HAL
archives-ouvertes.fr

Particles approximations of Vlasov equations with singular forces : Propagation of chaos.

Pierre-Emmanuel Jabin, Maxime Hauray

► To cite this version:

Pierre-Emmanuel Jabin, Maxime Hauray. Particles approximations of Vlasov equations with singular forces : Propagation of chaos.. Annales Scientifiques de l'École Normale Supérieure, Elsevier Masson, 2015. <hal-00609453v5>

HAL Id: hal-00609453
https://hal.archives-ouvertes.fr/hal-00609453v5

Submitted on 8 May 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives| 4.0 International License

Particles approximations of Vlasov equations with singular forces : Propagation of chaos

Maxime Hauray* and Pierre-Emmanuel Jabin †

Abstract. We obtain the mean field limit and the propagation of chaos for a system of particles interacting with a singular interaction force of the type $1/|x|^\alpha$, with $\alpha < 1$ in dimension $d \geq 3$. We also provide results for forces with singularity up to $\alpha < d-1$ but with large enough cut-off. This last result thus almost includes the most interesting case of Coulombian or gravitational interaction, but it is also interesting when the strength of the singularity α is larger but close to one, in which case it allows for very small cut-off.

Key words. Derivation of kinetic equations. Particle methods. Vlasov equation. Propagation of chaos.

1 Introduction

The N particles system. The starting point is the classical Newton dynamics for N point-particles. We denote by $X_i \in \mathbb{R}^d$ and $V_i \in \mathbb{R}^d$ the position and velocity of the i-th particle. For convenience, we also use the notation $Z_i = (X_i, V_i)$ and $Z = (Z_1, \dots, Z_n)$. Assuming that particles interact two by two with the interaction potential $\Phi(x)$, one finds the usual

$$\begin{cases} \dot{X}_i = V_i, \\ \dot{V}_i = E_N(X_i) = -\dfrac{1}{N} \displaystyle\sum_{j \neq i} \nabla \Phi(X_i - X_j). \end{cases} \tag{1.1}$$

The initial conditions Z^0 are given. We use the so-called mean-field scaling which consists in keeping the total mass (or charge) of order 1 thus formally enabling us to pass to the limit. This explains the $1/N$ factor in front of the force terms, and implies corresponding rescaling in position, velocity and time.

There are many examples of physical systems following (1.1). The best known concerns Coulombian or gravitational force $\Phi(x) = C/|x|^{d-2}$ with $C \in \mathbb{R}^*$, which serves as a guiding example and reference. Those describe the ions and electrons evolving in a plasma for $C > 0$, or gravitational interactions for $C < 0$. In the last case the system under study may be a galaxy, a cloud of star or galaxies (and thus particles can be "stars" or even "galaxies"). For simplicity we consider here only a basic form for the interaction. However the same techniques would apply to more complex models, for instance with several species (electrons and ions in a plasma), 3-particles (or more) interactions, models where the force depends also on the speed as in swarming models like the Cucker-Smale one [BCC11]...

*LATP, Université d'Aix-Marseille & CNRS UMR 7353, 13453 Marseille Cedex 13, France
†CSCAMM and Dpt of Mathematics, University of Maryland, College Park, MD 20742, USA

Indeed a striking feature of our analysis is that the Hamiltonian structure of (1.1) is never used. In fact the proofs and statements are identical for the more general

$$\begin{cases} \dot{X}_i = V_i, \\ \dot{V}_i = E_N(X_i) = \dfrac{1}{N} \displaystyle\sum_{j \neq i} F(X_i - X_j), \end{cases} \tag{1.2}$$

with a force kernel F not necessarily derived from a potential.

The potential and forces used in that article. Our first result apply to interaction potentials (and forces) that are "weakly" singular in the sense that they behave near zero like

$$(S^\alpha_\Phi) \qquad \Phi(x) \sim c|x|^{1-\alpha}, \qquad \text{for some } \ \alpha < 1, \ c \in \mathbb{R}^*.$$

As we never use the potential structure in our analysis, our result may in fact apply to more general force F, not associated to any potential, which behave near zero like

$$(S^\alpha) \qquad F(x) \sim \frac{c}{|x|^\alpha}, \qquad \text{for some } \ \alpha < 1, \ c \in \mathbb{R}^*.$$

The precise conditions are respectively given in (1.7) and (1.8). We refer to that conditions as the "weak" singular case because under these conditions with $\alpha < 1$, the potential is continuous and bounded near the origin. It sound in fact reasonable to expect that the analysis is much simpler in that case, than with a singular potential.

The second type of potentials or forces that we are dealing with are more singular, satisfying the (S^α_Φ) or the (S^α)-condition with $\alpha < d-1$, but with a additional cut-off η near the origin that will depends on N

$$\Phi_N(x) \sim \frac{c}{(|x|^2 + \eta(N)^2)^{\alpha-1}} \quad \text{or} \quad F_N(x) \sim \frac{c}{(|x|^2 + \eta(N)^2)^{\alpha}} \quad \text{with } \ \alpha < 1, \ c \in \mathbb{R}^*$$

The precise condition is given in equation (1.10) and (1.11). In that case, the potential Φ or the kernel F in fact depend on the number of particles. This might seem quite strange from the physical point of view but is in fact very common for numerical simulations (to regularize the interactions).

The Jeans-Vlasov equation. At first glance, the system (1.2) might seem quite reasonable. However many problems arise when one tries to use it for practical applications. In our case, the main issue is the number of particles, *i.e.* the dimension of the system. For example a plasma or a galaxy usually contains a very large number of 'particles", typically from 10^9 to 10^{23}, which makes solving (1.1) practically impossible.
As usual in this kind of situation, one would like to replace the discrete system (1.1) by a "continuous" model. In that case this model is posed in the space $\mathbb{R}^{2d}$, *i.e.* it involves the distribution function $f(t, x, v)$ in time, position and velocity. The evolution of that function $f(t, x, v)$ is given by the Jeans-Vlasov equation

$$\begin{cases} \partial_t f + v \cdot \nabla_x f + E(x) \cdot \nabla_v f = 0\,, \\ E(x) = -\nabla_x \int_{\mathbb{R}^d} \rho(t, y)\, \Phi(x-y)\, dy, \\ \rho(t, x) = \int_{\mathbb{R}^d} f(t, x, v)\, dv, \end{cases} \tag{1.3}$$

where here ρ is the spatial density and the initial density f^0 is given.
Our whole concern in this article is to understand when and in which sense, Eq. (1.3) can be seen as a limit of system (1.1). This question is of importance for theoretical reasons, to justify the validity of the Vlasov equation for example.
It also plays a role for numerical simulation, and especially Particles in Cells methods which introduce a large number (roughly around 10^6 or 10^8, to compare with the order 10^{10} to 10^{23} mentioned above) of "virtual" particles in order to obtain a particles system solvable numerically. The problem in that case is to explain why it is possible to correctly approximate the system by using much fewer particles. This would of course be ensured by the convergence of (1.1) to (1.3).
Of course, we shall make use of uniqueness results for the solution of equation (1.3). But this equation is now well understood, even when the interaction F is singular, including the Coulombian case. The existence of weak solutions goes back to [Dob79] or [Ars75]. Existence and uniqueness of global classical solutions is proved in [Pfa92], [Sch91] (see also [Hor93]) and at the same time in [LP91]. Of course those results require some smoothness on the initial data f^0 (for instance compact support and boundedness in [Pfa92]). We will state the precise result of existence and uniqueness we need in the Proposition 3 in Section 2.4.

Formal derivation of Eq. (1.3) **from** (1.1)**.** One of the simplest way to understand formally how to derive Eq. (1.3) is to introduce the empirical measure

$$\mu_N^Z(t) = \frac{1}{N}\sum_{i=1}^{N} \delta_{X_i(t),V_i(t)}.$$

In fact if $(X_i, V_i)_{1\le i\le N}$ is solution to (1.2), and if there is no self-interaction : $\nabla\Phi(0) = F(0) = 0$, then μ_N^Z solves (1.3) in the sense of distribution. Formally one may then expect that any limit of μ_N^Z still satisfies the same equation.

The question of convergence. The previous formal argument suggests a first way of rigorously deriving the Vlasov equation (1.3). Take a sequence of initial conditions Z^0 (to be given for every number N or a sequence of such numbers) and assume that the corresponding empirical measures converge (in the usual weak-$*$ topology)

$$\mu_N^Z(0) \longrightarrow f^0(x,v).$$

One would then try to prove that the empirical measures at later times $\mu_N^Z(t)$ weakly converge to a solution $f(t,x,v)$ to (1.3) with initial data f^0. In other words, is the following diagram commutative?

$$\begin{array}{ccc} \mu_N^Z(0) & \overset{\text{cvg}}{\rightsquigarrow} & f(0) \\ {\scriptstyle Npart}\downarrow & & \downarrow{\scriptstyle VP} \\ \mu_N^Z(t) & \overset{\text{cvg ?}}{\rightsquigarrow} & f(t) \end{array}$$

Note that for singular potentials Φ (among which Coulombian and gravitational interactions), one does not expect to be able to do that for all initial conditions. First of all solutions to (1.3) do not exist in general if the initial data f^0 is only a measure. And even if f^0 is smooth but a small amount of the particles are initially concentrated in a small region, problems will likely occur (as the interactions blow up). Of course it is not obvious what are admissible initial conditions and how to precise them. The simplest way is to give some properties that the initial conditions must satisfy. For instance one could ask the initial positions and velocities to be uniformly distributed on a grid for numerical simulations. However from the point of view of statistical physics, this is too restrictive. A natural question is then to determine what are acceptable assumptions on the initial conditions. This leads us to the notion of propagation of chaos.

