Author manuscript, published in "Mathematical Models and Methods in Applied Sciences 21, 11 (2011) 2179-2210"

hal-00521216, version 2 - 1 Dec 2010

Stochastic Mean-Field Limit: Non-Lipschitz Forces & Swarming

François Bolley José A. Cañizo José A. Carrillo

September 24, 2010

Abstract

We consider general stochastic systems of interacting particles with noise which are relevant as models for the collective behavior of animals, and rigorously prove that in the mean-field limit the system is close to the solution of a kinetic PDE. Our aim is to include models widely studied in the literature such as the Cucker-Smale model, adding noise to the behavior of individuals. The difficulty, as compared to the classical case of globally Lipschitz potentials, is that in several models the interaction potential between particles is only locally Lipschitz, the local Lipschitz constant growing to infinity with the size of the region considered. With this in mind, we present an extension of the classical theory for globally Lipschitz interactions, which works for only locally Lipschitz ones.

Keywords. Mean-field limit, diffusion, Cucker-Smale, collective behavior

1 Introduction

The formation of large-scale structures (patterns) without the need of leadership (self-organization) is one of the most interesting and not completely understood aspect in the collective behavior of certain animals, such as birds, fish or insects. This phenomena has attracted lots of attention in the scientific community, see [8, 14, 33, 37] and the references therein.

F. Bolley: Ceremade, UMR CNRS 7534, Université Paris-Dauphine, Place du Maréchal De Lattre De Tassigny, F-75016 Paris, France; e-mail: bolley@ceremade.dauphine.fr

J. A. Cañizo: Departament de Matemàtiques, Universitat Autònoma de Barcelona, E-08193 Bellaterra, Spain; e-mail: canizo@mat.uab.es

J. A. Carrillo: Institució Catalana de Recerca i Estudis Avançats and Departament de Matemàtiques, Universitat Autònoma de Barcelona, E-08193 Bellaterra, Spain; e-mail: carrillo@mat.uab.es

Mathematics Subject Classification (2010): Primary 82C22; Secondary 82C40, 35Q92

hal-00521216, version 2 - 1 Dec 2010

Most of the proposed models in the literature are based on particle-like description of a set of large individuals; these models are called Individual-Based Models (IBM). IBMs typically include several interactions between individuals depending on the species, the precise mechanism of interaction of the animals and their particular biological environment. However, most of these IBMs include at least three basic effects: a short-range repulsion, a long-range attraction and a "mimicking" behavior for individuals encountered in certain spatial regions. This so-called three-zone model was first used for describing fish schools in [2, 25] becoming a cornerstone of swarming modelling, see [3, 24].

The behavior of a large system of individuals can be studied through mesoscopic descriptions of the system based on the evolution of the probability density of finding individuals in phase space. These descriptions are usually expressed in terms of space-inhomogeneous kinetic PDEs and the scaling limit of the interacting particle system to analyze is usually called the *mean-field limit*. These kinetic equations are useful in bridging the gap between a microscopic description in terms of IBMs and macroscopic or hydrodynamic descriptions for the particle probability density. We refer to the review [12] for the different connections between these models and for a larger set of references.

The mean-field limit of deterministic interacting particle systems is a classical question in kinetic theory, and was treated in [6, 18, 32] in the case of the Vlasov equation. In these papers, the particle pairwise interaction is given by a globally bounded Lipschitz force field. Some of the recent models of swarming introduced in [19, 16, 22] do not belong to this class due to their growth at infinity leading to an interaction kernel which is only locally Lipschitz. These IBMs are kinetic models in essence since the interactions between individuals are at the level of the velocity variable to "align" their movements for instance or to impose a limiting "cruising speed". The mean-field limit for deterministic particle systems for some models of collective behavior with locally Lipschitz interactions was recently analysed in [9] showing that they follow the expected Vlasov-like kinetic equations.

On the other hand, noise at the level of the IBMs is an important issue since we cannot expect animals to react in a completely deterministic way. Therefore, including noise in these IBMs and thus, at the level of the kinetic equation is an important modelling ingredient. This stochastic mean-field limit formally leads to kinetic Fokker-Planck like equations for second order models as already pointed out in [10]. The rigorous proof of this stochastic mean-field limit has been carried out for globally Lipschitz interactions in [34, 30], see also [29].

This work is devoted to the rigorous analysis of the stochastic mean-field limit of locally Lipschitz interactions that include relevant swarming models in the literature such as those in [19, 16]. We will be concerned with searching the rate of convergence, as the number of particles $N \to \infty$, of the distribution of each of the particles and of the

hal-00521216, version 2 - 1 Dec 2010

empirical measure of the particle system to the solution of the kinetic equation. This convergence will also establish the propagation of chaos as $N \to \infty$ for the particle system and will be measured in terms of distances between probability measures. Here, we will not deal with uniform in time estimates since no stabilizing behaviour can be expected in this generality, such estimates were obtained only in a specific instance of Vlasov-Fokker-Planck equation, see [5]. The main price to pay to include possible growth at infinity of the Lipschitz constants of the interaction fields will be at the level of moment control estimates. Then, there will be a trade-off between the requirements on the interaction and the decay at infinity of the laws of the processes at the initial time.

The work is organized as follows: in the next two subsections we will make a precise descriptions of the main results of this work, given in Theorems 1.1 and 1.2 below, together with a small overview of preliminary classical well-known facts and a list of examples, variants and particular cases of applications in swarming models. The second section includes the proof of the stochastic mean-field limit of locally Lipschitz interacting particle systems under certain moment control assumptions (thus proving Theorem 1.1). Finally, the third section will be devoted to the proof of Theorem 1.2: a result of existence and uniqueness of the nonlinear partial differential equation and its associated nonlinear stochastic differential equation, for which the stochastic mean-field limit result can be applied. The argument will be performed in the natural space of probability measures by an extension to our diffusion setting of classical characteristics arguments for transport equations.

1.1 Main results

We will start by introducing the two instances of IBMs that triggered this research. The IBM proposed in [19] includes an effective pairwise potential $U : \mathbb{R}^d \longrightarrow \mathbb{R}$ modeling the short-range repulsion and long-range attraction. The only "mimicking" interaction in this model is encoded in a relaxation term for the velocity arising as the equilibrium speed from the competing effects of self-propulsion and friction of the individuals. We will refer to it as the D'Orsogna et al model in the rest. More precisely, this IBM for N-particles in the mean-field limit scaling reads as:

$$\begin{cases} \dfrac{dX^i}{dt} = V^i, \\ \dfrac{dV^i}{dt} = (\alpha - \beta\,|V^i|^2)V^i - \dfrac{1}{N}\displaystyle\sum_{j\neq i} \nabla U(|X^i - X^j|), \end{cases}$$

where $\alpha > 0$ measures the self-propulsion strength of individuals, whereas the term corresponding to $\beta > 0$ is the friction assumed to follow Rayleigh's law. A typical choice for

hal-00521216, version 2 - 1 Dec 2010

U is a smooth radial potential given by

$$U(x) = -C_A e^{-|x|^2/\ell_A^2} + C_R e^{-|x|^2/\ell_R^2}.$$

where C_A, C_R and ℓ_A, ℓ_R are the strengths and the typical lengths of attraction and repulsion, respectively.

The other motivating example introduced in [16] only includes an "alignment" or reorientation interaction effect and we will refer to it as the Cucker-Smale model. Each individual in the group adjust their relative velocity by averaging with all the others. This averaging is weighted in such a way that closer individuals have more influence than further ones. For a system with N individuals the Cucker-Smale model in the mean-field scaling reads as

$$\begin{cases} \dfrac{dX^i}{dt} = V^i, \\ \dfrac{dV_i}{dt} = \dfrac{1}{N}\displaystyle\sum_{j=1}^{N} w_{ij}\left(V^j - V^i\right), \end{cases}$$

with the *communication rate* matrix given by:

$$w_{ij} = w(|X^i - X^j|) = \frac{1}{(1 + |X^i - X^j|^2)^\gamma}$$

for some $\gamma \geq 0$. We refer to [16, 22, 11, 12] and references therein for further discussion about this model and qualitative properties. Let us remark that both can be considered particular instances of a general IBM of the form

$$\begin{cases} \dfrac{dX^i}{dt} = V^i \\ \dfrac{dV^i}{dt} = -F(X^i, V^i) - \dfrac{1}{N}\displaystyle\sum_{j=1}^{N} H(X^i - X^j, V^i - V^j)dt, \qquad 1 \leq i \leq N \end{cases} \tag{1.1}$$

where $F, H : \mathbb{R}^{2d} \longrightarrow \mathbb{R}$ are suitable functions: the D'Orsogna et al model with $F(x, v) = (\beta|v|^2 - \alpha)v$ and $H(x, v) = \nabla_x U(x)$ and the Cucker-Smale model with $F = 0$ and $H(x, v) = w(x)v$. Let us emphasize that F in the D'Orsogna et al model and H in the Cucker-Smale model are not globally Lipschitz functions in $\mathbb{R}^{2d}$.

Our aim is to deal with a general system of interacting particles of the type (1.1) with added noise and suitable hypotheses on F and H including our motivating examples. More precisely, we will work then with a general large system of N interacting $\mathbb{R}^{2d}$-valued processes $(X_t^i, V_t^i)_{t\geq 0}$ with $1 \leq i \leq N$ solution of

$$\begin{cases} dX_t^i = V_t^i dt, \\ dV_t^i = \sqrt{2}dB_t^i - F(X_t^i, V_t^i)dt - \dfrac{1}{N}\displaystyle\sum_{j=1}^{N} H(X_t^i - X_t^j, V_t^i - V_t^j)dt, \end{cases} \tag{1.2}$$

with independent and commonly distributed initial data (X_0^i, V_0^i) with $1 \leq i \leq N$. Here, and throughout this paper, the $(B_t^i)_{t\geq 0}$ are N independent standard Brownian motions in $\mathbb{R}^d$. More general diffusion coefficients will be considered in the next subsection. The asymptotic behavior of the Cucker-Smale system with added noise has been recently considered in [15], and eq. (1.2) includes as a particular case the continuous-time models discussed there. Our main objective will be to study the large-particle number limit in their mean-field limit scaling. It is sometimes usual to write $(X_t^{i,N}, V_t^{i,N})$ to track N individuals, but to avoid a cumbersome notation we will drop the superscript N unless the dependence on it needs to be emphasized.

By symmetry of the initial configuration and of the evolution, all particles have the same distribution on $\mathbb{R}^{2d}$ at time t, which will be denoted $f_t^{(1)}$. For any given $t > 0$ the particles get correlated due to the nonlocal term

$$-\frac{1}{N}\sum_{j=1}^{N} H(X_t^i - X_t^j, V_t^i - V_t^j)$$

in the evolution, though they are independent at initial time. But, since the pairwise action of two particles i and j is of order $1/N$, it seems reasonable that two of these interacting particles (or a fixed number k of them) become less and less correlated as N gets large: this is what is called propagation of chaos. The statistical quantities of the system are given by the empirical measure

$$\hat{f}_t^N = \frac{1}{N}\sum_{i=1}^{N} \delta_{(X_t^i, V_t^i)}.$$

It is a general fact, see Sznitman [34], that propagation of chaos for a symmetric system of interacting particles is equivalent to the convergence in N of their empirical measure. Following [34] we shall prove quantitative versions of these equivalent results.

We shall show that our N interacting processes $(X_t^i, V_t^i)_{t\geq 0}$ respectively behave as $N \to \infty$ like the processes $(\overline{X}_t^i, \overline{V}_t^i)_{t\geq 0}$, solutions of the kinetic McKean-Vlasov type equation on $\mathbb{R}^{2d}$

$$\begin{cases} d\overline{X}_t^i = \overline{V}_t^i \, dt \\ d\overline{V}_t^i = \sqrt{2}\, dB_t^i - F(\overline{X}_t^i, \overline{V}_t^i)dt - H * f_t(\overline{X}_t^i, \overline{V}_t^i)dt, \\ (\overline{X}_0^i, \overline{V}_0^i) = (X_0^i, V_0^i), \quad f_t = \mathrm{law}(\overline{X}_t^i, \overline{V}_t^i). \end{cases} \tag{1.3}$$

Here the Brownian motions $(B_t^i)_{t\geq 0}$ are those governing the evolution of the $(X_t^i, V_t^i)_{t\geq 0}$. Note that the above set of equations involves the condition that f_t is the distribution of $(\overline{X}_t^i, \overline{V}_t^i)$, thus making it nonlinear. The processes $(\overline{X}_t^i, \overline{V}_t^i)_{t\geq 0}$ with $i \geq 1$ are independent since the initial conditions and driving Brownian motions are independent. Moreover they

hal-00521216, version 2 - 1 Dec 2010

are identically distributed and, by the Itô formula, their common law f_t at time t should evolve according to

$$\partial_t f_t + v \cdot \nabla_x f_t = \Delta_v f_t + \nabla_v \cdot ((F + H * f_t) f_t), \quad t > 0, x, v \in \mathbb{R}^d. \tag{1.4}$$

Here $a \cdot b$ denotes the scalar product of two vectors a and b in $\mathbb{R}^d$ and $*$ stands for the convolution with respect to $(x, v) \in \mathbb{R}^{2d}$:

$$H * f(x) = \int_{\mathbb{R}^{2d}} H(x - y, v - w)\, f(y, w)\, dy\, dw.$$

Moreover, ∇_x stands for the gradient with respect to the position variable $x \in \mathbb{R}^d$ whereas ∇_v, $\nabla_v\cdot$ and Δ_v respectively stand for the gradient, divergence and Laplace operators with respect to the velocity variable $v \in \mathbb{R}^d$.

