$, and of course μ_0^N is exchangeable, then for any $k \le N$, $H(Q^{k,N}|\bar{Q}^{\otimes k}) = H_T^{k,N}$ satisfies (7.10).

Compared with Theorem 2.2 in [90] we simply have relaxed the assumption on the initial condition, and slightly improved the assumption (2) therein which is $\sup_t M_t^N = M^N < +\infty$.

Of course (7.6) and (7.7) are consequence of the more restrictive: there exists $0 < \gamma(K) < +\infty$ such that, for all $t \in (0,T)$, all $\omega \in C^0([0,T],\mathbb{R}^d)$ and all probability measure Q' on $C^0([0,T],\mathbb{R}^d)$ such that $\int |K(\omega_t - \omega'_t)| Q'(d\omega') < +\infty$, it holds for all ω,

$$\left|\int K(\omega_t - \omega'_t)(\bar{Q} - Q')(d\omega')\right|^2 \le \gamma(K)\, H(Q'|\bar{Q})\,. \tag{7.11}$$

Such an inequality seems very difficult to prove (due to the uniformity w.r.t. ω) except if K is bounded, where it amounts to Pinsker inequality (1.7) i.e. one can choose

$$\gamma(K) = 2\, ||K||_\infty^2\,.$$

In [71] the author proposes to use the improvement of Pinsker inequality shown in [14]

$$\left(\int g\, d|\mu - \nu|\right)^2 \le 4CH(\mu|\nu) \quad \text{with} \quad C = \frac{1}{6}\, \mu(g^2) + \frac{1}{3}\, \nu(g^2)\,.$$

This inequality has to be used with $\mu = [Q_t^{k+1,N}]_{|X_{\cdot\le t}^{1,N},\dots,X_{\cdot\le t}^{k,N}}$, hence once again one needs some uniform (in ω) control which is not easy to get. Notice that another weighted Pinsker inequality is shown in [14] namely

$$\left|\int g\, d|\mu - \nu|\right|^2 \le 2\left(1 + \ln \int e^{g^2} d\nu\right) H(\mu|\nu)$$

allowing us to slightly improve on the boundedness assumption of K replacing it by

$$\sup_{\omega'} \int e^{K^2(\omega' - \omega)}\, \bar{Q}(d\omega) < +\infty\,,$$

which is quite difficult to check, but this time only depends on $\bar{Q}$.

Remark 7.3. Why is the uniqueness assumption on the linearized McKean-Vlasov equation important ? Because it implies that the Girsanov transform up to the time τ_n introduced in the Remark 2.12 furnishes the unique solution of the SDE with drift $\hat{b}$ up to this time. Uniqueness of the solution up to T follows by the same argument as in Remark 2.12, so that the Girsanov transform of $\bar{Q}^{\otimes k}$ is exactly $Q^{k,N}$. ♢

Remark 7.4. If K is bounded, the main weakness of Lacker's results is the exponential behaviour w.r.t. the $\mathbb{L}^\infty$ squared norm of K. In order to use these results for an approximating sequence of K by some bounded K_N, it is necessary to use a logarithmic cut-off (as in [96]).
It is worth noticing that for a bounded K, a very similar quantitative bound (involving some $e^{c||K||^2_\infty}$) is obtained in [73] Theorem 1.1 (ii), where the relative entropy is replaced by some squared Wasserstein distance (see (1.19) in [73]). Recall that relative entropy controls the squared total variation distance. It thus seems that obtaining a better bound in a general bounded framework is not as easy.
Also notice that in the usual Lipschitz framework of McKean's Theorem 1.1 the squared W_2 Wasserstein distance is bounded by the initial squared distance plus

$$C_1\, T(||b||^2_{Lip} + ||K||^2_{Lip})\, e^{(C_2\, T\, (1+||b||^2_{Lip}+||K||^2_{Lip}))}\, \frac{k}{N}\,.$$

Hence here again some approximation method will require some logarithmic cut-off. ♢

7.2. **Reversing the entropy.**

Remark that during the discussion we have shown the following intermediate result: if (7.4) is satisfied and the linearized McKean-Vlasov equation associated to $\bar{Q}$ has a unique weak solution, then for all $k = 1, ..., N$ the SDE in $(\mathbb{R}^d)^k$,

$$dY^k_t = \sqrt{2}\, dW_t + \hat{b}^{\cdot,k}(Y_{\cdot\le t})\, dt$$

also has a unique solution, which is thus $Q^{k,N}$. This point was not known by Lacker in his study of the reverse entropy, so that he only looked at the case K bounded.

Assume now

$$\bar{I}_T = \int_0^T \bar{M}_t\, dt < +\infty\,; \;\text{ where }\; \bar{M}_t = \mathbb{E}^{\bar{Q}^{\otimes 2}}[|K(\omega^1_t - \omega^2_t)\, -\, (K * \bar{\rho}_t)(\omega^1_t)|^2]\,, \tag{7.12}$$

i.e. the analogue of (7.4) replacing $Q^{2,N}$ by $\bar{Q}^{\otimes 2}$. Notice that this time these quantities do not depend on N. This time, it holds

$$H(\bar{Q}^{\otimes N}|Q^N) \le H(\bar{Q}_0^{\otimes N}|Q^N_0) + \bar{I}_T\,. \tag{7.13}$$

If (7.12) is satisfied we thus have

$$H(\bar{Q}^{\otimes k}|Q^{k,N}) \le H(\bar{Q}^{\otimes N}|Q^N) < +\infty\,.$$

Since $\bar{Q}^{\otimes N} \ll Q^{N,N}$, (7.3), written in the form

$$E^{Q^{N,N}}\left[K(\omega^{i,N}_t - \omega^{k+1,N}_t)\Big|\omega^{1,N}_{\cdot\le t}, ..., \omega^{k,N}_{\cdot\le t}\right] = E^{Q^{N,N}}\left[K(\omega^{i,N}_t - \omega^{j,N}_t)\Big|\omega^{1,N}_{\cdot\le t}, ..., \omega^{k,N}_{\cdot\le t}\right]$$

for $j \ge k+1$ is still true $\bar{Q}^{\otimes N}$ almost surely.

Assuming now that, $\bar{Q}^{\otimes N}$ almost surely

$$|\mathbb{E}^{Q^{N,N}}[K(\omega_t^{1,N}-\omega_t^{k+1,N})|\omega_{\cdot\le t}^{1,N},...,\omega_{\cdot\le t}^{k,N}]-(K*\bar{\rho}_t)(\omega_t^{1,N})|^2$$
$$\le\ \gamma(K)\,H(\bar{Q}_t|[Q_t^{k+1,N}]_{|\omega_{\cdot\le t}^{1,N},...,\omega_{\cdot\le t}^{k,N}}) \tag{7.14}$$

we may follow the lines of the proof of Lacker's Theorem, defining $\bar{H}_t^{k,N}=H(\bar{Q}^{\otimes k}_{|\mathcal{F}_t}|Q^{k,N}_{|\mathcal{F}_t})$ and starting with

$$\frac{d}{dt}\,\bar{H}_t^{k,N}\le\frac{k(k-1)}{2N^2}\,\bar{M}_t\,+\,\frac{k}{2}\,\gamma(K)\,(\bar{H}_t^{k+1,N}-\bar{H}_t^{k,N})\,, \tag{7.15}$$

we can reverse the entropy and get the following slight generalization of Lacker's reverse Theorem 2.14 in [90]

Theorem 7.5. ***Lacker's reverse theorem***
Under the assumptions of Theorem 7.2, but this time with $H(\mu_0^{\otimes l}|\mu_0^{l,N})\le C_0\,l^2\,\varepsilon_N$, for any $l\le N$, replacing (7.6) by (7.14) and adding (7.4), it holds

$$\begin{aligned}H(\bar{Q}^{\otimes k}|Q^{k,N})\ &\le\ C_0\,\varepsilon_N\,2k^2\,e^{\gamma(K)T}+\frac{\bar{I}_T(k+2)^2}{4N^2}\,e^{\gamma(K)T}+\frac{k^2\,\bar{I}_T}{2N^2}\\&\quad+(C_0N^2\varepsilon_N+\bar{I}_T)\,\exp\left(-2N(e^{-\gamma(K)T/2}-\frac{k}{N})_+^2\right)\,.\end{aligned}$$

The previous result holds in particular if K is bounded, with $\gamma(K)=2||K||_\infty^2$ and bounding $\bar{I}_T$ by $4||K||_\infty^2\,T$.

7.3. Lacker's method and chaos for singular interactions.

Since it is delicate to verify (7.6) unless K is bounded, one can try to consider a sequence K_N of bounded kernels converging in some sense to K and the associated Q_N^N and $\bar{Q}_N$. Notice that

$$\max(I_{T,N}^N,\bar{I}_T^N)\ \le\ 4T\,||K_N||_\infty^2\,.$$

Let k be fixed. In order to simplify the discussion we will assume that

$$\varepsilon_N\le c\,\frac{T\,||K_N||_\infty^2}{N^2} \tag{7.16}$$

so that according to Theorem 7.5, $H(\bar{Q}_N^{\otimes k}|Q_N^{k,N})$ will go to 0 provided , as $N\to+\infty$

$$\ln N\,-\,||K_N||_\infty^2\,T\,-\,\ln(||K_N||_\infty^2\,T)\ \to\ +\infty\,. \tag{7.17}$$

It remains to control some distance between $Q_N^{k,N}$ and $Q^{k,N}$ on one hand, $\bar{Q}_N^{\otimes k}$ and $\bar{Q}^{\otimes k}$ on the other hand. The easiest part is the one concerned with the non linear S.D.E.'s. Actually the part concerned with the particle systems is much more intricate, and will require other tools. An alternate approach will be proposed in the next sections.

The $\mathbb{L}^q$ case. Consider first the situation of Theorem 4.1. The proof is precisely based on the convergence of $\bar{Q}_N$ to $\bar{Q}$ in the $\sigma(\mathbb{L}^1,\mathbb{L}^\infty)$ topology. Extension to the kth tensor product is immediate. Actually it is not difficult to reinforce the considered topology, because under the assumptions of Theorem 4.1, we may replace K^M-K by $|K^M-K|^2$, i.e. prove that $H(\bar{Q}|\bar{Q}_N)$ goes to 0. Some rates can be obtained in the various examples, but since $||K_N||_\infty$

is bounded by $\ln^{\frac{1}{2}} N$, these rates are disastrous. We may apply this result to the η relaxed Keller-Segel model. Using again Pinsker inequality we deduce converge in total variation distance for which we may apply the triangle inequality. We have thus obtained:

Proposition 7.6. *In the situation of Theorem 4.1, for an initial condition satisfying (7.16), for any cut-off $K \wedge A_N$ where A_N goes to infinity and satisfies (7.17), $d_{TV}(Q_N^{k,N}, \bar{Q}^{\otimes k})$ goes to 0 as N goes to infinity.*

This result has to be compared with [102] where some other cut-off is introduced, and with Theorem 1.1 in [73] where convergence in Wasserstein distance W_p $(1 \le p < 2)$ is obtained without cut-off (and also without rate). We shall come back to this later. ♢

More generally, it is enough to show that

$$\int_0^T \int |K_N * \rho_{t,N} - K * \rho_{t,N}|^2 \, \rho_t \, dx \, dt \, + \int_0^T \int |K * \rho_{t,N} - K * \rho_t|^2 \, \rho_t \, dx \, dt \, \to \, 0 \quad (7.18)$$

in order to prove that $H(\bar{Q}|\bar{Q}_N) \to 0$. Notice that even if we are able to show the previous convergence, it only furnishes a convergence in total variation distance for $\bar{Q}_N^{\otimes k}$ to $\bar{Q}^{\otimes k}$ as explained before. For this purpose it is actually enough to show convergence in total variation distance for $\bar{Q}_N^{\otimes k}$ to $\bar{Q}^{\otimes k}$. We shall study first the Keller-Segel model.

The Keller-Segel model (with or without confining potential). For the cut-off K_N we may use the one introduced in [10] subsection 2.5 (starting p.10) with the changes of notation corresponding to the present paper (mainly K_N here corresponds to ∇K^ε in [10]). We assume that the assumptions in Theorem 6.2 are satisfied.

According to [10] Lemma 2.11, almost all bounds one can think about are uniform in N. In particular (see (vii) in Lemma 2.11 of [10])

$$\sup_N \int_0^T |K_N * \rho_{t,N}|^2 \, \rho_{t,N} \, dx \, dt = \sup_N H(\bar{Q}_N|P) \, < +\infty \, .$$

The latter implies that the family $(\bar{Q}_N)_N$ is relatively compact for the weak topology $\sigma(\mathbb{L}^1, \mathbb{L}^\infty)$ and that any weak limit Q^* satisfies $H(Q^*|P) < +\infty$ thanks to the lower semi-continuity of relative entropy.

In addition according to [10] Lemma 2.12 any weak limit (in $\mathbb{L}^p(\mathbb{R}^+ \times \mathbb{R}^2)$) of the marginals flow $\rho_{.,N}$ of $\bar{Q}_N$ is a free energy solution of the Keller-Segel equation (6.2). According to Theorem 6.3 this solution is unique, so that the marginals flow $\rho_{.,N}$ weakly (in the previous sense) converges to $\rho_.$.

Consider $F(\omega) = \int_0^T f(\omega_t) \, dt$ for a bounded f. Since some subsequence $(\eta(N))$ of $\bar{Q}_N$ converges to Q^* it holds

$$\mathbb{E}^{Q^*}(F(\omega)) = \lim_N \mathbb{E}^{\bar{Q}_{\eta(N)}}(F(\omega)) = \lim_N \int_0^T \int f(x) \, \rho_{t,\eta(N)}(x) \, dx \, dt = \int_0^T \int f(x) \, \rho_t(x) \, dx \, dt \, .$$

This shows that the marginals flow of Q^* is $\rho_.$. A consequence is that for any $t > 0$, $\rho_{t,N}$ converges to ρ_t for the $\sigma(\mathbb{L}^1, \mathbb{L}^\infty)$ topology, slightly improving on the result in [10].

In order to show that $Q^* = \bar{Q}$, it remains to show that Q^* is a solution of the linear SDE,

$$dY_t = \sqrt{2} \, dB_t - (K * \rho_t)(Y_t) dt \, .$$

As for the proof of Theorem 3.8 this amounts to prove that

$$\lim_N \mathbb{E}^{\bar{Q}_N}\left[\left(\int_s^t \langle \nabla\varphi(\omega_u), (K_N * \rho_{u,N} - K * \rho_u)(\omega_u)\rangle\, du\right) H(\omega_{v\le s})\right] = 0\,, \tag{7.19}$$

for smooth φ and bounded H. We will not follow this way and come back to (7.18).

Thanks to Lemma 2.11 in [10], one can easily check that one can use the results of [54] replacing ρ_t by $\rho_{t,N}$. In particular Lemma 2.7 in [54] tells us that, for any $p \in [2, +\infty)$, as soon as $\rho_0 \in \mathbb{L}^p(\mathbb{R}^2)$ then

$$\sup_N \sup_{t\in[0,T]} ||\rho_{t,N}||_p < +\infty\,.$$

By interpolation, since we are dealing with probability densities, the same holds with any $1 \le p' \le p$. We will assume that ρ_0 satisfies such a condition in the sequel.