Propagation of chaos. In most physical settings, one expects the initial positions and velocities to be selected randomly and typically independently (or almost independently). In the case of total independence, the law of Z is initially given by $(f^0)^{\otimes N}$, i.e. each couple $Z_i = (X_i, V_i)$ is chosen randomly and independently with profile f^0. Note that by the empirical law of large number, also known as Glivenko-Cantelli theorem, the empirical measure $\mu_N^z(0)$ at time 0 is then close to f^0 with large probability in some weak norm, See the Proposition 6 for a precise version.
The propagation of chaos was formalized by Kac's in [Kac56] and goes back to Boltzmann and its Stosszahl ansatz. A standard reference is the famous course by Sznitman [Szn91]. It is mainly used in stochastic systems, where some randomness is introduced in the dynamics of the particles, but is also relevant in our context. Roughly speaking, the propagation of chaos will hold if when starting from (almost) i.i.d. particles with law f^0 at time 0, then the particles evolving according to (1.1) are almost independent at time t, with the law of one particle given by the solution $f(t)$ of the Vlasov equation (1.3). The meaning of "almost" in the last sentence is that independence becomes true only in the limit $N \to +\infty$.
Let us be more precise. Denoting by $f^N(t, x_1, v_1, \dots, x_N, v_N)$ the image by the dynamics (1.1) of the initial law $(f^0)^{\otimes N}$, one may define the k-marginals

$$f_k^N(t, z_1, \dots, z_k) = \int_{R^{2d(N-k)}} f^N(t, z_1, \dots, z_N)\, dz_{k+1} \dots dz_N.$$

According to the general definition, the propagation of chaos means that for any fixed k, $f_k^N(t)$ will converge weakly to $[f(t)]^{\otimes k}$ as $N \to \infty$. In fact it is sufficient that the convergence holds for $k = 2$. It is also equivalent to say that the empirical measures $\mu_N^z(t)$ converge in law towards the constant variable $f(t)$. If one of that equivalent properties is true, we say that the sequence $f^N(t)$ is $f(t)-$-chaotic. Shortly, this is possible because the marginals can be recovered from the expectations of moments of the empirical measure

$$f_k^N = \mathbb{E}(\mu_N^z(t, x_1, v_1) \dots \mu_N^z(t, x_k, v_k)) + O\left(\frac{k^2}{N}\right),$$

a result sometimes called Grunbaum lemma. For detailed explanations about quantification of the equivalence between convergence of the marginals f_k^N and the convergence of the law of the empirical distributions μ_N^z, we refer to [HM12]. This quantified equivalence was for instance used in the recent and important work of Mischler and Mouhot about Kac's program in kinetic theory [MM11]. So to sum up, the propagation of chaos holds if starting for particles initially distributed according to the law $(f^0)^{\otimes N}$ or more generally according to a f^0-chaotic sequence, then for any $t < T$ (the time of existence of the solution f with initial conditions f^0), the particles at time t are distributed according to $f(t)$-chaotic laws.
Here, we will give in theorem 1 and 2 a quantified version of the convergence in probability of $\mu_N^z(t)$ towards $f(t)$, starting from Z^0 with law $(f^0)^{\otimes N}$. Of course we do not expect

propagation of chaos to hold for any initial distribution f^0 : a result of propagation of chaos generally requires an appropriate uniqueness result on the limit equation. For that reason, we will restrict ourself to bounded and compactly supported initial condition f^0 , for which we known that there exists a unique strong (at least) local solution to the Vlasov equation (1.3), see Proposition 3.

Propagation of chaos: The notion of acceptable initial distributions. The previous description of the Propagation of chaos is statistical. However the System (1.1) is deterministic and so are the main techniques used in this article. The strategy that is followed consists in characterizing a certain set $\mathcal{S}_\mathcal{N}$ of initial conditions s.t. if the initial vector of positions and velocities Z^0 belongs to that set, then the empirical measure $\mu_N^Z(t)$ remains very close of the solution f to (1.3) (in some appropriate weak distance).
We then show that if the initial positions and velocities are chosen randomly according to $(f^0)^{\otimes N}$, then the vector Z^0 belongs to $\mathcal{S}_\mathcal{N}$ with probability going to 1 when N goes to ∞. This result directly implies the propagation of chaos.
See from a different point of view, the notion of propagation of chaos enables us to decide whether the set $\mathcal{S}_\mathcal{N}$ is large enough to be physically relevant: A set of initial conditions $\mathcal{S}_\mathcal{N}$ is acceptable if random initial distributions belong to it with large probability.

Well posedness for System (1.1)**: The notion of solutions that is used.** We have not mentioned yet the most basic question for System (1.1) or (1.2) with a singular force kernel, namely whether one can even expect to have solutions to the system for a fixed number of particles. Before even thinking about the limit $N \to \infty$, it is necessary to make sure that some sort of solutions to (1.1) exist. For this, we shall first precise what we mean by solutions to (1.2) in the following definition

Definition 1. *A solution to* (1.2) *or* (1.1) *with initial condition*

$$Z^0 = (X_1^0,\ V_1^0, \dots, X_N^0,\ V_N^0) \in \mathbb{R}^{2dN}$$

(at time 0*) is a continuous trajectory* $Z(t) = \big(X_1(t),\ V_1(t), \dots, X_N(t),\ V_N(t)\big)$ *such that*

$$\forall t \in \mathbb{R}^+,\ \forall i \le N, \qquad \begin{cases} X_i(t) = X_i^0 + \displaystyle\int_0^t V_i(s)\, ds \\ V_i(t) = V_i^0 + \displaystyle\frac{1}{N} \sum_{j \neq i} \int_0^t F(X_i(s) - X_j(s))\, ds. \end{cases} \tag{1.4}$$

The solution is said local if the trajectory is defined (at least) on a finite set $[0, T^*)$*, and global if it is defined on the whole* $[0, +\infty)$*.*

Since we only use forces that are singular only at the origin, the usual Cauchy-Lipschitz theory implies that starting from any initial conditions such that $X_i^N \neq X_j^N$ for all $i \neq j$, there exists a unique local solution, defined till the time of first collision time T^*, when for some couple i, j we have $X_i^N(T^*) = X_j^N(T^*)$. In the case where the interaction force F derives from a repulsive singular potential ϕ *strong enough*, i.e. if Φ satisfy $\lim_{x \to 0} \Phi(x) = +\infty$, then collisions can never occurs and the solutions given by the classical Cauchy-Lipschitz theory are global.

In the other cases, it is not possible to extend the local result in such a simple way, and one could try to apply the DiPerna Lions theory [DL89], that allows to handle vector-field that are locally in $W^{1,1}$. This sound useful since any force satisfying the condition (1.8), with $\alpha < d-1$ has the required local regularity. But unfortunately, the DiPerna-Lions theory also requires a condition on the growth of the vector-field at infinity, which is not satisfied in our case. However if the interaction forces F derives from a potential Φ which is bounded at the origin (without any sign condition), the DiPerna Lions theory still leads to global solution for almost every initial conditions. This is stated precisely in the following Proposition which is a consequence of [Hau04, Theorem 4].

Proposition 1. *Assume that $F = -\nabla\Phi$ with $\Phi \in W^{2,1}_{loc}$, and that $\Phi(x) \geq -C(1+|x|^2)$ for some constant $C > 0$. Then for any fixed N, there exists a unique measure preserving and energy preserving flow defined almost everywhere on $\mathbb{R}^{2dN}$ associated to* (1.1)*. Such a flow precisely satisfy*

i) there exists a set $\Omega \subset \mathbb{R}^{2dN}$ with $|\Omega| = 0$ s.t. for any initial data $Z^0 \in \mathbb{R}^{2dN} \setminus \Omega$, we have a trajectory $Z(t)$ solution to (1.1)*,*

ii) for a.e. trajectory the energy conservation is satisfied,

iii) the family of solutions defines a global flow, which preserves the measure on $\mathbb{R}^{2dN}$.

Remark that the condition on Φ are fulfilled by any potential satisfying an (S^{α}_{Φ})-condition. So this proposition implies the global existence of solution for almost all initial positions and velocities in that case, and this is completely sufficient for our results on propagation of chaos : Theorem 3 requires only the existence of a solution with given initial data and Theorem 1 requires the existence of solution for almost all initial data.

In the case of some particular more singular attractive potentials : the gravitational force ($\alpha = d-1$) in dimension 2 or 3, and also for some others power law forces, it is known [Saa73] that initial conditions leading to "standard" collisions (possibly multiple and simultaneous), is of zero measure. But, what is unknown even if it seems rather natural, is that the set of initial collisions leading to the so-called "non-collisions" singularities, which do exists [Xia92], is of zero measure for $N \geq 5$. Up to our knowledge it has only been proved for $N \leq 4$ [Saa77]. In fact, there is a large literature about this N body problem in physicist and mathematician community. However, as we shall see later, that discussion is not really relevant here since in the case of large singularity, we will use a regularization or cut-off of the force (see the condition (1.11), thanks to which the question of global existence becomes trivial.