Assuming the well-posedness of the stochastic differential system (1.2) and of the nonlinear equation (1.3) together with some uniform moment bounds, we will obtain our main result on the stochastic mean-field limit. Existence and uniqueness of solutions to (1.2), (1.3) and (1.4) verifying the assumptions of the theorem will also be studied but with more restrictive assumptions on F and H that we will comment on below.

Theorem 1.1. *Let f_0 be a Borel probability measure and (X_0^i, V_0^i) for $1 \le i \le N$ be N independent variables with law f_0. Assume that the drift F and the antisymmetric kernel H, with $H(-x, -v) = -H(x, v)$, satisfy that there exist constants $A, L, p > 0$ such that*

$$-(v - w) \cdot (F(x, v) - F(x, w)) \le A\,|v - w|^2 \tag{1.5}$$

$$|F(x, v) - F(y, v)| \le L \min\{|x - y|, 1\}(1 + |v|^p) \tag{1.6}$$

for all x, y, v, w in $\mathbb{R}^d$, and analogously for H instead of F. Take $T > 0$. Furthermore, assume that the particle system (1.2) *and the processes* (1.3) *have global solutions on $[0, T]$ with initial data (X_0^i, V_0^i) such that*

$$\sup_{0 \le t \le T} \Big\{ \int_{\mathbb{R}^{4d}} |H(x - y, v - w)|^2 df_t(x, v) df_t(y, w) + \int_{\mathbb{R}^{2d}} (|x|^2 + e^{a|v|^p}) df_t(x, v) \Big\} < +\infty, \tag{1.7}$$

with $f_t = law(\overline{X}_t^i, \overline{V}_t^i)$. Then there exists a constant $C > 0$ such that

$$\mathbb{E}\big[|X_t^i - \overline{X}_t^i|^2 + |V_t^i - \overline{V}_t^i|^2\big] \le \frac{C}{N^{e^{-Ct}}} \tag{1.8}$$

for all $0 \le t \le T$ and $N \ge 1$.

Moreover, if additionally there exists $p' > p$ such that

$$\sup_{0 \le t \le T} \int_{\mathbb{R}^{2d}} e^{a|v|^{p'}} df_t(x, v) < +\infty, \tag{1.9}$$

then for all $0 < \epsilon < 1$ there exists a constant C such that

$$\mathbb{E}\big[|X_t^i - \overline{X}_t^i|^2 + |V_t^i - \overline{V}_t^i|^2\big] \le \frac{C}{N^{1-\epsilon}} \tag{1.10}$$

for all $0 \le t \le T$ and $N \ge 1$.

hal-00521216, version 2 - 1 Dec 2010

This result classically ensures quantitative estimates on the mean field limit and the propagation of chaos. First of all, it ensures that the common law $f_t^{(1)}$ of any (by exchangeability) of the particles X_t^i at time t converges to f_t as N goes to infinity, as we have

$$W_2^2(f_t^{(1)}, f_t) \leq \mathbb{E}\big[|X_t^i - \overline{X}_t^i|^2 + |V_t^i - \overline{V}_t^i|^2\big] \leq \varepsilon(N) \tag{1.11}$$

Here W_2 stands for the Wasserstein distance between two measures μ and ν in the set $\mathcal{P}_2(\mathbb{R}^{2d})$ of Borel probability measures on $\mathbb{R}^{2d}$ with finite moment of order 2 defined by

$$W_2(\mu, \nu) = \inf_{(Z, \overline{Z})} \left\{\mathbb{E}\left[|Z - \overline{Z}|^2\right]\right\}^{1/2},$$

where the infimum runs over all couples of random variables $(Z, \overline{Z})$ in $\mathbb{R}^{2d} \times \mathbb{R}^{2d}$ with Z having law μ and $\overline{Z}$ having law ν (see [36] for instance). Moreover $\varepsilon(N)$ denotes the quantity in the right hand side of (1.8) or (1.10), depending on which part of Theorem 1.1 we are using.

Moreover, it proves a quantitative version of propagation of chaos: for all fixed k, the law $f_t^{(k)}$ of any (by exchangeability) k particles (X_t^i, V_t^i) converges to the tensor product $f_t^{\otimes k}$ as N goes to infinity, according to

$$\begin{aligned} W_2^2(f_t^{(k)}, f_t^{\otimes k}) &\leq \mathbb{E}\left[\left|(X_t^1, V_t^1, \cdots, X_t^k, V_t^k) - (\overline{X}_t^1, \overline{V}_t^1, \cdots, \overline{X}_t^k, \overline{V}_t^k)\right|^2\right] \\ &\leq k\, \mathbb{E}\left[|X_t^1 - \overline{X}_t^1|^2 + |V_t^1 - \overline{V}_t^1|^2\right] \leq k\varepsilon(N). \end{aligned}$$

It finally gives the following quantitative result on the convergence of the empirical measure $\hat{f}_t^N$ of the particle system to the distribution f_t : if φ is a Lipschitz map on $\mathbb{R}^{2d}$, then

$$\begin{aligned} &\mathbb{E}\left[\left|\frac{1}{N}\sum_{i=1}^N \varphi(X_t^i, V_t^i) - \int_{\mathbb{R}^{2d}} \varphi\, df_t\right|^2\right] \\ &\quad \leq 2\,\mathbb{E}\left[|\varphi(X_t^i, V_t^i) - \varphi(\overline{X}_t^i, \overline{V}_t^i)|^2 + \left|\frac{1}{N}\sum_{i=1}^N \varphi(\overline{X}_t^i, \overline{V}_t^i) - \int_{\mathbb{R}^{2d}} \varphi\, df_t\right|^2\right] \leq \varepsilon(N) + \frac{C}{N} \end{aligned}$$

by Theorem 1.1 and argument on the independent variables $(\overline{X}_t^i, \overline{V}_t^i)$ based on the law of large numbers, see [34].

The argument of Theorem 1.1 is classical for globally Lipschitz drifts [34, 30]. For space-homogeneous kinetic models it was extended to non-Lipschitz drifts by means of convexity arguments, first in one dimension in [4], then more generally in any dimension in [13, 27]. Here, in our space inhomogeneous setting, sole convexity arguments are hopeless, and we will replace them by moment arguments using hypothesis (1.7). We also refer to [31, 7] for related problems and biological discussions in space-homogeneous kinetic models with globally Lipschitz drifts but nonlinear diffusions.

hal-00521216, version 2 - 1 Dec 2010

hal-00521216, version 2 - 1 Dec 2010

Our proof will be written for $p > 0$, but one can simplify it with $p = 0$, by only assuming finite moments of order 2 in position and velocity. In this case our proof is the classical Sznitman's proof for existence, uniqueness, and mean-field limit for globally Lipschitz drifts, written in our kinetic setting and giving the classical decay rate in (1.8) as $1/N$, compared to (1.8)-(1.10). We will discuss further examples related to swarming models and extensions in subsection 1.2.

Section 3 will be devoted to the proof of existence, uniqueness, and moment propagation properties (1.7) and (1.9) for solutions to (1.2), (1.3) and (1.4). This well-posedness results and moment control for solutions will be obtained under more restrictive assumptions that those used in the proof of Theorem 1.1.

Theorem 1.2. *Assume that the drift F and the kernel H are locally Lipschitz functions satisfying that there exist $C, L \geq 0$ and $0 < p \leq 2$ such that*

$$-v \cdot F(x,v) \leq C(1+|v|^2) \tag{1.12}$$

$$-(v-w)\cdot(F(x,v)-F(x,w)) \leq L|v-w|^2(1+|v|^p+|w|^p), \tag{1.13}$$

$$|F(x,v)-F(y,v)| \leq L|x-y|(1+|v|^p), \tag{1.14}$$

$$|H(x,v)| \leq C(1+|v|), \tag{1.15}$$

$$|H(x,v)-H(y,w)| \leq L(|x-y|+|v-w|)(1+|v|^p+|w|^p), \tag{1.16}$$

for all $x, v, y, w \in \mathbb{R}^d$. Let f_0 be a Borel probability measure on $\mathbb{R}^{2d}$ such that

$$\int_{\mathbb{R}^{2d}} \left(|x|^2 + e^{a|v|^{p'}}\right) df_0(x,v) \; < +\infty.$$

for some $p' \geq p$. Finally, let (X_0^i, V_0^i) for $1 \leq i \leq N$ be N independent variables with law f_0. Then,

i) There exists a pathwise unique global solution to the SDE (1.2) *with initial data (X_0^i, V_0^i).*

ii) There exists a pathwise unique global solution to the nonlinear SDE (1.3) *with initial datum (X_0^i, V_0^i).*

iii) There exists a unique global solution to the nonlinear PDE (1.4) *with initial datum f_0.*

Moreover, for all $T > 0$ there exists $b > 0$ such that

$$\sup_{0\leq t\leq T} \int_{\mathbb{R}^{2d}} \left(|x|^2 + e^{b|v|^{p'}}\right) df_t(x,v) < +\infty.$$

Concerning the hypotheses on F, let us remark that we could also ask F to satisfy similar properties as H in (1.15)–(1.16), but (1.12)–(1.14) are slightly weaker.

1.2 Examples, extensions and variants

As discussed above the drift F models exterior or local effects, such as self propulsion, friction and confinement. In our motivating examples $F(x,v)=0$ in the Cucker-Smale model and $F(x,v)=(\beta|v|^2-\alpha)v$ in the D'Orsogna et al model. On the other hand, H models the interaction between individuals at (x,v) and (y,w) in the phase space being $H(x,v)=a(x)v$ with $a(x)=(1+|x|^2)^{-\gamma}$, $\gamma>0$ in the Cucker-Smale model and $H(x,v)=-\nabla U(x)$ in the D'Orsogna et al model. It is straightforward to check the assumptions of Theorems 1.1 and 1.2 in these two cases.

Of course, more general relaxation terms towards fixed "cruising speed" are allowed in the assumptions of Theorem 1.1, for instance: $F(x,v)=(\beta(x)|v|^\delta-\alpha(x))v$ with α,β globally Lipschitz bounded away from zero and infinity functions and $\delta>0$. Also, concerning the interaction kernel we may allow $H(x,v)=a(x)|v|^{q-2}v$ with $q\geq 1$ for a bounded and Lipschitz a in Cucker Smale as introduced in [23]. This has the effect of changing the equilibration rate towards flocking, see [11, 23] for details. However, the assumptions on existence and moment control in Theorem 1.2 are only verified for $q=2$. Other more general mechanisms can be included such as the one described in [26].

1.2.1 Variants on the assumptions

We first remark two simple extensions of the results in Theorem 1.1 by trading off growth control on F and H by moment control of the solutions to (1.2):

V1. Theorem 1.1 holds while weakening assumption (1.5) on F and H to

$$(v-w)\cdot(F(x,v)-F(x,w))\geq -A|v-w|^2(1+|v|^p+|w|^p)$$

both for F and H. Solutions in this case need to satisfy

$$\sup_{N\geq 1}\sup_{0\leq t\leq T}\mathbb{E}\left[e^{b|V_t^1|^p}\right]<+\infty$$

on the particle system or equivalent conditions on p' for the second estimate (1.10). Observe by lower continuity and weak convergence in N of the law of (X_t^1,V_t^1) to the law f_t of $(\overline{X}_t^1,\overline{V}_t^1)$ that this is a stronger assumption than part of the assumption (1.7) made in Theorem 1.1, more precisely

$$\sup_{0\leq t\leq T}\mathbb{E}\left[e^{b|\overline{V}_t^1|^p}\right]<+\infty\,.$$

Observe also that we may not have global existence in this case since for instance $F(x,v)=-v^3$ on $\mathbb{R}^2$, which leads to blow up in finite time, satisfies this new condition with $p=2$.

hal-00521216, version 2 - 1 Dec 2010

hal-00521216, version 2 - 1 Dec 2010

V2. One can remove the antisymmetry assumption on H in Theorem 1.1 by imposing

$$|H(x,v) - H(x,w)| \leq A\,|v-w|$$

instead of (1.5) for H. The reader can check that very little modifications are needed at the only point in the proof below where the symmetry of H is used, namely, when bounding term I_{21}. Actually, one can directly carry out the estimates instead of symmetrizing the term first. From the modeling point of view, it is important to include the non-antisymmetric case since some more refined swarming IBMs include the so-called "cone of vision" or "interaction region". In these models, individuals cannot interact with all the others but rather to a restricted set of individuals they actually see or feel, see [26, 12, 1]. From the mathematical point of view this implies that the interaction term $H * f_t$ need not always be a convolution but must be replaced by

$$H[f_t](x,v) = \int_{\mathbb{R}^{2d}} H(x,v;y,w)\,df_t(y,w);$$

here $H(x,v;\cdot,\cdot)$ is compactly supported in a region that depends on the value of (x,v) and $H(y,w;x,v)$ is not necessarily equal to $-H(x,v;y,w)$. Our results extend to this case.

