On The Mean Field limit for Cucker-Smale models

Roberto Natalini, Thierry Paul

ON THE MEAN FIELD LIMIT FOR CUCKER-SMALE MODELS

ROBERTO NATALINI AND THIERRY PAUL

ABSTRACT. In this note, we consider the Cucker-Smale dynamical system and we derive rigorously, using the Wasserstein distance, the Vlasov-type equation introduced in [9] in the mean-field limit without using empirical measures.

CONTENTS

1. INTRODUCTION

The Cucker-Smale model [2, 3] is a particles system which exhibit the so-called flocking phenomenon, namely the dynamical property of alignment of all the velocities and gathering of all the positions asymptotically when the time of evolution diverges. Its study has developed an intense activity these last years, see for example the intensive bibliography in the recent paper [7].

It consists in the following vector field on $\mathbf{R}^{2dN}$

$$(1) \qquad \begin{cases} \dot{x}_i &= v_i \\ \dot{v}_i &= \frac{1}{N}\sum_{j=1}^{N}\psi(x_i-x_j)(v_j-v_i). \end{cases} \qquad x_i, v_i \in \mathbf{R}^d, \quad i = 1, \dots, N$$

where $\psi : \mathbf{R}^d \to \mathbf{R}$ is a bounded Lipschitz continuous positive nonincreasing function.

Associated to (1), the following Vlasov type kinetic equation on $\mathbf{R}^{2d}$ was introduced in [9]:

$$(2) \quad \partial_t\rho^t(x,v) + v\cdot\nabla_x\rho^t(x,v) + \nabla_v\left(\rho^t(x,v)\int_{\mathbf{R}^{2d}}\psi(x-y)(v-w)\rho^t(y,w)dydw\right) = 0.$$

So far the kinetic non-linear equation (2) has been derived in a Lagrangian point of view, i.e. by following the trajectories solving the vector field (1) in the so-called empirical measure

$$\Pi_{(x_1,\dots,x_n;v_1,\dots,v_N)}(x,v) := \frac{1}{N}\sum_{i=1}^{N}\delta(x-x_i)\delta(v-v_i). \tag{3}$$

Indeed, let Φ^t be the flow generated by the system (1) and let us define

$$\Pi^t_{(x_1,\dots,x_n;v_1,\dots,v_N)} := \Pi_{\Phi^{-t}(x_1,\dots,x_n;v_1,\dots,v_N)}. \tag{4}$$

One shows easily that $\Pi^t_{(x_1,\dots,x_n;v_1,\dots,v_N)}(x,v)$ solves (2) for each N.

This point of view follows directly the Dobrushin way [4] of deriving the Vlasov equation for the large N limit of Hamiltonian vector fields. The Cucker-Smale model (1) and its Vlasov type associated equation (2) have been extensively studied in [7] and we quote the following of their results: the function ψ is supposed to be positive, bounded, uniformly Lipschitz continuous and nonincreasing.

(1) the solutions of (1) satisfy for all t

$$\sup_{i,j\leq N}\|x_i(t)-x_j(t)\|\leq C \quad\text{and}\quad \sup_{i,j\leq N}\|v_i(t)-v_j(t)\|\leq e^{-Dt}$$

where C and D are two positive constants depending only on $\sup_{i,j\leq N}\|x_i(0)-x_j(0)\|$ and $\sup_{i,j\leq N}\|v_i(0)-v_j(0)\|$.

(2) let $\Pi_{(x_1,\dots,x_n;v_1,\dots,v_N)}\to\mu(0)$ as $N\to\infty$, then

$$\overline{\lim_{N\to\infty}}\sup_{t\in\mathbf{R}} W_p(\Pi^t_{(x_1,\dots,x_n;v_1,\dots,v_N)}-\mu(t))=0$$

where W_p is the Wasserstein distance of exponent $p\in\mathbf{N}$ (see definition in Remark 3.5 below).

(3) Let $\int_{\mathbf{R}^{2d}}\mu(0)=1,\ \int_{\mathbf{R}^{2d}}v\mu(0)=:\bar v,\ \int_{\mathbf{R}^{2d}}v^2\mu(0)<\infty$. Then, for some $E,F>0$ and every p,

$$\left(\int\|v-\bar v\|^p d\mu(t)\right)^{\frac{1}{p}}\leq Ee^{-Ft}.$$

Recently, a more Eulerian way of deriving the Vlasov equation associated to Hamiltonian vector fields was introduced ([6], after [5]). It consists in considering the movement of a generic particle solution to the Liouville equation associated to an Hamiltonian system, rather than considering the empirical measure associated to the Hamiltonian flow. Transposed in the (non Hamiltonian) present situation, the method leads to the following.

We will consider the pushforward[1] by the flow Φ^tassociated to (1) of a compactly supported, symmetric by permutations of the variables N-body probability density ρ_N^{in} on phase space $\mathbf{R}^{2dN}$.

[1] We recall that the pushforward of a measure μ by a measurable function Φ is $\Phi\#\mu$ defined by $\int\varphi d(\Phi\#\mu):=\int(\varphi\circ f)d\mu$ for every measurable function f.

In order to describe the movement of a "generic" particle in the limit of diverging number of particles, we want to perform an average on the N particles but one. This means that we consider the first marginal of

$$\Phi^t \# \rho_N^{in} := \rho_N^t \tag{5}$$

that is

$$\rho_{N;1}^t(x,\xi) := \int_{\mathbf{R}^{2(d-1)N}} \rho_N^t(x,\xi,x_2,\xi_2,\dots,x_N,\xi_N)dx_2d\xi_2\dots dx_Nd\xi_N. \tag{6}$$

Then, when ρ_N^{in} is factorized as $\rho_N^{in} = (\rho^{in})^{\otimes N}$, where ρ^{in} is a probability measure on $\mathbf{R}^{2d}$, we will prove that the marginal of ρ_N^t tends, as $N \to \infty$, to the solution of the effective non-linear Vlasov type equation (2) with $\rho^{t=0} = \rho^{in} \in L^1(\mathbf{R}^{2d}, dxdv)$.

In fact, we can prove (see Remark 3.5 below) that the marginals at every order n of ρ_N^t as defined by

$$\rho_{N;n}^t(x_1,\xi_1,\dots,x_n,\xi_n) := \int_{\mathbf{R}^{2dN(-n)}} \rho_N^t(x_i,\xi_1,x_2,\xi_2,\dots,x_N,\xi_N)dx_{n+1}d\xi_{n+1}\dots dx_Nd\xi_N, \tag{7}$$

tends, in Wasserstein topology of any exponent $p \geq 1$, to the nth tensorial power of the solution ρ^t of the Vlasov equation, i.e. $\rho_{N;n}^t \to (\rho^t)^{\otimes n}$. Nevertheless, for sake of clarity of this short note, we will present in detail only the case $n = 1$.

This result (for $n = 1$), i.e. Theorem 2.2 below, has to be put in correspondence with the item 2 of the results of [7] just described, but there are several differences. First, our results hold true for more general initial data than the empirical measures. Second, our result will not be uniform in time as in item 1. Third, we will get an explicit rate of convergence in the asymptotic $N \to \infty$.