Once remarked that $|K_N - K| = |K_N - K|\,\mathbf{1}_{|K|\ge A_N}$ for some A_N going to infinity, we see that $K_N - K \in \mathbb{L}^r(\mathbb{R}^2)$ for all $1 \le r < 2$ and that $||K_N - K||_r \to 0$ as N growths to infinity. Using $r = 1$, it follows that $||(K_N - K) * \rho_{t,N}||_{p'} \to 0$ for all $1 \le p' \le p$. Taking $p' = 2 + 2\alpha$ for some $\alpha > 0$ and using Hôlder inequality we have

$$\begin{aligned}\int_0^T \int |K_N * \rho_{t,N} - K * \rho_{t,N}|^2\, \rho_t\, dx\, dt &\le T\, ||(K_N - K) * \rho_{t,N}||^2_{p'} \sup_{0\le t\le T} ||\rho_t||_{1+(1/\alpha)} \\ &\le T\, ||(K_N - K)||^2_1 \sup_{N,t} ||\rho_{t,N}||^2_{p'} \sup_t ||\rho_t||_{1+(1/\alpha)}\end{aligned}$$

that goes to 0 provided $\frac{1+\alpha}{\alpha} \le p$ and $2(1+\alpha) \le p$, which is always possible as soon as $p \ge 3$. This tackles the first term in (7.18).

For the second term, first choose some A such that $||K\,\mathbf{1}_{|K|>A}||_1 \le \varepsilon$. As for the first term we have, taking $\alpha = 1/2$ in the computations,

$$\int_0^T \int |K\mathbf{1}_{|K|>A} * (\rho_{t,N} - \rho_t)|^2\, \rho_t\, dx\, dt \le T\,\varepsilon^2\, 4(\sup_{N,t} ||\rho_{t,N}||^2_3 + \sup_t ||\rho_t||^2_3) \sup_t ||\rho_t||^2_3$$

which is a constant (independent of N and ε of course) times ε^2.
If we look now at the remaining term

$$\int_0^T \int |K\mathbf{1}_{|K|\le A} * (\rho_{t,N} - \rho_t)|^2\, \rho_t\, dx\, dt$$

we may argue as follows. First, since for each t, $\rho_{t,N}$ converges to ρ_t for the $\sigma(\mathbb{L}^1, \mathbb{L}^\infty)$ topology, $(K\mathbf{1}_{|K|\le A} * (\rho_{t,N} - \rho_t))(x)$ goes to 0 for all (t, x). In addition, since $K \in \mathbb{L}^1$, the same estimate as before shows that $|K\mathbf{1}_{|K|\le A} * (\rho_{t,N} - \rho_t)|^2$ is bounded in $\mathbb{L}^{3/2}(\rho_t\, dx\, dt)$, hence uniformly integrable. We may thus apply Vitali's convergence theorem to get that

$$\lim_N \int_0^T \int |K\mathbf{1}_{|K|\le A} * (\rho_{t,N} - \rho_t)|^2\, \rho_t\, dx\, dt = 0\,.$$

We have finally obtained

$$\limsup_N \int_0^T \int |K * (\rho_{t,N} - \rho_t)|^2\, \rho_t\, dx\, dt \le C\,\varepsilon^2$$

from which we deduce the convergence to 0 since ε is arbitrary. We have thus shown:

Proposition 7.7. *In the situation of Theorem 6.2, for an initial condition satisfying (7.16), for the cut-off K_N introduced in* [10] *and satisfying (7.17), if in addition $\rho_0 \in \mathbb{L}^3(\mathbb{R}^2)$, $d_{TV}(Q_N^{k,N}, \bar{Q}^{\otimes k})$ goes to* 0 *as N goes to infinity.*

Remark 7.8. The weak convergence to 0 of $\rho_{t,N}$ for all t is a key element of the proof, since we have to manage the square of $|K * (\rho_{t,N} - \rho_t)|$. It does not follow from the results in [10] nor the boundedness in an $\mathbb{L}^p$ as shown in [54], the latter furnishing for each t convergence of subsequences.

We have shown convergence in entropy to get at the end some convergence in total variation distance, which is stronger than the weak convergence one can expect if we follow the proof of Theorem 3.8 as we mentioned in (7.19). In addition the fact that the expectation to consider is the one w.r.t. $\bar{Q}_N$ introduces some intricacies.

The second part of the proof prevents us from getting a quantitative rate of convergence. ♢

Remark 7.9. Proposition 7.7 can be compared with [102] but also, more directly with [96]. In the latter reference a very similar result is obtained in Theorem 1.3, but in Wasserstein W_1 distance. It requires in particular to show that the non linear S.D.E. in Theorem 6.20 has a strong solution, which is shown by using the same cut-off as the one we are using and convergence. The main difference is that ρ_0 is assumed to be bounded in [96] and it is shown that under this assumption $\sup_t ||\rho_t||_\infty < +\infty$ (Thm 1.1 (ii)). We confess that we do not understand all the arguments in the Appendix of [96] in order to get this result, which is crucial for this approach, even if we can easily believe that the result is true. Nevertheless, it is shown in [54] Lemma 2.8 that $\rho_. \in C_b^\infty([\varepsilon, T] \times \mathbb{R}^2)$ for all $\varepsilon > 0$, so that all the job consists in the understanding of what happens for small t's. ♢

The 2D vortex model. Under the simplifying assumptions of Proposition 5.9 it is know that $\sup_t ||\rho_t||_\infty \leq ||\rho_0||_\infty$ (see [69] proof of Theorem 2 for a clever argument). Actually in [98] the authors are using a regularizing cut-off (see the top of p.487 therein), say K_N, and prove that the same result is true for the regularized equation i.e.

$$\sup_N \sup_t ||\rho_{t,N}||_\infty \leq ||\rho_0||_\infty \,.$$

One can thus easily adapt the proof we have done for the Keller-Segel model, since this time we know the uniqueness of the solution for the non linear P.D.E. in the set of bounded solutions (see e.g [69] Theorem 2). Details are left to the reader. It follows

Proposition 7.10. *In the situation of Proposition 5.9, for an initial condition satisfying (7.16), for the cut-off K_N introduced in* [98] *and satisfying (7.17), $d_{TV}(Q_N^{k,N}, \bar{Q}^{\otimes k})$ goes to* 0 *as N goes to infinity.*

Remark 7.11. Compared with the existing literature the previous proposition only improves on the convergence type, since we are looking at the full law of the process and not only at the marginals flow as in [62, 69, 55]. It is however weaker for at least three reasons: we are using a cut-off and not looking at the *true* particle system, the initial condition is more general in [62], [69, 55] contain quantitative results. ♢

We firmly believe that the situation is exactly the same for the sub-Coulombic repulsive model.

8. Large number of particles: more on convergence and chaos.

We continue with the notations of the previous section. In the latter section, the point was to obtain convergence to 0 for the relative entropy $H(Q^{k,N}|\bar{Q}^{\otimes k})$. Another standard approach is first to show the weaker tightness of the family $(Q^{k,N})_N$, then to show that there is only one possible weak limit and identify this weak limit as $\bar{Q}^{\otimes k}$. This is the strategy adopted for instance in [113] for the $\mathbb{L}^p - \mathbb{L}^q$ case, in [62] at the level of the marginals flow and in [63, 112] for the Keller-Segel model. In these two references only tightness is shown together with a partial identification of the possible limits. This will be discussed later.

8.1. Tightness.

Theorem 8.1. *Let P denotes the reversible probability measure on $\mathbb{R}^d$, with invariant probability measure $\gamma_0(dx) = Z^{-1}\, e^{-\sum_{j=1}^d V(x^j)} dx$. Here V is smooth, non-negative, and $|V'|$ is assumed to be bounded. Assume that $H(\mu_0^N|\gamma_0^{\otimes N}) \le CN$ (μ_0^N being exchangeable as before). In each of the following situations*

(1) ***$\mathbb{L}^p$-case.*** *$K\mathbf{1}_{|K|>A} \in \mathbb{L}^p(\mathbb{R}^d)$ for some $p \ge d$ if $d \ge 3$ or $p > 2$ if $d = 2$ and some $A > 0$. The additional drift b is bounded.*

(2) ***sub-Coulombic case.*** *$d \ge 3$, $\chi < 0$, b is bounded and Lipschitz, and*

$$K(x) = \chi\, \frac{x}{|x|^{s+2}}\, \mathbf{1}_{x\neq 0} \quad \text{for} \quad 0 < s \le d-2\,.$$

In addition for some $q > d/(d-s)$, $\sup_N ||\tilde{\rho}^N_{0,1,2}||_q < +\infty$, where $\tilde{\rho}^N_{0,1,2}$ is the density of $X_0^{1,N} - X_0^{2,N}$.

(3) ***The 2D vortex case with confinement.*** *$d = 2$, $K(x) = \chi\, \frac{x^\perp}{|x|^2}$ and $b = -\nabla\tilde{V}$ derives from the confining potential as in Theorem 5.7.*

Then for all fixed k, $\sup_{N>k} H(Q^{k,N}|\bar{P}^{\otimes k}) \le C\, k(k+1) < +\infty$ for some constant C that does not depend on k.
Consequently, the family $(Q^{k,N})_N$ is tight.

Proof. It holds

$$H(Q^N|P) = H(Q^N|\frac{d\mu_0^N}{\gamma_0^{\otimes N}}\, P) + H(\mu_0^N|\gamma_0^{\otimes N})$$

so that, in each situation, thanks to the assumption on the initial distribution and to the calculation in the proof of Lacker's theorem for case (1) (since (7.4) is satisfied in this case), to Remark 5.2 for case (2) and to Theorem 5.7 for case (3), we have $H(Q^N|P) \le CN$ for some C that does not depend on N. Since P is a product measure we may apply the trick (7.1) yielding $H(Q^{k,N}|\bar{P}^{\otimes k}) \le C\, \frac{kN}{N-k} \le C\, k(k+1)$. □

In the $\mathbb{L}^p$-case, this result is new for the critical $p = d$. Another nice proof, based on the $\mathbb{L}^q$ (for any $q < +\infty$) integrability of some Girsanov density for $p > d$, is made in [113] Lemma 3.2 (see Remark 8.2 for a comment on this result).

In the 2D vortex case without confinement, the previous result fails since we only knows that $H(Q^N|P) \le CN^2$. Of course when collisions occur, as for the Keller-Segel model, the previous approach also fails. One can thus come back to the standard approach using

Prokhorov's tightness criterion. Namely, tightness of $(\mathbb{Q}^{1,N})_N$ will follow from the following estimates: for some N_0,

$$(i) \qquad \lim_{a\to+\infty} \limsup_{N\geq N_0} \mathbb{P}(|X_0^{1,N}|>a) = 0\,,$$

$$(ii) \qquad \forall\varepsilon>0\,, \lim_{\eta\to 0} \limsup_{N\geq N_0} \mathbb{P}\left(\sup_{|t-s|<\eta} |X_t^{1,N}-X_s^{1,N}|\geq\varepsilon\right) = 0\,. \tag{8.1}$$

An analogue is true for $Q_N^{k,N}$, k being fixed and $N_0\geq k$.

For condition (ii) we may start with

$$\begin{aligned}\mathbb{P}\left(\sup_{|t-s|<\eta} |X_t^{1,N}-X_s^{1,N}|\geq\varepsilon\right) \;\leq\; & \mathbb{P}\left(\sup_{|t-s|<\eta} |B_t^{1,N}-B_s^{1,N}|\geq\varepsilon/3\sqrt{2}\right)\\ & +\mathbb{P}\left(\sup_{|t-s|<\eta}\left(\int_s^t |b(X_u^{1,N})|du\right)\geq\varepsilon/3\right)\\ & +\mathbb{P}\left(\sup_{|t-s|<\eta}\left|\int_s^t \frac{1}{N}\sum_{j=1}^N K(X_u^{1,N}-X_u^{j,N})\,du\right|\geq\varepsilon/3\right)\,.\end{aligned}$$

We then have

$$\begin{aligned}\mathbb{P}\left(\sup_{|t-s|<\eta} |B_t^{1,N}-B_s^{1,N}|\geq\varepsilon/3\sqrt{2}\right) \;&\leq\; \frac{3\sqrt{2}}{\varepsilon}\,\mathbb{E}\left(\sup_{|t-s|<\eta} |B_t^{1,N}-B_s^{1,N}|\right)\\ &\leq\; \frac{3\sqrt{2}}{\varepsilon}\,c_{univ}\,(\eta\,\ln(T/\eta))^{\frac12}\,.\end{aligned}$$

provided $\eta<T/e$, where c_{univ} is an universal constant. The last inequality follows from $|z|\leq|z^1|+|z^2|$ in $\mathbb{R}^2$ and [56] Lemma 3 or Lemma 4 where two different proofs are given for the moments controls of the modulus of continuity of a linear Brownian motion.

Next

$$\mathbb{P}\left(\sup_{|t-s|<\eta}\left(\int_s^t |b(X_u^{1,N})|du\right)\geq\varepsilon/3\right) \leq \mathbf{1}_{3||b||_\infty\eta>\varepsilon}$$

when b is bounded (in particular $b=0$ for the $2D$ vortex model). If b is a-Lipschitz, non necessarily bounded, we have $|b(y)|\leq|b(0)|+a|y|$ so that we may write as soon as $\eta|b(0)|<\varepsilon/6$

$$\mathbb{P}\left(\sup_{|t-s|<\eta}\left(\int_s^t |b(X_u^{1,N})|du\right)\geq\varepsilon/3\right) \;\leq\; \mathbb{P}\left(\sup_s |X_s^{1,N}|>\frac{\varepsilon}{6a\eta}\right)\,.$$

Now

$$\sup_{s\leq t}|X_s^{1,N}|\leq\;|X_0^{1,N}|+\sup_{s\leq T}|B_s^{1,N}|+\int_0^T\frac{1}{N}\sum_j|K(X_u^{1,N}-X_u^{j,N})|du+|b(0)|T+a\int_0^t\sup_{u\leq s}|X_u^{1,N}|du$$

so that using Gronwall lemma we have

$$\mathbb{E}(\sup_{s\leq T}|X_s^{1,N}|)=C<+\infty$$

where C does not depend on N as soon as $\mathbb{E}(|X_0^{1,N}|) < +\infty$ and, using exchangeability, $\int_0^T |K(X_u^{1,N} - X_u^{2,N})|du = c < +\infty$ independent of N, yielding

$$\mathbb{P}\left(\sup_s |X_s^{1,N}| > \frac{\varepsilon}{6a\eta}\right) \leq C\,\frac{6a\eta}{\varepsilon}\,.$$

For the last term

$$\mathbb{P}\left(\sup_{|t-s|<\eta}\left|\int_s^t \frac{1}{N}\sum_{j=1}^N K(X_u^{1,N} - X_u^{j,N})\,du\right| \geq \varepsilon/3\right) \leq \frac{3}{\varepsilon}\,\mathbb{E}\left(\sup_{|t-s|<\eta}\left|\frac{1}{N}\sum_{j=1}^N \int_s^t K(X_u^{1,N} - X_u^{j,N})\,du\right|\right)\,.$$

For $1 < \gamma$,

$$\int_s^t \left|\frac{1}{N}\sum_{j=1}^N K(X_u^{1,N} - X_u^{j,N})\right| du \;\leq\; (\frac{t-s}{T})^{\frac{\gamma-1}{\gamma}} \left(\int_0^T \left|\frac{1}{N}\sum_{j=1}^N K(X_u^{1,N} - X_u^{j,N})\right|^\gamma du\right)^{\frac{1}{\gamma}}$$

so that, if

$$\int_0^T \mathbb{E}\left(\left|\frac{1}{N}\sum_{j=1}^N K(X_u^{1,N} - X_u^{j,N})\right|^\gamma\right) du \leq M(\gamma), \tag{8.2}$$

in particular, thanks to convexity and exchangeability, if

$$\int_0^T \mathbb{E}\left[\left|K(X_u^{1,N} - X_u^{j,N})\right|^\gamma\right] du \leq M'(\gamma) \tag{8.3}$$

for all $N > N_0$ large enough, condition (ii) in (8.1) is satisfied. Notice that this condition is stronger than the $\mathbb{L}^1$ integrability required to control the second term for a Lipschitz additional drift.