V3. Theorem 1.1 also holds when F is an exterior drift in position only, non globally Lipschitz, for instance satisfying

$$|F(x) - F(y)| \leq A|x-y|(1+|x|^q+|y|^q)$$

with $q > 0$. Now, the moment control condition (1.7) has to be reinforced by assuming

$$\sup_{0\leq t\leq T} \mathbb{E}\left[e^{b|\overline{X}_t^1|^q}\right] < +\infty \quad \text{and} \quad \sup_{N\geq 1}\sup_{0\leq t\leq T} \mathbb{E}\left[e^{b|X_t^1|^q}\right] < +\infty$$

on the particle system or equivalent conditions on p' for the second estimate (1.10). Observe again by weak convergence in N of the law of (X_t^1, V_t^1) to the law f_t of $(\overline{X}_t^1, \overline{V}_t^1)$ that the latter new moment control assumption is stronger.

1.2.2 Extensions to nonlinearly dependent diffusion coefficient

Some researchers have recently argued that the diffusion coefficient at a given point x may depend on the neighbours of the point to be considered [20, 37]. More precisely, they can depend on local in space averaged quantities of the swarm, such as the averaged local density or velocity. The averaged local density at x in the particle system (X_t^i, V_t^i) for $1 \leq i \leq N$ is defined as

$$\frac{1}{N}\sum_{j=1}^N \eta_\varepsilon(x - X_t^j);$$

from which its corresponding continuous version is

$$\int_{\mathbb{R}^{2d}} \eta_\varepsilon(x-y) df_t(y,w).$$

Here $\eta_\varepsilon(x) = \frac{1}{\varepsilon^d}\eta\left(\frac{x}{\varepsilon}\right)$ where η is a nonnegative radial nonincreasing function with unit integral but non necessarily compactly supported and ε measures the size of the interaction. The name of "local average" comes from the smearing of choosing η_ε instead of a Dirac delta at 0, which would be meaningless in the setting of a particle system. Such a diffusion coefficient is considered in [20] with $\eta(x) = \frac{Z}{1+|x|^2}$ and $\varepsilon = 1$, from the point of view of the long-time behaviour of solutions to the kinetic equation, not of the mean-field limit: there the particle system evolves according to the diffusive Cucker-Smale model

$$\begin{cases} dX_t^i = V_t^i dt \\ dV_t^i = \sqrt{\dfrac{1}{N}\displaystyle\sum_{j=1}^N a(X_t^i - X_t^j)}\, dB_t^i - \dfrac{1}{N}\displaystyle\sum_{j=1}^N a(X_t^i - X_t^j)(V_t^i - V_t^j)dt, \qquad 1 \le i \le N \end{cases}$$

with $a(x) = \frac{Z}{1+|x|^2}$.

Other local quantities upon which the diffusion coefficient may depend on is the averaged local velocity at x defined as

$$\bar{u}(x) := \frac{1}{N}\sum_{j=1}^N V_t^j\, \eta_\varepsilon(x - X_t^j)$$

in the particle system, and

$$\int_{\mathbb{R}^{2d}} w\, \eta_\varepsilon(x-y) df_t(y,w)$$

in the continuous setting. More generically, we can consider diffusion coefficients in the particle system such as

$$g\left(\frac{1}{N}\sum_{j=1}^N h(V_t^j)\, \eta_\varepsilon(x - X_t^j)\right). \tag{1.17}$$

Here, η_ε controls which individuals we should take into account in the average and with which strength; among these X^j, how each velocity influences at x is controlled by h; finally, after averaging over those j, g controls how we should compute the diffusion coefficient.

For instance, given the diffusion coefficient $g(\overline{u}(x))$, we could argue that g should be large for small $\overline{u}(x)$ (large noise for small velocity), and conversely; there we see g as an even function, nonincreasing on $\mathbb{R}^+$. Similar coefficients were used in [37] of the form

$$g\left(\frac{1}{N(x)}\sum_{j=1}^{N(x)} V_t^j\, \eta_\varepsilon(x - X_t^j)\right)$$

hal-00521216, version 2 - 1 Dec 2010

hal-00521216, version 2 - 1 Dec 2010

with $\eta(x) = Z\mathbb{1}_{|x|\leq 1}$ and $N(x) = \sharp\{j; |x - X_t^j| \leq \varepsilon\}$. However, we cannot include this scaling in the mean-field setting.

On the other hand, mean-field limits such as those in Theorem 1.1 have been obtained in [30, 34] with the diffusion coefficient

$$\frac{1}{N}\sum_{j=1}^{N} \sigma(x, v; X_t^j, V_t^j)$$

where σ is a $2d \times 2d$ matrix with globally Lipschitz coefficients.

We include the two variants above by considering diffusion coefficients of the form

$$\sigma[X_t^i, V_t^i; \hat{f}_t^N]$$

where, for a probability measure f on $\mathbb{R}^{2d}$, $\sigma[z; f]$ is a $2d \times 2d$ matrix with coefficients

$$\sigma_{kl}[z; f] = g\Big(\int_{\mathbb{R}^{2d}} \sigma_{kl}(z, z')\, df(z')\Big)$$

in the notation $z = (x, v), z' = (x', v') \in \mathbb{R}^{2d}$. We shall assume that g is globally Lipschitz on $\mathbb{R}$ and

$$\begin{aligned} &|\sigma_{kl}(z, z') - \sigma_{kl}(\bar{z}, \bar{z}')| \\ &\leq C\Big(\min\{|x - \bar{x}| + |x' - \bar{x}'|, 1\} + |v - \bar{v}| + |v' - \bar{v}'|\Big)(1 + |v|^q + |v'|^q + |\bar{v}|^q + |\bar{v}'|^q.) \end{aligned}$$

In this notation, [34] corresponds to $g(x) = x$ and σ_{kl} bounded and Lipschitz, and (1.17) to $\sigma_{kl}(z, z') = h(v')\,\eta_\varepsilon(x - x')$ where the kernel η_ε is bounded and Lipschitz and

$$|h(v) - h(v')| \leq C|v - v'|(1 + |v|^q + |v'|^q).$$

Observe that this framework does not include the model considered in [20] for which the diffusion coefficient is given by a non locally Lipschitz g.

In this notation and assumption, if furthermore there exists $b > 0$ such that

$$\sup_{0\leq t\leq T} \mathbb{E}\left[e^{b|\overline{X}_t^1|^{2q}}\right] < +\infty \quad \text{and} \quad \sup_{N\geq 1}\sup_{0\leq t\leq T} \mathbb{E}\left[e^{b|X_t^1|^{2q}}\right] < +\infty$$

on the nonlinear process and particle system, then (1.8) holds in Theorem 1.1, and correspondingly with p' for the second estimate (1.10) (see Remark 2.1).

1.2.3 One-variable formulation

We now give a formulation of the mean-field limit in one variable $z \in \mathbb{R}^D$, to be thought of as $z = (x, v) \in \mathbb{R}^{2d}$ as in our examples above or as $z = v \in \mathbb{R}^d$ in a space-homogeneous setting. We consider the particle system

$$dZ_t^i = \sigma\, dB_t^i - F(Z_t^i)\, dt - \frac{1}{N}\sum_{j=1}^{N} H(Z_t^i - Z_t^j)\, dt, \quad 1 \leq i \leq N$$

where σ is a (for instance) constant $D \times D$ matrix, the $(B_t^i)_{t\geq 0}$ are N independent standard Brownian motions on $\mathbb{R}^D$ and the initial data Z_0^i are independent and identically distributed. We also consider the nonlinear processes $(\overline{Z}_t^i)_{t\geq 0}$ defined by

$$\begin{cases} d\overline{Z}_t^i = \sigma dB_t^i - F(\overline{Z}_t^i)dt - H * f_t(\overline{Z}_t^i)dt, \\ \overline{Z}_0^i = Z_0^i, \ f_t = \text{law}(Z_t^i). \end{cases}$$

Assume now that there exists $C > 0$ such that

$$(z - z') \cdot (F(z) - F(z')) \geq -C|z - z'|^2(1 + |z|^p + |z'|^p)$$

for all $z, z' \in \mathbb{R}^D$ with $p > 0$. Assume also global existence and uniqueness of these processes, with

$$\sup_{N\geq 1} \sup_{0\leq t\leq T} \mathbb{E}\left[e^{b|Z_t^1|^p}\right] < +\infty$$

and

$$\sup_{0\leq t\leq T} \left\{ \int_{\mathbb{R}^D} e^{b|z|^p} df_t(z) + \int_{\mathbb{R}^{2D}} |H(z - z')|^2 df_t(z)\, df_t(z') \right\} < +\infty,$$

or equivalent conditions on p'. Then (1.8) and (1.10) in Theorem 1.1 holds.

2 Mean-field limit: proof

This section is devoted to the proof of Theorem 1.1. We follow the coupling method [34, 30, 35]. Given $T > 0$, we will use C to denote diverse constants depending on T, the functions F and H, and moments of the solution f_t on $[0, T]$, but not on the number of particles N.

Proof of Theorem 1.1. Let us define the fluctuations as $x_t^i := X_t^i - \overline{X}_t^i$, $v_t^i := V_t^i - \overline{V}_t^i$, $i = 1, \dots, N$. For notational convenience, we shall drop the time dependence in the stochastic processes. As the Brownian motions $(B_t^i)_{t\geq 0}$ considered in (1.2) and (1.3) are equal, for all $i = 1, \dots, N$, we deduce

$$dx^i = v^i\, dt\,, \tag{2.1}$$

$$dv^i = -\big(F(X^i, V^i) - F(\overline{X}^i, \overline{V}^i)\big)dt - \frac{1}{N}\sum_{j=1}^{N} \Big(H(X^i - X^j, V^i - V^j) - (H * f_t)(\overline{X}^i, \overline{V}^i) \Big)\, dt. \tag{2.2}$$

Let us consider the quantity $\alpha(t) = \mathbb{E}\left[|x^i|^2 + |v^i|^2\right]$ (independent of the label i by symmetry), which bounds the distance $W_2^2(f_t^{(1)}, f_t)$ as remarked in (1.11). Then, by using (2.1)-(2.2), we readily get

$$\frac{1}{2}\frac{d}{dt}\mathbb{E}\left[|x^i|^2\right] = \mathbb{E}\left[x^i \cdot v^i\right] \leq \frac{1}{2}\alpha(t) \tag{2.3}$$

hal-00521216, version 2 - 1 Dec 2010

hal-00521216, version 2 - 1 Dec 2010

and

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\mathbb{E}\left[|v^i|^2\right] &= -\mathbb{E}\left[v^i\cdot\left(F(X^i,V^i)-F(\overline{X}^i,\overline{V}^i)\right)\right]\\ &\quad -\frac{1}{N}\mathbb{E}\left[\sum_{j=1}^N v^i\cdot\Big(H(X^i{-}X^j,V^i{-}V^j)-H*f_t(\overline{X}^i,\overline{V}^i)\Big)\right] =: I_1+I_2. \qquad (2.4)\end{aligned}$$

Step 1.- Estimate I_1 by moment bounds: We decompose I_1 in (2.4) as

$$I_1 = -\mathbb{E}\left[v^i\cdot\left(F(X^i,V^i)-F(X^i,\overline{V}^i)\right)\right]-\mathbb{E}\left[v^i\cdot\left(F(X^i,\overline{V}^i)-F(\overline{X}^i,\overline{V}^i)\right)\right].$$

By assumption (1.5)-(1.6) on F, I_1 can be controlled by

$$I_1 \le A\,\mathbb{E}\left[|v^i|^2\right]+L\,\mathbb{E}\left[|v^i|\,\min\{|x^i|,1\}\,(1+|\overline{V}^i|^p)\right] := I_{11}+L\,I_{12}\,.$$

Given $R>0$, the second term I_{12} is estimated according to

$$\begin{aligned}I_{12} &\le \mathbb{E}[|v^i|\,|x^i|]+\mathbb{E}\Big[\mathbb{1}_{|\overline{V}^i|\le R}\,|v^i|\,\min\{|x^i|,1\}\,|\overline{V}^i|^p\Big]+\mathbb{E}\Big[\mathbb{1}_{|\overline{V}^i|>R}\,|v^i|\,\min\{|x^i|,1\}\,|\overline{V}^i|^p\Big]\\ &\le (1+R^p)\mathbb{E}\left[|v^i|\,|x^i|\right]+\frac{1}{2}\mathbb{E}\left[|v^i|^2\right]+\frac{1}{2}\mathbb{E}\left[\mathbb{1}_{|\overline{V}^i|>R}\,|\overline{V}^i|^{2p}\right]\\ &\le (1+R^p)\alpha(t)+\frac{1}{2}\left(\mathbb{E}\left[|\overline{V}^i|^{4p}\right]\right)^{1/2}\left(\mathbb{E}\left[\mathbb{1}_{|\overline{V}^i|>R}\right]\right)^{1/2}\end{aligned}$$

by the Young and the Cauchy-Schwarz inequalities. Invoking the Markov inequality, hypothesis (1.7) implies that there exists $C>0$ such that

$$\mathbb{E}\left[\mathbb{1}_{|\overline{V}^i_t|>R}\right]\le e^{-aR^p}\,\mathbb{E}\left[e^{a|\overline{V}^i_t|^p}\right]\le C\,e^{-a\,R^p} \qquad (2.5)$$

for all i and $0\le t\le T$. By defining $r=aR^p/2$, we conclude that given $T>0$, there exists $C>0$ such that

$$I_1 \le C(1+r)\,\alpha(t)+C\,e^{-r} \qquad (2.6)$$

holds for all $r>0$ and all $0\le t\le T$.