Our methods will also apply to systems of the general form

$$\begin{cases} \dot{x_i} &= v_i \\ \dot{v_i} &= = \frac{1}{N}\sum_{j=1}^{N} \gamma(x_i - x_j, v_i - v_j). \end{cases} \qquad x_i, v_i \in \mathbf{R}^d, \quad i = 1,\dots,N \tag{8}$$

where $\gamma(x,v) : \mathbf{R}^{2d} \to \mathbf{R}^d$ is a Lipschitz continuous function, bounded in (x,v), Corollary 2.5 or bounded in x and sublinear in v, Theorem 2.4 and Corollary 2.6.

In the following, we will denote by $\mathrm{Lip}(\gamma)_{(x,v)}$ the (local) Lipschitz constant of γ at the point $(x,v) \in \mathbf{R}^{2d}$ and we will suppose, without loss of generality, that there exists $\gamma_0 > 0$ such that, for all $(x,v) \in \mathbf{R}^{2d}$,

$$\begin{aligned} \gamma(x,v) &\leq \gamma_0|v| \\ \mathrm{Lip}(\gamma)_{(x,v)} &\leq \gamma_0|v| \end{aligned}$$

where $|\cdot|$ denotes the Euclidean norm on $\mathbf{R}^d$.

We will show in Appendix A the gobal existence of the solutions to (8), with an exponential growth in time with respect to the initial conditions. Of course, uniqueness is a consequence of (the iteration in time of) the Cauchy-Lipschitz Theorem.

The Cucker-Smale vector fields satisfies this assumptions with $\gamma_0 = \|\psi\|_{L^\infty}$.

Note that the following Vlasov type kinetic equation on $\mathbf{R}^d$ is obviously associated to (8), in a natural way:

$$\partial_t\rho^t(x,v)+v\cdot\nabla_x\rho^t(x,v)+\nabla_v\left(\rho^t(x,v)\int_{\mathbf{R}^{2d}}\gamma(x-y,v-w)\rho^t(y,w)dydw\right)=0. \tag{9}$$

Let us immediately notice that our assumptions on γ contain the two assumptions in Theorem 2.3 in [10]. The first one because of the Lipschitz and sublinearity properties of γ, and the second one thanks to the fact, see [11, 12], that

$$\sup_{|\operatorname{Lip}(\varphi)|\leq 1}\left|\int\varphi(\mu-\nu)\right|\leq W_1(\mu,\nu)$$

where W_1 is the Wasserstein distance of exponent 1, as defined in Remark 3.5 in the present paper. Therefore, thanks to the results of [10], there exist a unique continuous solution to (9) in $C_0(\mathbf{R},\mathcal{P}_c(\mathbf{R}^{2d}))$, where $\mathcal{P}_c(\mathbf{R}^{2d})$ is the space of compactly supported probability measures on $\mathbf{R}^{2d}$.

2. The results

In this section we will state our main results, but let us start by recalling the definition of the second order Wasserstein distance $\operatorname{dist}_{\mathrm{MK},2}$ (see [1, 11, 12]).

Definition 2.1 (quadratic Wasserstein distance)**.** *The Wasserstein distance of order two between two probability measures μ,ν on $\mathbf{R}^m$ with finite second moments is defined as*

$$W_2(\mu,\nu)^2=\inf_{\gamma\in\Gamma(\mu,\nu)}\int_{\mathbf{R}^m\times\mathbf{R}^m}|x-y|^2\gamma(dx,dy)$$

where $\Gamma(\mu,\nu)$ is the set of probability measures on $\mathbf{R}^m\times\mathbf{R}^m$ whose marginals on the two factors are μ and ν. E

That is to say that elements γ of $\Gamma(\mu,\nu)$, called couplings or transportation maps or transport maps or just maps, depending of the authors, satisfy, for every test functions a and b in $C_0(\mathbf{R}^d)$,

$$\int_{\mathbf{R}^{2m}}a(x)\gamma(dx,dy)=\int_{\mathbf{R}^m}a(x)\mu(dx),\quad \int_{\mathbf{R}^{2m}}b(y)\gamma(dx,dy)=\int_{\mathbf{R}^m}b(y)\nu(dy).$$

The symmetry property in μ,ν is obvious and the separability property is easily proven by taking the optimal coupling between μ and itself equal to

$$\gamma=\mu\delta(x-y). \tag{10}$$

Conversely, if $W_2(\mu,\nu)^2=0$, then $\int_{\mathbf{R}^m\times\mathbf{R}^m}|x-y|^2\gamma(dx,dy)=0$ for some γ so that $x=y$ γ a.e. and, for every Borel function φ,

$$\int\varphi(x)\mu(dx)=\int\varphi(x)\gamma(dx,dy)=\int\varphi(y)\gamma(dx,dy)=\int\varphi(y)\nu(dy)\Rightarrow\mu=\nu.$$

The proof of the triangular inequality is much more involved, see again [1, 11, 12]).

We can now state the main result of this note, proven in Section 3.4.

Theorem 2.2 (Cucker-Smale model)**.**
Let Φ^t be the flow generated by the system (1), ψ bounded positive nonincreasing Lipschitz continuous, and let ρ^t be the solution to (2) with an initial condition $\rho^{in} \in L^1(\mathbf{R}^{2d})$ compactly supported.
Let moreover $\rho^t_{N;1}$ be the first marginal of $\rho^t_N := \Phi^t\#(\rho^{in})^{\otimes N}$, as defined in (6).
Then, for all $N > 1,\ t \in \mathbf{R}$,

$$W_2(\rho^t_{N;1}, \rho^t) \leq 4\|\psi\|^2_\infty(2|\bar{v}| + |supp[\rho^{in}]|)\left(\frac{e^{Lt}-1}{L}\right)^{\frac{1}{2}} N^{-\frac{1}{2}}$$

with

$$L := 2(1 + 8\|\psi\|^2_\infty\|v\|^2_{L^\infty(supp[\rho^{in}])}) \text{ and } \bar{v} = \int v\rho^{in} dxdv,$$

where $\|v\|_{L^\infty(supp[\rho^{in}])} := \sup_{(x,v)\in supp[\rho^{in}]} |v|$.

Remark 2.3. *There exists an equivalent result for higher orders marginals of $\rho(t)$ and other Monge-Kantorovich distances but we prefer in this note to concentrate on the case of first marginal and quadratic Wasserstein distance. The proof in the more general situation is very close to the one presented here. See Remark 3.5 for some details.*

Theorem 2.2 is actually a corollary of the following more general result, proven in Section 3.2

Theorem 2.4 (General Cucker-Smale model with general sublinear force)**.**
Let Φ^t the flow defined by the system (8), $\gamma(x, v)$ Lipschitz continuous bounded in x and sublinear in v, and let ρ^t be the solution to (9), γ bounded on $\mathbf{R}^{2d}$, with an initial condition $\rho^{in}_1 \in L^1(\mathbf{R}^{2d})$ compactly supported.
Finally, let $\rho^t_{N;1}$ be the first marginal of $\rho^t_N := \Phi^t\#(\rho^{in})^{\otimes N}$, as defined in (6).
Then, for all $N > 1, t \in \mathbf{R}$,

$$W_2(\rho^t_{N;1}, \rho^t) \leq C(t)N^{-\frac{1}{2}}$$

with

$$\begin{aligned} C(t) &:= \left(4\int_0^t \sup_{(x,v),(x',v')\in supp[\rho^t]} |\gamma(x-x', v-v')|^2 e^{\int_s^t L(u)du} ds\right)^{\frac{1}{2}}, \\ \underline{L}(u) &:= 2(1 + 2\min(\sup_{\substack{i,l=1,\dots,N \\ (Y,\Xi)\in supp[\rho^u_N]}} \mathrm{Lip}\,(\gamma)^2_{(y_i-y_l,\xi_i-\xi_l)}, \sup_{(x,v),(x',v')\in supp[\rho^u]} \mathrm{Lip}\,(\gamma)^2_{(x-x',v-v')}))). \end{aligned}$$

The following result is elementarly derived from Theorem 2.4.