It is immediate to see that the previous discussion extends to $(X_\cdot^{1,N}, ..., X_\cdot^{k,N})$ for all fixed k.

Remark 8.2. The previous proof applies in all the cases of Theorem 8.1 since condition (8.2) is satisfied with any $\gamma < d/(d-1)$. There are some differences on the assumptions on the initial condition (only tightness of the initial μ_0^N is required here). Of course the relative compactness in Theorem 8.1 holds for the weak $\sigma(\mathbb{L}^1, \mathbb{L}^\infty)$ topology and is thus a little bit stronger.

Notice that in Lemma 3.2 in [113], the supremum over (s,t) is outside the expectation, but its size $(t-s)^2$ is small enough thanks to Kolmogorov continuity criterion. We cannot use this tightness result here since the power of $|t-s|$ in our controls can be smaller than 1.

The line of reasoning used for this derivation is simpler but close to the one in [62] proof of Lemma 5.2. ♢

We may thus state

Theorem 8.3. *Assume that the initial measure is exchangeable and satisfies*

$$\sup_{N\geq N_0} \int |x^1|^2\,\rho_0^N(x)\,dx \;< +\infty\,.$$

Also assume that b is bounded and Lipschitz. In each of the following situations

(1) ***The 2D vortex case.*** $d = 2$, $K(x) = \chi\,\frac{x^\perp}{|x|^2}$,

(2) ***The Keller-Segel case.*** $d = 2$, $K(x) = \chi \frac{x}{|x|^2}$, $\chi < 4$,

the family $(Q^{k,N})_N$ *is tight.*

Proof. The Keller-Segel case is an immediate consequence of the previous line of reasoning and Lemma 6.14, once remarked that the additional drift only needs to be bounded or Lipschitz and not necessary gradient for this Lemma.
For the Biot-Savart kernel, it is enough to look at the calculations of subsection 5.3, in particular equation (5.5) who shows that

$$\sup_{N \geq N_0} \int_0^T \mathbb{E}\left[\left|K(X_u^{1,N} - X_u^{2,N})\right|^\gamma\right] du < +\infty$$

for all $0 \leq \gamma < 2$. □

Remark 8.4. For the Keller-Segel model a similar result is stated in [63] for $\chi < 1$. Despite the fact that the authors are proving at the same time convergence of the regularized model we discussed in section 6 and tightness, the proofs are very close to what we have done (and conversely), see e.g. the first inequality at the top of p. 2821 therein. Our use of [56] simplifies the argument, and the extension of Lemma 6.14 to all values of χ allows us to cover the larger range $\chi < 4$.
A general tightness result for the full range $\chi < 4$ at the marginals flow level (which is weaker but not as easy) is shown in Theorem 4 (i) of [112].
For the 2D vortex model, tightness is shown in Lemma 5.2 of [62] also using very similar arguments. ♢

8.2. **Towards consistency.**

In this subsection we shall show that any weak limit in Theorems 8.3, 8.1, is a solution of the corresponding non linear equation. The method of proof is almost standard (one can find some similar arguments in [62] for instance). We shall nevertheless give the main elements of proof, in order to fill some gaps in previous works, and to show how the use of entropy on the path space simplifies many arguments.

Proposition 8.5. *In any of the situations of Theorem 8.1 and Theorem 8.3 (assuming in addition that* b *is continuous in the* $\mathbb{L}^p$ *case), introduce*

$$\mathcal{Q}^N = \frac{1}{N} \sum_{j=1}^N \delta_{X_\cdot^{j,N}}$$

where $X_\cdot^{\cdot,N}$ *denotes the whole path of the process on the time interval* $[0, T]$. *Then* $\mathcal{Q}^N$ *is tight and any weak limit* $\mathcal{Q}$ *is almost surely a solution of the non linear S.D.E. (1.3).*

Proof. The statement on tightness follows from proposition 2.2 (ii) in [110].
Take some subsequence, still denoted by $\mathcal{Q}^N$, that weakly converges to some $\mathcal{Q}$. Recall that $\mathcal{Q}$ is a random variable taking values in the set of probability measures on the path space. In order to prove that $\mathcal{Q}$ almost surely solves (1.3), it is enough to look at the associated martingale problem. Take some $0 \leq s < t \leq T$. Following a now standard method (see

[15] for example) for any given continuous and bounded h defined on $C^0([0,s],\mathbb{R}^d)$, any $\varphi \in C^0_b(\mathbb{R}^d)$ and any probability measure ν defined on $C^0([0,T],\mathbb{R}^d)$ we introduce

$$\mathcal{F}(\nu) = \int\int \left[\varphi(\omega_t) - \varphi(\omega_s) - \int_s^t \Delta\varphi(\omega_u)du \, - \int_s^t \langle b(\omega_u) + K(\omega_u - \omega'_u), \nabla\varphi(\omega_u)\rangle \, du\right] \\ h(\omega_.)\, \nu(d\omega)\, \nu(d\omega')\,.$$

The statement of the Theorem amounts to

$$\mathcal{F}(\mathcal{Q}) = 0 \text{ almost surely }. \tag{8.4}$$

In order to prove (8.4) the first step is that

$$\mathbb{E}[(\mathcal{F}(\mathcal{Q}^N))^2] \le C(\varphi, h, s, t)\, \frac{1}{N}\,. \tag{8.5}$$

It turns out that

$$\mathcal{F}(\mathcal{Q}^N) = \frac{1}{N}\left(\sum_{j=1}^N h(X^{\cdot,N}_{\cdot})\,(A^{j,N}(t) - A^{j,N}(s))\right)$$

where

$$\begin{aligned} A^{j,N}(t) - A^{j,N}(s) &= \varphi(X^{j,N}_t) - \varphi(X^{j,N}_s) - \int_s^t \Delta\varphi(X^{j,N}_u)du \\ &\quad - \int_s^t \left\langle b(X^{j,N}_u) + \frac{1}{N}\sum_{i\neq j} K(X^{j,N}_u - X^{i,N}_u), \nabla\varphi(X^{j,N}_u)\right\rangle du \\ &= \int_s^t \langle \nabla\varphi(X^{j,N}_u), dB^{j,N}_u\rangle \end{aligned}$$

is thus a square integrable martingale increment. (8.5) immediately follows using standard stochastic calculus.

Since we always assume that b is bounded and continuous, if K is also bounded and continuous, then $\mathcal{F}$ is continuous and bounded for the topology of weak convergence, so that $\mathbb{E}(|\mathcal{F}(\mathcal{Q})|) = \lim_N \mathbb{E}(|\mathcal{F}(\mathcal{Q}^N)|) = 0$. But in all the cases we are interested in, K is not bounded.

For $\varepsilon > 0$, let $\psi_\varepsilon(u)$ be a continuous function on $\mathbb{R}^+$ such that $\mathbf{1}_{u\le 2\varepsilon} \le \psi_\varepsilon(u) \le \mathbf{1}_{u\le\varepsilon}$. In all the cases except the $\mathbb{L}^p$ case, we replace K by $K_\varepsilon = K\,\psi_\varepsilon(1/|K|)$ in order to define $\mathcal{F}_\varepsilon$. In the $\mathbb{L}^p$ case we have first to regularize K into a continuous function, this regularization inducing a small error term in all what follows and then pass to the limit. We will not detail this step, assume that K is always continuous in the sequel, and make the previous approximation.

Since $\mathcal{F}_\varepsilon$ is continuous and bounded, it holds

$$\mathbb{E}(|\mathcal{F}_\varepsilon(\mathcal{Q})|) = \lim_N \mathbb{E}(|\mathcal{F}_\varepsilon(\mathcal{Q}^N)|)\,. \tag{8.6}$$

It remains to control

$$\mathbb{E}(|\mathcal{F}(\mathcal{Q}^N) - \mathcal{F}_\varepsilon(\mathcal{Q}^N)|) \text{ and } \mathbb{E}(|\mathcal{F}(\mathcal{Q}) - \mathcal{F}_\varepsilon(\mathcal{Q})|)\,.$$

As we already remarked, in all these cases, there exists some $\gamma > 1$ such that

$$\sup_N \int_0^T \int\int |K|^\gamma(\omega_u - \omega'_u)\, \mathcal{Q}^N(d\omega_u)\, \mathcal{Q}^N(d\omega'_u)\, du \, < +\infty\,.$$

Using Hölder and Markov inequalities, we deduce that

$$\sup_N \int_0^T \int\int |K-K_\varepsilon|(\omega_u-\omega'_u)\,\mathcal{Q}^N(d\omega_u)\,\mathcal{Q}^N(d\omega'_u)\,du \;\le\; C(T,\gamma)\,\varepsilon^{\gamma-1}\,.$$

The same is true for $|K_\eta - K_\varepsilon|$ for any $\eta>0$, so that taking the limit in N, it is still true for $|K_\eta - K_\varepsilon|$ replacing $\mathcal{Q}^N$ by $\mathcal{Q}$, and finally for $|K-K_\varepsilon|$ using monotone convergence.
We have thus shown that both $\sup_N \mathbb{E}(|\mathcal{F}(\mathcal{Q}^N)-\mathcal{F}_\varepsilon(\mathcal{Q}^N)|)$ and $\mathbb{E}(|\mathcal{F}(\mathcal{Q})-\mathcal{F}_\varepsilon(\mathcal{Q})|)$ go to 0 as $\varepsilon\to 0$.
Finally, using the triangle inequality

$$\mathbb{E}(|\mathcal{F}(\mathcal{Q})| \le C\,(\frac{1}{\sqrt{N}} + 2\varepsilon^\theta + \mathbb{E}(|\mathcal{F}_\varepsilon(\mathcal{Q}^N)-\mathcal{F}_\varepsilon(\mathcal{Q})|)$$

so that choosing first ε small enough, and then N large enough, the left hand side is arbitrarily small, i.e. $\mathbb{E}(|\mathcal{F}(\mathcal{Q})|)=0$ so that $\mathcal{F}(\mathcal{Q})=0$ almost surely. □

Remark 8.6. In order to simplify several proofs we almost always assumed that b is bounded. Actually, provided $\sup_t \mathbb{E}(X_t^{1,N}) \le C < +\infty$ where C does not depend on N, we may skip the bounded assumption and only assume that b is Lipschitz (hence with linear growth), using another cut-off if necessary. Part of the arguments have been developped at the end of the previous subsection. This is the case in particular for the Keller-Segel and the 2D vortex models with a gaussian confining potential. ♢

Proposition 8.5 contains an interesting existence corollary, extending some existence results we have already recalled or proved

Corollary 8.7. *Assume that $H(\bar\rho_0 dx|\gamma_0) < +\infty$ where $\bar\rho_0$ is the density of $\bar X_0$. The non linear S.D.E. (1.3) has a solution, with initial distribution $\bar\rho_0 dx$ in the following cases*

(1) ***sub-Coulombic case.*** *$d\ge 3$, $\chi<0$, , and*

$$K(x) = \chi\,\frac{x}{|x|^{s+2}}\,\mathbf{1}_{x\neq 0} \;\; \textit{for} \;\; 0<s\le d-2\,.$$

In addition for some $q>d/(d-s)$, $||\bar\rho_0||_q<+\infty$ and and b is bounded and Lipschitz.

(2) ***The 2D vortex case with confinement.*** *$d=2$, $K(x)=\chi\,\frac{x^\perp}{|x|^2}$, the additional drift b is a confining potential as in Theorem 5.7 and $\int |x|^2\,\bar\rho_0(x)dx<+\infty$.*

(3) ***The 2D vortex case.*** *$d=2$, $K(x)=\chi\,\frac{x^\perp}{|x|^2}$, and $\int |x|^2\,\bar\rho_0(x)dx<+\infty$.*

Consequently the non linear P.D.E. (1.1) has a solution with initial condition $\bar\rho_0$ in all these cases.

It is enough to use what we achieved to do with an initial $\rho_0^N = \bar\rho_0^{\otimes N}$, since any limiting $\mathcal{Q}$ will do the job. Recall that in subsection 5.2 in the sub-coulombic situation, or in Proposition 5.9 we assumed that $\bar\rho_0$ is bounded.
Notice that in the $\mathbb{L}^p$ case and for the Keller-Segel model (with $\chi<4$), we also recover the existence part of Theorem 4.1 and Theorem 6.20. In the latter case, we have thus obtained the existence of a solution to the Keller-Segel equation (with $\chi<4$), but we do not know (at this stage) that it is a free energy solution as in Theorem 6.2.

8.3. **Convergence.**

We start by tackling the $\mathbb{L}^p$ situation.

Theorem 8.8. *Assume that $K\mathbf{1}_{|K|>A} \in \mathbb{L}^p(\mathbb{R}^d)$ for some $p \geq d$ if $d \geq 3$ or $p > 2$ if $d = 2$ and some $A > 0$ and that the additional drift b is continuous and bounded. Also asume that the initial condition $\mu_0^N = \rho^N\, dx$ is chaotic so that $\mu_0^{k,N} \to (\bar\rho_0\, dx^1)^{\otimes k}$ and satisfies $H(\mu_0^N|\gamma_0^{\otimes N}) \leq CN$.*
Then Q^N is chaotic and for each k, $Q^{k,N}$ weakly converges to $(\bar Q)^{\otimes k}$ where $\bar Q$ is the unique solution of the non linear SDE (1.3) with initial condition $\bar\rho_0\, dx$.