Step 2.- Estimate I_2 by moment bounds: We decompose the second term in (2.4) as

$$\begin{aligned}I_2 = &-\frac{1}{N}\mathbb{E}\left[\sum_{j=1}^N v^i\cdot\Big(H(X^i{-}X^j,V^i-V^j)-H(\overline{X}^i{-}\overline{X}^j,\overline{V}^i{-}\overline{V}^j)\Big)\right]\\ &-\frac{1}{N}\mathbb{E}\left[v^i\cdot\Big(H(0,0)-(H*f_t)(\overline{X}^i,\overline{V}^i)\Big)\right]\\ &-\frac{1}{N}\mathbb{E}\left[\sum_{j\neq i}^N v^i\cdot\Big(H(\overline{X}^i{-}\overline{X}^j,\overline{V}^i{-}\overline{V}^j)-(H*f_t)(\overline{X}^i,\overline{V}^i)\Big)\right]\\ =:&\ I_{21}+I_{22}+I_{23}.\end{aligned} \qquad (2.7)$$

Since all particles are equally distributed and H is antisymmetric, we rewrite I_{21} as

$$I_{21} = -\frac{1}{2N^2} \sum_{i,j=1}^{N} \mathbb{E}\left[(v^i - v^j) \cdot \Big(H(X^i - X^j, V^i - V^j) - H(\overline{X}^i - \overline{X}^j, \overline{V}^i - \overline{V}^j)\Big)\right].$$

Analogously to the argument used to bound I_1 in the first step, for each (i, j) we introduce the intermediate term $H(X^i - X^j, \overline{V}^i - \overline{V}^j)$, split the expression in two terms, and estimate the corresponding expectations using (1.5)-(1.6) on H by

$$I_{21} \le A\, \mathbb{E}\left[|v^i - v^j|^2\right] + L\, \mathbb{E}\left[|v^i - v^j|\, \min\{|x^i - x^j|, 1\}\, (1 + |\overline{V}^i - \overline{V}^j|^p)\right]. \quad (2.8)$$

For a given $R > 0$, and fixed (i, j), consider the event $\mathcal{R} := \{|\overline{V}_i| \le R,\ |\overline{V}_j| \le R\}$ and the random variable $\mathcal{Z} := |v^i - v^j|\, \min\{|x^i - x^j|, 1\}\, (1 + |\overline{V}^i - \overline{V}^j|^p)$. Then the last expectation in (2.8) can be estimated as follows, using again the Young and Cauchy-Schwarz inequalities:

$$\begin{aligned}
\mathbb{E}[\mathcal{Z}] &= \mathbb{E}[\mathbb{1}_{\mathcal{R}} \mathcal{Z}] + \mathbb{E}[\mathbb{1}_{\mathcal{R}^C} \mathcal{Z}] \\
&\le (1 + 2^p R^p)\, \mathbb{E}\left[|v^i - v^j|\, |x^i - x^j|\right] + \frac{1}{2} \mathbb{E}\left[|v^i - v^j|^2\right] \\
&\quad + \mathbb{E}\left[\mathbb{1}_{\mathcal{R}^C}\, (1 + |\overline{V}^i - \overline{V}^j|^p)^2\right] \\
&\le 2(1 + 2^p R^p)\, \alpha(t) + 2\alpha(t) + (\mathbb{E}[\mathbb{1}_{\mathcal{R}^C}])^{1/2} \left(\mathbb{E}\left[(1 + |\overline{V}^i - \overline{V}^j|^p)^4\right]\right)^{1/2} \\
&\le 2(2 + 2^p R^p)\, \alpha(t) + C\left(\mathbb{E}\left[\mathbb{1}_{|\overline{V}^i| > R}\right] + \mathbb{E}\left[\mathbb{1}_{|\overline{V}^j| > R}\right]\right)^{1/2} \left(1 + \mathbb{E}\left[|\overline{V}^i|^{4p}\right]\right)^{1/2} \\
&\le C(1 + R^p)\alpha(t) + C\, e^{-a\, R^p/2}
\end{aligned} \quad (2.9)$$

by hypothesis (1.7). Inserting (2.9) into (2.8) and defining $r = aR^p/2$, we conclude that given $T > 0$, there exists $C > 0$ such that

$$I_{21} \le C(1 + r)\, \alpha(t) + C\, e^{-r} \quad (2.10)$$

holds for all $r > 0$ and all $0 \le t \le T$.

We now turn to estimate I_{22}, i.e., the second term in (2.7). Using that $H(0, 0) = 0$, we get

$$I_{22} \le \frac{1}{N} \left(\mathbb{E}\left[|v^i|^2\right]\right)^{1/2} \left(\mathbb{E}\left[\left|(H * f_t)(\overline{X}^i, \overline{V}^i)\right|^2\right]\right)^{1/2} \le \frac{C}{N} \sqrt{\alpha(t)}\,. \quad (2.11)$$

The latter inequality follows from

$$\mathbb{E}\left[\left|(H * f_t)(\overline{X}^i, \overline{V}^i)\right|^2\right] = \int_{\mathbb{R}^{4d}} |H(x - y, v - w)|^2\, df_t(x, v)\, df_t(y, w), \quad (2.12)$$

which is bounded on $[0, T]$ due to hypothesis (1.7).

hal-00521216, version 2 - 1 Dec 2010

hal-00521216, version 2 - 1 Dec 2010

The last term I_{23} is treated as in the classical case in [34, Page 175] by a law of large numbers argument. We include here some details for the sake of the reader. By symmetry we assume that $i = 1$. We start by applying the Cauchy-Schwarz inequality to obtain

$$I_{23} \leq \frac{1}{N} \left(\mathbb{E}\left[|v^1|^2\right]\right)^{1/2} \left(\mathbb{E}\left[\Big|\sum_{j=2}^{N} Y^j\Big|^2\right]\right)^{1/2}$$

where $Y^j := H(\overline{X}^1 - \overline{X}^j, \overline{V}^1 - \overline{V}^j) - (H * f_t)(\overline{X}^1, \overline{V}^1)$ for $j \geq 2$. Note that, for $j \neq k$,

$$\mathbb{E}\left[Y^j \cdot Y^k\right] = \mathbb{E}\left[\mathbb{E}\left[Y^j \cdot Y^k | (\overline{X}^1, \overline{V}^1)\right]\right] = \mathbb{E}\left[\mathbb{E}\big[Y^j | (\overline{X}^1, \overline{V}^1)\big] \cdot \mathbb{E}\big[Y^k | (\overline{X}^1, \overline{V}^1)\big]\right]$$

by independence of the N processes $(\overline{X}^j_t, \overline{V}^j_t)_{t \geq 0}$, where

$$\mathbb{E}\left[Y^j | (\overline{X}^1, \overline{V}^1)\right] = \int_{\mathbb{R}^{2d}} [H(\overline{X}^1_t - y, \overline{V}^1_t - w) - (H * f_t)(\overline{X}^1_t, \overline{V}^1_t)]\, f_t(y, w)\, dy\, dw = 0$$

since $(\overline{X}^j_t, \overline{V}^j_t)$ has probability distribution f_t. Hence,

$$\begin{aligned}\mathbb{E}\left[\Big|\sum_{j=2}^{N} Y^j\Big|^2\right] &= (N-1)\mathbb{E}\left[|Y^2|^2\right] \\ &\leq (N-1) \int_{\mathbb{R}^{4d}} |H(x-y, v-w)|^2\, df_t(x,v)\, df_t(y,w) \leq C\,(N-1)\end{aligned}$$

as in (2.12) due to hypothesis (1.7). Therefore, we get

$$I_{23} \leq \frac{C}{\sqrt{N}} \sqrt{\alpha(t)}\,. \tag{2.13}$$

Hence, combining the estimates (2.10), (2.11) and (2.13) to estimate I_2, we get that there exists $C > 0$ such that

$$I_2 \leq C(1+r)\,\alpha(t) + C\,e^{-r} + \frac{C}{\sqrt{N}} \sqrt{\alpha(t)} \tag{2.14}$$

holds for all $r > 0$ and all $0 \leq t \leq T$.

Step 3.- Proof of (1.8)*:* It follows from (2.3), (2.4), (2.6), (2.14), the above estimates and the Young inequality that

$$\alpha'(t) \leq C\,(1+r)\,\alpha(t) + Ce^{-r} + \frac{C}{\sqrt{N}}\sqrt{\alpha(t)} \leq C\,(1+r)\,\alpha(t) + Ce^{-r} + \frac{C}{N}$$

for all $t \in [0, T]$, all $N \geq 1$ and all $r > 0$. From this differential inequality and Gronwall's lemma, we can first deduce that the quantity $\alpha(t)$ is bounded on $[0, T]$, uniformly in N, by a constant $D > 0$. Hence, the function $\beta(t) := \alpha(t)/(eD)$ is bounded by $1/e$, so that

$1-\ln\beta \leq -2\ln\beta$. Now, whenever $\beta(t) > 0$, take $r := -\ln\beta(t) > 0$. This choice proves that, for any t such that $\beta(t) > 0$,

$$\beta'(t) \leq C(1-\ln\beta(t))\,\beta(t) + \frac{C}{N} \leq -C\,\beta(t)\,\ln\beta(t) + \frac{C}{N}. \tag{2.15}$$

Actually, the above inequality is also true whenever $\beta(t) = 0$ (with the convention that $z\log z = 0$ for $z = 0$), as can be seen by choosing $r := \log N$ in that case. Hence, (2.15) holds for all $t \in [0,T]$. Now, the function $u(t) := \beta(Ct)$ satisfies $u(0) = 0$ and

$$u' \leq -u\,\ln u + \frac{1}{N}$$

on $[0, T/C]$. Let finally $a(t)$ be a function on $[0, T/C]$, to be chosen later on. Then the map $v(t) = u(t)\,N^{a(t)}$ satisfies $v(0) = 0$ and

$$v' \leq -v\ln v + N^{a-1} + v\,\ln N(a+a') \leq -v\ln v + 1 \leq \frac{1}{e} + 1$$

on $[0, T/C]$ provided we choose $a(t) = e^{-t} \leq 1$. Hence, this choice of $a(t)$ implies the bound

$$v(t) \leq \Big(\frac{1}{e} + 1\Big)\frac{T}{C}$$

for $0 \leq t \leq T/C$, that is,

$$\mathbb{E}\big[|X_t^i - V_t^i|^2 + |\overline{X}_t^i - \overline{V}_t^i|^2\big] = \alpha(t) \leq C N^{-e^{-Ct}}$$

for $0 \leq t \leq T$, and thus, (1.8) is proven.