Eventually, the only case in which we are not covered by the existing literature is the case of non potential force satisfying the (S^{α})-condition for some $\alpha < 1$, for which we claimed a result without cut-off. In that case, we follow the following simple strategy. As in the case with larger singularity we use a cut-off or regularization of the interaction force. The existence of global solution is then straightforward. And our results of convergence are valid independently of the size of cut-off (or smoothing parameter) used. It can be any positive function of the number of particles N.
Note that this suggests in fact that for a fixed N, the analysis done in this article should almost implies the existence of solutions for almost all initial conditions. If one checks precisely, ours proofs show that trajectories may be extended after a collision where the relative velocities between the two particles goes to a non zero limit. Hence the only collisions

that remain problematic are those where the relative speed of the colliding particles vanishes, but our result implies more or less that this never happens. We have tried to make this more precise in remark 7 after Theorem 1.

Previous results with cut-off or for smooth interactions. The mean-field limit and the propagation of chaos are known to hold for smooth interaction forces ($\Phi \in C^2$ in general or at least $W^{2,\infty}_{loc}$) since the end of the seventies and the works of Braun and Hepp [BH77], Neunzert and Wick [NW80] and Dobrushin [Dob79]. Those articles introduce the main ideas and the formalism behind mean field limits, we also refer to the nice book by Spohn [Spo91]. Their proofs however rely on Gronwall type estimates and are connected to the fact that Gronwall estimates are actually true for (1.1) uniformly in N if $\nabla\Phi = F \in W^{1,\infty}$. Those makes them impossible to generalize to any case where F is singular (including Coulombian interactions and many other physically interesting models).
Instead, by keeping the same general approach, it is possible to deal with singular interactions with cut-off. For instance for Coulombian interactions, one could consider

$$F_N(x) = C\,\frac{x}{(|x|^2 + \varepsilon(N)^2)^{d/2}},$$

or other type of regularization at the scale $\varepsilon(N)$. The system (1.2) does not have much physical meaning but the corresponding studies are crucial to understand the convergence of numerical methods. For particles initially on a regular mesh, we refer to the works of Ganguly and Victory [GV89], Wollman [Wol00] and Batt [Bat01] (the latter gives a simpler proof, but valid only for larger cut-off). Unfortunately they had to impose that $\lim_{N\to\infty} \varepsilon(N)/N^{-1/d} = +\infty$, meaning that the cut-off for convergence results is usually larger than the one used in practical numerical simulations . Note that the scale $N^{-1/d}$ is the average distance between two neighboring particles in position.
These “numerically oriented” results do not imply the propagation of chaos, as the particles are on a mesh initially and hence cannot be taken randomly. Moreover, we emphasize that the two problems with initial particles on a mesh, or with initial particles not equally distributed seems to be very different. In the last case, Ganguly, Lee, and Victory [GLV91] prove the convergence only for a very large cut-off $\varepsilon(N) \approx (\ln N)^{-1}$.

Previous results for $2d$ Euler or other macroscopic equations. A well known system, very similar at first sight with the question here, is the vortices system for the $2d$ incompressible Euler equation. One replaces (1.1) by

$$\dot{X}_i = \frac{1}{N}\sum_{j\neq i} \alpha_i\,\alpha_j\,\nabla^{\perp}\Phi(X_i - X_j), \tag{1.5}$$

where Φ is still the Coulombian kernel (in 2 dimensions here) and $\alpha_i = \pm 1$. One expects this system to converge to the Euler equation in vorticity formulation

$$\partial_t\omega + \mathrm{div}\,(u\,\omega) = 0, \qquad \mathrm{div}\,u = 0, \qquad \mathrm{curl}\,u = \omega. \tag{1.6}$$

The same questions of convergence and propagation of chaos can be asked in this setting. Two results without regularization for the true kernel are already known. The work of Goodman, Hou and Lowengrub, [GHL90] and [GH91], has a numerical point of view but

uses the true singular kernel in a interesting way. The work of Schochet [Sch96] uses the weak formulation of Delort of the Euler equation and proves that empirical measures with bounded energy converge towards measures that are weak solutions to (1.6). Unfortunately, the possible lack of uniqueness of Euler equation in the class of measures does not allows to deduce the propagation of chaos.
As equations like (1.6) are notoriously harder to deal with than kinetic equations like (1.3), one could expect similar results for our problem. Unfortunately, the mean field limits are more difficult for kinetic-like equations, with space and velocity variables. There are several reasons for that, in particular the fact that system (1.1) is second order while (1.5) is first order. This implies that collisions or near collisions (in physical space) between particles are very common for (1.1) even for repulsive interactions and rare for (1.5), at least for vortices of opposite sign.
For example, the references mentioned above use the symmetry of the forces in the vortex case, a symmetry which cannot exist in our kinetic problem; independently of the additional structural assumptions like $F = -\nabla\Phi$. The force is still symmetric with respect to the space variable, but there is now a velocity variable which breaks the argument used in the vortices case. For a more complete description of the vortices system, we refer to the references already quoted or to [Hau09], which introduces in that case techniques similar to the one used here.

A previous result in singular cases without cut-off. To our knowledge, the only mean field limit result available up to now in this case is [HJ07]. This proves the convergence (not the propagation of chaos) provided that

- The interaction kernel F (or $\nabla\Phi$) satisfy a (S^α)-condition with $\alpha < 1$.
- The particles are initially well distributed, meaning that the minimal inter-distance in $\mathbb{R}^{2d}$ is of the same order as the average distance between neighboring particles $N^{-1/2d}$.

The second assumption is all right for numerical purposes but does not allow to consider physically realistic initial conditions, as per the propagation of chaos property. This assumption is indeed not generic for empirical measures chosen with law $(f^0)^{\otimes N}$, *i.e.* it is satisfied with probability going to 0 in the large N limit.

Our result without cut-off. In the present article, we keep the same conditions on the interaction kernel, but require only a much weaker assumption on the minimal distance between particles. This allows us to prove the propagation of chaos, for potentials satisfying a (S^α_Φ)-condition, (S^α_Φ)-condition

$$(S^\alpha_\Phi) \qquad \exists\, C > 0, \quad \forall\, x \in \mathbb{R}^d\backslash\{0\}, \quad |\nabla\Phi(x)| \le \frac{C}{|x|^\alpha}, \quad |\nabla^2\Phi(x)| \le \frac{C}{|x|^{\alpha+1}}, \tag{1.7}$$

or in the more general (1.2), the corresponding (S^α)-condition

$$(S^\alpha) \qquad \exists\, C > 0, \quad \forall\, x \in \mathbb{R}^d\backslash\{0\}, \quad |F(x)| \le \frac{C}{|x|^\alpha}, \quad |\nabla F(x)| \le \frac{C}{|x|^{\alpha+1}}, \tag{1.8}$$

with $\alpha < 1$ in both cases.

But since the dynamic is deterministic, the key result of this article is also deterministic and gives as a corollary the propagation of chaos. Here we will only give the main lines of that deterministic result, a precise statement is made in Theorem 3, after some technical notions are first introduced.

A sketch of Theorem 3. *Assume that $d \geq 2$ and that the potential Φ satisfies the (S^α_Φ) (1.7) condition (or (1.8) for (1.2)), for some $\alpha < 1$ and let $0 < \gamma < 1$. Assume also that the initial distribution of particles satisfy:*

i) $\inf_{i\neq j} |(X_i^0, V_i^0) - (X_j^0, V_j^0)| \geq N^{-\gamma(1+r)/2d}$ *for some $r \in (1, r^*)$ where $r^* := \frac{d-1}{1+\alpha}$,*

ii) the initial positions and velocities are relatively "well distributed" in every ball of diameter $N^{-\gamma/2d}$ in some sense to be precise,

iii) For some $R > 0$, $\forall N \in \mathbb{N}$, $\mathrm{Supp}\,\mu_N^0 \subset B(0, R)$, the ball of radius R and center 0 of R^{2d}.

Then there exists a strong solution f_N to (1.3) defined on a time interval $[0, T^*)$, and for any $T \in [0, T^*)$, there exists two constants C_0 and C_1 such that for $N \geq e^{C_1 T}$, an appropriate weak distance between μ_N and f_N is bounded by $\frac{e^{C_0 t}}{N^{\gamma/2d}}$.

Note that in this result, we compare μ_N with a solution f_N to (1.3) whose initial data f_N^0 is different f^0, but in most case close from it. In fact f_N^0 is obtained from μ_N^0 by replaced the Dirac masses by small "blops". So that thanks to that deterministic result we can also construct explicit sequences of initial conditions for which the convergence towards the Vlasov equation will holds (for instance, particles well chosen on a mesh, but not only).

Combined with a statistical analysis showing that conditions $(i) - (iii)$ are almost always satisfied, and that f_N^0 can almost always be chosen close from f^0, this deterministic result allows us to deduce the main consequence of this article namely a quantitative version of propagation of chaos

Theorem 1. *Assume that $d \geq 3$ and that Φ satisfies a (S^α_Φ)-condition (1.7) with $\alpha < 1$ (or (1.8) for (1.2). Then there exist a positive real number $\gamma^* \in (0, 1)$ depending only on (d, α) and a function $s^* : \gamma \in (\gamma^*, 1) \to s^*_\gamma \in (0, \infty)$ s.t. the following is true.*

- For any non negative initial data $f^0 \in L^\infty(\mathbb{R}^{2d})$ with compact support and total mass one, we consider the unique global, bounded, and compactly supported solution f of the Vlasov equation (1.3), see Proposition 3.