Proof. According to Proposition 4.8 (1.3) has a unique weak solution $\bar Q$. Any weak limit in Proposition 8.5 is thus equal to the deterministic measure $\bar Q$. According to [110] Proposition 2.2. (i), Q^N is thus chaotic. □

In the framework of a *convolution kernel* K, this result extends Theorem 1.6 in [113] to more general initial conditions and to the critical case $d = p$. Still for $p < d$, [73] Theorem 1.1 (i) contains a similar statement, but replacing weak convergence by convergence in Wasserstein W_2 distance for one particle, while propagation of chaos is shown in [73] Theorem 5.1 by the same method as [113] (itself inspired by [15]). The fact that the Girsanov density belongs to all the $\mathbb{L}^q$ is crucial in these proofs.

In order to study the other singular models we start with a simple remark. For $N = 2n$, consider the pair empirical measure

$$\mathcal{Q}_2^N = \frac{1}{N}\sum_{i=1}^n (\delta_{(X_\cdot^{2i-1,N},X_\cdot^{2i,N})} + \delta_{(X_\cdot^{2i,N},X_\cdot^{2i-1,N})})$$

which is a symmetric probability measure on $C^0([0,T],\mathbb{R}^d\otimes\mathbb{R}^d)$. In a sense we are looking at pairs of coordinates. Of course, if N is odd, we are loosing one term (the last X^N for instance), but it is immediate to see that this error term will disappear in the limit in what follows.
Introduce the pair functional defined for probability measures ν on $C^0([0,T],\mathbb{R}^d\otimes\mathbb{R}^d)$ by

$$\mathcal{F}_2(\nu) = \int \left[\varphi(\omega_t,\eta_t) - \varphi(\omega_s,\eta_s) - \int_s^t \Delta\varphi(\omega_u,\eta_u)du - \int_s^t \langle b(\omega_u,\eta_u),\nabla\varphi(\omega_u,\eta_u)\rangle\, du\right] h(\omega_\cdot,\eta_\cdot)\,\nu(d\omega,d\eta) -$$
$$- \int\int \left[\int_s^t \left(\langle K(\omega_u-\omega'_u),\nabla_1\varphi(\omega_u,\eta_u)\rangle + \langle K(\eta_u-\eta'_u),\nabla_2\varphi(\omega_u,\eta_u)\rangle\right) du\right] h(\omega_\cdot,\eta_\cdot)\,\nu(d\omega,d\eta)\nu(d\omega',d\eta')\,.$$

Notice that

$$\int K(\omega_u-\omega'_u)\mathcal{Q}_2^N(d\omega',d\eta') = \frac{1}{N}\sum_j K(\omega_u - X_u^{j,N})$$

so that the same calculations as in the proof of Proposition 8.5, show that

$$\mathcal{F}_2(\mathcal{Q}_2^N) = h(X_\cdot^N)\,\frac{\sqrt{2}}{N}\sum_{i=1}^n \int_s^t \langle \nabla_1\varphi(X_u^{2i-1,N},X_u^{2i,N}) + \nabla_2\varphi(X_u^{2i,N},X_u^{2i-1,N}), dB_u^{2i-1,N}\rangle$$

$$+h(X^N_\cdot)\,\frac{\sqrt{2}}{N}\sum_{i=1}^n\int_s^t\langle\nabla_2\varphi(X^{2i-1,N}_u,X^{2i,N}_u)+\nabla_1\varphi(X^{2i,N}_u,X^{2i-1,N}_u),dB^{2i,N}_u\rangle\,.$$

Hence $\mathbb{E}[(\mathcal{F}_2(\mathcal{Q}_2^N))^2]\le C(\varphi,h,s,t)\,\frac{1}{N}$, and we may thus argue as in the proof of Proposition 8.5.

The tightness of the distribution $\mathcal{L}\mathcal{Q}_2^N$ of $\mathcal{Q}_2^N$ is a consequence of the one of $Q^{2,N}$ since the intensity measure $I(\mathcal{L}\mathcal{Q}_2^N)$ is given by

$$\int f\,dI(\mathcal{L}\mathcal{Q}_2^N)=\mathbb{E}\left[\frac{1}{N}\sum_{i=1}^n(f(X^{2i-1,N}_\cdot,X^{2i,N}_\cdot)+f(X^{2i,N}_\cdot,X^{2i-1,N}_\cdot))\right]=\mathbb{E}[f(X^{1,N}_\cdot,X^{2,N}_\cdot)]\,,$$

and as shown in the proof of Proposition 2.2 (ii) in [110] tightness of $\mathcal{L}\mathcal{Q}_2^N$ is equivalent to tightness of $I(\mathcal{L}\mathcal{Q}_2^N)$. We then deduce as before that any weak limit $\mathcal{Q}_2$ satisfies $\mathcal{F}_2(\mathcal{Q}_2)=0$ almost surely.

Arguing as in the proof of Proposition 2.2 (i) in [110], introducing the same truncature K_ε as before, we have

$$|\mathcal{F}_{2,\varepsilon}(Q^{2,N})-\mathbb{E}[\mathcal{F}_{2,\varepsilon}(\mathcal{Q}_2)]|\;\le\;|\mathcal{F}_{2,\varepsilon}(Q^{2,N})-\mathbb{E}[\mathcal{F}_{2,\varepsilon}(\mathcal{Q}_2^N)]|+|\mathbb{E}[\mathcal{F}_{2,\varepsilon}(\mathcal{Q}_2)]-\mathbb{E}[\mathcal{F}_{2,\varepsilon}(\mathcal{Q}_2^N)]|$$

where the second term goes to 0 using convergence and the first one goes to 0 using symmetry (see [110] (2.3)). Taking another subsequence if necessary, we may assume that $Q^{2,N}$ weakly converges to $\bar{Q}_2$. The same argument as the one used at the end of the proof of Proposition 8.5 shows that $|\mathcal{F}_{2,\varepsilon}(\bar{Q}_2)-\mathcal{F}_2(\bar{Q}_2)|$ goes to 0 as $\varepsilon\to 0$. hence taking limits first w.r.t. N and second w.r.t. ε we have obtained that

$$\mathcal{F}_2(\bar{Q}_2)=0\,.$$

We similarly have

$$\mathcal{F}(\bar{Q})=0\,,$$

for any weak limit of $Q^{1,N}$.

One classically deduce the existence of a standard Brownian motion $(\bar{B}^1_\cdot,\bar{B}^2_\cdot)$ such that $\bar{Q}_2$ is the law of

$$\begin{aligned}
d\bar{X}^1_t &= \sqrt{2}\,d\bar{B}^1_t\;-\;b(\bar{X}^1_t)\,dt\;-\;(K*\bar{\rho}^1_t)(\bar{X}^1_t)\,dt\,,\\
\bar{\rho}^1_t(x)\,dx &= \mathcal{L}(\bar{X}^1_t)\,,\\
d\bar{X}^2_t &= \sqrt{2}\,d\bar{B}^2_t\;-\;b(\bar{X}^2_t)\,dt\;-\;(K*\bar{\rho}^2_t)(\bar{X}^2_t)\,dt\,,\\
\bar{\rho}^2_t(x)\,dx &= \mathcal{L}(\bar{X}^2_t)\,,
\end{aligned}$$

with initial condition $\bar{\rho}_0(x_1)\,dx^1\otimes\bar{\rho}_0(x_2)\,dx^2$ provided the initial condition ρ_0^N is chaotic.

Even if the two non-linear SDE's are *autonomous*, it is not clear that $\bar{X}^1_\cdot$ and $\bar{X}^2_\cdot$ are independent. Thanks to exchangeability we know that $\tilde{Q}_1=\tilde{Q}_2$, when $\tilde{Q}_i$ denotes the law of $\bar{X}^i_\cdot$. In particular $\tilde{Q}_1\otimes\tilde{Q}_2$ is a solution of the previous system, so that if we know that this system admits only one (weak) solution, any weak limit of $Q^{2,N}$ is given by the law of two independent processes.

Notice that if $\bar{Q}_2$ is absolutely continuous w.r.t. the Wiener measure (on the product path space), Girsanov theory tells us that its drift is given by the sum of the two drifts $g_1(t,x^1)+g_2(t,x^2)$ so that the Girsanov density factorizes. Since the Wiener measure itself is a product measure, we deduce as before that $\bar{Q}_2$ is a product measure. According to what we have already recalled $\tilde{Q}_1\ll P$ iff $\int_0^T|K*\bar{\rho}^1_t|^2(\bar{X}^1_t)\,dt$ is almost surely finite. Since both marginals

of $\bar{Q}_2$ are equal to $\tilde{Q}_1$ as soon as $\int_0^T |K * \bar{\rho}_t^1|^2(\bar{X}_t^1)\,dt$ is almost surely finite, $\int_0^T (|K * \bar{\rho}_t^1|^2 + K * \bar{\rho}_t^2|^2)(\bar{\omega}_t)\,dt$ is also $\bar{Q}_2$ almost surely finite, so that $\bar{Q}_2$ is absolutely continuous, i.e. the latter property is equivalent to the absolute continuity of the marginal $\tilde{Q}_1$.
We may state

Theorem 8.9. ***The 2D vortex case with confinement.*** *Let $d = 2$, $K(x) = \chi\,\frac{x^\perp}{|x|^2}$.*
Assume that the initial condition $\mu_0^N = \rho_0^N\,dx$ is chaotic so that $\mu_0^{k,N} \to (\bar{\rho}_0\,dx^1)^{\otimes k}$, satisfies $H(\mu_0^N|\gamma_0^{\otimes N}) \le CN$ and finally satisfies $\int |x|^2\,\bar{\rho}_0(x)dx < +\infty$. Also assume that the additional drift b is a confining potential as in Theorem 5.7.
Then Q^N is chaotic and for each k, $Q^{k,N}$ weakly converges to $(\bar{Q})^{\otimes k}$ where $\bar{Q}$ is the unique solution of the non linear SDE (1.3) with initial condition $\bar{\rho}_0\,dx$ satisfying $\int_0^T I(\bar{Q}\circ\bar{\omega}_t^{-1})dt < +\infty$.

Proof. Conditions on b and the initial measure ensure that $\sup_N H(Q^{2,N}|\bar{P}^{\otimes 2}) < +\infty$ according to Theorem 8.1, and we deduce from the lower semi-continuity of the relative entropy that $H(\bar{Q}^2|\bar{P}^{\otimes 2}) < +\infty$. It follows that $\int_0^T I(\bar{\rho}_t)dt < +\infty$. Uniqueness of the solution is thus given by Proposition 5.9 with a slight difference: if there is no confinement one has to assume that the Boltzmann entropy of $\bar{\rho}_0$ is finite, while here, lower semi continuity of the relative entropy shows that $H(\bar{\rho}_0|\gamma_0) < +\infty$, which is enough for our purpose. Notice that this uniqueness shows that $\rho^1 = \rho^2$ in the pair particles non-linear SDE. It is also classical that chaos for the first two coordinates implies chaos for any finite set of coordinates. □

In the Biot-Savart situation the proof we have given is simpler than the one in [62], but we have to add a confinement potential.

Actually, in what precedes, we used relative entropy to get uniqueness for the non linear SDE. In the 2D vortex case, the uniqueness part in Proposition 5.9 only requires that $\int_0^T I(\bar{\rho}_s)ds < +\infty$. Again, according to super-additivity of the Fisher information, and lower semi continuity, this condition will be satisfied provided

$$\int_0^T I(\rho_s^N)ds \le C(T)N\,.$$

This bound is exactly what is obtained in Proposition 5.1 of [62] but, as we discussed in Remark 5.8 it requires to prove some uniform (in N) lower bound for the normalized Boltzmann entropy. We thus have the following result, which is the one obtained in [62]

Theorem 8.10. *For the 2D vortex case (without confinement), under the same assumptions, the same conclusion as in Theorem 8.9 is true if the initial condition satisfies the additional moment condition* $\sup_{N\ge N_0}\int |x^1|^2\,\rho_0^N(x)dx < +\infty$.

We turn now to the sub-Coulombic model. All what we have done for the $2D$ vortex model in the proof of Theorem 8.9 is still available, assuming this time the hypotheses in Corollary 8.7 (1), except the final uniqueness result for the non linear S.D.E. As explained at the end of subsection 5.2, our goal will be to find conditions for the flow of densities $\bar{\rho}_\cdot \in \mathbb{L}^1([0,T], \mathbb{L}^{d/d-s-2}(\mathbb{R}^d))$ and apply Proposition 4.10.

We start with a very natural Lemma we found in [101]. We recall the proof for completeness.

Lemma 8.11. *Let $\tilde{Q}$ be the law of a diffusion process*

$$X_t = X_0 + \sqrt{2}B_t + \int_0^t g(s, X_s)\, ds\,.$$

Assume that $g \in C_b^\infty([0,T]\times\mathbb{R}^d)$. Denote by m_t the density of X_t for $t \geq 0$. Assume that $m_0 \in \mathbb{L}^q(\mathbb{R}^d)$. Then for $1 \leq q \leq +\infty$,

$$||m_t||_q \leq e^{\frac{q-1}{q}\, t\, \sup_{u\leq t} ||(\nabla.g(u,.))_-||_\infty}\, ||m_0||_q\,.$$

Proof. Assume first that m_0 is bounded with compact support and $q < +\infty$. Due to our assumptions we know that $m_t \in C^1(]0,T], C^\infty(\mathbb{R}^d)) \cap \mathbb{L}^q(\mathbb{R}^d)$, has gaussian tails and that $m_t \to m_0$ as $t \to 0$ in $\mathbb{L}^q(\mathbb{R}^d)$. The following calculation is thus rigorous for $t > 0$

$$\begin{aligned}
\frac{d}{dt}\int m_t^q\, dx &= \int q\, m_t^{q-1}\, \partial_t\, m_t\, dx \\
&= q \int m_t^{q-1}\, (\Delta m_t - \nabla.(g(t,.)m_t))\, dx \\
&= -\, q(q-1) \int m_t^{q-2}\, |\nabla m_t|^2\, dx \,+\, q(q-1) \int m_t^{q-1} \nabla.m_t\, g(t,.)\, dx \\
&= -\, q(q-1) \int m_t^{q-2}\, |\nabla m_t|^2\, dx \,-\, (q-1) \int m_t^q\, \nabla.g(t,.)\, dx \\
&\leq (q-1)\, ||(\nabla.g(t,.))_-||_\infty \int m_t^q\, dx\,.
\end{aligned}$$

By Gronwall's lemma, for any $t \geq t_0 > 0$ we thus have

$$||m_t||_q \leq e^{\frac{q-1}{q}\, \sup_{t_0\leq u\leq t} ||(\nabla.g(u,.))_-||_\infty\, (t-t_0)}\, ||m_{t_0}||_q$$

and we may let $t_0 \to 0$ thanks to the continuity at the origin of the $\mathbb{L}^q$ norm. Extension to a general initial m_0 in $\mathbb{L}^q$ follows using approximation. For $q = +\infty$, we have

$$((1-\varepsilon)||m||_\infty)^{\frac{q-1}{q}}\, ||m||_q\, m^{1/q}(\mathbf{1}_{|m|\geq(1-\varepsilon)||m||_\infty}) \,\leq\, ||m||_\infty^{\frac{q-1}{q}}\, ||m||_1$$

and since for all $\varepsilon > 0$, $m(\mathbf{1}_{m\geq(1-\varepsilon)||m||_\infty}) > 0$, the result is obtained by using first $q \to +\infty$ and then $\varepsilon \to 0$, the left hand side of the previous inequality with m_t and the right hand side with m_0. □

An immediate consequence is the following. Let $\tilde{Q}$ be the law of a diffusion process as before for a general g. If there exists a sequence g_n of smooth drifts as before such that, on one hand the associated $\tilde{Q}_n$ weakly converges to $\tilde{Q}$ (or more generally the convergence holds true at the level of the marginals), on the other hand

$$\sup_n \sup_{u\leq t} ||(\nabla.g_n(u,.))_-||_\infty = S_t < +\infty\,,$$

then the time marginals of $\tilde{Q}$ have densities $m_.$ satisfying

$$||m_t||_q \leq e^{\frac{q-1}{q}\, t\, S_t}\, ||m_0||_q\,.$$

This is an immediate consequence of Riesz representation theorem and the density of continuous and bounded functions in $\mathbb{L}^{q/q-1}$ for $q > 1$, the result for $q = 1$ being obvious.