Step 4.- Proof of (1.10)*:* If additionally there exists $p' > p$ such that hypothesis (1.9) holds, then by the Markov inequality, estimate (2.5) turns into

$$\mathbb{E}\left[\mathbb{1}_{|\overline{V}_t^i|>R}\right] \leq e^{-aR^{p'}}\,\mathbb{E}\left[e^{a|\overline{V}_t^i|^{p'}}\right] \leq C\,e^{-a\,R^{p'}}.$$

Hence, following the same proof, the quantity $\alpha(t)$ finally satisfies the differential inequality

$$\alpha'(t) \leq C(1+r)\,\alpha(t) + C\,e^{-r^{p'/p}} + \frac{C}{N}$$

for all $N \geq 1$ and all $r > 0$. If we choose $r = (\ln N)^{p/p'}$, and since $\alpha(0) = 0$, this integrates to

$$\alpha(t) \leq \frac{2}{N(1+r)}\Big(e^{C(1+r)t} - 1\Big) \leq \frac{2}{N}\,e^{C(1+r)T} = 2\,e^{CT}\,e^{C(\ln N)^{p/p'}T - \ln N}.$$

Given $\epsilon > 0$, there exists a constant D such that

$$C(\ln N)^{p/p'}T - \ln N \leq D - (1-\epsilon)\ln N$$

for all $N \geq 1$, so that

$$\alpha(t) \leq 2\,e^{CT+D}\,N^{-(1-\epsilon)}$$

for all $N \geq 1$. This concludes the proof of Theorem 1.1. □

hal-00521216, version 2 - 1 Dec 2010

Remark 2.1. In the setting of subsection 1.2.2 when the constant diffusion coefficient $\sqrt{2}$ governing the evolution of the particle (X_t^i, V_t^i) is replaced by a more general $\sigma[X_t^i, V_t^i; \hat{f}_t]$, then, by the Itô formula, we have to control the extra term

$$\sum_{k,l} \mathbb{E}\big[\big|\sigma_{kl}[X_t^i, V_t^i; \hat{f}_t^N] - \sigma_{kl}[\overline{X}_t^i, \overline{V}_t^i; f_t]\big|^2\big].$$

For that purpose we use the Lipschitz property of g, introduce the intermediate term $\displaystyle\frac{1}{N}\sum_{j=1}^N \sigma_{kl}\big((\overline{X}_t^i, \overline{V}_t^i), (\overline{X}_t^j, \overline{V}_t^j)\big)$ and adapt the argument used above to bound the term I_2.

hal-00521216, version 2 - 1 Dec 2010

3 Existence and uniqueness

This section is devoted to the proof of Theorem 1.2 on existence, uniqueness and propagation of moments for solutions of the particle system (1.2), the nonlinear process (1.3) and the associated PDE (1.4). This provides a setting under which Theorem 1.1 holds, showing that the existence and moment bound hypotheses are satisfied under reasonable conditions on the coefficients of the equations and the initial data alone.

Given $T > 0$, we will denote by b and C constants, that may change from line to line, depending on T, the functions F and H, and moments of the initial datum f_0.

3.1 Existence and uniqueness of the particle system

Let us start by proving point *i)* of Theorem 1.2. In this section we let $f_0 \in \mathcal{P}_2(\mathbb{R}^{2d})$ and consider the particle system for $1 \leq i \leq N$:

$$\begin{cases} dX_t^i = V_t^i dt \\ dV_t^i = \sqrt{2} dB_t^i - F(X_t^i, V_t^i)dt - \dfrac{1}{N}\displaystyle\sum_{j=1}^N H(X_t^i - X_t^j, V_t^i - V_t^j)dt, \end{cases} \tag{3.1}$$

with initial data (X_0^i, V_0^i) for $1 \leq i \leq N$ distributed according to f_0. Here the $(B_t^i)_{t\geq 0}$, for $i = 1, \ldots, N$, are N independent standard Brownian motions on $\mathbb{R}^d$.

Lemma 3.1. *Let $f_0 \in \mathcal{P}_2(\mathbb{R}^{2d})$, and assume that F, H are locally Lipschitz and satisfy* (1.12) *and* (1.15)*. For $1 \leq i \leq N$, take random variables (X_0^i, V_0^i) with law f_0. Then* (3.1) *admits a pathwise unique global solution with initial datum (X_0^i, V_0^i) for $1 \leq i \leq N$.*

Proof. The system (3.1) can be written as the SDE

$$d\mathbf{Z}_t^N = \sigma^N \, d\mathbf{B}_t^N + \mathbf{b}(\mathbf{Z}_t^N)\, dt$$

in $\mathbb{R}^{2dN}$, where $\mathbf{Z}_t^N = (X_t^1, V_t^1, \ldots, X_t^N, V_t^N)$. Here σ^N is a constant $2dN \times 2dN$ matrix, $(\mathbf{B}_t^N)_{t\geq 0}$ is a standard Brownian motion on $\mathbb{R}^{2dN}$, and $\mathbf{b} : \mathbb{R}^{2dN} \to \mathbb{R}^{2dN}$ is a locally Lipschitz function defined in the obvious way. Moreover, letting $\langle \cdot, \cdot \rangle$ be the scalar product and $\|\cdot\|$ the Euclidean norm on $\mathbb{R}^{2dN}$, then for all $\mathbf{Z}^N = (X^1, V^1, \ldots, X^N, V^N)$,

$$\begin{aligned}\left\langle \mathbf{Z}^N, \mathbf{b}(\mathbf{Z}^N)\right\rangle &= \sum_{i=1}^N X^i \cdot V^i - \sum_{i=1}^N V^i \cdot F(X^i, V^i) - \frac{1}{N}\sum_{i,j=1}^N V^i \cdot H(X^i - X^j, V^i - V^j) \\ &\leq (C + \frac{1}{2})(N + \|\mathbf{Z}^N\|^2) + C\frac{1}{N}\sum_{i,j=1}^N |V^i|(1 + |V^i - V^j|) \\ &\leq C(N + \|\mathbf{Z}^N\|^2) + \frac{C}{2N}\sum_{i,j=1}^N (1 + |V^i|^2 + |V^j|^2) \leq C(N + \|\mathbf{Z}^N\|^2).\end{aligned}$$

Here we have used the elementary inequality $2ab \leq a^2 + b^2$, and the bounds (1.12) and (1.15). This is a sufficient condition for global existence and pathwise uniqueness, see [21, Chapter 5, Theorems 3.7 and 3.11] for instance. □

Remark 3.2. For the existence we do not use any properties of symmetry of the system, and in particular we do not need the initial data to be independent. On the other hand, the condition (1.15) in Lemma 3.1 can be relaxed to

$$-v \cdot H(x, v) \leq C(1 + |v|^2),$$

if we impose that H is antisymmetric, i.e., $H(-x, -v) = -H(x, v)$ for all $x, v \in \mathbb{R}^d$. Actually, in this case we can perform a symmetrization in (i, j) to estimate the term involving H by

$$\begin{aligned}-\frac{1}{N}\sum_{i,j=1}^N V^i \cdot H(X^i - X^j, V^i - V^j) &= -\frac{1}{2N}\sum_{i,j=1}^N (V^i - V^j) \cdot H(X^i - X^j, V^i - V^j) \\ &\leq \frac{C}{2N}\sum_{i,j=1}^N (1 + |V^i - V^j|^2) \leq C(N + \|\mathbf{Z}^\mathbf{N}\|^2).\end{aligned}$$

3.2 Existence and uniqueness for the nonlinear process and PDE

In this section we prove points *ii)* and *iii)* in Theorem 1.2, namely, the existence and uniqueness of solutions to the nonlinear SDE (1.3):

$$\begin{cases} d\overline{X}_t = \overline{V}_t\, dt \\ d\overline{V}_t = \sqrt{2}\, dB_t - F(\overline{X}_t, \overline{V}_t)dt - H * f_t(\overline{X}_t, \overline{V}_t)dt, \\ f_t = \mathrm{law}(\overline{X}_t, \overline{V}_t), \quad \mathrm{law}(\overline{X}_0, \overline{V}_0) = f_0. \end{cases}$$

hal-00521216, version 2 - 1 Dec 2010

hal-00521216, version 2 - 1 Dec 2010

and to the associated nonlinear PDE (1.4):

$$\partial_t f_t + v \cdot \nabla_x f_t = \Delta_v f_t + \nabla_v \cdot ((F + H * f_t) f_t), \quad t > 0,\ x, v \in \mathbb{R}^d,$$

under the hypotheses of Theorem 1.2. Notice that we drop the superscript i for the SDE (1.3), as the problem is solved independently for each i. For the PDE (1.4), we always consider solutions in the sense of distributions:

Definition 3.3. Assume that $F, H : \mathbb{R}^{2d} \to \mathbb{R}^{2d}$ are continuous, and that (1.15) holds. Given $T > 0$, a function $f : [0, T] \to \mathcal{P}_2(\mathbb{R}^{2d})$, continuous in the W_2 topology, is a solution of equation (1.4) with initial data $f_0 \in \mathcal{P}_2(\mathbb{R}^{2d})$ if for all $\varphi \in \mathcal{C}_0^\infty([0, T) \times \mathbb{R}^{2d})$ it holds that

$$\int_{\mathbb{R}^{2d}} \varphi_0 \, df_0 = -\int_0^T \int_{\mathbb{R}^{2d}} \big(\partial_s \varphi_s + \Delta_v \varphi_s - \nabla_v \varphi_s \cdot (F + H * f_s) + \nabla_x \varphi_s \cdot v\big) \, df_s \, ds. \quad (3.2)$$

Notice that all terms above make sense due to the continuity of F, H, equation (1.15), and the bound on the second moment of f_t on bounded time intervals (needed for $H * f_t$ to make sense). For the purpose of this definition, condition (1.15) can actually be relaxed to $|H(x, v)| \leq C(1 + |x|^2 + |v|^2)$, but we will always work under the stronger hypothesis below.

The proof spans several steps, that we split in several subsections. Since some parts of the proof hold under weaker conditions on F and H, we will specify the hypotheses needed in each part.

3.2.1 Existence and uniqueness of an associated linear SDE

Let us first consider a related linear problem: we want to solve the SDE

$$\begin{cases} dX_t = V_t dt \\ dV_t = \sqrt{2} dB_t - F(X_t, V_t) dt - (H * g_t)(X_t, V_t) dt \end{cases} \quad (3.3)$$

for given $F, H : \mathbb{R}^{2d} \to \mathbb{R}^d$ and $g : [0, T] \to \mathcal{P}_2(\mathbb{R}^{2d})$.

Lemma 3.4. *Assume that F and H are locally Lipschitz functions satisfying* (1.12), (1.15), *and* (1.16). *Let $f_0 \in \mathcal{P}_2(\mathbb{R}^{2d})$ and $g : [0, T] \to \mathcal{P}_2(\mathbb{R}^{2d})$ be a continuous curve in the W_2 topology. Then the equation* (3.3) *with inital datum (X_0, V_0) distributed according to f_0 has a global pathwise unique solution. Moreover this solution has bounded second moment on $[0, T]$.*

Proof. We rewrite (3.3) as

$$dZ_t = \sigma \, dB_t + b(t, Z_t) \, dt$$

on $\mathbb{R}^{2d}$, where $Z_t = (X_t, V_t)$. Here σ is a $2d \times 2d$ matrix, $(B_t)_{t\geq 0}$ is a standard Brownian motion on $\mathbb{R}^{2d}$ and

$$b(t,x,v) = (v, -F(x,v) - (H * g_t)(x,v)).$$

Let us first observe that the growth condition (1.15) on H implies

$$|(H * g_t)(x,v)| \leq C \int_{\mathbb{R}^{2d}} (1 + |v-w|)\, dg_t(y,w) \leq C\left(1 + |v| + \int_{\mathbb{R}^{2d}} |w|\, dg_t(y,w)\right). \tag{3.4}$$

This together with the Cauchy-Schwarz inequality results in

$$|v \cdot (H * g_t)(x,v)| \leq C\left(1 + |v|^2 + \int_{\mathbb{R}^{2d}} |w|^2\, dg_t(y,w)\right). \tag{3.5}$$

Estimate (3.4) ensures that, for fixed $x, v \in \mathbb{R}^d$, the map $t \mapsto (H * g_t)(x,v)$ is bounded on $[0,T]$, using the fact that the second moment of g_t (and hence its first moment) is uniformly bounded on $[0,T]$. Then, the same applies to $t \mapsto b(t,x,v)$.

Also, the map $(x,v) \mapsto (H * g_t)(x,v)$ is locally Lipschitz, uniformly on $t \in [0,T]$; indeed, for $x, v, y, w \in \mathbb{R}^d$, using (1.16) we get

$$\begin{aligned}
&\big|(H * g_t)(x,v) - (H * g_t)(y,w)\big| \\
&\qquad \leq \int_{\mathbb{R}^{2d}} \Big|H(x-X, v-V) - H(y-X, w-V)\Big|\, dg_t(X,V) \\
&\qquad \leq L\left(\big|x-y\big| + \big|v-w\big|\right) \int_{\mathbb{R}^{2d}} (1 + |v-V|^p + |w-V|^p) dg_t(X,V) \\
&\qquad \leq C(\big|x-y\big| + \big|v-w\big|)\Big[1 + |v|^p + |w|^p + \int_{\mathbb{R}^{2d}} |V|^p dg_t(X,V)\Big].
\end{aligned}$$

The moment of g_t above is bounded on $[0,T]$ since $p \leq 2$ and the curve is continuous in the W_2-metric. As F is also locally Lipschitz, we conclude that $t \mapsto b(t,x,v)$ is locally Lipschitz.