- For each $N \in \mathbb{N}^*$, we consider the particles system (1.1) with initial positions $Z^0 = (X_i^0, V_i^0)_{i\leq N}$ chosen randomly according to the probability $(f^0)^{\otimes N}$.

Then, for all $T > 0$, any

$$\gamma^* < \gamma < 1 \quad \text{and} \quad 0 < s < s^*_\gamma,$$

there exists three positive constants $C_0(T, f, \Phi)$, $C_1(\gamma, s, T, f, \Phi)$ and $C_2(f^0, \gamma)$ such that for $N \geq e^{C_1 T}$

$$\mathbb{P}\left(\exists t \in [0, T],\ W_1(\mu_N^Z(t), f(t)) \geq \frac{3e^{C_0 t}}{N^{\gamma/(2d)}}\right) \leq \frac{C_2}{N^s}, \tag{1.9}$$

where W_1 denotes the 1 Monge-Kantorovitch Wasserstein distance. The constants C_1 and C_2 blow up when γ or s approach their maximum value.

Remark 1. *The notation* $C(T, f)$ *means that the constant depends on the function* f *(essentially via conserved quantities like* $\|f\|_\infty$ *and also the size of its support) on the whole time interval* $[0, T]$.

Remark 2. *We have explicit formulas for* γ^* *and* s^*_γ *namely*

$$\gamma^* := \frac{2 + 2\alpha}{d + \alpha} \quad \text{and} \quad s^*_\gamma := \frac{\gamma d - (2 - \gamma)\alpha - 2}{2(1 + \alpha)}.$$

Those conditions are not completely obvious, but it can be checked that if $\alpha < 1$ *and* $d \geq 3$, $\gamma^* < 1$ *so that admissible* γ *exist. And for an admissible* γ, s^*_γ *is also positive, so that admissible* s *also exists. The best choices for* γ *and* s *would be* $\gamma = 1$ *and* $s = \frac{d-\alpha-2}{2(1+\alpha)}$ *as those give the fastest convergence. Unfortunately the constant* C_1 *and* C_2 *would then be* $+\infty$ *hence the more complicated formulation.*

Remark 3. *Roughly speaking, under the assumption of Theorem 1, except for a small set of initial conditions* $\mathcal{S}^c_N$, *the deviation between the empirical measure and the limit is at most of the same order as the average inter-particle distance* $N^{-1/2d}$.

Remark 4. *Unfortunately, that result do not apply in dimension* $d = 2$, *even if* α *is very small. It can be seen that the condition on* γ *in theorem 1 is empty in that case.*

Remark 5. *By embedding and interpolation, the* W_1 *distance could be replaced by essentially any weak distance on measures, with maybe a change in the exponent* γ. *It is however the natural distance for the statistical studies of the initial conditions. Unfortunately it does not seem to be the right distance for the previous deterministic result where a more complicated infinite Monge-Kantorovitch-Wasserstein distance is needed.*

Remark 6. *Here we prove the propagation of chaos starting with independent initial particles, i.e. particles with law* $(f^0)^{\otimes N}$. *However, since the independence is not preserved by the particles dynamic, this give a very special role to the time* $t = 0$. *Because of this, people sometimes prefers to state the propagation of chaos starting from any laws on the initial particles that are* f^0*-chaotic, in the sense recalled in the paragraph on propagation of chaos. But we cannot prove that the propagation of chaos will hold for any initial* f^0*-chaotic sequence of initial distribution for the particles, because the chaoticity does not imply any control on the minimal inter-particles distance in position-velocity space, which is crucial in our case. However, we can state a more general result if we include the control on the minimal inter-particles distance and another on the infinite norm on the* f^N_0 *in the requirements. Shortly it may be stated as follows : If the initial distributions of particles are* f^0 *chaotic and there exists a control on the two quantities mentioned above at the appropriate scale, then the chaoticity and the control on the two quantities are preserved by the dynamics.*

Remark 7. *As mentioned in the introduction, the arguments in the proof of Theorem 1 prove that, at fixed* N, *there exists a global solution to* (1.4) *for almost all initial conditions. In fact, in a very sketchy way, this theorem also propagates a control on the minimal inter-particles distance in position-velocity space. Used as is, it only says that asymptotically, the control is good with large probability. However for fixed* N, *if we let some constants increase as much as needed, it is possible to modify the argument and obtain a control for almost all initial configurations. Since the proof also implies that the only bad collisions are the collisions with vanishing relative velocities, we can obtain some existence (and also) uniqueness for almost all initial data of the ODE* (1.2) *in that case.*

The improvements with respect to [HJ07]. The major improvement is of course the much weaker condition on the initial distribution of positions and velocities. We are hence able to show the propagation of chaos, which is again the crucial property for applications to physics. We also managed to simplify the proof of the deterministic result, considerably so in the long time case which was quite intricate before and does not require any special treatment here.
Moreover, the deterministic part of our analysis is quantitative. For large enough N, Theorem 3 gives a precise rate of convergence in Monge-Kantorovitch-Wasserstein distance, which is of course quite useful from the point of view of numerical analysis.
Unfortunately, the condition on the potential Φ (or the force F) is still the same and does not allow to treat Coulombian interactions. There are some physical reasons for this condition, which are discussed in a later paragraph. We refer to [BHJ10] for some ideas in how to go beyond this threshold in the repulsive case.

The result with cut-off. The result presented here is in one sense slightly weaker than the previously known result [GLV91], since we just miss the critical case $\alpha = d-1$. But in that last work the cut-off used is very large: $\varepsilon(N) \approx (\ln N)^{-1}$, and is equivalent to the absence of singularity from a physical point of view. Here we will use cut-off that are some power of N. They will still be quite larger than the average distance between neighboring particles for α close to $d-1$, but in the case where α is larger but close to 1 the cut-off parameter may be even smaller than the minimal inter-particles distance.
The result is stated for potentials or forces depending on N and satisfying the following condition

$$(S^\alpha_{\Phi,m}) \qquad \begin{array}{ll} i) & \Phi \text{ satisfies a } (S^\alpha_\Phi)-\text{condition with possibly } \alpha \geq 1, \\ ii) & \forall\,|x| \geq N^{-m},\ \Phi_N(x) = \Phi(x) \\ iii) & \forall\,|x| \leq N^{-m},\ |\nabla \Phi_N(x)| \leq N^{m\alpha}. \end{array} \tag{1.10}$$

Even if one starts with the Hamiltonian structure of (1.1), it is not necessary to preserve it when applying the cut-off and it is possible to consider (1.2) with F_N s.t.

$$(S^\alpha_m) \qquad \begin{array}{ll} i) & F \text{ satisfy a } (S^\alpha)-\text{condition} \\ ii) & \forall\,|x| \geq N^{-m},\ F_N(x) = F(x) \\ iii) & \forall\,|x| \leq N^{-m},\ |F_N(x)| \leq N^{m\alpha}, \end{array} \tag{1.11}$$

which essentially means that the interaction kernel is regularized at scales lower than N^{-m}. Note that in fact we would not need any estimate on $\nabla^2 \Phi_N$ or ∇F_N for very small x. The result would still be true if F_N only converges to F for large enough x, with an error satisfying $\|F_N - F\|_1 \leq N^{-1/2d}$. The proof could be adapted to that case, but for simplicity we choose this presentation. We also point out that one would like to take m as large as possible if we want to be close to the dynamics without cut-off.

Theorem 2. *Assume that $d \geq 3$, $\gamma \in (0,1)$ and that Φ_N satisfies a $(S^\alpha_{\Phi,m})$-condition (or F_N the (S^α_m)-condition) for some $1 \leq \alpha < d-1$ with a cut-off order m such that*

$$m < m^* := \frac{\gamma}{2d} \min\left(\frac{d-2}{\alpha-1}, \frac{2d-1}{\alpha}\right).$$

Choose any initial condition $f^0 \in L^\infty$ with compact support and total mass one for the Vlasov equation (1.3), and denote by f the unique (at least) local strong solution of the

Vlasov equation(1.3) *with initial condition* f^0*, given by Proposition 3. For each* $N \in \mathbb{N}^*$*, consider the particles system* (1.1) *for* Φ_N *or* (1.2) *for* F_N *with initial positions* $(X_i, V_i)_{n\leq N}$ *chosen randomly according to the probability* $(f^0)^{\otimes N}$*.*
Then there exists a time T^**, smaller than the time of existence of the strong solution* f*, and equals to* $+\infty$ *in the case where the strong solution to the Vlasov equation* (1.3) *are global such that for all* $T < T^*$ *there exists positive constants* $C_0(T, f, \Phi)$*,* $C_1(\gamma, m, T, f, \Phi)$ $C_2(f)$ *and* $C_3(f)$ *such that for* $N \geq e^{C_1 T}$

$$\mathbb{P}\left(\exists t \in [0,T],\ W_1(\mu_N(t), f(t)) \geq \frac{4\, e^{C_0 t}}{N^{\gamma/2d}}\right) \leq C_2 N^{\gamma}\, e^{-C_3 N^{\lambda}},$$

where $\lambda = 1 - \max\left(\gamma, \frac{1}{d}\right)$*.*

Remark 8. *Our result is valid only locally in time in the case where blow-up may occur in the Vlasov equation, as for instance in dimension larger than* 4*. But it is valid for any time in dimension three or less, since in that case the strong solutions of the Vlasov equations we are dealing with are global, see Proposition 3 in section 2.4.*