We will apply this result with $g(t,x) = -(K * \bar{\rho}_t)(x) - b(x)$ and an approximating sequence

$$g_n(t,x) = -(K_n * \bar{\rho}_t)(x) - b(x)$$

with

$$K_n(y) = \chi \, \frac{y}{|y|^{s+2} + \psi_n(|y|^2)}$$

where $\psi_n(u) = \frac{1}{n}\,\psi(nu)$ for some smooth non-increasing and non-negative ψ with $\psi(0) = 1$ and $\psi(1) = 0$. We thus have

$$\nabla.K_n(y) = \chi \, \frac{(d-s-2)|y|^{s+2} + \psi_n(|y|^2) - 2\psi_n'(|y|^2)\,|y|^2}{(|y|^{s+2} + \psi_n(|y|^2))^2} \, \leq \, 0$$

if $\chi < 0$. It follows that

$$||(\nabla.g_n(u,.))_-||_\infty \, \leq \, ||(\nabla.b)_+||_\infty \, ,$$

and since K_n, hence g_n is smooth that for all n and all $1 \leq q \leq +\infty$,

$$||m_{n,t}||_q \leq e^{\frac{q-1}{q}\,t\,||(\nabla.b)_+||_\infty} \, ||m_0||_q \, . \tag{8.7}$$

It is worth to notice that here $m_{n,t}$ denotes the density of the solution of the *linear* S.D.E with drift g_n.

If we denote by $\tilde{Q}_n$ the distribution of the solution of this linear S.D.E., it remains to show that $\tilde{Q}_n$ converges to the solution of the non-linear S.D.E. $\bar{Q}$ whose marginals flow is $\bar{\rho}_.$. Recall that, under the assumptions of Corollary 8.7 we know that $H(\bar{Q}|P) < +\infty$, in particular

$$\mathbb{E}^{\bar{Q}} \left(\int_0^T \, |K * \bar{\rho}_t|^2(\omega_t) \, dt \right) < +\infty \, .$$

It is immediate that $|K_n| \leq |K|$ and $K_n \to K$ almost everywhere as $n \to +\infty$. Applying Lebesgue bounded convergence theorem we deduce that

$$\mathbb{E}^{\bar{Q}} \left(\int_0^T \, |(K - K_n) * \bar{\rho}_t|^2(\omega_t) \, dt \right) \to 0$$

i.e. that $H(\bar{Q}|\tilde{Q}_n) \to 0$. Thanks to Pinsker inequality again, we thus have that $\tilde{Q}_n$ goes to $\bar{Q}$ in total variation distance, hence weakly.

Gathering all we have done, we deduce that any weak limit of $Q^{k,N}$ satisfies the non linear S.D.E. (1.3), and, provided the initial density $\bar{\rho}_0 \in \mathbb{L}^q$, has a marginals flow $\rho_. \in \mathbb{L}^1([0,T], \mathbb{L}^q(\mathbb{R}^d))$. According to Proposition 4.10 if $q > d/d-s-2$, there exists a unique solution of (1.3) such that its marginals flow satisfies this integrability condition, so that there exists a unique weak limit of $Q^{k,N}$. Let us summarize

Theorem 8.12. *Consider the sub-Coulombic case:* $d \geq 3$, $\chi < 0$, , *and*

$$K(x) = \chi \, \frac{x}{|x|^{s+2}} \, \mathbf{1}_{x \neq 0} \;\; \textit{for} \;\; 0 < s \leq d-2 \, ,$$

also assume that the additional drift b is bounded and Lipschitz. Assume that the initial condition $\mu_0^N = \rho_0^N \, dx$ is chaotic so that $\mu_0^{k,N} \to (\bar{\rho}_0 \, dx^1)^{\otimes k}$, and satisfies $H(\mu_0^N | \gamma_0^{\otimes N}) \leq CN$. Finally assume that for some $q > d/d-s-2$, $\bar{\rho}_0 \in \mathbb{L}^q(\mathbb{R}^d)$.

Then Q^N is chaotic and for each k, $Q^{k,N}$ weakly converges to $(\bar{Q})^{\otimes k}$ where $\bar{Q}$ is the unique solution of the non linear SDE (1.3) with initial condition $\bar{\rho}_0 \, dx$ satisfying $\int_0^T \, ||\bar{Q} \circ \omega_t^{-1}||_{\mathbb{L}^q} \, dt < +\infty$.

9. More on the $2D$ situation and the Keller-Segel model.

9.1. More on the Keller-Segel model.

In this section we consider (first) the Keller-Segel model with confinement, $d = 2$, $K(x) = \chi \frac{x}{|x|^2}$, $\chi < 4$. We have seen in the previous section that the family $Q^{k,N}$ is tight and that any weak limit $\bar{Q}$ of $Q^{1,N}$ satisfies the non linear SDE (1.3). The only remaining problem is to get some uniqueness.

In all this section we assume (at least) that $\rho_0^{N,U}$ is exchangeable and satisfies the integrability assumptions in Theorem 6.17. We generically denote by $\bar{\rho}_.$ the marginals flow of some weak limit $\bar{Q}$ of $Q^{1,N}$.

If we add a confining potential Theorem 6.17 tells us that

$$\int_0^T \left| \nabla_1 \rho_t^{N,U} + \frac{\chi}{N} \sum_{j \neq 1} \frac{x^1 - x^j}{|x^1 - x^j|^2} \right|^2 \rho_t^{N,U}(x)\, dx\, dt \leq C$$

for a constant that does not depend on N. Unfortunately we were unable to translate this property in terms of the behaviour of the first marginal. In particular we did not succeed in proving that the limiting marginal $\bar{\rho}_.$ is a free energy solution, ensuring uniqueness.

Another approach would be to assume more integrability for $\bar{\rho}_0$ as in the sub-Coulombic situation and try to show that this stronger integrability is preserved for $\bar{\rho}_t$. This would again imply uniqueness thanks to Corollary 6.22. Once again we were not able to rigorously prove this property.

We nevertheless have some integrability result for $\bar{\rho}_.$. Indeed, if the initial condition is exchangeable it is easy to see that the left hand side in the inequality of Theorem 6.17 is of size N, so that using again the variational definition of the relative entropy and the fact that $e^{-\tilde{U}} dx$ is a product measure, there exists a constant (that does not depend on N) such that

$$\int_0^T H(\rho_t^{1,N,U} dx^1 | e^{-U} dx^1)\, dt \leq C\,.$$

Using the lower semi continuity of relative entropy, Fatou's lemma and according to the discussion in Remark 6.19 we thus have, for some other C

$$\int_0^T \int_{\mathbb{R}^2} \bar{\rho}_t\, |\ln(\bar{\rho}_t)|\, dx\, dt \leq C < +\infty\,. \tag{9.1}$$

Actually one can obtain a stronger result.

Proposition 9.1. *Assume that $\rho_0^{N,U}$ is exchangeable and satisfies*

$$\int |x|^2 \rho_0^{N,U} dx = CN \quad \text{and} \quad \int \rho_0^{N,U} |\ln(\rho_0^{N,U})|\, dx \leq C'N$$

for some (finite) constants C and C'. Then

$$\sup_{t \in [0,T]} \int \bar{\rho}_t |\ln(\bar{\rho}_t)|\, dx < +\infty$$

where $\bar{\rho}_.$ denotes the marginals flow of any weak limit $\bar{Q}$ of $Q^{1,N}$ in Theorem 8.3 (2).

Proof. According to (6.14) it holds

$$H(\rho_0^{N,U}dx|e^{-\tilde{U}}dx) + \int \rho_0^{N,U}(x)\,\frac{\chi}{2N}\sum_{i<j}\ln\left(\varepsilon+|x^i-x^j|^2\right)dx \;\geq$$

$$\geq\; H(\rho_t^{N,U,\varepsilon}dx|e^{-\tilde{U}}dx) + \int \rho_t^{N,U,\varepsilon}(x)\,\frac{\chi}{2N}\sum_{i<j}\ln\left(\varepsilon+|x^i-x^j|^2\right)dx\,.$$

After simple manipulations and using exchangeability it follows

$$\begin{aligned} H(\rho_t^{N,U,\varepsilon}dx|e^{-\tilde{U}}dx) \;\leq\; & H(\rho_0^{N,U}dx|e^{-\tilde{U}}dx) + \int \rho_0^{N,U}(x)\,\frac{\chi(N-1)}{4N}\sum_{j>1}\left|\ln\left(\varepsilon+|x^1-x^j|^2\right)\right|dx \\ & + \int \rho_t^{N,U,\varepsilon}(x)\,\frac{\chi(N-1)}{4N}\sum_{j>1}\ln^-\left(\varepsilon+|x^1-x^j|^2\right)dx\,, \end{aligned} \tag{9.2}$$

where g^- denotes the negative part of g.

According to the variational formulation of relative entropy, and after a standard approximation by a bounded function, we have for any $\beta>0$

$$\beta\int \rho_t^{N,U,\varepsilon}(x)\sum_{j>1}\ln^-\left(\varepsilon+|x^1-x^j|^2\right)dx \leq H(\rho_t^{N,U,\varepsilon}dx|e^{-\tilde{U}}dx) + \tag{9.3}$$

$$+\ln\left(\int e^{\beta\sum_{j>1}\ln^-\left(\varepsilon+|x^1-x^j|^2\right)}\prod_{j>1}e^{-U(x^j)}\,e^{-U(x^1)}dx^2\,dx^N dx^1\right).$$

The last term is equal to

$$(N-1)\,\ln\left(\int\frac{1}{(\varepsilon+|x^1-x^2|^2)^\beta}\,\mathbf{1}_{\varepsilon+|x^1-x^2|^2\leq 1}\,e^{-U(x^1)}e^{-U(x^2)}dx^1\,dx^2\right) = (N-1)C(\beta)$$

where $C(\beta)$ does not depend on ε, provided $\beta<1$. We have obtained

$$\left(\frac{1}{2}-\frac{\chi(N-1)}{4\beta N}\right)\,H(\rho_t^{N,U,\varepsilon}dx|e^{-\tilde{U}}dx)\leq C(\beta,\rho_0^{N,U},N)\,.$$

If $\chi<4$ we thus get a first desired result, namely $\sup_t H(\bar{\rho}_t dx^1|e^{-U}dx^1)\leq C$, using first $H(\rho_t^{1,N,U,\varepsilon}dx^1|e^{-U}dx^1)\leq\frac{1}{N}H(\rho_t^{N,U,\varepsilon}dx|e^{-\tilde{U}}dx)$, letting ε go to 0 thanks to the lower semi continuity of the relative entropy, dividing by N and letting N go to infinity, using again the lower semi continuity of the relative entropy.

Note that (9.3)also holds for $t=0$ so that the three assumptions on $\rho_0^{N,U}$ in Theorem 6.17 reduce to the ones in the statement of this proposition, and actually one can replace $|\ln(\rho_0^{N,U})|$ by $\ln(\rho_0^{N,U})$ since the negative part of the ln is controlled by the second moment of $\rho_0^{N,U}$. Similarly $\int U\bar{\rho}_t dx^1\leq C$ is obtained, uniformly in t, using Lemma 6.14 (1) and Fatou's lemma. The final statement in the Proposition is a consequence of Remark 6.19. □

A similar argument is used in [17].

The previous Proposition is still far from the required $\mathbb{L}^q$ bound for $q > 1$ ensuring uniqueness of the non linear SDE in Corollary 6.22. Remember that the proof of this Corollary is based on the study of the *linear* SDE

$$dX_t = \sqrt{2}\, dB_t - \nabla U(X_t)dt - (K * \bar{\rho}_t)(X_t)dt$$

and the application of Theorem 3.8 in the time dependent situation $K*\bar{\rho}_t \in \mathbb{L}^\infty([0,T], \mathbb{L}^p(\mathbb{R}^2))$ for some $p > 2$.

In the general situation we are facing (where $\bar{Q}$ is a weak limit of the particle system), one may ask about the integrability of $K * \bar{\rho}_t$. This requires to work in some Orlicz spaces. We shall briefly recall the material of this theory we need for the sequel.

9.2. **The (super) basics for Orlicz spaces.**

A function $\Phi : \mathbb{R}^+ \to \mathbb{R}$ is a Young function if it is non-decreasing, convex, satisfies $\Phi(0) = 0$ and $\lim_{u\to+\infty} \Phi(u)/u = +\infty$. The convex conjugate $\Phi^*(a) = \sup_{u>0} (au - \Phi(u))$ is also a Young function.

The associated Orlicz space $\mathbb{L}_\Phi(\mathbb{R}^n)$ is the set of functions f defined on $\mathbb{R}^n$ and satisfying $\int \Phi(|f|/c)dx < +\infty$ for some $c > 0$. It is a vector space and can be equipped with two norms: the Orlicz norm (or dual norm) is defined by

$$|f|_\Phi = \sup\{ \int |g\, f|\, dx \; ; \; \int \Phi^*(|g|)\, dx \le 1\} \,,$$

and the Luxemburg norm (gauge norm) is defined as

$$N_\Phi(f) = \inf\{c > 0 \; ; \; \int \Phi(|f|/c)dx \; \le 1\} \,.$$

These norms are equivalent, more precisely $N_\Phi(f) \le |f|_\Phi \le 2\, N_\Phi(f)$. $\mathbb{L}_\phi$ equipped with these norms is a Banach space.

We shall say that a function Φ is *moderate* if, for any $\alpha > 0$ there exist $c(\alpha)$ and $C(\alpha)$ such that $c(\alpha)\Phi(u) \le \Phi(\alpha\, u) \le C(\alpha)\Phi(u)$. It is immediate that for a moderate Young function and $\alpha > 0$, $\mathbb{L}_\Phi = \mathbb{L}_{\alpha\,\Phi}$ and that the norms $N_{\alpha\Phi}$ and N_Φ are equivalent.