Finally, for the scalar product $\langle \cdot, \cdot \rangle$ on $\mathbb{R}^{2d}$, we deduce

$$\langle (x,v), b(t,x,v) \rangle = x \cdot v - v \cdot F(x,v) - v \cdot (H * g_t)(x,v) \leq C\Big(1 + |x|^2 + |v|^2\Big), \tag{3.6}$$

by (1.12), (3.5), and again the fact that the second moment of g_t is uniformly bounded on $[0,T]$. As in the proof of Lemma 3.1, these are sufficient conditions for global existence and pathwise uniqueness for solutions to (3.4) with square-integrable initial data.

Moreover, by (3.6),

$$\frac{d}{dt}\mathbb{E}[|X_t|^2 + |V_t|^2] = 2d + 2\,\mathbb{E}\,\langle (X_t, V_t), b(t, X_t, V_t) \rangle \leq 2d + C\mathbb{E}\big[1 + |X_t|^2 + |V_t|^2\big],$$

so that by integration the second moment $\mathbb{E}[|X_t|^2 + |V_t|^2]$ is bounded on $[0,T]$. □

hal-00521216, version 2 - 1 Dec 2010

hal-00521216, version 2 - 1 Dec 2010

3.2.2 Existence and uniqueness of an associated linear PDE

By Itô's formula, the law f_t of the solution of (3.3) at time t is a solution of the following linear PDE:

$$\partial_t f_t + v \cdot \nabla_x f_t = \Delta_v f_t + \nabla_v \cdot (f_t \,(F + H * g_t)) \,, \tag{3.7}$$

in the distributional sense as in (3.2) of Definition 3.3. Moreover, the curve $t \mapsto f_t$ is continuous for the W_2 topology. Indeed, on the one hand

$$W_2^2(f_t, f_s) \leq \mathbb{E}\left[|X_t - X_s|^2 + |V_t - V_s|^2\right].$$

On the other hand, the paths $t \mapsto X_t(\omega)$ are continuous in time for *a.e.* ω, and (X_t, V_t) has bounded second moment on $[0,T]$; hence, by the Lebesgue continuity theorem, for fixed s the map $t \mapsto \mathbb{E}\left[|X_t - X_s|^2 + |V_t - V_s|^2\right]$ is continuous, and hence converges to 0 as t tends to s. (Alternatively, one can obtain quantitative bounds on the time continuity by estimating $\mathbb{E}\left[|X_t - X_s|^2 + |V_t - V_s|^2\right]$ in the spirit of the last equation in the proof of Lemma 3.4; we do not follow this approach here).

Then one can follow a duality argument in order to show that solutions to (3.7) are unique, which we sketch now. Take a solution f_t of (3.7) with $f_0 = 0$; we wish to show that $f_t = 0$ for any $t > 0$. For fixed $t_0 > 0$ and φ smooth with compact support in $\mathbb{R}^{2d}$, consider the solution h_t defined for $t \in [0, t_0]$ of the dual problem

$$\partial_t h_t + v \cdot \nabla_x h_t = -\Delta_v h_t + (F + H * g_t) \cdot \nabla_v h_t,$$
$$h_{t_0} = \varphi.$$

This is a linear final value problem, and by considering $h_{t_0 - t}$ one can show that it has a solution by classical arguments. In addition, for each t, this solution h_t is a continuous function, as can be seen through classical results on propagation of regularity. Then, as h_t solves the dual equation of (3.7), it holds that

$$\frac{d}{dt}\int f_t \, h_t = 0 \qquad (t \in (0, t_0)),$$

from which $\int f_{t_0} \varphi = \int f_0 h_0 = 0$. Since φ is arbitrary, this shows that $f_{t_0} = 0$ and proves the uniqueness.

3.2.3 Existence and uniqueness for the nonlinear PDE and SDE

We are now ready to finish proving points *ii)* and *iii)* in Theorem 1.2.

Step 1.- Iterative scheme: Take $f^0 \in \mathcal{P}_2(\mathbb{R}^{2d})$ and random variables (X^0, V^0) with law f_0, and let $(B_t)_{t \geq 0}$ be a given standard Brownian motion on $\mathbb{R}^d$. We define the stochastic

hal-00521216, version 2 - 1 Dec 2010

processes $(X^n_t, V^n_t)_{t\geq 0}$ recursively by

$$\begin{cases} dX^n_t = V^n_t\,dt \\ dV^n_t = \sqrt{2}\,dB_t - F(X^n_t, V^n_t)\,dt - (H * f^{n-1}_t)(X^n_t, V^n_t)\,dt, \\ (X^n_0, V^n_0) = (X^0, V^0) \end{cases}$$

for $n \geq 1$, where $f^n_t := \mathrm{law}(X^n_t, V^n_t)$ and it is understood that $f^0_t := f^0$ for all $t \geq 0$. Observe that these are linear SDEs for which existence and pathwise uniqueness are given by Lemma 3.4 since all $t \mapsto f^{n-1}_t$ is continuous for the W_2 topology. We also know from section 3.2.2 that the f^n_t are weak solutions to the PDE

$$\partial_t f^n_t + v \cdot \nabla_x f^n_t = \Delta_v f^n_t + \nabla_v \cdot (f^n_t(F + H * f^{n-1}_t)).$$

with initial condition f_0. More precisely, the following holds for $n \geq 1$ and all $\varphi \in \mathcal{C}^\infty_0([0,T) \times \mathbb{R}^{2d})$:

$$\int_{\mathbb{R}^{2d}} \varphi_0\,df^0 = -\int_0^T \int_{\mathbb{R}^{2d}} (\partial_s\varphi_s + \Delta_v\varphi_s - \nabla_v\varphi_s \cdot (F + H * f^{n-1}_s) + \nabla_x\varphi_s \cdot v)\,df^n_s\,ds. \quad (3.8)$$

Step 2.- Uniform estimates on moments of f^n_t: We shall prove the following lemma:

Lemma 3.5. *Assume Hypothesis* (1.12)–(1.16) *on F and H. Let f_0 be a probability measure on $\mathbb{R}^2$ such that*

$$\int_{\mathbb{R}^{2d}} \left(|x|^2 + e^{a|v|^p}\right) df_0(x,v) < +\infty$$

for a positive a. Then for all T there exists a positive constant b such that, for the laws f^n_t of the processes (X^n_t, V^n_t),

$$\sup_{n\geq 1} \sup_{0\leq t\leq T} \int_{\mathbb{R}^{2d}} \left(|x|^2 + e^{b|v|^p}\right) df^n_t(x,v) < +\infty.$$

Proof. We prove this lemma in two steps.

Step 1.- Bound for moments of order 2: Let

$$e_n(t) = \int_{\mathbb{R}^{2d}} |v|^2\,df^n_t(x,v)$$

for $n \geq 1$ and $t \geq 0$. Using (1.12) and (3.5) applied to the measure f^{n-1}_t, we get

$$\begin{aligned} e'_n(t) &= \frac{d}{dt}\int_{\mathbb{R}^{2d}} |v|^2\,df^n_t = 2d - 2\int_{\mathbb{R}^{2d}} v \cdot ((F + H * f^{n-1}_t)\,df^n_t(x,v) \\ &\leq 2d + 2C\int_{\mathbb{R}^{2d}} \left(1 + |v|^2 + \int_{\mathbb{R}^{2d}} |w|^2 df^{n-1}_t(y,w)\right) df^n_t(x,v) \\ &\leq C\big(1 + e_n(t) + e_{n-1}(t)\big)\,, \end{aligned}$$

hal-00521216, version 2 - 1 Dec 2010

for diverse constants C depending on F and H but not on t or n. Since moreover

$$e_n(0) = e_0(t) = e_0(0) = \int_{\mathbb{R}^{2d}} |v|^2 \, df_0(x,v),$$

for all t and n, then one can prove by induction that

$$\sup_{n \geq 0} \int_{\mathbb{R}^{2d}} |v|^2 df_t^n \leq \Big(D + \frac{1}{2}\Big) e^{2Dt} - \frac{1}{2}$$

where $D = \max\{C, e_0(0)\}$. Moreover the bound

$$\frac{d}{dt} \int_{\mathbb{R}^{2d}} |x|^2 \, df_t^n(x,v) = 2 \int_{\mathbb{R}^{2d}} x \cdot v \, df_t^n(x,v) \leq \int_{\mathbb{R}^{2d}} |x|^2 \, df_t^n(x,v) + e_n(t)$$

ensures that also $\int_{\mathbb{R}^{2d}} |x|^2 \, df_t^n(x,v)$ is bounded on $[0,T]$, uniformly in n.

Step 2.- Bound for exponential moments: Let $\alpha = \alpha(t)$ be a smooth positive function to be chosen later on and $\langle v \rangle = (1+|v|^2)^{1/2}$. Then we have the following *a priori* estimate:

$$\begin{aligned}\frac{d}{dt} \int_{\mathbb{R}^{2d}} e^{\alpha(t)\langle v \rangle^p} df_t^n(x,v) = \int_{\mathbb{R}^{2d}} \big[d\alpha p \langle v \rangle^{p-2} + \alpha p(p-2)|v|^2 \langle v \rangle^{p-4} + \alpha^2 p^2 |v|^2 \langle v \rangle^{2p-4} \\ + \alpha' \langle v \rangle^p - \alpha p \langle v \rangle^{p-2} v \cdot (F + H * f_t^{n-1}) \big] e^{\alpha \langle v \rangle^p} df_t^n(x,v).\end{aligned}$$

But, by (1.12), (3.5) applied to f_t^{n-1}, and the bound on the moment of order 2 in Step 1,

$$-v \cdot (F + H * f_t^{n-1}) \leq C \left(1 + |v|^2 + \int_{\mathbb{R}^{2d}} |w|^2 df_t^{n-1}(y,w) \right) \leq C \langle v \rangle^2$$

uniformly on $n \geq 1$ and $t \in [0,T]$, so

$$\begin{aligned}&\frac{d}{dt} \int_{\mathbb{R}^{2d}} e^{\alpha(t)\langle v \rangle^p} df_t^n(x,v) \\ &\qquad \leq \int_{\mathbb{R}^{2d}} \big[C\alpha \langle v \rangle^{p-2} + C\alpha^2 \langle v \rangle^{2p-2} + \alpha'(t) \langle v \rangle^p + C\alpha \langle v \rangle^p \big] e^{\alpha \langle v \rangle^p} df_t^n(x,v) \\ &\qquad \leq \int_{\mathbb{R}^{2d}} \big[C\alpha + C\alpha^2 + \alpha' \big] \langle v \rangle^p \, e^{\alpha \langle v \rangle^p} df_t^n(x,v).\end{aligned}$$

since $p \leq 2$. Choosing α such that $C\alpha + C\alpha^2 + \alpha' \leq 0$ (for instance, $\alpha(t) = Me^{-2Ct}$, with $0 < M \leq 1$) we conclude that

$$\frac{d}{dt} \int_{\mathbb{R}^{2d}} e^{\alpha(t)\langle v \rangle^p} df_t^n \leq 0$$

Hence, we obtain

$$\int_{\mathbb{R}^{2d}} e^{\alpha(t)\langle v \rangle^p} df_t^n(x,v) \leq \int_{\mathbb{R}^{2d}} e^{\alpha(0)\langle v \rangle^p} df_0^n(x,v) = \int_{\mathbb{R}^{2d}} e^{\alpha(0)\langle v \rangle^p} df_0(x,v),$$

which is finite provided $\alpha(0) \leq a$, which can be satisfied by taking $M = \min\{a, 1\}$ above. Then, taking $b = \alpha(T)$ we conclude that

$$\sup_{n \geq 0} \sup_{0 \leq t \leq T} \int_{\mathbb{R}^{2d}} e^{b|v|^p} df_t^n < \infty. \tag{3.9}$$

□

We notice for later use that as a direct consequence of (3.9) using the Markov inequality, there exists $C \geq 0$ such that

$$\sup_{0 \leq t \leq T} \sup_{n \geq 0} \mathbb{E}\left[\mathbb{1}_{|V_t^n| > R}\right] \leq e^{-bR^p} \sup_{0 \leq t \leq T} \sup_{n \geq 0} \mathbb{E}\left[e^{b|V_t^n|^p}\right] \leq C\, e^{-b\, R^p}\,. \tag{3.10}$$

Step 3.- Existence for the nonlinear PDE. We intend to carry out an argument analogous to the one for the existence and uniqueness of solutions for the 2D Euler equation in fluid mechanics, found for example in [28]. We will prove that the f_t^n converge to a limit, and that this limit is a solution to the nonlinear PDE. To simplify notation we drop time subscripts and use the following shortcuts:

$$v^n := V^{n+1} - V^n, \quad x^n := X^{n+1} - X^n, \quad Z^n := (X^n, V^n), \quad z := (x, v).$$