In dimension $d = 3$, the minimal cut-off is given by the order of $m^* = \frac{\gamma}{6}\min((\alpha-1)^{-1}, 5\alpha^{-1})$. As γ can be chosen very close to one, for α larger but close to one, the previous bound tells us that we can choose cut-off of order almost $N^{-5/6}$, i.e. much smaller than the likely minimal inter-particles distance in position space (of order $N^{-2/3}$, see the third section). With such a small cut-off, one could hope that it is almost never used when we calculate the interaction forces between particles. Only a negligible number of particles will become so close to each others during the time T. This suggests that there should be some way to extend the result of convergence without cut-off at least to some $\alpha > 1$.
Unfortunately, we do not know how to make rigorous the previous probabilistic argument on the close encounters. First it is highly difficult to translate for particles system that are highly correlated. To state it properly we need infinite bounds on the 2 particles marginal. But obtaining such a bound for singular interaction seems difficult. Moreover, it remains to neglect the influence of particles that have had a close encounters (its trajectory after a encounter is not well controlled) on the other particles.
Let us also mention that astro-physicists doing gravitational simulations ($\alpha = d-1$) with tree codes usually use small cut-off parameters, lower than $N^{-1/d}$ by some order. See [Deh00] for a physical oriented discussion about the optimal length of this parameter.

The structure of the force term: Potential, repulsion, attraction? In the particular case where the force derives from a potential $F = -\nabla\Phi$, The system (1.1) is endowed with some important additional structure, for example the conservation of energy

$$\frac{1}{N}\sum_i \frac{|V_i|^2}{2} + \frac{1}{2N^2}\sum_{i\neq j}\Phi(X_i - X_j) = const.$$

When the forces are repulsive, *i.e.* $\Phi \geq 0$, this immediately bounds the kinetic energy and separately the potential energy. However this precise structure is never used in this article, which may seem weird at first glance. We present here some arguments that can explain this fact.
First, for the interactions considered in the case without cut-off, again satisfying a (S^α_Φ) condition with $\alpha < 1$, the potential Φ is continuous (hence locally bounded). In that case the

singularity in the force term is too weak to really see or use a difference between repulsive and attractive interactions. Two particles having a close encounter cannot have a strong influence onto each other, both in the attractive or repulsive case. Similarly the fact that the interaction derive from a potential is not really useful, hence our choice later on of the slightly more general setting.
It should here be noted that the previous discussion applies to every previous result on the mean field limit or propagation of chaos in the kinetic case: They all require assumptions (typically $\nabla^2\Phi$ locally bounded) implying that the attractive or repulsive nature of the interaction does not matter; the situation is different for the macroscopic "Euler-like" cases, see the comments in the paragraph devoted to that case. The present contribution shows that mean field limits and propagation of chaos are essentially valid at least as long as the potential is bounded (instead of only $W^{2,\infty}_{loc}$ as before). This corresponds to the physical intuition that nothing should go wrong as long as the local interaction between two very close particles is too weak to impact the dynamics.

The exact structure of the interaction kernel should become crucial once this threshold is passed, *i.e.* for $\Phi(x) \sim C|x|^{1-\alpha}$ at the origin with $\alpha \geq 1$. But here we use in that case a cut-off, which weaken the effect of the interaction between two very close particles. In fact in order to prove the mean-field limit, we are able to show that if the cut-off is large enough, these local interactions may be neglected. So our technics still do not make any difference between the repulsive or attractive cases.
However in the case where that "strong" singularity is repulsive, the potential energy is bounded, and if we were able to use this fact, we will obtain results depending of the attractive-repulsive character of the interaction. In that respect, we point out that the information contained in a bounded potential energy is actually quite weak and clearly insufficient, at least with our technics. Assume for instance that $\Phi(x) \sim |x|^{1-\alpha}$ for some $\alpha > 1$. Then the boundedness of the potential energy implies that the minimal distance in physical space between any two particles is of order $N^{-2/(\alpha-1)}$, which is at best N^{-2} in the Coulomb case, $\alpha = 2$. But it can be checked that the cut-off parameter N^{-m} given in Theorem 2 as a power m which is always much lower than $\frac{2}{\alpha-1}$, i.e. that the cut-off we use is always much larger than the minimal distance provided by the bound on the potential energy. To go further, an interesting idea is to compare the dynamics of the N particles with or without cut-off. But even if the difference between the original force and its mollified version is well localized, it is quite difficult to understand how we can control the difference between the two associated dynamics. We refer to [BJ08] for a first attempt in that direction, in which well-localized and singular perturbation of the free transport are investigated.

Therefore in those singular settings, the repulsive or potential structure of the interaction will only help in a more subtle (and still unidentified) manner. An interesting comparison is the stability in average proved in [BHJ10]: This requires repulsive interaction not to control locally the trajectory but in order to use the statistical properties of the flow (through the Gibbs equilibrium).

A short sketch of the proof. As mentioned above, the Vlasov equation (1.3) is satisfied by the empirical distribution μ_N of the interacting particle system provided that $F(0)$ is set to 0. Hence the problem of convergence can be reformulated into a problem of stability of the empirical measures - seen initially as a measure valued perturbation of the smooth profile f^0 - around the solution $f(t)$ of the Vlasov equation.

Our proof uses 3 ingredients to obtain this stability

- Show that with large probability the empirical measure at the initial time $\mu_N^z(0)$ is very close to f^0 and that the particles are not too badly distributed in the space $\mathbb{R}^{2d}$. In particular, we need that the regularization f_N^0 obtained from $\mu_N^z(0)$ by replacing all the Dirac masses by small "blobs" is bounded in L^∞.
- Compare the solution $f(t)$ to (1.3) with f^0 as initial data to the solution $f_N(t)$ to (1.3) with f_N^0 as initial data.
- Control the distance (in some appropriate Monge-Kantorovitch-Wasserstein distance) between $\mu_N^z(t)$ and $f_N(t)$.

The first two steps are not overly complicated because they rely on previous known results and rather simple probabilistic estimates. The second is a standard stability result for Vlasov-Poisson for example. The difficulties are hence concentrated in the third step. This uses a deterministic result that has to be more precise than the one in [HJ07]. First of all it has to allow for a much smaller minimal distance in $\mathbb{R}^{2d}$ between the particles at the initial time and thus as well at later times. The other complication comes from that we want an explicit bound on the distance between $f_N(t)$ and $\mu_N^z(t)$ in the infinite MKW distance, defined in Section 2. We emphasize that the use of the infinite MKW distance is important. We were not able to perform it with other MKW distance of order $p < +\infty$. It may seem strange to propagate a stronger norm for a problem with low regularity but in fact it turns out to be the only MKW distance with which we can handle a localized singularity in the force and Dirac masses in the distribution.

In essence controlling the distance between $f_N(t)$ and $\mu_N^z(t)$ requires to prove that the difference between the force terms acting on f_N and on $\mu_N^z(t)$ is bounded by this distance plus a small correction; then one may conclude by a Gronwall-like argument. There are several major complications in the proof.
First of all one has to deal with possible singularities in the discrete force term. This is due to particles that may come very close one to another in physical space $\mathbb{R}^d$ or even collide, but with non-vanishing relative speed. The force term may become very large in that case but only for a very short time as the particles do not remain close long. To overcome this difficulty, we have to average those singularities over a short time interval.
Then when one compares the force terms it is necessary to distinguish between three domains of interacting particles:

- Interaction between particles far enough in the physical space and that remain far enough over the short time interval where we average. This is the simplest case as one does not see the discrete nature of the problem at that level. The estimates need to be adapted to the distance used here but are otherwise very similar in spirit to the continuous problem or other previous works for mean field limits.
- Interaction between particles close enough in the physical space $\mathbb{R}^d$, but with sufficiently different velocities. Here we start to see the discrete level of the problem and in fact we cannot compare anymore the force term for the continuous Vlasov equation with the one for the particles' system. Instead we just show that both are small. This is simple for the continuous term as the size of the domain is itself quite small. For the discrete force term, it is considerably more complicated. First of all one has to

bound the average over the short time interval of the interaction term between any two particles in that situation. Roughly speaking if those have a relative velocity of order v one expects the interaction to behave like

$$\int_t^{t+\tau} \frac{ds}{|\delta + (s-s_0)v|^\alpha},$$

where τ is the size of the average in time and δ is the minimum distance between the two particles on the time interval $[t,\ t+\tau]$, which is reached at time s_0, and v is the relative velocity at this time. This is where the condition $\alpha < 1$ first comes into play as it allows to bound the previous integral independently of δ, provided that v is large enough. To conclude this part it is finally necessary to sum the previous bound over all particles in the corresponding domain. In [HJ07], we could directly bound the number of particles there; here as we do not have the same lower bound on the minimal distance in $\mathbb{R}^{2d}$, we have to use the distance between f_N and μ_N^z. This can be done if we use the W_∞ distance.