To any C^2 function Φ only convex at infinity and such that $\Phi(u)/u \to +\infty$ at infinity, one can associate several Young functions as follows.

First, if Φ is convex on $[c_0, +\infty)$ (we may always assume that $c_0 \ge 1$) using convexity and Taylor formula we have for $u \ge c_0$, $\Phi(c_0) - \Phi(u) \ge (c_0 - u)\, \Phi'(u)$ so that $\Phi'(u) \ge \frac{\Phi(u)}{u} + \frac{1}{u}\,(c_0\Phi'(u) - \Phi(c_0))$. Since $\Phi(u)/u \to +\infty$, $\Phi'(u) \to +\infty$ as $u \to +\infty$, so that one can find $c_1 \ge c_0$ such that $c_0\, \Phi'(c_1) \ge \Phi(c_0)$. It follows that for all $u \ge c_1$, $\Phi'(u) \ge \frac{\Phi(u)}{u}$. We can thus define for all $c \ge c_1$,

$$\tilde{\Phi}_c(u) = \frac{\Phi(c)}{c}\, u\, \mathbf{1}_{u\le c} + \Phi(u)\, \mathbf{1}_{u\ge c} \tag{9.4}$$

which is a Young function. If Φ is moderate, so is $\tilde{\Phi}_c$.

Another natural way consists in first choosing $c_0 \ge 1$ such that $\Phi'(c_0) \ge 1$ and introduce for any $c \ge c_0$

$$\bar{\Phi}_c(u) = u\, \mathbf{1}_{u\le c} + (\Phi(u) - \Phi(c))\, \mathbf{1}_{u\ge c} \tag{9.5}$$

which is again a Young function, moderate if Φ is.

If we choose the same large enough c such that $\Phi(c)/c \geq 1$ one of course has $\bar{\Phi}_c \leq \tilde{\Phi}_c$ so that

$$N_{\bar{\Phi}_c} \leq N_{\tilde{\Phi}_c}\,.$$

Conversely if we choose c large enough such that $\Phi(c)/c \geq 2$ one has $(\Phi(c)/c)\,\bar{\Phi}_c \geq \tilde{\Phi}_c$. If $\bar{\Phi}_c$ is moderate, one can find $\alpha(c)$ and $\beta(c)$ such that $\bar{\Phi}_c(u/\alpha(c)) \leq (\Phi(c)/c)\,\bar{\Phi}_c(u) \leq \bar{\Phi}_c(u/\beta(c))$ so that the norms $N_{(\Phi(c)/c)\bar{\Phi}_c}$ and $N_{\bar{\Phi}_c}$ are equivalent, hence,

$$\text{for some } a(c) > 0\,,\ a(c)\,N_{\bar{\Phi}_c} \geq N_{\tilde{\Phi}_c} \text{ provided } \Phi \text{ is moderate}\,.$$

Orlicz spaces are generalizing $\mathbb{L}^p$ spaces (corresponding to $\Phi(u) = u^p/p$ and $\Phi^*(a) = a^q/q$ with $\frac{1}{p} + \frac{1}{q} = 1$). One can thus expect that several properties are preserved. The first one immediately follows from the above construction, namely the Orlicz-Hölder inequality

$$\int |f\,g|\,dx \leq |f|_\Phi\,N_{\Phi^*}(g)\ \leq\ 2\,N_\Phi(f)\,N_{\Phi^*}(g)\,.$$

An interesting generalization is proved in Theorem 2.3 of [103], namely

$$\text{If for all } u,\ \Phi^{-1}(u)\,\Psi^{-1}(u)\ \leq\ \Theta^{-1}(u)\ \text{ then }\ N_\Theta(fg) \leq 2\,N_\Phi(f)\,N_\Psi(g)\,. \tag{9.6}$$

In the same spirit, one can extend Young's inequality for a convolution product (Theorem 2.5 in [103])

$$\text{If for all } u,\ \Phi^{-1}(u)\,\Psi^{-1}(u)\ \leq\ u\,\Theta^{-1}(u)\ \text{ then }\ N_\Theta(f*g) \leq 2\,N_\Phi(f)\,N_\Psi(g)\,. \tag{9.7}$$

Let now Φ,Ψ and Θ be convex at infinity and satisfying the condition in (9.7) for u large i.e $u \geq c$. This condition is trivially satisfied for $\bar{\Phi}_c$, $\bar{\Psi}_c$ and $\bar{\Theta}_c$ for $u \leq c$ where all inverse functions are identity.
If now condition (9.6) is satisfied for $u \geq c$, we have for $u \leq c$, $\bar{\Phi}_c^{-1}(u)\,\bar{\Psi}_c^{-1}(u)\ \leq\ c\,\bar{\Theta}_c^{-1}(u)$, so that

$$N_{\bar{\Theta}_c}(fg) \leq \kappa(c)\,N_{\bar{\Phi}_c}(f)\,N_{\bar{\Psi}_c}(g)\ \text{ for some } \kappa(c), \text{ provided } \Psi \text{ or } \Phi \text{ or } \Theta \text{ is moderate.}$$

9.3. $2D$ singular models and Orlicz spaces.

The following functions will play an important role in the sequel

Definition 9.2. For α and β both in $\mathbb{R}$, and $\gamma \geq 0$ we define

$$\Phi_{\alpha,\beta}(u) = u^2\,|\ln|^\alpha(u)\,|\ln_2|^\beta(u)\ \text{ and }\ \Psi_\gamma(u) = u\,|\ln|^\gamma(u)\,,$$

where $\ln_2(u) = \ln(|\ln(u)|)$ for large u's. These functions are C^2 and convex for large u's.

We may state a first application of (9.7)

Lemma 9.3. *Let $\bar{\rho}_.$ be a limiting flow of marginals in the Keller-Segel model. Then if $\beta > 1$, there exists a constant $k(\beta)$ such that*

$$\sup_{T \geq t > 0} \Big(N_{\bar{\Phi}_{1,-\beta}}(K * \bar{\rho}_t) + ||K * \bar{\rho}_t||_1 \Big)\ \leq k(\beta)\,.$$

Proof. On one hand $|(K\,\mathbf{1}_{|K|\le 1})*\bar\rho_t|\le 1$ so that $\sup_t ||(K\,\mathbf{1}_{|K|\le 1})*\bar\rho_t||_1\le 1$. It follows, using interpolation, that $(K\,\mathbf{1}_{|K|\le 1})*\bar\rho_t$ belongs to all Orlicz spaces with a norm uniformly bounded in t.
On the other hand it is easily seen that $K\mathbf{1}_{|K|\ge 1}\in\mathbb{L}_{\bar\Phi_{-1,-\beta}}$ for $\beta>1$ since

$$\int_{|z|\le 1}|K||\ln|^{-1}(|K|)\,|\ln_2|^{-\beta}(|K|)\,dz<+\infty$$

provided $\beta>1$. It is also immediate that $\int_{|z|\le 1}|K|dz<+\infty$ so that $\sup_t ||(K\,\mathbf{1}_{|K|\ge 1})*\bar\rho_t||_1<+\infty$ according to Young's inequality for a convolution product.

We also know that $\sup_t N_{\bar\Psi_1}(\bar\rho_t)=C<+\infty$.
A simple computation shows that for some positive constants $d(\beta)$ and $D(\beta)$ and large v's,

$$d(\beta)\,\Phi^{-1}_{-1,-\beta}(v)\le\sqrt{v}\,|\ln|^{1/2}(v)\,|\ln_2|^{\beta/2}(v)\;\le D(\beta)\,\Phi^{-1}_{-1,-\beta}(v)\,,$$

and similarly

$$d_1\,\Psi_1^{-1}(v)\le v\,|\ln|^{-1}(v)\;\le D_1\,\Psi_1^{-1}(v)\,.$$

It follows that, for v large,

$$\Phi^{-1}_{-1,-\beta}(v)\,\Psi_1^{-1}(v)\le C\,v\,\Phi^{-1}_{1,-\beta}(v)\,.$$

Since $\Phi_{\alpha,\beta}$ and Ψ_1 are moderate, the result follows from the previous subsection. □

Remark 9.4. According to the properties described in the previous subsection, $K*\bar\rho_t$ belongs to all the $\mathbb{L}_{\bar\Phi_{\alpha,\beta}}$ for $\alpha<1$ and any β. ♢

The next step is to extend Theorem 3.8 in dimension $d=2$ to the case where the drift $g\in\mathbb{L}_\Theta(\mathbb{R}^2)$ for some $\Theta(u)\gg u^2$ but $\ll u^p$ for $p>2$, and large u's.

To this end, following the proof of this Theorem we see that it is enough to choose A such that $\sup_{0<t\le T}\,N_\Theta(|g_t|\,\mathbf{1}_{|g_t|>A})$ is small enough (depending on universal constants like the one in the equation below) and to prove that for any probability density ρ

$$N_{\Theta^*}(\rho)\;\le\;C\,(I(\rho)+1)\,. \tag{9.8}$$

Remark 9.5. The Keller-Segel case.
Denote $g_t=K*\bar\rho_t$. Then $g_t^M=(K\,\mathbf{1}_{|K|\le M})*\bar\rho_t$ can be used in the proof of Theorem 3.8. One can however slightly modify the proof of (3.8) in order to get uniform estimates in t. Indeed for $M>A$,

$$\begin{aligned}\int|g_t^M|^2\rho_t^M dx\;&\le\;2\int|g_t^A|^2\rho_t^M dx\;+\;2\int|(K\,\mathbf{1}_{A<|K|\le M})*\bar\rho_t|^2\rho_t^M\,dx\\&\le\;2A^2+C\,N_{\bar\Phi_{1,-\beta}}((K\,\mathbf{1}_{A<|K|\le M})*\bar\rho_t)\,N_{\Theta^*}(\rho_t^M)\end{aligned}$$

for $\Theta(u)=u\,|\ln|(u)\,|\ln_2|^{-\beta}(u)$ which is convex for large u's and $\beta>1$, C being a universal constant. According to the proof of Lemma 9.3, it thus suffices to choose A in such a way that $N_{\bar\Phi_{-1,-\beta}}(K\,\mathbf{1}_{A<|K|})$ is small enough to get the desired uniform in t bound. ♢

It remains to come back to Lemma 3.1 and its proof to get the desired (9.8).
Let ρ be a density of probability on $\mathbb{R}^2$. In the sequel C denotes a generic constant that may change from a line to the other.

First for $f \in \mathbb{L}_\Upsilon$ it holds

$$\int |f\,\nabla\rho|\,dx \quad \leq \quad (I(\rho))^{\frac{1}{2}}\,\left(\int \rho\,f^2\,dx\right)^{\frac{1}{2}}$$

so that we would apply Orlicz-Hölder inequality with f^2 and $M(u) = \Upsilon(\sqrt{u})$. Since in general the latter is not a Young function we have to modify it provided M is convex at infinity, i.e. $u\Upsilon''(u) - \Upsilon'(u) > 0$ for large u, and such that $M(u)/u \to +\infty$ at infinity, i.e. $\Upsilon(u)/u^2 \to +\infty$.
For simplicity we will consider $\Upsilon(u) = \kappa^{-1}\,\Phi_{\alpha,\beta}(u) = \kappa^{-1}\,u^2\,|\ln|^\alpha(u)\,|\ln_2|^\beta(u)$ for $\alpha > 0$ and large u's. In this case

$$M(u) \sim \kappa^{-1}\,u|\ln|^\alpha(u)|\ln_2|^\beta(u)$$

where $h(u) \sim m(u)$ means that for some constants c and C and all large enough u, $c\,h(u) \leq m(u) \leq C\,h(u)$.
If we consider any Young function A_κ which is equal to $\kappa^{-1}\,u|\ln|^\alpha(u)|\ln_2|^\beta(u)$ for large u, we have, still for large u's,

$$A^*_\kappa(u) \sim \Theta_{\kappa,\alpha,\beta}(u) = u\,e^{\kappa^{1/\alpha}\,\alpha^{\beta/\alpha}\,u^{1/\alpha}\,|\ln|^{-\beta/\alpha}(u)}\,,$$

which is not moderate so that we have to be careful. The good way to overcome this difficulty is to remark that $\Theta^*_{\kappa,\alpha,\beta} \sim A_\kappa$, so that we will go the other way, starting with $M^*_\kappa = \tilde{\Theta}_{\kappa,\alpha,\beta}$, obtaining then $M = (M^*_\kappa)^*$ and finally $\Upsilon(u) = M(u^2)$ which is a Young function since M is. Of course $\Upsilon \sim \tilde{\Phi}_{\alpha,\beta}$ is moderate so that we do not change $\mathbb{L}_\Upsilon$ replacing it by $\mathbb{L}_{\tilde{\Phi}_{\alpha,\beta}}$ and only replace the corresponding norm by an equivalent one.

Notice that $c\,N_\Upsilon(f) \leq N_M^{\frac{1}{2}}(f^2) \leq C\,N_\Upsilon(f)$ for some positive constants c and C.
We have thus obtained, for any $\kappa > 0$, and $f \in \mathbb{L}_\Upsilon$

$$\int |f\,\nabla\rho|\,dx \quad \leq \quad C(\kappa,\alpha,\beta)\,(I(\rho))^{\frac{1}{2}}\,|f|_{\tilde{\Phi}_{\alpha,\beta}}\,N^{\frac{1}{2}}_{(\tilde{\Theta}_{\kappa,\alpha,\beta})}(\rho)$$

Since $\Phi^*_{\alpha,\beta} \sim \Phi_{-\alpha,-\beta}$, we have thus obtained an extension of the first inequality in the proof of Lemma 3.1, i.e. for $\alpha > 0$ and $\kappa > 0$, there exists some $C(\kappa,\alpha,\beta)$ such that

$$N_{\tilde{\Phi}_{-\alpha,-\beta}}(\nabla\rho) \leq C(\kappa,\alpha,\beta)\,(I(\rho))^{\frac{1}{2}}\,\left(N_{\tilde{\Theta}_{\kappa,\alpha,\beta}}(\rho)\right)^{\frac{1}{2}}\,, \tag{9.9}$$

where for large u's

$$\Theta_{\kappa,\alpha,\beta}(u) = ue^{\kappa^{1/\alpha}\,\alpha^{\beta/\alpha}\,u^{1/\alpha}\,|\ln|^{-\beta/\alpha}(u)}\,.$$

The second tool in the proof of Lemma 3.1 is Sobolev inequality (we did not write with the optimal constant but with an easy to read reference). Several extensions to Orlicz spaces are known, under the name of Orlicz-Sobolev inequalities. Let us mention the early [52] (with a nice recent proof in [84]) and the more recent and sharp results in [48, 49]. We will use Theorem 1. in [48], we recall below in the case of the dimension $n = 2$, so that $n' = \frac{n}{n-1} = 2$ in (1.2) of [48].