Corollary 2.5 (General Cucker-Smale model with bounded global Lipschitz force)**.**
Let ρ^t be the solution to (9), γ globaly Lipschitz and bounded on $\mathbf{R}^{2d}$, with an initial condition $\rho^{in} \in L^1(\mathbf{R}^{2d})$ and Φ^t the flow deinfed by the system (8). Finally, let $\rho^t_{N;1}$ be the first marginal of $\rho^t_N := \Phi^t\#(\rho^{in})^{\otimes N}$, as defined in (6).

Then, for all $N > 1, t \in \mathbf{R}$,

$$W_2(\rho^t_{N;1}, \rho^t) \le C(t) N^{-\frac{1}{2}}$$

with

$$C(t) := \left(4\|\gamma\|_\infty \frac{e^{\Lambda t}-1}{\Lambda}\right)^{\frac{1}{2}} \qquad \Lambda := 2(1+2\operatorname{Lip}(\gamma)^2).$$

Note that no need of compacity of the support of ρ^{in} is needed any more in Corollary 2.5.

Theorem 2.4 gives a precise estimate involving, through the expression of the functions $C(t)$ and $L(t)$, the knowledge of the size of the support of the initial data propagated by the particle flow $\Phi(t)$ driven by (9) and the kinetic flow induced by (9). This information might be given by the explicit models, i.e. the function γ, as it is the case for the Cucker-Smale model (see also the very end of this section).

In the general case, the boundness in space and sublinearity in velocities of γ allow to control the increasing of the flow $\Phi(t)$, and leads to our next result, whose proof is given in Section 3.3 below.

Corollary 2.6. *Under the same hypothesis as in Theorem 2.4 we have that, for all $N > 1, t \in \mathbf{R}$,*

$$W_2(\rho^t_{N;1}, \rho^t) \le C(t) N^{-\frac{1}{2}}$$

with

$$\begin{aligned} C(t) &= 2\gamma_0(\|v\|_{L^\infty(supp[\rho^{in}])} + \|v\|_{L^1(supp[\rho^{in}])}) \\ &\quad \times \left(2e^{\left(e^{4\gamma_0 t}4\gamma_0(\|v\|_{L^\infty(supp[\rho^{in}])}+\|v\|_{L^1(supp[\rho^{in}])})^2\right)} e^{4t}(e^{4\gamma_0 t}-1)\right)^{\frac{1}{2}}. \end{aligned} \tag{11}$$

Let us finish this section of results by a simple remark. Flocking properties of the solution of the Vlasov kinetic equation (2) can be derived from the corresponding properties of the particle system (1), as in [8, 7]. But they can also be derived by a direct PDE study of (2), as in [10]. This suggest a kind of inverse questioning: suppose one determines some flocking properties for the solution of a general kinetic Vlasov equation, e.g. (9). Does this infer on flocking properties for the corresponding particle system, e.g. (8)? Our Corollary 2.6 gives some insight on this problem, as it tells us quantitatively how close is the solution of the kinetic equation (and its tensorial powers) to the marginals of any orders of the pushforward obey the corresponding particle flow of a general N particle density.

More precisely, suppose that the solution ρ^t to (9) satisfies the following property:

$$supp[\rho^t] \subset B(\bar{x}+t\bar{v}, X) \times B(\bar{v}, Ve^{-\alpha t}) \tag{12}$$

for some positive constants $\bar{x}, X, \bar{v}, V, \alpha$.

This implies easily that $L(u)$ in Theorem 2.4 can be easily estimated by

$$L(u) \le 2(1+8\gamma_0^2(\bar{v}^2+V^2e^{-2\alpha u})) \le 2(1+8\gamma_0^2(\bar{v}^2+V^2))$$

and, therefore, in the same Theorem,

$$C(t) \leq 4\gamma_0(\bar{v}^2+V^2)\frac{e^{2(1+8(\bar{v}^2+V^2))t}-1}{2(1+8(\bar{v}^2+V^2))}. \tag{13}$$

Note that in (13), $C(t)$ has an exponential growth, and not a double exponential one as in Corollary 2.6. In particular, since (12) holds true for the Cucker-Smale models by Theorem 3.1, equation (3.2), in [10], (13) provides an alternative proof of Theorem 2.2 with (slightly) different values of the constants involved in its statement.

Of course the first marginal $\rho^t_{N;1}$ of the pushforward of $(\rho^{in})^{\otimes N}$ by the flows induced by the system (8) has no reason for being compactly supported, as we did not impose anything on the particle flow. Nevertheless, the following result provides for $\rho^t_{N;1}$ a weak version of the support property (12).

Proposition 2.7. *With the same notations and hypothesis as in Theorem 2.4, let us suppose moreover that ρ^t satisfies* (12)*..*

For all $t,\epsilon>0$ let us define

$$N_{t,\epsilon} := \left(\frac{C(t)}{\epsilon}\right)^2,$$

where $C(t)$ is the function defined by (13)*.*

Then, for every $N \geq N_{t,\epsilon}$ and every Lipschitz function φ on $\mathbf{R}^{2d}$ of Lipschitz constant smaller than 1*,*

$$\int_{\mathbf{R}^{2d}\setminus B(\bar{x}+t\bar{v},X)\times B(\bar{v},Ve^{-\alpha t})} \varphi(x,v)\rho^t_{N;1}dxdv \leq \epsilon.$$

Proof. It is an immediate consequence of Theorem 2.4 or Theorem 2.6, and the fact that the Wasserstein distance induces a metric for the weak topology in the sense that

$$\sup_{\mathrm{Lip}\,(\varphi)\leq 1}\int(\mu-\nu)\varphi(x,v)dxdv \leq W_2(\mu.\nu)$$

for every probability measures μ and ν $\square$

3. Proofs

3.1. **Preliminaries.** Let us first recall the general situation we are dealing with, in order also to fix the notations.

We consider on $\mathbf{R}^{2dN}$ the following Cucker-Smale type vector field

$$\begin{aligned} \dot{x}_i &= v_i \\ \dot{v}_i &= G_i(X,V),\ i=1,\dots,N \end{aligned} \tag{14}$$

where

$$G_i(X,V) = \frac{1}{N}\sum_{j=1}^N \gamma(x_i-x_j,v_i-v_j). \tag{15}$$

Here the function $\gamma(x,v):\mathbf{R}^{2d}\to\mathbf{R}^d$ is a Lipschitz, bounded in x and sublinear in v, continuous function., such that $\gamma(x,v),\mathrm{Lip}(\gamma)_{(x,v)}\leq\gamma_0|v|$.