In many sense, this step is the crucial one: It obviously only works for $\alpha < 1$. The physical explanation is clear: If $\alpha < 1$ the deviation in velocity due to a collision (another particle coming very close) is small. In particular there cannot be any fast variation in the velocities of the particles. In contrast when $\alpha > 1$, a particle coming very close to another one can change its velocity over a very short time interval (even if their relative velocities remain of order 1). Therefore for $\alpha < 1$, it is enough to control the distance in $\mathbb{R}^{2d}$ between particles. For $\alpha > 1$, it would be necessary to have in addition a good control for the distance in the physical space. Note that this problem for $\alpha > 1$ is present even in the case of repulsive interaction, since the deviation in velocity can be of order 1 in any case.

- Interaction between particles close enough in $\mathbb{R}^{2d}$ over the short time interval where we average. In [HJ07] this case was relatively simple as the diameter of the corresponding domain in $\mathbb{R}^{2d}$ was of the same order as the lower bound on the minimal distance (still in $\mathbb{R}^{2d}$) that we were propagating. Here this lower bound is much smaller, of the order of the square of the diameter. This is where the main technical improvement lies with respect to [HJ07].

Organization of the paper. In the next section, we introduce the notations, and state the deterministic results on which the propagation of chaos relies. In the third section, we explain how to obtain the propagation of chaos from the deterministic results. The fourth section is devoted to the proof of the two deterministic theorems.
As the Hamiltonian structure is never used in the following, from now on we focus on the slightly more general (1.2).

2 Notations and other important theorems

2.1 Notations and useful results

We first need to introduce some notations and to define different quantities in order to state the result.

In the sequel, we shall always use the euclidian distance on $\mathbb{R}^d$ for positions or velocity, or on $\mathbb{R}^{2d}$ for the couples position-velocity. In all case, it will be denoted by $|x|$, $|v|$, $|z|$. The notation $B_n(a,R)$ will always stand for the ball of center a and radius R in dimension $n=d$ or $2d$. Be careful that the Lebesgue measure of a measurable set A will also be denoted by $|A|$.

- **Empirical distribution μ_N and minimal inter-particle distance d_N**

Given a configuration $Z=(X_i,V_i)_{i\le N}$ of the particles in the phase space $\mathbb{R}^{2dN}$, the associated empirical distribution is the measure

$$\mu_N^Z = \frac{1}{N}\sum \delta_{X_i,V_i}.$$

An important remark is that if $(X_i(t,),V_i(t))_{i\le N}$ is a solution of the system of ODE (1.2), then the measure $\mu_N^Z(t)$ is a solution of the Vlasov equation (1.3), provided that the interaction force satisfies $F(0)=0$. This condition is necessary to avoid self-interaction of Dirac masses. It means that the interaction force is defined everywhere, but discontinuous and has a singularity at 0. In that conditions, the previously known results [BH77], [NW80] cannot be applied.
For every empirical measure, we define the minimal distance d_N^Z between particles in $\mathbb{R}^{2d}$

$$d_N^Z(\mu_N) := \min_{i\neq j} |Z_i - Z_j| = \min_{i\neq j}\big(|X_i-X_j|^2+|V_i-V_j|^2\big)^{\frac{1}{2}}. \tag{2.1}$$

This is a non physical quantity, but it is crucial to control the possible concentrations of particles and we will need to bound that quantity from below.
In the following we often omit the Z superscript, in order to keep "simple" notations.

- **Infinite MKW distance**

First, we use many times the Monge-Kantorovitch-Wasserstein distance of order one and infinite. The order one distance, denoted by W_1, is classical and we refer to the very clear book of Villani for definition and properties [Vil03]. The second one denoted W_∞ is not widely used, so we recall its definition. We start with the definition of transference plane

Definition 2. *Given two probability measures μ and ν on $\mathbb{R}^n$ for any $n\ge 1$, a transference plane π from μ to ν is a probability measure on $X\times X$ s.t.*

$$\int_X \pi(x,dy)=\mu(x), \qquad \int_X \pi(dx,y)=\nu(y),$$

that is the first marginal of π is μ and the second marginal is ν.

With this we may define the W_∞ distance

Definition 3. *For two probability measures μ and ν on $\mathbb{R}^n$, with $\Pi(\mu,\nu)$ the set of transference planes from μ to ν:*

$$W_\infty(\mu,\nu)=\inf\{\pi - esssup\,|x-y| \; |\pi\in\Pi\}.$$

There is also another notion, called the transport map. A transport map is a measurable map $T:\operatorname{Supp}\mu\to\mathbb{R}^n$ such that $(Id,T)_\#\mu\in\Pi$. This means in particular that $T_\#\mu=\nu$, where the push forward of a measure m by a transform L is defined by

$$L_\# m(O) = m(L^{-1}(O)), \qquad \text{for any measurable set } O.$$

In one of the few works on the subject [CDPJ08] Champion, De Pascale and Juutineen prove that if μ is absolutely continuous with respect to the Lebesgue measure $\mathcal{L}$, then at least one optimal transference plane for the infinite Wasserstein distance is given by a optimal transport map, i.e. there exists T s.t. $(Id, T)_{\#}\mu \in \Pi$ and

$$W_\infty(\mu, \nu) = \mu - \text{esssup}_x\, |Tx - x|.$$

Although that is not mandatory (we could actually work with optimal transference planes), we will use this result and work in the sequel with transport maps. That will greatly simplify the notation in the proof.

Optimal transport is useful to compare the discrete sum appearing in the force induced by the N particles to the integrals of the mean-field force appearing in the Vlasov equation. For instance, if f is a continuous distribution and μ_N an empirical distribution we may rewrite the interaction force of μ_N using a transport map $T = (T_x, T_v)$ of f onto μ_N

$$\frac{1}{N}\sum_{i\neq j} F(X_i^0 - X_j^0) = \int F(X_i^0 - T_x(y,w)) f(y,w)\, dydw.$$

Note that in the equality above, the function F is singular at $x = 0$, and that we impose $F(0) = 0$. The interest of the infinite MKW distance is that the singularity is still localized “in a ball“ after the transport : the term under the integral in the right-hand-side has no singularity out of a ball of radius $W_\infty(f, \nu_N)$ in x. Others MKV distance of order $p < +\infty$ destroys that simple localization after the transport, which is why it seems more difficult to use them.

- **The scale ε.** We also introduce a scale

$$\varepsilon(N) = N^{-\gamma/2d}\,, \tag{2.2}$$

for some $\gamma \in (0,1)$ to be fixed later but close enough from 1. Remark that this scale is larger than the average distance between a particle and its closest neighbor, which is of order $N^{-1/2d}$. We shall do a wide use of that scale in the sequel, and will often define quantities directly in term of ε rather than N. For instance, the cut-off order m used in the (S_m^α)-condition may be rewritten in term of ε, with $\bar m := \frac{2d}{\gamma} m \in \left(1, \min(\frac{d-2}{\alpha-1}, \frac{2d-1}{\alpha})\right)$.

$$(S_m^\alpha) \qquad \begin{array}{ll} i) & F \text{ satisfy a } (S^\alpha) - \text{condition} \\ ii) & \forall\, |x| \geq \varepsilon^{\bar m},\ F_N(x) = F(x) \\ iii) & \forall\, |x| \leq \varepsilon^{\bar m},\ |F_N(x)| \leq \varepsilon^{-\bar m \alpha}, \end{array} \tag{2.3}$$

- **The solution f_N of Vlasov equation with blob initial condition.**

Now we defined a smoothing of μ_N at the scale $\varepsilon(N)$. For this, we choose a kernel ϕ : $\mathbb{R}^{2d} \to \mathbb{R}$ radial with compact support in $B_{2d}(0,1)$ and total mass one, and denote $\phi_\varepsilon(\cdot) = \varepsilon^{-2d}\phi(\cdot/\varepsilon)$. The precise choice of ϕ is not very relevant, and the simplest one is maybe $\phi = \frac{1}{|B_{2d}(0,1)|}\mathbf{1}_{B_{2d}(0,1)}$. We use this to smooth μ_N and define

$$f_N^0 = \mu_N^0 * \phi_{\varepsilon(N)}, \tag{2.4}$$

and denote by $f_N(t,x,v)$ the solution to the Vlasov Eq. (1.3) for the initial condition f_N^0.

The interest of f_N is that we may assume that f_N^0 belongs to L^∞, see Proposition 8 in the Appendix for deviations estimates on that quantity. And this also holds for any time since infinite bound are propagated by the Vlasov equation. That L^∞ bound allows to use standard stability estimates to control its W_1 distance to another solution of the Vlasov equation, see Loeper result [Loe06] recalled in Proposition 4.
A key point in the rest of the article is that f_N^0 and μ_N^0 are very close in W_∞ distance as per

Proposition 2. *For any $\phi : \mathbb{R}^{2d} \to \mathbb{R}$ radial with compact support in $B_{2d}(0,1)$ and total mass one we have for any $\mu_N^0 = \frac{1}{N}\sum_{i=1}^N \delta_{(X_i^0, V_i^0)}$*

$$W_\infty(f_N^0, \mu_N^0) = c_\phi \varepsilon(N)$$

where c_ϕ is the smallest c for which $\operatorname{Supp}\phi \subset B_{2d}(0,c)$.