Theorem 9.6. *Let A be a Young function such that $\int_0 \frac{A^*(t)}{t^3}\,dt < +\infty$. Define successively*

$$B(s) = \int_0^s \frac{A^*(t)}{t^3}\,dt$$

and, provided $B(s) \to +\infty$ as $s \to +\infty$,

$$D(s) = \int_0^s t\,(B^{-1}(t^2))^2\,dt\,.$$

Then there exists a constant C such that

$$N_D(\rho) \;\le\; C\,N_A(\nabla\rho)\,.$$

We will apply this Theorem with $A = \tilde{\Phi}_{-\alpha,-\beta}$. Notice that $A^*(u) = 0$ for u close to 0 so that the necessary condition in the Theorem is satisfied. We already know that $A^* \sim \Phi_{\alpha,\beta}$. We thus have

$$B(s) = C + \int_c^s \frac{|\ln|^\alpha(t)\,|\ln_2|^\beta(t)}{t}\,dt\,,$$

for some constants C and c.

Integrating by parts it is easily seen that for $\alpha > 0$, $B(s) \sim |\ln|^{1+\alpha}(s)\,|\ln_2|^\beta(s)$ so that $B^{-1}(u)$ is sandwiched between two functions of the form

$$M'(\alpha,\beta)\,\exp\left(M(\alpha,\beta)\,u^{\frac{1}{1+\alpha}}\,|\ln|^{-\frac{\beta}{1+\alpha}}(u)\right)\,.$$

One next checks that

$$D(t) \sim t^{\frac{2\alpha}{1+\alpha}}\,|\ln|^{2\beta/(1+\alpha)}(t)\,\exp\left(2M(\alpha,\beta)\,t^{\frac{2}{1+\alpha}}\,|\ln|^{-\frac{2\beta}{1+\alpha}}(t)\right)\,.$$

It remains to compare $\tilde{\Theta}_{\kappa,\alpha,\beta}$ with D. First write

$$N_{\tilde{D}}(\rho) \le N_{\tilde{D}}(\rho\,\mathbf{1}\rho \le c) + N_{\tilde{D}}(\rho\,\mathbf{1}_{\rho\ge c}) \le C(\alpha,\beta) + N_D(\rho\,\mathbf{1}_{\rho\ge c})$$

since ρ is a probability density, c being given in the construction of $\tilde{D}$ i.e. $D(u) = \tilde{D}(u)$ for $u \ge c$. For $\alpha \ge 1$, we may choose κ small enough in order to have for large t's

$$\begin{aligned}
\Theta_{\kappa,\alpha,\beta}(t) \;&\le\; t\,\exp\left(M(\alpha,\beta)\,t^{\frac{1}{\alpha}}\,|\ln|^{-\frac{\beta}{\alpha}}(t)\right)\\
&\le\; c(\alpha,\beta)\,t^{\frac{2\alpha}{1+\alpha}}\,|\ln|^{2\beta/(1+\alpha)}(t)\exp\left(2M(\alpha,\beta)\,t^{\frac{2}{1+\alpha}}\,|\ln|^{-\frac{2\beta}{1+\alpha}}(t)\right)\\
&\le\; D(t)\,.
\end{aligned}$$

We have to use some $c(\alpha,\beta)$ since D is only similar to the function without this constant. We thus have $\tilde{\Theta}_{\kappa,\alpha,\beta} \le \tilde{D}$ everywhere by construction, so that $N_{\tilde{\Theta}_{\kappa,\alpha,\beta}} \le N_{\tilde{D}}$. Gathering all what precedes we thus have

$$N_{\tilde{\Theta}_{\kappa,\alpha,\beta}}(\rho) \le C(\alpha,\beta) + N_D(\rho) \tag{9.10}$$

and finally using (9.10), (9.9) and Theorem 9.6

$$N_{\tilde{\Phi}_{-\alpha,\beta}}(\nabla\rho) \le C\,(I(\rho)+1)$$

and finally

$$N_D(\rho) \le C\,(I(\rho)+1)$$

as expected.

We may thus state the following improvement of Theorem 3.8

Theorem 9.7. *In dimension $d = 2$ consider the SDE, $X_t = X_0 + \sqrt{2}\, B_t + \int_0^t 2g_s(X_s)ds$. Assume that $H(\nu_0|\gamma_0) < +\infty$ (recall Theorem 3.8) and that for A large enough,*

$$\sup_{0<t\leq T} N_{\Phi_{\alpha,\beta}}(g_t\, \mathbf{1}_{|g_t|>A}) < +\infty$$

for some $\alpha \geq 1$ and $\beta \in \mathbb{R}$. Then the previous SDE has a unique weak solution Q and this Q satisfies $H(Q|P) < +\infty$.

As a consequence the results in Theorem 8.8 extend for $d = 2$ if the interaction kernel satisfies the previous integrability condition and the additional drift b is bounded.

Proof. In the proof of Theorem 3.8 we write for $M > A$,

$$\begin{aligned}\int |g_t^M|^2 \rho_t^M dx &= \int |g_t^M\, \mathbf{1}_{|g_t^M|>A}|^2 \rho_t^M dx + \int |g_t^M\, \mathbf{1}_{|g_t^M|\leq A}|^2 \rho_t^M dx \\ &\leq A^2 + C\, N_{\tilde{\Phi}_{\alpha,-\beta}}(g_t^M\, \mathbf{1}_{|g_t^M|>A})\, N_{\Theta_{\alpha,\beta}}(\rho_t^M)\end{aligned}$$

apply (9.8) with $\Theta^* = \Theta_{\alpha,\beta}$, and conclude as for the proof of Theorem 3.8. □

According to Remark 9.5 the case $g_t = K * \bar{\rho}_t$ enters the framework of the previous Theorem with $\alpha = 1$ and $\beta < -1$. It follows that the *linear* SDE

$$dX_t = \sqrt{2}\, dB_t - \nabla U(X_t)dt - (K * \bar{\rho}_t)(X_t)dt$$

has an unique weak solution which is of finite relative entropy (assuming that $\int |\nabla U|^2\, \bar{\rho}_t dx < +\infty$.) This proves that any weak limit $\bar{Q}$ of $Q^{1,N}$ has finite relative entropy, hence that the weak limit is unique according to Theorem 6.20. In addition, according to the discussion at the end of subsection 8.3, since $\bar{Q}$ is absolutely continuous w.r.t. P, propagation of chaos holds true. Let us summarize all this

Theorem 9.8. *Consider the Keller-Segel case $d = 2$, $0 < \chi < 4$ and*

$$K(x) = \chi\, \frac{x}{|x|^2}\, \mathbf{1}_{x\neq 0}$$

with a confining potential U such that $\int e^{-U} dx < +\infty$, i.e. an additional drift $b = \nabla U$ and assuming that b is Lipschitz.

Assume that the initial condition $\mu_0^N = \rho_0^N dx$ is chaotic so that $\mu_0^{k,N} \to (\bar{\rho}_0\, dx^1)^{\otimes k}$, satisfies $\int \rho_0^N\, |\ln \rho_0^N| dx < CN$ and $\int |x^1|^2\, \rho_0^{1,N}\, dx^1 < +\infty$.

Then Q^N is chaotic and for each k, $Q^{k,N}$ weakly converges to $(\bar{Q})^{\otimes k}$ where $\bar{Q}$ is the unique solution of the non linear SDE (1.3) with initial condition $\bar{\rho}_0 dx$ whose marginals flow is the (unique) free energy solution of the non linear PDE (1.1).

Actually this non linear EDS admits only one solution such that its marginals flow is bounded in entropy uniformly in time.

The assumptions on ρ_0^N allow us to use Proposition 9.1, Theorem 6.17 (recall that the proof of Proposition 9.1 shows that one can reduce the assumptions on ρ_0^N in Theorem 6.17 to the ones in the previous one.

The previous Theorem holds with a quadratic confining potential, so that using the "self similar" coordinates as discussed in Remark 6.21 (at the process level) we can state

Corollary 9.9. *Theorem 9.8 is still true for the Keller-Segel model without confinement ($U = 0$).*

References

[1] M. Arnaudon, A. B. Cruzeiro, C. Léonard, and J. C. Zambrini. An entropic interpolation problem for incompressible viscous fluids. *Ann. Inst. Henri Poincaré Probab. Stat.*, 56(3):2211–2235, 2020.

[2] R. F. Bass and Z-Q. Chen. Brownian motion with singular drift. *Ann. Probab.*, 31(2):791–817, 2003.

[3] L. Beck, F. Flandoli, M. Gubinelli, and M. Maurelli. Stochastic ODEs and stochastic linear PDEs with critical drift: regularity, duality and uniqueness. *Electron. J. Probab.*, 24:Paper No. 136, 72, 2019.

[4] J. Bedrossian and N. Masmoudi. Existence, uniqueness and Lipschitz dependence for Patlak-Keller-Segel and Navier-Stokes in $\mathbb{R}^2$ with measure-valued initial data. *Arch. Ration. Mech. Anal.*, 214(3):717–801, 2014.

[5] G. Ben Arous and O. Zeitouni. Increasing propagation of chaos for mean field models. *Ann. Inst. H. Poincaré Probab. Statist.*, 35(1):85–102, 1999.

[6] M. Ben-Artzi. Global solutions of two-dimensional Navier-Stokes and Euler equations. *Arch. Ration. Mech. Anal.*, 128(4):329–358, 1994.

[7] S. Benachour, B. Roynette, D. Talay, and P. Vallois. Nonlinear self-stabilizing processes. I. Existence, invariant probability, propagation of chaos. *Stochastic Process. Appl.*, 75(2):173–201, 1998.

[8] S. Benachour, B. Roynette, and P. Vallois. Nonlinear self-stabilizing processes. II. Convergence to invariant probability. *Stochastic Process. Appl.*, 75(2):203–224, 1998.

[9] A. Blanchet, J. A. Carrillo, and N. Masmoudi. Infinite time aggregation for the critical Patlak-Keller-Segel model in $\mathbb{R}^2$. *Comm. Pure Appl. Math.*, 61(10):1449–1481, 2008.

[10] A. Blanchet, J. Dolbeault, and B. Perthame. Two-dimensional Keller-Segel model: optimal critical mass and qualitative properties of the solutions. *Electron. J. Differential Equations*, page No. 44, 2006.

[11] S. G. Bobkov. Upper bounds for Fisher information. *Electron. J. Probab.*, 27:Paper No. 115, 44, 2022.

[12] F. Bolley. Limite de champ moyen de systèmes de particules. In *Seminaire: Equations aux Dérivées Partielles. 2009–2010*, Sémin. Équ. Dériv. Partielles, pages Exp. No. XXXI, 15. École Polytech., Palaiseau, 2012.

[13] F. Bolley, D. Chafaï, and J. Fontbona. Dynamics of a planar Coulomb gas. *Ann. Appl. Probab.*, 28(5):3152–3183, 2018.

[14] F. Bolley and C. Villani. Weighted Csiszár-Kullback-Pinsker inequalities and applications to transportation inequalities. *Ann. Fac. Sci. Toulouse Math. (6)*, 14(3):331–352, 2005.

[15] M. Bossy and D. Talay. Convergence rate for the approximation of the limit law of weakly interacting particles: application to the Burgers equation. *Ann. Appl. Probab.*, 6(3):818–861, 1996.

[16] J. Boursier, D. Chafaï, and C. Labbé. Universal cutoff for Dyson Ornstein Uhlenbeck process. *Probab. Theory Related Fields*, 185(1-2):449–512, 2023.

[17] D. Bresch, P-E. Jabin, and Z. Wang. Mean field limit and quantitative estimates with singular attractive kernels. *Duke Math. J.*, 172(13):2591–2641, 2023.

[18] E. A. Carlen. Conservative diffusions. *Comm. Math. Phys.*, 94(3):293–315, 1984.

[19] E. A. Carlen. Superadditivity of Fisher's information and logarithmic Sobolev inequalities. *J. Funct. Anal.*, 101(1):194–211, 1991.

[20] J. A. Carrillo, L. Chen, J-G. Liu, and J. Wang. A note on the subcritical two dimensional Keller-Segel system. *Acta Appl. Math.*, 119:43–55, 2012.

[21] J. A. Carrillo, S. Lisini, and E. Mainini. Uniqueness for Keller-Segel-type chemotaxis models. *Discrete Contin. Dyn. Syst.*, 34(4):1319–1338, 2014.

[22] J. A. Carrillo, R. J. McCann, and C. Villani. Kinetic equilibration rates for granular media and related equations: entropy dissipation and mass transportation estimates. *Rev. Mat. Iberoamericana*, 19(3):971–1018, 2003.

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

[24] P. Cattiaux. Hypoellipticité et hypoellipticité partielle pour les diffusions avec une condition frontière. *Ann. Inst. H. Poincaré Probab. Statist.*, 22(1):67–112, 1986.