Also, we write

$$\gamma^n(t) := \mathbb{E}\left[|x^n|^2\right] + \mathbb{E}\left[|v^n|^2\right] = \mathbb{E}\left[|Z^{n+1} - Z^n|^2\right].$$

We compute, by Itô's formula, and for any $n \geq 1$,

$$\frac{d}{dt}\mathbb{E}\left[|x^n|^2\right] = 2\,\mathbb{E}\left[x^n \cdot v^n\right] \leq \mathbb{E}\left[|x^n|^2\right] + \mathbb{E}\left[|v^n|^2\right] = \gamma^n(t)\,, \tag{3.11}$$

with

$$\begin{aligned} \frac{1}{2}\frac{d}{dt}\mathbb{E}\left[|v^n|^2\right] = &- \mathbb{E}\left[v^n \cdot \left(F(Z^{n+1}) - F(Z^n)\right)\right] \\ &- \mathbb{E}\left[v^n \cdot \Big((H * f_t^n)(Z^{n+1}) - (H * f_t^{n-1})(Z^n)\Big)\right] =: T_1 + T_2. \end{aligned} \tag{3.12}$$

Estimate for T_1. We decompose the term T_1 as

$$\begin{aligned} T_1 = &- \mathbb{E}\left[(V^{n+1} - V^n) \cdot \left(F(X^{n+1}, V^{n+1}) - F(X^{n+1}, V^n)\right)\right] \\ &- \mathbb{E}\left[(V^{n+1} - V^n) \cdot \left(F(X^{n+1}, V^n) - F(X^n, V^n)\right)\right], \end{aligned}$$

which by (1.13)–(1.14) is bounded above by

$$\begin{aligned} T_1 &\leq L\,\mathbb{E}\left[|v^n|^2(1 + |V^n|^p + |V^{n+1}|^p)\right] + L\,\mathbb{E}\left[|v^n|\,|x^n|\,(1 + |V^n|^p)\right] \\ &=: T_{11} + T_{12}. \end{aligned}$$

hal-00521216, version 2 - 1 Dec 2010

Given $R > 0$, we bound T_{11} as follows:

$$\begin{aligned} T_{11} &\leq L(1+2R^p)\,\mathbb{E}\left[|v^n|^2\right] + L\,\mathbb{E}\left[|v^n|^2|V^n|^p \mathbb{1}_{|V^n|>R}\right] + L\,\mathbb{E}\left[|v^n|^2|V^{n+1}|^p \mathbb{1}_{|V^{n+1}|>R}\right] \\ &\leq L\,(1+2R^p)\gamma^n(t) + L\,\mathbb{E}\left[|v^n|^4|V^n|^{2p}\right]^{1/2}\,\mathbb{E}\left[\mathbb{1}_{|V^n|>R}\right]^{1/2} \\ &\quad + L\,\mathbb{E}\left[|v^n|^4|V^{n+1}|^{2p}\right]^{1/2}\,\mathbb{E}\left[\mathbb{1}_{|V^{n+1}|>R}\right]^{1/2} \\ &\leq C\,(1+R^p)\gamma^n(t) + C\,\mathbb{E}\left[\mathbb{1}_{|V^n|>R}\right]^{1/2} + C\,\mathbb{E}\left[\mathbb{1}_{|V^{n+1}|>R}\right]^{1/2}, \end{aligned}$$

where we have used the uniform-in-n bound on moments of f^n obtained in (3.9). For the term T_{12}, we get

$$\begin{aligned} T_{12} &\leq L(1+R^p)\,\mathbb{E}\left[|v^n||x^n|\right] + L\,\mathbb{E}\left[|v^n|\,|x^n|\,|V^n|^p \mathbb{1}_{|V^n|>R}\right] \\ &\leq \frac{L}{2}\,(1+R^p)\gamma^n(t) + \frac{L}{2}\,\mathbb{E}\left[|x^n|^2\right] + \frac{L}{2}\,\mathbb{E}\left[|v^n|^2\,|V^n|^{2p}\mathbb{1}_{|V^n|>R}\right] \\ &\leq L\,(1+R^p)\gamma^n(t) + \frac{L}{2}\,\mathbb{E}\left[|v^n|^4\,|V^n|^{4p}\right]^{1/2}\mathbb{E}\left[\mathbb{1}_{|V^n|>R}\right]^{1/2} \\ &\leq L\,(1+R^p)\gamma^n(t) + C\,\mathbb{E}\left[\mathbb{1}_{|V^n|>R}\right]^{1/2}, \end{aligned}$$

using again the bound on moments of f^n in (3.9). Finally, using (3.10), there exist constants b and C such that for all $0 \leq t \leq T$

$$T_1 \leq C(1+R^p)\,\gamma^n(t) + C\,e^{-bR^p} \tag{3.13}$$

for all n and $R > 0$.

Estimate for T_2. On the other hand, for T_2,

$$\begin{aligned} T_2 &= -\mathbb{E}\left[v^n \cdot \Big((H * f_t^n)(Z^{n+1}) - (H * f_t^n)(Z^n)\Big)\right] - \mathbb{E}\left[v^n \cdot \Big(H * (f_t^n - f_t^{n-1})(Z^n)\Big)\right] \\ &=: T_{21} + T_{22}. \end{aligned} \tag{3.14}$$

For the first term T_{21}, we proceed analogously to the estimates of T_{11} and T_{12} to obtain

$$\begin{aligned} T_{21} &= -\,\mathbb{E}\left[v^n \cdot \int_{\mathbb{R}^{2d}} (H(Z^{n+1}-z) - H(Z^n - z))\,f_t^n(x,v)\,dx\,dv\right] \\ &\leq L\,\mathbb{E}\left[|v^n| \int_{\mathbb{R}^{2d}} |Z^{n+1}-Z^n|(1+|V^n|^p+|V^{n+1}|^p+|v|^p) f_t^n(x,v)\,dx\,dv\right] \\ &\leq C\,\mathbb{E}\left[|v^n||Z^{n+1}-Z^n|\left(1+|V^n|^p+|V^{n+1}|^p\right)\right] \\ &\leq C(1+R^p)\,\gamma^n(t) + C\,e^{-bR^p} \end{aligned} \tag{3.15}$$

where the last steps where not detailed since they are very similar to the estimates of T_{11} and T_{12}, and the uniform moment bounds (3.9) and (1.16) were used. Now, for T_{22}, observe that, taking $g^n := \mathrm{law}((X^n, V^n, X^{n-1}, V^{n-1}))$, we can write the following identity

$$A := H * (f_t^n - f_t^{n-1})(Z^n) = \int_{\mathbb{R}^{4d}} \big(H(X^n - x, V^n - v) - H(X^n - y, V^n - w)\big)\,dg^n$$

hal-00521216, version 2 - 1 Dec 2010

where we used the shortcut notation dg^n for the measure $dg^n(x,v,y,w)$. By the Cauchy-Schwarz inequality and the uniform moment bounds (3.9), we get

$$
\begin{aligned}
|A| &\le L \int_{\mathbb{R}^{4d}} (|x-y| + |v-w|)(1 + |V^n|^p + |v|^p + |w|^p)\, dg^n(x,v,y,w) \\
&\le L(1+|V^n|^p)\,\mathbb{E}\left[|x^{n-1}| + |v^{n-1}|\right] \\
&\quad + L\left(\int_{\mathbb{R}^{4d}} (|x-y|+|v-w|)^2\, dg^n\right)^{1/2} \left(\int_{\mathbb{R}^{4d}} (|v|^p + |w|^p)^2\, dg^n\right)^{1/2} \\
&\le C(1+|V^n|^p)\mathbb{E}\left[|x^{n-1}| + |v^{n-1}|\right] + C\left(\mathbb{E}\left[|Z^n - Z^{n-1}|^2\right]\right)^{1/2} \\
&\le C(1+|V^n|^p)\,\gamma^{n-1}(t)^{1/2}\,.
\end{aligned}
$$

Using the expression of T_{22}, we deduce

$$
\begin{aligned}
T_{22} &\le \mathbb{E}\left[|v^n||H * (f^n_t - f^{n-1}_t)(Z^n)|\right] \le C\gamma^{n-1}(t)^{1/2}\,\mathbb{E}\left[|v^n|(1+|V^n|^p)\right] \\
&\le C\gamma^{n-1}(t) + \mathbb{E}\left[|v^n|^2\right]\mathbb{E}\left[|V^n|^{2p}\right] \le C\gamma^{n-1}(t) + C\gamma^n(t)\,,
\end{aligned} \tag{3.16}
$$

again by the Cauchy-Schwarz inequality and the uniform moment bounds (3.9).

Hence, putting (3.13), (3.14), (3.15) and (3.16) in (3.11) and (3.12), we conclude

$$
\frac{d}{dt}\gamma^n(t) \le C(1+R^p)\gamma^n(t) + C\gamma^{n-1}(t) + Ce^{-bR^p}.
$$

Induction Argument. Taking $R > 1$ we may write that

$$
\frac{d}{dt}\gamma^n(t) \le C\left(r\gamma^n(t) + \gamma^{n-1}(t) + e^{-r}\right), \tag{3.17}
$$

for some other constant $C > 0$ and for any all $r > 1$. Gronwall's Lemma then proves that

$$
\gamma^n(t) \le C\int_0^t e^{Cr(t-s)}\gamma^{n-1}(s)\, ds + Ce^{-r}te^{Crt},
$$

and iterating this inequality gives

$$
\begin{aligned}
\gamma^n(t) &\le C^n \int_0^t e^{Cr(t-s)}\gamma^0(s)\frac{(t-s)^{n-1}}{(n-1)!}\, ds + Cte^{-r}e^{Crt}\sum_{i=0}^{n-1}\frac{C^i t^i}{(i+1)!} \\
&\le C^n e^{Crt}\frac{t^{n-1}}{(n-1)!}\int_0^t \gamma^0(s)\, ds + Cte^{-r}e^{Crt}e^{Ct} \\
&\le C^n e^{Crt} t^n \sup_{s\in[0,t]} \gamma^0(s) + Cte^{Ct}e^{r(Ct-1)}.
\end{aligned}
$$

Taking $r = n$, we obtain

$$
\gamma^n(t) \le \exp\left(n(\ln(Ct) + Ct)\right) \sup_{s\in[0,t]} \gamma^0(s) + Cte^{Ct}e^{n(Ct-1)}.
$$

hal-00521216, version 2 - 1 Dec 2010

Choosing $0 < T_* < T$ small enough such that $\lambda := \max\{CT_* - 1, \ln(CT_*) + CT_*\} < 0$, then

$$\sup_{t\in[0,T_*]} \gamma^n(t) \le e^{\lambda n} \left(\sup_{s\in[0,T_*]} \gamma^0(s) + CTe^{CT} \right). \tag{3.18}$$

Since by definition $W_2(f_t^{n+1}, f_t^n)^2 \le \mathbb{E}|Z_t^{n+1} - Z_t^n|^2 = \gamma^n(t)$, we conclude that the sequence of curves $\{t \in [0,T_*] \mapsto f_t^n\}_{n\ge 0}$ is a Cauchy sequence in the metric space $C([0,T_*], \mathcal{P}_2(\mathbb{R}^{2d}))$ equipped with the distance

$$\mathcal{W}_2(f,g) := \sup_{t\in[0,T_*]} W_2(f_t, g_t).$$

By completeness of this space, we define $f \in C([0,T_*], \mathcal{P}_2(\mathbb{R}^{2d}))$ by $f_t := \lim_{n\to+\infty} f_t^n$ for $t \in [0,T_*]$.

This convergence and the uniform moment bounds on f^n in (3.9) allow us to pass to the limit in (3.8). Let us point out how to deal with the nonlinear term in the equation: observe first that for fixed s, it is given by

$$\int_{\mathbb{R}^{2d}} \nabla_v \varphi_s \cdot H * f_s^{n-1}\, df_s^n = \int_{\mathbb{R}^{4d}} \nabla_v \varphi_s(x,v) \cdot H(x-y, v-w) df_s^{n-1}(y,w)\, df_s^n(x,v). \tag{3.19}$$

But on the one hand f_s^{n-1} and f_s^n converge to f_s for the W_2 topology, hence so does $f_s^{n-1} \otimes f_s^n$ to $f_s \otimes f_s$ (in $\mathbb{R}^{4d}$). On the other hand

$$|\nabla_v \varphi_s(x,v) \cdot H(x-y, v-w)| \le \|\nabla_v \varphi_s\|_{L^\infty}(1 + |v| + |w|)$$

by (1.15). Therefore (3.19) converges to

$$\int_{\mathbb{R}^{4d}} \nabla_v \varphi_s(x,v) \cdot H(x-y, v-w) df_s(y,w)\, df_s(x,v) = \int_{\mathbb{R}^{2d}} \nabla_v \varphi_s \cdot H * f_s\, df_s$$

for all s. Uniform-in-s bounds finally allow to pass to the limit in the integral in s.