Proof. Unfortunately even in such a simple case, it is not possible to give a simple explicit formula for the optimal transport map. But there is a rather simple optimal transference plan. Define

$$\pi(x,v,y,w) = \frac{1}{N}\sum_i \phi_\varepsilon(x-y, v-w)\, \delta_{(X_i^0, V_i^0)}(y,w).$$

Note that

$$\int_{\mathbb{R}^{2d}} \pi(x,v,dy,dw) = [\mu_N^0 * \phi_\varepsilon](x,v) = f_N^0(x,v),$$

and since ϕ_ε has mass 1

$$\int_{\mathbb{R}^{2d}} \pi(dx,dv,y,w) = \frac{1}{N}\sum_i \delta_{(X_i^0, V_i^0)}(y,w) = \mu_N^0(y,w).$$

Therefore π is a transference plane between f_N^0 and μ_N^0. Now take any (x,v,y,w) in the support of π. By definition there exists i s.t. $y = X_i^0$, $w = V_i^0$ and (x,v) is in the support of $\phi_\varepsilon(. - X_i^0, v - V_i^0)$. Hence by the assumption on the support of ϕ

$$|x-y|^2 + |v-w|^2 \le c_\phi[\varepsilon(N)]^2,$$

which gives the upper bound.
We turn to the lower bound. Remark that the assumptions imply that $\phi > 0$ on the interior of $B(0, c_\phi)$. Choose X_i^0, V_i^0 any extremal point of the cloud (X_j^0, V_j^0). Denote $u_i \in S^{2d-1}$ a vector separating the cloud at X_i^0, V_i^0, *i.e.*

$$u_i \cdot (X_j^0 - X_i^0,\ V_j^0 - V_i^0) < 0, \qquad \forall j \neq i.$$

Now define $(x,v) = (X_i^0,\ V_i^0) + c_\phi\, \varepsilon(N)\, u_i$. Since $\phi_\varepsilon > 0$ on $B(0, c_\phi\, \varepsilon)$ then $f_N^0(x,v) > 0$. Denote by T the optimal transference map. $T(x,v)$ has to be one of the $(X_j^0,\ V_j^0)$. Hence by the definition of u_i, $|(x,v) - T(x,v)| \ge c_\phi\, \varepsilon(N)$. Since it is true for any u in a neighborhood of u, it implies that $f_N^0 - \operatorname{esssup}|T - Id| \ge c_\phi \varepsilon(N)$. That last argument may be adapted if we use an optimal transference plane, rather than a map. This means in particular than the plane π defined above is optimal. □

2.2 Statement of the deterministic result without cut-off

A broad version of the following theorem was already stated in the introduction. Now that all the required quantities have been defined, we shall state a precise version. It shows that the particles system may be approximated by the solution of the Vlasov equation with the "blob" distribution f_N^0 as initial conditions, provided that two conditions on the minimal inter-particle distance $d_N(0)$ and the infinite norm of f_N^0 are satisfied.

Theorem 3. *Assume that $d \geq 2$ and that the interaction force F satisfies a (S^α) condition, for some $\alpha < 1$ and let $0 < \gamma < 1$. Assume that the initial condition Z_N^0 are such that for each N, there exists a global solution Z_N to the N particle system (1.1) or (1.2). Assume also that the initial empirical distribution μ_N^0 of the particles and its ε-enlargement f_N^0 satisfy*

i) $d_N^0 := d_N(\mu_N(0)) \geq \varepsilon^{1+r} = N^{-\gamma(1+r)/2d}$ *for some* $r \in (1, r^*)$ *where* $r^* := \frac{d-1}{1+\alpha}$,

ii) $\|f_N^0\|_\infty \leq C_\infty$*, a constant independent of* N,

iii) For some $R_0 > 0$, $\forall N \in \mathbb{N}$, $\operatorname{Supp} \mu_N^0 \subset B_{2d}(0, R_0)$.

Then for any $T > 0$, there exist two constants $C_0(R_0, C_\infty, F, T)$ and $C_1(R_0, C_\infty, F, \gamma, r, T)$ such that for $N \geq e^{C_1 T}$ the following estimate is true

$$\forall t \in [0,\ T], \quad W_\infty(\mu_N(t), f_N(t)) \leq \frac{e^{C_0 t}}{N^{\gamma/2d}}. \tag{2.5}$$

Remark 9. *For the hypothesis about the existence of the solution to the N particles system, we refer to the discussion of that point in the introduction. The bound (2.5) is a inequality of the type $W_\infty(t) \leq W_\infty(0)e^{Ct}$, where the value of $W_\infty(0) = W_\infty(f_N^0, \mu_N^0)$ has been bounded by $N^{-\gamma/2d}$. But that last bound follows from Proposition 2.*

The previous theorem is valid in dimension 2. But unfortunately, its conditions are not generic in that case if the initial conditions are chosen independently. This is why we cannot conclude to propagation of chaos for $d = 2$.

2.3 Statement of the deterministic result with cut-off

As in the case without cut-off, the probabilistic result 2 relies on a deterministic result, much simpler with cut-off since it does not need any control on the minimal inter-particles distance. The precise deterministic result is the following

Theorem 4. *Assume that $d \geq 2$ and that the interaction force F_N satisfies a condition (S_m^α), for some $1 \leq \alpha < d-1$, with a cut-off order satisfying*

$$\frac{\gamma}{2d}\bar{m} < m^* := \frac{\gamma}{2d} \min\left(\frac{d-2}{\alpha-1}, \frac{2d-1}{\alpha}\right).$$

Assume also that for any N, the initial empirical distribution of the particles μ_N^0 and its ε enlargement f_N^0 satisfy

i) $\|f_N^0\|_\infty \leq C_\infty$*, a constant independent of* N,

ii) For some $R_0 > 0$, $\forall N \in \mathbb{N}$, $\operatorname{Supp} \mu_N^0 \subset B_{2d}(0, R_0)$.

We denote by f_N the unique strong solution of the Vlasov equation (1.3)*, given by Proposition 3, which is defined till a time T^*_N.*
Then, the time T^*_N is bounded by below by some time T^* depending only on the constants C_∞ and R_0 appearing above, and for any $T < T^*$, there exist two constants $C_0(R_0, C_\infty, F, T)$ and $C_1(R_0, C_\infty, F, \gamma, r, T)$ such that for $N \geq e^{C_1 T}$ the following estimate is true

$$\forall t \in [0,\ T], \quad W_\infty(\mu_N(t), f_N(t)) \leq \frac{e^{C_0 t}}{N^{\gamma/2d}}. \tag{2.6}$$

Theorem 4 result has also an interest for numerical simulation because one obvious way to fulfill the hypothesis on the infinite norm of f^0_N is is to put particles initially on a mesh (with a grid length of $N^{-1/2d}$ in $\mathbb{R}^{2d}$). In that case, the result is even valid with $\gamma = 1$.

Before going to the proof of our results on the mean field limit, we shall give some results about the existence and uniqueness of strong solution of the Vlasov equation (1.3).

2.4 Uniqueness, Stability of solutions to the Vlasov equation 1.3.

The already known results about the well-posedness (in the strong sense) of the Vlasov equation that we are considering are gathered in the following proposition.

Proposition 3. *For any dimension d, and any $\alpha \leq d-1$, and any bounded initial there exists a unique local (in time) strong solution to the Vlasov equation* (1.3) *that remains bounded and compactly supported. In general, the maximal time of existence T^* of this solution maybe finite, but in the two particular case below we have $T^* = +\infty$*

- $\alpha < 1$ *(and any d),*
- $d \leq 3$*, and* $\alpha \leq d - 1$.

In the other cases, the maximal time of existence of the strong solution may be bounded by below by some constant depending only on the infinite norm and the size of the support of the initial condition. And in any case, the size of the support at any time t may also be bounded by a constant depending on the same quantities.

The local existence part in Proposition 3 is a consequence of the following Lemma which is proved in the Appendix and the following Proposition 4

Lemma 1. *Let $f \in L^\infty([0,\ T],\ \mathbb{R}^{2d})$ with compact support be a solution to* (1.3) *in the sense of distribution with an F satisfying* (1.8) *with $\alpha \leq d - 1$. Then if we denote by $R(t)$ and $K(t)$ the size of the supports of f in space and velocity, they satisfy for a numerical constant C*

$$R(t) \leq R(0) + \int_0^t K(s)\, ds,$$
$$K(t) \leq K(0) + C\, \|f(0)\|_{L^\infty}^{\alpha/d}\, \|f(0)\|_{L^1}^{1-\alpha/d} \int_0^t K(s)^\alpha\, ds.$$

The local uniqueness part in Proposition 3 is a consequence of the following stability estimate proved in [Loe06] for $\alpha = d-1$. Its proof may be adapted to less singular case. For instance, the adaptation is done in [Hau09] in the Vortex case.

Proposition 4 (From Loeper). *If f_1 and f_2 are two solutions of Vlasov Poisson equations with different interaction forces F_1 and F_2 both satisfying a (S^α)-condition, with $\alpha < d-1$, then*

$$\frac{d}{dt} W_1(f_1(t), f_2(t)) \leq C \max(\|\rho_1\|_\infty, \|\rho_2\|_\infty) \left[W_1(f_1(t), f_2(t)) + \|F_1 - F_2\|_1 \right]$$

In the case $\alpha = d-1$, Loeper only obtain in [Loe06] a "log-Lip" bound and not a linear one, but it still implies the stability.