We used the notation $X=(x_1,\dots,x_N),V=(v_1,\dots,v_N)$.

In fact, we are rather interested in the Liouville equation associated to (14) [9], namely

$$\partial_t\rho_N^t+v\cdot\nabla_x\rho_N^t+\sum_{i=1}^N\nabla_{v_i}(G_i\rho_N^t)=0,\ \ \rho_N^{t=0}=\rho_N^{in} \tag{16}$$

with $\rho_N^{in}\in\mathcal{P}(\mathbf{R}^{2dN})$.

Although the argument is standard, let us recall why the solution of (16) is equal to the pushforward of ρ_N^{in} by the flow Φ_t generated by (14): integrating ρ_N^{in} against a test φ function composed by Φ_t gives

$$\partial_t\int\varphi(\Phi_t(X,V))\rho_N^{in}(X,V)dXdV=\int\varphi(X,V)\partial_t(\Phi_t\#\rho_N^{in}(X,V))dXdV \tag{17}$$

On the other side

$$\begin{aligned}\partial_t\int\varphi(\Phi_t(X,V))\rho_N^{in}(X,V)dXdV &= \int(\dot\Phi_t\cdot\nabla_{(X,V)}\varphi)(\Phi_t(X,V))\rho_N^{in}(X,V)dXdV\\ &= \int((V,G)\cdot\nabla_{(X,V)}\varphi)(\Phi_t(X,V)))\rho_N^{in}(X,V)dXdV\\ &= \int((V,G)\cdot\nabla_{(X,V)}\varphi)(X,V))(\Phi_t\#\rho_N^{in}(X,V)dXdV\\ &= -\int\varphi(X,V)\big(\nabla_{(X,V)}\cdot(V,G)\big)\Phi_t\#\rho_N^{in}(X,V)dXdV \qquad (18)\end{aligned}$$

So that (17) and (18) implies that $\Phi_t\#\rho_N^{in}$ solves (16).

We want to prove that the marginals of ρ_N^t tend, as $N\to\infty$, to the solution of a Vlasov type equation.

Let us recall that such Vlasov-type equation associated to (16), introduced in [9] for the Cucker-Smale model, reads

$$\partial_t\rho^t(x,v)+v.\nabla_x\rho^t(x,v)+\nabla_v(G_{\rho^t}\rho^t(x,v))=0,\quad \rho^{t=0}=\rho_1^{in}\in L^1(\mathbf{R}^{2d},dxdv), \tag{19}$$

with

$$G_\rho(x,v)=\int_{\mathbf{R}^{2d}}\gamma(x-y,v-w)\rho(y,w)dydw. \tag{20}$$

We can now prove the main results of this paper.

3.2. **Proof of Theorem 2.4.** The proof will be articulated around the four lemmas which follow.

We will denote $X=(x_1,\dots,x_N),V=(v_1,\dots,v_N),Y=(y_1,\dots,y_N),\Xi=(\xi_1,\dots,\xi_N)$.

Let π_N^{in} be defined by

$$\pi_N^{in}(Y,\Xi,X,V):=(\rho^{in})^{\otimes N}(X,V)\delta(X-Y)\delta(V-\Xi),$$

Obviously $\pi_N^{in}\in\Pi((\rho^{in})^{\otimes N},(\rho^{in})^{\otimes N})$. Moreover, as mentioned before,

$$\int_{\mathbf{R}^{2dN}\times\mathbf{R}^{2dN}}(|Y-X|^2+|\Xi-V|^2)\pi_N^{in}(dY,d\Xi,dX,dV)=0 \tag{21}$$

so that π_N^{in} is an optimal coupling between $(\rho^{in})^{\otimes N}$ and itself.

The following first Lemma will be one of the keys of the proof of Theorem 2.2. It consists in considering in evolving a coupling π_N^{in} of two initial conditions of the Liouville (16) and Vlasov (19) equations by the two dynamics of each factor of π_N^{in}.

Lemma 3.1. *Let $\pi_N(t)$ be the unique (measure) solution to the following linear transport equation*

$$\partial_t\pi_N+V.\nabla_X\pi_N+\Xi.\nabla_Y\pi_N+\sum_{i=1}^N(\nabla_{\xi_i}.(G_i(Y,\Xi)\pi_N)+\nabla_{v_i}.(G_{\rho^t}(x_i,v_i))\pi_N))=0 \tag{22}$$

with $\pi_N(0)=\pi_N^{in}$.

Then, for all $t\in\mathbf{R}$, $\pi_N(t)$ is a coupling between ρ_N^t and $(\rho^t)^{\otimes N}$.

Proof. By taking the two marginals of the two sides of the equality, one get that they satisfy the two Liouville and Vlasov equations. The result is then obtained by uniqueness of the solutions of both equations. □

Lemma 3.2. *Let*

$$\begin{aligned} D_N(t) &:= \int\frac{1}{N}\sum_{j=1}^N(|x_j-y_j|^2+|v_j-\xi_j|^2)d\pi_N(t)\\ &= \frac{1}{N}\int((X-Y)^2+(V-\Xi)^2)d\pi_N(t)\,. \end{aligned}$$

Then

$$\frac{dD_N}{dt}\leq\mathbf{L}(t)D_N+\frac{1}{N}\sum_{j=1}^N\int|G_{\rho^t}(x_i,v_i)-G_i(X,V)|^2\,(\rho^t)^{\otimes N}dXdV\,,$$

with

$$\mathbf{L}(t):=2(1+2\min(\sup_{\substack{i,l=1,\dots,N\\ X,V,Y,\Xi\in supp[\pi_N(t)]}}\mathrm{Lip}\,(\gamma)^2_{(x_i-x_l,v_i-v_l)},\sup_{\substack{i,l=1,\dots,N\\ X,V,Y,\Xi\in supp[\pi_N(t)]}}\mathrm{Lip}\,(\gamma)^2_{(y_i-y_l,\xi_i-\xi_l)})\quad). \tag{23}$$

Proof. We first notice that

$$\frac{dD_N}{dt}=\frac{2}{N}\int\left((V-\Xi).(X-Y)+\sum_{i=1}^N(v_i-\xi_i).(G_{\rho^t}(x_i,v_i)-G_i(Y,\Xi))\right)d\pi_N$$

Using $2uv \leq u^2+v^2$ we get

$$
\begin{aligned}
\frac{dD_N}{dt} &\leq \frac{1}{N}\int\left((X-Y)^2+2(V-\Xi)^2)+\sum_{i=1}^N|G_i(Y,\Xi)-G_{\rho^t}(x_i,v_i)|^2\right)d\pi_N\\
&\leq 2D_N(t)+\frac{1}{N}\int\left(\sum_{i=1}^N|G_i(Y,\Xi)-G_{\rho^t}(x_i,v_i)|^2\right)d\pi_N. \qquad (24)
\end{aligned}
$$