[25] P. Cattiaux. Calcul stochastique et opérateurs dégénérés du second ordre. I. Résolvantes, théorème de Hörmander et applications. *Bull. Sci. Math.*, 114(4):421–462, 1990.
[26] P. Cattiaux. Stochastic calculus and degenerate boundary value problems. *Ann. Inst. Fourier (Grenoble)*, 42(3):541–624, 1992.
[27] P. Cattiaux. A pathwise approach of some classical inequalities. *Potential Anal.*, 20(4):361–394, 2004.
[28] P. Cattiaux. Hypercontractivity for perturbed diffusion semigroups. *Ann. Fac. Sci. Toulouse Math. (6)*, 14(4):609–628, 2005.
[29] P. Cattiaux, G. Conforti, I. Gentil, and C. Léonard. Time reversal of diffusion processes under a finite entropy condition. *Ann. Inst. Henri Poincaré Probab. Stat.*, 59(4):1844–1881, 2023.
[30] P. Cattiaux and M. Fradon. Entropy, reversible diffusion processes, and Markov uniqueness. *J. Funct. Anal.*, 138(1):243–272, 1996.
[31] P. Cattiaux and A. Guillin. On quadratic transportation cost inequalities. *J. Math. Pures Appl. (9)*, 86(4):341–361, 2006.
[32] P. Cattiaux and A. Guillin. Semi log-concave Markov diffusions. In *Séminaire de Probabilités XLVI*, volume 2123 of *Lecture Notes in Math.*, pages 231–292. Springer, Cham, 2014.
[33] P. Cattiaux, A. Guillin, and F. Malrieu. Probabilistic approach for granular media equations in the non uniformly convex case. *Probab. Theory Relat. Fields*, 140(1):19–40, 2008.
[34] P. Cattiaux and C. Léonard. Minimization of the Kullback information of diffusion processes. *Ann. Inst. H. Poincaré Probab. Statist.*, 30(1):83–132, 1994.
[35] P. Cattiaux and C. Léonard. Correction to: "Minimization of the Kullback information of diffusion processes" [Ann. Inst. H. Poincaré Probab. Statist. **30** (1994), no. 1, 83–132; MR1262893 (95d:60056)]. *Ann. Inst. H. Poincaré Probab. Statist.*, 31(4):705–707, 1995.
[36] P. Cattiaux and C. Léonard. Large deviations and Nelson processes. *Forum Math.*, 7(1):95–115, 1995.
[37] P. Cattiaux and C. Léonard. Minimization of the Kullback information for some Markov processes. In *Séminaire de Probabilités, XXX*, volume 1626 of *Lecture Notes in Math.*, pages 288–311. Springer, Berlin, 1996.
[38] P. Cattiaux and L. Pédèches. The 2-D stochastic Keller-Segel particle model: existence and uniqueness. *ALEA Lat. Am. J. Probab. Math. Stat.*, 13(1):447–463, 2016.
[39] P. Cattiaux and F. Petit. Singular diffusions, time reversal and applications to Fokker-Planck equations. Not to be published. Available on https://perso.math.univ-toulouse.fr/cattiaux/publications-2/, 2001.
[40] P. Cattiaux, S. Roelly, and H. Zessin. Une approche gibbsienne des diffusions browniennes infinidimensionnelles. *Probab. Theory Related Fields*, 104(2):147–179, 1996.
[41] E. Cépa and D. Lépingle. Diffusing particles with electrostatic repulsion. *Probab. Theory Related Fields*, 107(4):429–449, 1997.
[42] E. Cépa and D. Lépingle. Brownian particles with electrostatic repulsion on the circle: Dyson's model for unitary random matrices revisited. *ESAIM Probab. Statist.*, 5:203–224, 2001.
[43] E. Cépa and D. Lépingle. No multiple collisions for mutually repelling Brownian particles. In *Séminaire de Probabilités XL*, volume 1899 of *Lecture Notes in Math.*, pages 241–246. Springer, Berlin, 2007.
[44] L. P. Chaintron and A. Diez. Propagation of chaos: a review of models, methods and applications. I. Models and methods. *Kinet. Relat. Models*, 15(6):895–1015, 2022.
[45] L. P. Chaintron and A. Diez. Propagation of chaos: a review of models, methods and applications. II. Applications. *Kinet. Relat. Models*, 15(6):1017–1173, 2022.
[46] T. Chan. The Wigner semi-circle law and eigenvalues of matrix-valued diffusions. *Probab. Theory Related Fields*, 93(2):249–272, 1992.
[47] A. Chodron de Courcel, M. Rosenzweig, and S. Serfaty. The attractive log gas: stability, uniqueness, and propagation of chaos. Available on ArXiv: 2311.14560 [math.AP], 2023.
[48] A. Cianchi. A sharp embedding theorem for Orlicz-Sobolev spaces. *Indiana Univ. Math. J.*, 45(1):39–65, 1996.
[49] A. Cianchi. Optimal Orlicz-Sobolev embeddings. *Rev. Mat. Iberoamericana*, 20(2):427–474, 2004.
[50] L. Colombani and P. Le Bris. Propagation of chaos in mean field networks of FitzHugh-Nagumo neurons. *Mathematical Neuroscience and Appl.*, 3, 2023.
[51] E. B. Davies. *Heat kernels and spectral theory*, volume 92 of *Camb. Tracts Math.* Cambridge etc.: Cambridge University Press, 1989.

[52] T. K. Donaldson and N. S. Trudinger. Orlicz-Sobolev spaces and imbedding theorems. *J. Functional Analysis*, 8:52–75, 1971.

[53] A. Eberle, A. Guillin, and R. Zimmer. Quantitative Harris-type theorems for diffusions and McKean-Vlasov processes. *Trans. Amer. Math. Soc.*, 371(10):7135–7173, 2019.

[54] G. Egaña Fernández and S. Mischler. Uniqueness and long time asymptotic for the Keller-Segel equation: the parabolic-elliptic case. *Arch. Ration. Mech. Anal.*, 220(3):1159–1194, 2016.

[55] X. Feng and Z. Wang. Quantitative propagation of chaos for $2d$ vortex model on the whole space. Available on ArXiv: 2310.05156 [math.AP], 2023.

[56] M. Fischer and G. Nappo. On the moments of the modulus of continuity of Itô processes. *Stoch. Anal. Appl.*, 28(1):103–122, 2010.

[57] H. Föllmer. An entropy approach to the time reversal of diffusion processes. In *Stochastic differential systems (Marseille-Luminy, 1984)*, volume 69 of *Lect. Notes Control Inf. Sci.*, pages 156–163. Springer, Berlin, 1985.

[58] H. Föllmer. Time reversal on Wiener space. In *Stochastic processes—mathematics and physics (Bielefeld, 1984)*, volume 1158 of *Lecture Notes in Math.*, pages 119–129. Springer, Berlin, 1986.

[59] J. Fontbona. A probabilistic interpretation and stochastic particle approximations of the 3-dimensional Navier-Stokes equations. *Probab. Theory Related Fields*, 136(1):102–156, 2006.

[60] J. Fontbona and B. Jourdain. A trajectorial interpretation of the dissipations of entropy and Fisher information for stochastic differential equations. *Ann. Probab.*, 44(1):131–170, 2016.

[61] J. Fontbona and S. Méléard. A random space-time birth particle method for 2D vortex equations with external field. *Math. Comp.*, 77(263):1525–1558, 2008.

[62] N. Fournier, M. Hauray, and S. Mischler. Propagation of chaos for the 2D viscous vortex model. *J. Eur. Math. Soc. (JEMS)*, 16(7):1423–1466, 2014.

[63] N. Fournier and B. Jourdain. Stochastic particle approximation of the Keller-Segel equation and two-dimensional generalization of Bessel processes. *Ann. Appl. Probab.*, 27(5):2807–2861, 2017.

[64] N. Fournier and Y. Tardy. Collisions of the supercritical Keller-Segel particle system. To appear in J. Eur. Math. Soc. Available on ArXiv: 2110.08490 [math.PR], 2021.

[65] N. Fournier and Y. Tardy. A simple proof of non-explosion for measure solutions of the Keller-Segel equation. *Kinet. Relat. Models*, 16(2):178–186, 2023.

[66] M. Fukushima, Y. Ōshima, and M. Takeda. *Dirichlet forms and symmetric Markov processes*, volume 19 of *De Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, 1994.

[67] I. Gentil, C. Léonard, L. Ripani, and L. Tamanini. An entropic interpolation proof of the HWI inequality. *Stochastic Process. Appl.*, 130(2):907–923, 2020.

[68] D. Godinho and C. Quininao. Propagation of chaos for a sub-critical Keller-Segel model. *Ann. Inst. Henri Poincaré. Prob. Stat.*, 51(3):965–992, 2015.

[69] A. Guillin, P. Le Bris, and P. Monmarché. Uniform in time propagation of chaos for the $2d$ vortex model and other singular stochastic systems. Available on ArXiv: 2108.08675 [math.PR], 2022.

[70] A. Guillin, P. Le Bris, and P. Monmarché. On systems of particles in singular repulsive interaction in dimension one: log and Riesz gas. *J. Éc. polytech. Math.*, 10:867–916, 2023.

[71] Y. Han. Entropic propagation of chaos for mean field diffusion with L^p interactions via hierarchy, linear growth and fractional noise. Available on ArXiv: 2205.02772v4 [math.PR], 2022.

[72] Y. Han. Solving McKean-Vlasov SDES via relative entropy. Available on ArXiv: 2204.05709v3 [math.PR], 2022.

[73] Z. Hao, M. Röckner, and X. Zhang. Strong convergence of propagation of chaos for McKean-Vlasov SDE's with singular interactions. Available on ArXiv: 2204.07952v2 [math.PR], 2022.

[74] M. Hauray and S. Mischler. On Kac's chaos and related problems. *J. Funct. Anal.*, 266(10):6055–6157, 2014.

[75] X. Huang and F-Y. Wang. Distribution dependent SDEs with singular coefficients. *Stochastic Process. Appl.*, 129(11):4747–4770, 2019.

[76] P-E. Jabin and Z. Wang. Mean field limit and propagation of chaos for Vlasov systems with bounded forces. *J. Funct. Anal.*, 271(12):3588–3627, 2016.

[77] P-E. Jabin and Z. Wang. Mean field limit for stochastic particle systems. In *Active particles. Vol. 1. Advances in theory, models, and applications*, Model. Simul. Sci. Eng. Technol., pages 379–402. Birkhäuser/Springer, Cham, 2017.

[78] P-E. Jabin and Z. Wang. Quantitative estimates of propagation of chaos for stochastic systems with $W^{-1,\infty}$ kernels. *Invent. Math.*, 214(1):523–591, 2018.
[79] P. Jin. Brownian motion with singular time-dependent drift. *J. Theoret. Probab.*, 30(4):1499–1538, 2017.
[80] E. Keller and L. A. Segel. Initiation of slime mold aggregation viewed as an instability. *J. theor. Biol.*, 26:399–415, 1970.
[81] E. Keller and L. A. Segel. Model for chemotaxis. *J. theor. Biol.*, 30:225–234, 1971.
[82] D. Kinzebulatov. Form boundedness and SDES with singular drift. Available on ArXiv: 2305.00146 v.3 [math.PR], 2023.
[83] D. Kinzebulatov. On particle systems and critical strengths of general singular interactions. Available on ArXiv: 2402.17009 v.1 [math.PR], 2024.
[84] H. Kone. A short proof of the Orlicz-Sobolev inequality. *Adv. in Appl. Math.*, 107:116–124, 2019.
[85] N. Krylov and M. Röckner. Strong solutions of stochastic equations with singular time dependent drift. *Probab. Theory Relat. Fields*, 131(2):154–196, 2005.
[86] N. V. Krylov. On diffusion processes with drift in L_d. *Probab. Theory Related Fields*, 179(1-2):165–199, 2021.
[87] N. V. Krylov. On stochastic equations with drift in L_d. *Ann. Probab.*, 49(5):2371–2398, 2021.
[88] N. V. Krylov. On stochastic Itô processes with drift in L_d. *Stochastic Process. Appl.*, 138:1–25, 2021.
[89] D. Lacker. On a strong form of propagation of chaos for McKean-Vlasov equations. *Electron. Commun. Probab.*, 23:Paper No. 45, 11, 2018.
[90] D. Lacker. Hierarchies, entropy and quantitative propagation of chaos for mean field diffusions. *Probability and Mathematical Physics.*, 4(2):377–432, 2023.
[91] D. Lacker and L. Le Flem. Sharp uniform in time propagation of chaos. *Probab. Theory Relat. Fields*, 187:443–480, 2023.
[92] C. Léonard. Girsanov theory under a finite entropy condition. In *Séminaire de Probabilités XLIV*, volume 2046 of *Lecture Notes in Math.*, pages 429–465. Springer, Heidelberg, 2012.
[93] C. Léonard. Some properties of path measures. In *Séminaire de Probabilités XLVI*, volume 2123 of *Lecture Notes in Math.*, pages 207–230. Springer, Cham, 2014.
[94] C. Léonard. A survey of the Schrödinger problem and some of its connections with optimal transport. *Discrete Contin. Dyn. Syst.*, 34(4):1533–1574, 2014.
[95] C. Léonard. Feynman-Kac formula under a finite entropy condition. *Probab. Theory Related Fields*, 184(3-4):1029–1091, 2022.
[96] J-G. Liu and R. Yang. Propagation of chaos for the Keller-Segel equation with a logarithmic cut-off. *Methods Appl. Anal.*, 26(4):319–347, 2019.
[97] Z. M. Ma and M. Röckner. *Introduction to the theory of (nonsymmetric) Dirichlet forms*. Universitext. Springer-Verlag, Berlin, 1992.
[98] C. Marchioro and M. Pulvirenti. Hydrodynamics in two dimensions and vortex theory. *Comm. Math. Phys.*, 84(4):483–503, 1982.
[99] S. Méléard. Asymptotic behaviour of some interacting particle systems: Mc Kean-Vlasov and Boltzmann models. *Probabilistic models for non linear partial differential equations. ed. Talay and Tubaro, Lecture Notes in Mathematics*, 1627:42–95, 1995.
[100] P.-A. Meyer and W. A. Zheng. Construction de processus de Nelson réversibles. In *Séminaire de probabilités, XIX, 1983/84*, volume 1123 of *Lecture Notes in Math.*, pages 12–26. Springer, Berlin, 1985.
[101] P. Monmarché, Z. Ren, and S. Wang. Time-uniform log-sobolev inequalities and applications to propagation of chaos. Available on ArXiv: 2401.07966 [math.PR], 2024.
[102] C. Olivera, A. Richard, and M. Tomašević. Quantitative particle approximation of nonlinear Fokker-Planck equations with singular kernel. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 24(2):691–749, 2023.
[103] R. O'Neil. Fractional integration in Orlicz spaces. I. *Trans. Amer. Math. Soc.*, 115:300–328, 1965.
[104] H. Osada. A stochastic differential equation arising from the vortex problem. *Proc. Japan Acad. Ser. A Math. Sci.*, 61(10):333–336, 1985.
[105] H. Osada. Propagation of chaos for the two-dimensional Navier-Stokes equation. *Proc. Japan Acad. Ser. A Math. Sci.*, 62(1):8–11, 1986.
[106] R. T. Rockafellar. Integrals which are convex functionals. *Pacific J. Math.*, 24:525–539, 1968.
[107] M. Röckner and X. Zhang. Well-posedness of distribution dependent SDEs with singular drifts. *Bernoulli*, 27(2):1131–1158, 2021.

[108] L. C. G. Rogers and Z. Shi. Interacting Brownian particles and the Wigner law. *Probab. Theory Related Fields*, 95(4):555–570, 1993.

[109] M. Rosenzweig and S. Serfaty. Global in time mean-field convergence for singular Riesz-type diffusive flows. *Ann. Appl. Probab.*, 33(2):954–998, 2023.

[110] A. S. Sznitman. Topics in propagation of chaos. In *École d'Été de Probabilités de Saint-Flour XIX—1989*, volume 1464 of *Lecture Notes in Math.*, pages 165–251. Springer, Berlin, 1991.

[111] S. Takanobu. On the existence and uniqueness of SDE describing an n-particle system interacting via a singular potential. *Proc. Japan Acad. Ser. A Math. Sci.*, 61(9):287–290, 1985.

[112] Y. Tardy. Convergence of the empirical measure for the Keller-Segel model in both subcritical and critical cases. *Kinet. Relat. Models.*, 16:178–186, 2023.

[113] M. Tomašević. Propagation of chaos for stochastic particle systems with singular mean-field interaction of $L^q - L^p$ type. *Electron. Commun. Probab.*, 28:1–13, 2023.

[114] D. Trevisan. Well-posedness of multidimensional diffusion processes with weakly differentiable coefficients. *Electron. J. Probab.*, 21:Paper No. 22, 41, 2016.

[115] B. Tschiderer and L. Chun Yeung. A trajectorial approach to relative entropy dissipation of McKean-Vlasov diffusions: gradient flows and HWBI inequalities. *Bernoulli*, 29(1):725–756, 2023.

[116] D. Wei. Global well-posedness and blow-up for the 2-D Patlak-Keller-Segel equation. *J. Funct. Anal.*, 274(2):388–401, 2018.

Patrick CATTIAUX,, Institut de Mathématiques de Toulouse. Université de Toulouse. CNRS UMR 5219., 118 route de Narbonne, F-31062 Toulouse cedex 09.

Email address: `patrick.cattiaux@math.univ-toulouse.fr`