With this, we have shown that f_t is a solution on $[0,T_*]$ of the nonlinear PDE (1.4). Now, one can extend the solution to the whole interval $[0,T]$ by iterating this procedure, starting from T_*. This can be done since the additional time T_*' for which we can extend a solution starting at T_* depends only on moment bounds on f_{T_*}, for which we have the bound (3.9), valid up to T.

Step 4.- Existence for the nonlinear SDE: Now we use f_t to define the process (X_t, V_t) by

$$\begin{cases} dX_t = V_t\, dt \\ dV_t = \sqrt{2}\, dB_t - F(X_t, V_t)\, dt - (H * f_t)(X_t, V_t)\, dt, \\ (X_0, V_0) = (X^0, V^0) \end{cases} \tag{3.20}$$

thanks to Lemma 3.4. Observe that for all t, f_t is the W_2-limit of f_t^n and $p \leq 2$, so that

$$\int_{\mathbb{R}^{2d}} |v|^p df_t(x,v) = \int_{\mathbb{R}^{2d}} |v|^p df_t^n(x,v) \leq C,$$

uniformly in $t \in [0,T]$ since the f_t^n have second moments bounded according to Lemma 3.5. If g_t is the law of (X_t, V_t), then, as in section 3.2.2, g_t is a weak solution of the linear PDE

$$\partial_t g_t + v \cdot \nabla_x g_t = \Delta_v g_t + \nabla_v((F + H * f_t) g_t).$$

Of course, f_t is also a solution of the same linear PDE; by uniqueness of solutions to this linear PDE (see again section 3.2.2), we deduce that $f_t = g_t$, and hence (X_t, V_t) is a solution to the nonlinear SDE (1.3) on $[0,T]$.

Step 5.- Uniqueness for the nonlinear PDE (1.4)**:** Now, take two solutions f^1, f^2 of the nonlinear PDE, and define the processes (X_t^1, V_t^1) and (X_t^2, V_t^2) by (3.20), putting f^1 and f^2 in the place of f, respectively. As the law of (X_t^i, V_t^i) ($i = 1, 2$) solves the linear PDE (3.7) with f_t^i instead of g_t, so this law must in fact be f_t^i by uniqueness of the linear PDE. Then, if we follow the same calculation we did in step 3 above, we obtain the following instead of (3.17):

$$\frac{d}{dt}\gamma(t) \leq Cr\gamma(t) + C\, e^{-r}, \tag{3.21}$$

for a constant C and any $r \geq 1$, with $\gamma(t) := \mathbb{E}\left[|X_t^1 - X_t^2|^2\right] + \mathbb{E}\left[|V_t^1 - V_t^2|^2\right]$. For the above to be valid, we need a bound on exponential moments of f_t^1 and f_t^2. This estimate can be obtained in a similar way as in Lemma 3.5 and therefore, we omit the proof:

Lemma 3.6. *Assume hypotheses* (1.12)–(1.16) *on F and H. Let $(f_t)_{t\geq 0}$ be a solution to* (1.4) *with initial datum a probability measure f_0 on $\mathbb{R}^{2d}$ such that*

$$\int_{\mathbb{R}^{2d}} \left(|x|^2 + e^{a|v|^p}\right) f_0(x,v)\, dx\, dv < +\infty$$

for a positive a. Then for all T there exists $b > 0$ which depends only on f_0 and T, such that

$$\sup_{0\leq t\leq T} \int_{\mathbb{R}^{2d}} \left(|x|^2 + e^{b|v|^p}\right) f_t(x,v)\, dx\, dv < +\infty.$$

Observe now that

$$\gamma(0) = \mathbb{E}\left[|X_0^1 - X_0^2|^2 + |V_0^1 - V_0^2|^2\right] = 0$$

since $X_0^1 = X_0^2 = X^0$, and similarly for V.

hal-00521216, version 2 - 1 Dec 2010

hal-00521216, version 2 - 1 Dec 2010

Assume now that γ is non identically 0. Then, with the same argument as in step 3 of the proof of Theorem 1.1, whenever $0 < \gamma(t) < 1/e$ we can choose $r := -\ln \gamma(t)$ in (3.21) to obtain

$$\frac{d}{dt}\gamma(t) \leq -C\gamma(t)\ln\gamma(t) + C\gamma(t) \leq -2C\gamma(t)\ln\gamma(t). \tag{3.22}$$

If $\gamma(t) = 0$ at some point, then one can see that $\frac{d}{dt}\gamma(t) \leq 0$ by letting $r \to +\infty$ in (3.21), and in that case the inequality (3.22) holds trivially (again setting $z \ln z = 0$ at $z = 0$ by continuity). Hence, (3.22) holds as long as $0 \leq \gamma(t) < 1/e$. By Gronwall's Lemma, this implies that $\gamma(t) = 0$ for $t \in [0,T]$ and shows that f_t^1 and f_t^2 coincide, proving that solutions to the nonlinear PDE are unique.

Step 6.- Uniqueness for the nonlinear SDE (1.3)**:** Take two pairs of stochastic processes (X_t^1, V_t^1) and (X_t^2, V_t^2) which are solutions to the nonlinear SDE (1.3). Then, their laws are solutions to the nonlinear PDE (1.4), and by the previous step we know that they must be the same. If we call f_t their common law, then both (X_t^1, V_t^1) and (X_t^2, V_t^2) are solutions to the linear SDE (3.3) with f_t instead of g_t, and by uniqueness of this linear SDE (see section 3.2.1), they must coincide.

This concludes the proof of Theorem 1.2.

Acknowledgments. The last two authors acknowledge support from the project MTM2008-06349-C03-03 DGI-MCI (Spain) and 2009-SGR-345 from AGAUR-Generalitat de Catalunya. All authors were partially supported by the ANR-08-BLAN-0333-01 Projet CBDif-Fr. This work was initiated while the first author was visiting UAB; it is a pleasure for him to thank this institution for its kind hospitality.

References

[1] M. Agueh, R. Illner and A. Richardson. Analysis and simulations of a refined flocking and swarming model of Cucker-Smale type. Preprint, 2010.

[2] I. Aoki. A Simulation Study on the Schooling Mechanism in Fish. *Bull. Jap. Soc. Sci. Fisheries* 48:1081–1088, 1982.

[3] A. Barbaro, K. Taylor, P. F. Trethewey, L. Youseff and B. Birnir. Discrete and continuous models of the dynamics of pelagic fish: application to the capelin. *Math. and Computers in Simulation* 79:3397–3414, 2009.

[4] S. Benachour, B. Roynette, D. Talay and P. Vallois. Nonlinear self-stabilizing processes. I. Existence, invariant probability, propagation of chaos. *Stoch. Proc. Appl.* 75(2):173–201, 1998.

hal-00521216, version 2 - 1 Dec 2010

[5] F. Bolley and A. Guillin and F. Malrieu. Trend to equilibrium and particle approximation for a weakly selfconsistent Vlasov-Fokker-Planck equation. *M2AN* 44(5):867–884, 2010.

[6] W. Braun and K. Hepp. The Vlasov Dynamics and Its Fluctuations in the 1/N Limit of Interacting Classical Particles. *Commun. Math. Phys.* 56:101–113, 1977.

[7] M. Burger, V. Capasso and D. Morale. On an aggregation model with long and short range interactions. *Nonlinear Analysis. Real World Appl.* 8:939–958, 2007.

[8] S. Camazine, J.-L. Deneubourg, N. R. Franks, J. Sneyd, G. Theraulaz and E. Bonabeau. Self-Organization in Biological Systems. Princeton University Press (2003).

[9] J. A. Cañizo, J. A. Carrillo and J. Rosado. A well-posedness theory in measures for some kinetic models of collective motion. *M3AS*, to appear.

[10] J. A. Carrillo, M. R. D'Orsogna and V. Panferov. Double milling in self-propelled swarms from kinetic theory, *Kinetic and Related Models* 2:363-378, 2009.

[11] J. A. Carrillo, M. Fornasier, J. Rosado and G. Toscani. Asymptotic Flocking Dynamics for the kinetic Cucker-Smale model, *SIAM J. Math. Anal.* 42:218–236, 2010.

[12] J. A. Carrillo, M. Fornasier, G. Toscani and F. Vecil, Particle, Kinetic, and Hydrodynamic Models of Swarming, to appear as a book chapter in Modeling and Simulation in Science, Engineering and Technology, Birkhauser, Boston.

[13] P. Cattiaux, A. Guillin and F. Malrieu. Probabilistic approach for granular media equations in the non uniformly convex case. *Prob. Theor. Rel. Fields* 140(1-2):19–40, 2008.

[14] I. D. Couzin, J. Krause, N. R. Franks, S. A. Levin. Effective leadership and decision making in animal groups on the move. *Nature* 433:513–516, 2005.

[15] F Cucker and E. Mordecki. Flocking in noisy environments. *J. Math. Pures Appl.* 89(3):278–296, 2008.

[16] F. Cucker and S. Smale. Emergent behavior in flocks. *IEEE Trans. Automat. Control* 52:852-862, 2007.

[17] P. Degond and S. Motsch. Continuum limit of self-driven particles with orientation interaction. *Math. Mod. Meth. Appl. S.* 18(1):1193–1215, 2008.

[18] R. Dobrushin. Vlasov equations. *Funct. Anal. Appl.* 13:115–123, 1979.

hal-00521216, version 2 - 1 Dec 2010

[19] M. R. D'Orsogna, Y. L. Chuang, A. L. Bertozzi and L. Chayes. Self-propelled particles with soft-core interactions: patterns, stability, and collapse. *Phys. Rev. Lett.* 96, 2006.

[20] R. Duan, M. Fornasier and G. Toscani. A kinetic flocking model with diffusion. *Comm. Math. Phys.*, to appear, 2010.

[21] S. N. Ethier and T. G. Kurtz. *Markov Processes: Characterization and Convergence*. Wiley Series in Probability and Statistics. Wiley, 1986.

[22] S.-Y. Ha and E. Tadmor. From particle to kinetic and hydrodynamic descriptions of flocking. *Kinetic and Related Models* 1:415–435, 2008.

[23] S.-Y. Ha, T. Ha and J. Ho Kim. Emergent behavior of a Cucker-Smale type particle model with nonlinear velocity couplings. Preprint, 2009.

[24] C. K. Hemelrijk and H. Hildenbrandt. Self- Organized Shape and Frontal Density of Fish Schools. *Ethology* 114, 2008.

[25] A. Huth and C. Wissel. The Simulation of the Movement of Fish Schools. *J. Theo. Bio.*, 1992.

[26] Y. X. Li, R. Lukeman and L. Edelstein-Keshet, Minimal mechanisms for school formation in self-propelled particles, *Physica D* 237:699–720, 2008.

[27] F. Malrieu. Logarithmic Sobolev inequalities for some nonlinear PDE's. *Stoc. Proc. Appl.* 95(1):109–132, 2001.

[28] C. Marchioro and M. Pulvirenti. *Mathematical Theory of Incompressible Nonviscous Fluids*, Springer, Berlin, 1993.

[29] H. P. McKean. Propagation of chaos for a class of non-linear parabolic equations. In *Lecture Series in Differential Equations, Session 7*, Catholic Univ., 1967.

[30] S. Méléard. Asymptotic behaviour of some interacting particle systems; McKean-Vlasov and Boltzmann models. In *Probabilistic models for nonlinear partial differential equations (Montecatini Terme, 1995)*, Lecture Notes in Math. 1627. Springer, Berlin, 1996.

[31] D. Morale, V. Capasso and K. Oelschläger. An interacting particle system modelling aggregation behavior: from individuals to populations, *J. Math. Biol.* 50:49–66, 2005.

[32] H. Neunzert. An introduction to the nonlinear Boltzmann-Vlasov equation. In *Kinetic theories and the Boltzmann equation* (Montecatini Terme, 1981), Lecture Notes in Math. 1048. Springer, Berlin, 1984.

[33] J. Parrish, and L. Edelstein-Keshet. Complexity, pattern, and evolutionary trade-offs in animal aggregation. *Science* 294: 99–101, 1999.

[34] A.-S. Sznitman. Topics in propagation of chaos. In *Ecole d'Eté de Probabilités de Saint-Flour XIX — 1989*, Lecture Notes in Math. 1464. Springer, Berlin, 1991.

[35] C. Villani. Limite de champ moyen. Lecture notes, 2002.

[36] C. Villani. *Optimal transport, Old and new.* Springer-Verlag, Berlin, 2009.

[37] C. Yates, R. Erban, C. Escudero, L. Couzin, J. Buhl, L. Kevrekidis, P. Maini and D. Sumpter. Inherent noise can facilitate coherence in collective swarm motion. *Proc. Nat. Acad. Sci.* 106(14):5464–5469, 2009.

hal-00521216, version 2 - 1 Dec 2010