Let us add to $G_i(Y,\Xi)-G_{\rho^t}(x_i,v_i)$ the null term $\big(G_i(X,V)-G_i(X,V)\big)$ so that

$$
|G_i(Y,\Xi)-G_{\rho^t}(x_i,v_i)|^2\leq 2\big(|G_i(Y,\Xi)-G_i(X,V)|^2+|G_i(X,V)-G_{\rho^t}(x_i,v_i)|^2\big)
$$

Let us first estimate

$$
|G_i(Y,\Xi)-G_i(X,V)|^2\leq \min\left(\mathrm{Lip}(G_i)^2_{(X,V)},\mathrm{Lip}(G_i)^2_{(Y,\Xi)}\right)\big(|X-Y|^2+|V-\Xi|^2\big).
$$

By (15), we have

$$
\begin{aligned}
&\min\left(\mathrm{Lip}(G_i)^2_{(X,V)},\mathrm{Lip}(G_i)^2_{(Y,\Xi)}\right)\\
\leq\ &\min\Big(\sup_{l=1,\dots,N}\mathrm{Lip}\left(\gamma(x_i-x_l,v_i-v_l)^2,\sup_{l=1,\dots,N}\mathrm{Lip}\left(\gamma(y_i-y_l,\xi_i-\xi_l)^2\right)\right.
\end{aligned}
$$

Therefore, by convexity,

$$
\begin{aligned}
&\frac{2}{N}\int\sum_{i=1}^N|G_i(Y,\Xi)-G_i(X,V)|^2\pi_N(dX,dV,dY,d\Xi)\\
\leq\ & 4\min\Big(\sup_{\substack{i,l=1,\dots,N\\ X,V,Y,\Xi\in supp[\pi_N(t)]}}\mathrm{Lip}\left(\gamma(x_i-x_l,v_i-v_l)^2,\sup_{\substack{i,l=1,\dots,N\\ X,V,Y,\Xi\in supp[\pi_N(t)]}}\mathrm{Lip}\left(\gamma(y_i-y_l,\xi_i-\xi_l)^2\right)\right.\\
&\times D_N(t)
\end{aligned}
$$

And the lemma follows by (24). $\square$

It remains to estimate in (24) the term

$$
\begin{aligned}
&\frac{1}{N}\int\sum_{i=1}^N|G_i(X,V)-G_{\rho^t}(x_i,v_i)|^2\pi^t_N(dX,dV,dY,d\Xi)\\
=\ &\frac{1}{N}\int\sum_{i=1}^N|G_i(X,V)-G_{\rho^t}(x_i,v_i)|^2(\rho^t)^{\otimes N}dXdV
\end{aligned}
$$

The following result is a variant of Lemma 3.3 in [6] with the special value $p=2$ and d replaced by $2d$.

Lemma 3.3. *Let ρ be a compactly supported probability density on $\mathbf{R}^{2d}$ and let F be a locally bounded vector field on $\mathbf{R}^{2d}$.*

For each $j = 1, \dots, N$, one has

$$\int \left| F \star \rho(x_j, v_j) - \frac{1}{N} \sum_{k=1}^{N} F(x_j - x_k, v_j - v_k) \right|^2 \rho^{\otimes N} dX dV \leq \frac{4}{N} \sup_{(x,v),(x',v') \in supp[\rho]} |F(x - x', v - v')|^2. \tag{25}$$

Proof. Let us denote, for $x_j, x, v_j, v \in \mathbf{R}^d$, $j = 1, \dots, N$,

$$\nu_{x_j,v_j}(x, v) := F \star \rho(x_j, v_j) - F(x_j - x, v_j - v) \tag{26}$$

(note that $F \star \rho(x_j, v_j)$ doesn't depend on (x, v)).

One has (let us remind the notation $X := (x_1, \dots, x_N),\ V := (v_1, \dots, v_N)$)

$$\begin{aligned}
& \int \left| F \star \rho(x_j, v_j) - \frac{1}{N} \sum_{k=1}^{N} F(x_j - x_k, v_j - v_k) \right|^2 \rho^{\otimes N} dX dV \\
= & \int \left| \frac{1}{N} \sum_{k=1}^{N} \nu_{x_j,v_j}(x_k, v_k) \right|^2 \rho^{\otimes N} dX dV \\
= & \frac{1}{N^2} \sum_{k,l=1,\dots N} \int \nu_{x_j,v_j}(x_k, v_k) \nu_{x_j,v_j}(x_l, v_l) \rho^{\otimes N} dX dV \\
= & \frac{1}{N^2} \sum_{k \neq l \neq j \neq k} \int \left(\int \nu_{x_j,v_j}(x_k, v_k) \rho(x_k, v_k) dx_k dv_k \right) \nu_{x_j,v_j}(x_l, v_l) \rho(x_l, v_l) \rho(x_j, v_j) dx_l dv_l dx_j dv_j \\
+ & \frac{1}{N^2} \sum_{k \neq l = j} \int \left(\int \nu_{x_j,v_j}(x_k, v_k) \rho(x_k, v_k) dx_k dv_k \right) \nu_{x_j,v_j}(x_j, v_j) \rho(x_j, v_j) dx_j dv_j \\
+ & \frac{1}{N^2} \sum_{k \neq j} \int \nu_{x_j,v_j}(x_k, v_k)^2 \rho(x_k, v_k) dx_k dv_k \rho(x_j, v_j) dx_j dv_j \\
+ & \frac{1}{N^2} \int \nu_{x_j,v_j}(x_j, v_j)^2 \rho(x_j, v_j) dx_j dv_j \\
\leq & \frac{N+1}{N^2} \sup_{(x,v),(x',v') \in supp[\rho]} \nu_{x',v'}(x, v)^2
\end{aligned}$$

since, by (26), $\int \nu_{x',v'}(x, v) \rho(x, v) dx dv = 0.$ for all $x', v' \in \mathbf{R}^d$, and $\int \rho(x, v) dx dv = 1$.

By (26) again,

$$|\nu_{x',v'}(x, v)| \leq 2 \sup_{(x,v) \in supp[\rho]} |F(x' - x, v' - v)| \leq 2 \sup_{(x,v),(x',v') \in supp[\rho]} |F(x' - x, v' - v)|$$

and the lemma is proved. □

Lemma 3.3 with $F(x,v)=\gamma(x,v)$ together with Lemma 3.2 gives immediately that

$$\frac{dD_N}{dt}\leq L(t)D_N+\frac{4}{N}\sup_{(x,v),(x',v')\in supp[\rho^t]}|\gamma(x-x',v-v')|^2$$

and, by Gronwall's inequality,

$$D_N(t)\leq\frac{4}{N}\int_0^t\sup_{(x,v),(x',v')\in supp[\rho^t]}|\gamma(x-x',v-v')|^2e^{\int_s^t L(u)du}ds \tag{27}$$

since $D_N(0)=0$ by (21).

Let us denote by $(\pi_N(t))_1$ the measure on $\mathbf{R}^{2d}\times\mathbf{R}^{2d}$ defined, for every test function $\varphi(x_1,v_1;y_1,\xi_1)$, by

$$\int_{\mathbf{R}^{2dN}\times\mathbf{R}^{2dN}}\varphi\pi_N(dX,dV;dYd\Xi)=\int_{\mathbf{R}^{2d}\times\mathbf{R}^{2d}}\varphi(x_1,v_1;y_1,\xi_1)(\pi_N(t))_1(dx_1,dv_1;dy_1,d\xi_1)$$

We now notice the following straightforward fact.

Lemma 3.4. *$(\pi_N(t))_1$ is a coupling between $\rho^t_{N;1}$ and ρ^t.*

Let us note that π_N^{in} is obviously symmetric by permutation of the phase-space variables, that is

$$T_\sigma\#\pi_N^{in}=\pi_N^{in}\,,\qquad\text{for each } \sigma\in\mathfrak{S}_N\,, \tag{28}$$

where $\mathfrak{S}_N$ is the group of permutations of N elements and

$$\begin{aligned}&T_\sigma(x_1,v_1,\dots,x_N,v_N,y_1,\xi_1,\dots,y_N,\xi_N)\\&=(x_{\sigma(1)},v_{\sigma(1)},\dots,x_{\sigma(N)},v_{\sigma(N)},y_{\sigma(1)},\xi_{\sigma(1)},\dots,y_{\sigma(N)},\xi_{\sigma(N)})\,.\end{aligned}$$

Therefore, $\pi_N(t)$ is also symmetric by permutations for all $t\in\mathbf{R}$, as being the unique solution to the equation (22), being itself, by construction, symmetric by permutations,
.

Consequently, one has easily that

$$D_N(t)=\int(|x_1-y_1|^2+|v_1-\xi_1|^2)d(\pi_N(t))_1 \tag{29}$$

and Lemma 3.4 immediately implies that

$$D_N(t)\geq W_2(\rho_1(t),(\rho(t))_1)^2\,, \tag{30}$$

so that, by (27),

$$W_2(\rho_1(t),(\rho(t))_1)^2\leq\frac{4}{N}\int_0^t\sup_{(x,v),(x',v')\in supp[\rho^t]}|\gamma(x-x',v-v')|^2e^{\int_s^t L(u)du}ds. \tag{31}$$

Finally, recalling that

$$\pi_N^{in}(Y,\Xi,X,V):=(\rho^{in})^{\otimes N}(X,V)\delta(X-Y)\delta(V-\Xi),$$

one get immediately that

$$\begin{aligned}(Y,\Xi;X,V)\in supp[\pi_N(t)]\ &\Rightarrow\ (Y,\Xi)\in supp[\rho^t_N]\text{ and}\\&\Rightarrow\ (X,V)\in supp[(\rho^t)^{\otimes_N}]\Leftrightarrow(x_i,v_i)\in supp[\rho^t],\ i=1,\dots,N.\end{aligned}$$

Therefore, after (23)

$$\mathbf{L}(t) \le L(t) := 2\Big(1 + 2\min\big(\sup_{\substack{i,l=1,\dots,N\\(Y,\Xi)\in supp[\rho_N^t]}} \mathrm{Lip}\,(\gamma)^2_{(y_i-y_l,\xi_i-\xi_l)}, \sup_{(x,v),(x',v')\in supp[\rho^t]} \mathrm{Lip}\,(\gamma)^2_{(x-x',v-v')}\big)\Big). \quad (32)$$

Theorem 2.4 is proven.

Remark 3.5 (Higher order Wasserstein and marginals)**.** *As we wanted to leap this note as short as possible, we expressed our results only for the first marginal $\rho_{N;1}$ in the 2-Wasserstein topology, but the method developed in [6] allows as well, with the same kind of modification than the ones used before in this section, to the higher cases. Let us remind, for $p \ge 1$, the definition of $W_p(\mu,\nu)$ for two positive measures μ,ν (cf. Definition 2.1*

$$W_p(\mu,\nu)^p = \inf_{\gamma\in\Gamma(\mu,\nu)} \int_{\mathbf{R}^m\times\mathbf{R}^m} |x-y|^p \gamma(dx,dy),$$

and, for any probability measure ρ on $\mathbf{R}^{2dN}$ and $n = 1,\dots,N$, the definition of the nth marginal $\rho_{N;n}$ defined on $\mathbf{R}^{2dn}$

$$\rho_{N;n}(x_1,\dots,x_n,v_1,\dots,v_n) := \int_{\mathbf{R}^{2d(N-n)}} \rho((x_1,\dots,x_n,x_{n+1},\dots,x_N,v_1,\dots,v_n,v_{n+1},\dots,v_N)dx_{n+1}\dots dx_N dv_{n+1}\dots dv_N.$$

One gets, for each $p \ge 1$, $N \ge 1$ and $n = 1,\dots,N$,

$$\frac{1}{n} W_p(\rho^t_{N;n},(\rho^t)^{\otimes n})^p \le D_{p,n}(t) N^{-\min(p/2,1)}$$

with

$$D_{p,n}(t) = 2^{2p}\max(1,p-1)([p/2]+1)\int_0^t \sup_{\substack{k,l=1,dots,N\\(Y,\Xi)\in supp[\rho_N^t]}} |\gamma(y_k-y_l,\xi_k-\xi_l)|^p e^{2\max(1,p-1)\int_s^t L(u))du} ds$$

where

$$L(u) = 1 + 2^{p-1} \sup_{(x,v),(y,\xi)\in supp[\rho^t]} \mathrm{Lip}(\gamma)^p_{(x-y,v-\xi)}.$$

The main changes in the proof are the use of the Young inequality

$$puv^{p-1} \le u^p + (p-1)v^p \le \max(1,p-1)(u^p+v^p)$$

instead of $2uv \le u^2+v^2$ before (24), the convexity of $|\cdot|^p$ for $p \ge 1$ and a variant of Lemma 3.3, similar to Lemma 3.3 in [6], which reads

$$\int \left| F\star\rho(x_j,v_j) - \frac{1}{N}\sum_{k=1}^N F(x_j-x_k,v_i-v_j)\right|^p \rho^{\otimes N} dX dV$$
$$\le \frac{2[p/2]+2}{N^{\min(p/2,1)}} \sup_{\substack{k,l1,dots,N\\(Y,\Xi)\in supp[\rho_N^t]}} |F(y_k-y_l,\xi_k-\xi_l)|^p.$$

Finally, the statement of Lemma 3.4 becomes now easily that $\pi_N(t)_n$ *is a coupling between* $\rho^t_{N;n}$ *and* $(\rho^t)^{\otimes n}$, *where* $\pi_N(t)_n$ *is defined through by*

$$\int_{\mathbf{R}^{2dN}\times\mathbf{R}^{2dN}} \varphi\pi_N(dX,dV;dYd\Xi) = \int_{\mathbf{R}^{2d}\times\mathbf{R}^{2d}} \varphi\pi_N(t)_n(d(x,v;y,\xi)_1\dots d(x,v;y,\xi)_n)$$

for every test function $\varphi((x,v;y,\xi)_1,\dots,(x,v;y,\xi)_n)$,

3.3. **Proof of Corollary 2.6.** We get to estimates $L(t)$ and $C(t)$ as given in Theorem 2.4 out of the estimates established in Appendices A and B.

Let us recall that

$$\begin{aligned}
L(t) &:= 2(1+2\min(\sup_{\substack{i,l=1,\dots,N\\(Y,\Xi)\in supp[\rho^t_N]}}\mathrm{Lip}\,(\gamma)^2_{(y_i-y_l,\xi_i-\xi_l)},\sup_{(x,v),(x',v')\in supp[\rho^t]}\mathrm{Lip}\,(\gamma)^2_{(x-x',v-v')}))).\\
&\le 2(1+2\sup_{(x,v),(x',v')\in supp[\rho^t]}\mathrm{Lip}\,(\gamma)^2_{(x-x',v-v')}).\\
&\le 2(1+2\gamma_0^2\sup_{(x,v),(x',v')\in supp[\rho^t]}|v-v'|^2)\\
&\le 2(1+8\gamma_0^2\sup_{(x,v)\in supp[\rho^t]}|v|^2)\\
&\le 2(1+8\gamma_0^2\sup_{(x,v)\in supp[\rho^{in}]}|\Phi_v(t)(x,v)|^2)
\end{aligned}$$

Thanks to (37), we get

$$L(t) \;\le\; 2(1+16\gamma_0^2e^{4\gamma_0t}(\|v\|_{L^\infty(supp[\rho^{in}])}+\|v\|_{L^1(supp[\rho^{in}])})^2)$$

By the same argument, we get

$$\begin{aligned}
C(t)^2 &:= 4\int_0^t\sup_{(x,v),(x',v')\in supp[\rho^t]}|\gamma(x-x',v-v')|^2e^{\int_s^tL(u)du}ds,\\
&\le 32\gamma_0^2(\|v\|_{L^\infty(supp[\rho^{in}])}+\|v\|_{L^1(supp[\rho^{in}])})^2\\
&\quad\times\int_0^t e^{4(\gamma_0s+t-s)}e^{16\gamma_0^2(\|v\|_{L^\infty(supp[\rho^{in}])}+\|v\|_{L^1(supp[\rho^{in}])})^2(e^{4\gamma_0t}-e^{4\gamma_0s})/4\gamma_0}ds\\
&\le 8\gamma_0^2(\|v\|_{L^\infty(supp[\rho^{in}])}+\|v\|_{L^1(supp[\rho^{in}])})^2e^{e^{4\gamma_0t}4\gamma_0(\|v\|_{L^\infty(supp[\rho^{in}])}+\|v\|_{L^1(supp[\rho^{in}])})^2}e^{4t}(e^{4\gamma_0t}-1).
\end{aligned}$$

Corollary 2.6 is proven.

3.4. **Proof of Theorem 2.2.** In order to prove Theorem 2.2, we need to estimate, in the Cucker-Smale particular case, that is when

$$\gamma(x,v)=\psi(x)v,$$

the two quantities

$$\begin{aligned}
C(t) &:= \left(4\int_0^t \sup_{(x,v),(x',v')\in supp[\rho^t]} |\gamma(x-x',v-v')|^2 e^{\int_s^t L(u)du} ds\right)^{\frac{1}{2}},\\
&\leq 2\|\psi\|_\infty \left(\int_0^t \sup_{(x,v),(x',v')\in supp[\rho^t]} |v-v'|^2 e^{\int_s^t L(u)du} ds\right)^{\frac{1}{2}},\\
L(t) &:= 2(1+2\min(\sup_{\substack{i,l=1,\dots,N\\ (Y,\Xi)\in supp[\rho_N^t]}} \mathrm{Lip}\,(\gamma)^2_{(y_i-y_l,\xi_i-\xi_l)}, \sup_{(x,v),(x',v')\in supp[\rho^t]} \mathrm{Lip}\,(\gamma)^2_{(x-x',v-v')}))).\\
&\leq 2(1+2\sup_{\substack{i,l=1,\dots,N\\ (Y,\Xi)\in supp[\rho_N^t]}} \mathrm{Lip}\,(\gamma)^2_{(y_i-y_l,\xi_i-\xi_l)})\\
&\leq 2(1+2\|\psi\|_\infty^2 \sup_{\substack{i,l=1,\dots,N\\ (Y,\Xi)\in supp[\rho_N^t]}} |\xi_i-\xi_l|^2)
\end{aligned}$$

We will need just a very little part of the stability results expressed by Ha, Kim and Zhuang in [7], namely the following inequality:

$$\frac{d}{dt}\sup_{\substack{i,l=1,\dots,N\\ (X,V)\in supp[\rho_N^{in}]}} |v_k(t)-v_l(t)| \leq 0,\ \forall t,$$

Indeed, formula (8) in Lemma 2.2 in [7] stipulates that $\frac{d}{dt}\mathcal{D}_V(t)\leq 0$, where $\mathcal{D}_V(t)$, as defined in Corollary 1 of [7], is precisely $\sup_{\substack{i,l=1,\dots,N\\ (X,V)\in supp[\rho_N^{in}]}} |v_k(t)-v_l(t)|$ in the case $p=2$ of the definition (4) of $\mathcal{D}_V(0):=\mathcal{D}_V$ in [7].

This inequality leads naturally to

$$\sup_{\substack{i,l=1,\dots,N\\ (X,V)\in supp[\rho_N^t]}} |v_i-v_l|^2 \leq 4\|v\|^2_{L^\infty(supp[\rho^{in}])},$$

which implies

$$L(t)\leq 2(1+8\|\psi\|_\infty^2\|v\|^2_{L^\infty(supp[\rho^{in}])}) := L$$

We will estimate $\sup_{(x,v),(y,\xi)\in supp[\rho^t]} |v-\xi|^2 \leq 2\sup_{(x,v),(y,\xi)\in supp[\rho^t]} (|v|^2+|\xi|^2)$ thanks to Lemma 3.2 in [10] which stipulates that

$$supp_v[\rho^t]\subset B(\bar{v},V(t))$$

with $\bar{v}=\int v\rho^{in}dxdv$ and $\frac{d}{dt}V(t)\leq 0$.

Therefore $\|v\|_{L^\infty(supp[\rho^t])}\leq|\bar{v}|+V(0)$ so that, since one can take $V(0)=|\bar{v}|+|supp[\rho^{in}]|$,

$$\sup_{(x,v),(x'v')\in supp[\rho^t]} |v-v'|^2 \leq 2\|\psi\|_\infty^2|(2|\bar{v}|+|supp[\rho^{in}]|)^2.$$

and

$$\begin{aligned} C(t) &\leq 2\|\psi\|_\infty^2 \left(2\int_0^t (2|\bar v| + |supp[\rho^{in}]|)^2 e^{L(t-s)} ds\right)^{\frac12} \\ &= 2\|\psi\|_\infty^2 \left(2(2|\bar v| + |supp[\rho^{in}]|)^2 \frac{e^{Lt}-1}{L}\right)^{\frac12} \end{aligned}$$

Theorem 2.2 is proved.

Appendix A. Dynamical estimates for general Cucker-Smale particle systems

In this section,we give global estimates on the flow $\Phi^N(t)$ generated by (8).

We have, for each $i = 1, \dots, N$,

$$\begin{aligned} \frac{d}{dt}|v_i| &\leq |\dot v_i| \\ &\leq \frac1N \sum_{j=1}^N |\gamma(x_i - v_j, v_i - v_j)| \leq \frac1N \sum_{j=1}^N \gamma_0 |v_i - v_j| \\ &\leq \frac1N \sum_{j=1}^N \gamma_0(|v_j| + |v_i|), \end{aligned} \tag{33}$$

so that

$$\frac{d}{dt}\sum_{i=1}^N |v_i| \leq 2\gamma_0 \sum_{i=1}^N |v_i|.$$

and, by Gronwall inequality,

$$\sum_{i=1}^N |v_i(t)| \leq \sum_{i=1}^N |v_i(0)| e^{2\gamma_0 t}. \tag{34}$$

Turning back to (33), we get by (34), for each $i = 1, \dots, N$,

$$\begin{aligned} \frac{d}{dt}(|v_i|) &\leq \gamma_0|v_i| + \frac{\gamma_0}{N}\sum_{j=1}^N |v_j| \leq \gamma_0|v_i| + \gamma_0 e^{2\gamma_0 t}\frac1N \sum_{j=1}^N |v_j(0)| \\ &\leq \gamma_0|v_i| + \gamma_0 e^{2\gamma_0 t} \max_{j=1,\dots,N} |v_j(0)|. \end{aligned}$$

Therefore, again by Gronwall Lemma and uniformly in N,

$$\begin{aligned} |v_i(t)| &\leq e^{\gamma_0 t}\big(|v_i(0) + \gamma_0 \max_{j=1,\dots,N} |v_j(0)| \frac{e^{\gamma_0 t}-1}{\gamma_0}\big) \\ &\leq \max_{j=1,\dots,N} |v_j(0)| e^{2\gamma_0 t}, \quad i = 1, \dots, N. \end{aligned} \tag{35}$$

Finally, (8) gives immediatly that

$$||x_1(t)| - |x_i(0)|| \leq \max_{j=1,\dots,N} |v_j(0)| \frac{e^{2\gamma_0 t}-1}{2\gamma_0}, \quad i = 1, \dots, N.$$

Appendix B. Dynamical estimates for general Cucker-Smale kinetic systems

Let us recall that, according to Theorem 2.3 in [10], there exist a diffeomorphism $\Phi(t)$ on $\mathbf{R}^{2d}$ such that the solution ρ^t of (9) is given by

$$\rho(t) = \Phi(t)\#\rho^{in}.$$

Moreover, $\Phi(t)$ solves the system

$$\dot{\Phi}(t) := \begin{pmatrix} \dot{\Phi}_x(t) \\ \dot{\Phi}_v(t) \end{pmatrix} = \begin{pmatrix} \Phi_v(t) \\ (\gamma * \rho^t)(\Phi(t)) := \int \gamma(\Phi_x(t) - y, \Phi_v(t) - \xi)\rho^t(y,\xi)dxd\xi \end{pmatrix}$$

Therefore,

$$\frac{d}{dt}|\Phi_v(t)| \leq |\dot{\Phi}_v(t)| \leq \gamma_0(\|\rho^t\|_1|\Phi_v(t)| + \int |\xi|\rho^t(y,\xi)dyd\xi). \tag{36}$$

Obviously $\|\rho^t\|_{L^1} = \|\rho^{in}\|_{L^1} = 1$.

Lemma B.1.

$$\int |v|\rho^t dxdv \leq e^{2\gamma_0 t}\int |v|\rho^{in}dxdv.$$

Proof.

$$\begin{aligned}
\frac{d}{dt}\int |v|\rho^t dxdv &\leq |\int |v|\dot{\rho}^t dxdv| \\
&\leq |\int |v|\nabla_v(\int \gamma(x-y.v-\xi)\rho^t(y,\xi)\rho^t(x,v)dxdvdyd\xi| \\
&\leq \int |\frac{v}{|v|}\gamma_0((v| + |\xi|)\rho^t(y,\xi)\rho^t(x,v)dxdvdyd\xi \\
&\leq 2\gamma_0\int |v|\rho^t dxdv
\end{aligned}$$

The result follows by Gronwall inequality. □

Thanks to Lemma B.1, (36) becomes

$$\frac{d}{dt}|\Phi_v(t)| \leq \gamma_0|\Phi_v(t)| + \gamma_0 e^{2\gamma_0 t}\int |v|\rho^{in}dxdv,$$

and,by Gronwall Lemma, we have

$$\begin{aligned}
|\Phi_v(t)(x,v)| &\leq e^{\gamma_0 t}(|v| + (e^{\gamma_0 t} - 1)\int |v|\rho^{in}dxdv) \\
&\leq e^{2\gamma_0 t}(\|v\|_{L^\infty(supp[\rho^{in}])} + \|v\|_{L^1(supp[\rho^{in}])}). \qquad (37)
\end{aligned}$$

Acknowledgements. The work of Thierry Paul was partly supported by LIA LYSM (co-funded by AMU, CNRS, ECM and INdAM) and IAC (Istituto per le Applicazioni del Calcolo "Mauro Picone") from CNR.

References

[1] L. Ambrosio, N.Gigli, G. Savaré, "Gradient Flows in Metric Spaces and in the Space of Probability Measures", Second Edition, Birlhäuser Verlag, Berlin, 2008.

[2] F. Cucker and S. Smale, On the mathematics of emergence. Japan. J. Math. **2**, 197227, (2007).

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

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

[5] F. Golse, C. Mouhot, V. Ricci: *Empirical measures and Vlasov hierarchies*, Kinetic and Related Models **6** (2013), 919–943.

[6] F. Golse, C. Mouhot, T. Paul, On the Mean-Field and Classical Limits of Quantum Mechanics, , Communication in Mathematical Physics **343**, 165-205 (2016).

[7] S.-Y. Ha, J. Kim, X. Zhang, Uniform stability of the Cucker-Smale model and applications to the mean-field limit, Kinetic and Related Models **11** (5), 1157–1181 (2018)

[8] S.-Y. Ha and J.-G. Liu, A simple proof of Cucker-Smale flocking dynamics and mean-field limit, Commun. Math. Sci., **7** (2009), 297–325.

[9] S.-Y. Ha, E. Tadmor, From particle to kinetic and hydrodynamic descriptions of flocking, Kinet. Relat. Models **1** (2008) 415-435.

[10] B. Piccoli, F. Rossi, E. Trélat, *Control to flocking of the kinetic Cucker-Smale model*, SIAM J. MATH. ANAL . Vol. 47, No. 6, pp. 46854719.

[11] C. Villani: "Topics in Optimal Transportation", Amer. Math. Soc., Providence (RI), 2003.

[12] C. Villani: "Optimal Transport. Old and New", Springer-Verlag, Berlin, Heidelberg, 2009.

(R.N) IAC, Via dei Taurini, 19, 00185 Roma RM Italy
E-mail address: roberto.natalini@cnr.it

(T.P.) CNRS & LJLL Sorbonne Universit 4 place Jussieu 75005 Paris, France France
E-mail address: thierry.paul@upmc.fr