Finally, the global character of the solution in Proposition 3 is a consequence

- of the lemma 1 if $\alpha < 1$, since in that case, the estimates obtained in that lemma show that $R(t)$ and $K(t)$ cannot explode,
- a much more delicate issue in the case $d \leq 3$, and $\alpha = d-1$, finally solved in [LP91], [Sch91] and [Pfa92]. Their proof may also be extended to the less singular case $\alpha < d-1$.

3 From deterministic results (Theorem 3 and 4) to propagation of chaos.

The assumptions made in Theorem 3 may seem a little bit strange, but they are in some sense generic, when the initial positions and speed are chosen with the law $(f^0)^{\otimes N}$. Therefore, to prove Theorem 1 form Theorem 3, we need to

- Estimate the probability that empirical measure chosen with the law $(f^0)^{\otimes N}$, do not satisfy the conditions $i)$ and $ii)$ of the deterministic theorem 3, and are far away from f^0 in W_1 distance (the last conditions is important for the previous point on the distance between f and f_N).
- Use the bound obtained on the initial distribution f_N^0 to control the W_1 distance between $f(t)$ and $f_N(t)$, which are two solutions of the Vlasov equation 1.3.

For the first point, we will use known results detailed in the next two sections. The second point is then a simple application of Proposition 4. After that, an application of the deterministic Theorem 3 with a good choice of the parameters γ and r will allow us to conclude the proof.

3.1 Estimates in probability on the initial distribution.

Deviations on the infinite norm of the smoothed empirical distribution f_N. The precise result we need is given by the Proposition 8 in the Appendix. It tells us that if the approximating kernel is $\phi = \mathbf{1}_{[-\frac{1}{2},\frac{1}{2}]^{2d}}$, then

$$\mathbb{P}\left(\|f_N^0\|_\infty \geq 2^{1+2d} \|f^0\|_\infty \right) \leq C_2 N^\gamma e^{-C_1 N^{1-\gamma}}.$$

with $C_2 = (2R^0 + 2)^{2d}$, R^0 the size of the support of f, and $C_1 = (2\ln 2 - 1)\, 2^{2d} \|f\|_\infty$.

We would like to mention that we were first aware of the possibility of getting such estimates in a paper of Bolley, Guillin and Villani [BGV07], where the authors obtain quantitative concentration inequality for $\|f^N - f\|_\infty$ in infinite norm under the additional assumption that f^0 and ϕ are Lipschitz. Unfortunately, they cannot be used in our setting because they would require a too large smoothing parameter. Gao obtain in [Gao03] very precise large deviations estimates for $\|f^N - f\|_\infty$, but as usual with large deviations estimates, they are only asymptotic and therefore less convenient for our problem. Finally, we choose to prove a more simple estimate that is well adapted to our problem.

Deviations for the minimal inter-particle distance. It may be proved with simple arguments that the scale η_m is almost surely larger than $N^{-1/d}$ when $f^0 \in L^\infty$. A precise result is stated in the Proposition below, proved in [Hau09]

Proposition 5. *There exists a constant c_{2d} depending only on the dimension such that if $f^0 \in L^\infty(\mathbb{R}^{2d})$, then*

$$\mathbb{P}\left(d_N(Z) \geq \frac{l}{N^{1/d}}\right) \geq e^{-c_{2d}\|f^0\|_\infty l^d}.$$

We point out that this is not a large deviation result : the inequalities are in the wrong direction. This is quite natural because d_N is not a average quantity, but an infimum. It is that condition that prevents us from obtaining a "large deviation" type result in Theorem 1, contrarily to the cut-off case of Theorem 2. In fact, the only bound it provides on the "bad" set is

$$\mathbb{P}\left(d_N(Z) \leq \frac{l}{N^{1/d}}\right) \leq 1 - e^{-c_{2d}\|f^0\|_\infty l^d} \leq c_{2d}\|f^0\|_\infty l^d.$$

With the notation of Theorem 3 it comes that if $s = \gamma\frac{1+r}{2} - 1 > 0$ then

$$\mathbb{P}\left(d_N(Z) \leq \varepsilon^{1+r}\right) = \mathbb{P}\left(d_N(Z) \leq \frac{N^{-s/d}}{N^{1/d}}\right) \leq c_{2d}\|f^0\|_\infty N^{-s}. \tag{3.1}$$

Deviations for the W_1 MKW distance. It is more or less classical that if the Z_i are independent random variable with identical law f, the empirical measure μ_N^Z goes in probability to f. This theorem is known as the empirical law of large number or Glivenko-Cantelli theorem and is due in this form to Varadarajan [Var58]. But, the convergence may be quantified in Wasserstein distance, and recently upper bound on the large deviations of $W_1(\mu_N^Z, f)$ where obtained by Bolley, Guillin and Villani [BGV07] and Boissard [Boi11]. But the first one concern only very large deviations, so we will quote only the last result which is more interesting for us.

Proposition 6 (Boissard [Boi11] Annexe A, Proposition 1.2)**.** *Assume that f is a non negative measure compactly supported on $B_{2d}(0,R) \subset \mathbb{R}^{2d}$. If $d \geq 2$, and the $Z = (Z_1, \ldots, Z_N)$ are chosen according to the law $(f^0)^{\otimes N}$, then there is an explicit constant $C_1 = 2^{-(2d+1)}R^{-2d}$, such that the associated empirical measures μ_N^Z satisfy*

$$\mathbb{P}\Big(W_1(\mu_N^Z, f) \geq \mathbb{E}[W_1(\mu_N^Z, f)] + L\Big) \leq e^{-C_1 N L^2}.$$

Since it is already known (see [Boi11] or [DY95] and references therein) that there exists a numerical constant $C_2(d)$ such that

$$\mathbb{E}[W_1(\mu_N^Z, f)] \leq C_2 \frac{R}{N^{1/2d}},$$

the previous result with $L = C_2 \frac{R}{N^{1/(2d)}}$ *implies that for* $C_3(R,d) := C_1(R) C_2(d)^2 R^2$,

$$\mathbb{P}\left(W_1(\mu_N^Z, f) \geq 2\frac{C_2 R}{N^{1/2d}}\right) \leq e^{-C_3 N^{1-1/d}}. \tag{3.2}$$

3.2 Conclusion

Now take the assumptions of Theorem 1 : F satisfies a (S^α) condition for $\alpha < 1$ and $f^0 \in L^\infty$ with support in some ball $B_{2d}(0, R_0)$ in dimension $d \geq 3$. We chose

$$\gamma \in \left(\gamma^* = \frac{2+2\alpha}{d+\alpha}, 1\right), \text{ and } \quad r \in \left(\frac{2}{\gamma} - 1, r^* = \frac{d-1}{1+\alpha}\right),$$

the condition on γ ensuring that the second interval is non empty. We also define

$$s := \gamma\,\frac{1+r}{2} - 1 > 0, \quad \lambda = 1 - \max\left(\gamma, \frac{1}{d}\right)$$

Denote by ω_1, ω_2 the sets of initial conditions s.t. respectively (i), and (ii) of Theorem 3 hold and ω_3 s.t. $W_1(\mu_N, f^0) \leq \frac{1}{N^{\gamma/(2d)}}$. Precisely

$$\omega_1 := \{Z^0 \text{ s.t. } d_N(Z^0) \geq \varepsilon^{1+r}\}, \quad \omega_2 := \{Z^0 \text{ s.t. } \|f_N^0\|_\infty \leq 2^{1+2d}\|f^0\|_\infty\}$$
$$\omega_3 := \{Z^0 \text{ s.t. } W_1(\mu_N^0, f^0) \leq \varepsilon\}$$

By the results stated in the previous section, one knows that for $N \geq (2C_2R)^{2d/(1-\gamma)}$

$$\mathbb{P}(\omega_1^c) \leq C\,N^{-s}, \quad \mathbb{P}(\omega_2^c) \leq CN^\gamma e^{-CN^{1-\gamma}}, \quad \mathbb{P}(\omega_3^c) \leq \; e^{-CN^{1-\frac{1}{d}}}. \tag{3.3}$$

Denote $\omega = \omega_1 \cap \omega_2 \cap \omega_3$. Hence $|\omega^c| \leq |\omega_1^c| + |\omega_2^c| + |\omega_3^c|$ and for N large enough

$$\mathbb{P}(\omega^c) \leq C\,N^{-s} + C\,N^\gamma\,e^{-C\,N^{1-\gamma}} + e^{-CN^{1-\frac{1}{d}}} \leq C\,N^{-s}, \tag{3.4}$$

and checking carefully the dependence, we can see that the constant C depends only on $d, R, \|f^0\|_\infty, \gamma$.
Since we known that global solutions to the N particles system (1.1) or (1.2) exist for almost all initial conditions (see the discussion on this point in the introduction), one may apply Theorem 3 to a.e. initial condition in ω and get on $[0,\ T]$

$$W_1(f_N, \mu_N) \leq W_\infty(f_N, \mu_N) \leq \frac{e^{C_0 t}}{N^{\gamma/(2d)}}.$$

Now apply the stability around solution of Vlasov equation given by Proposition4 and get

$$W_1(f, f_N) \leq W_1(f^0, f_N^0)\,e^{C_0 t} \leq \frac{2}{N^{\gamma/(2d)}}\,e^{C_0 t}.$$

The factor 2 comes from the fact that $W_1(f^0, f_N^0) \leq W_1(f^0, \mu_N^0) + W_1(\mu_N^0, f_N^0)$. We conclude that

$$W_1(f, \mu_N) \leq \frac{3}{N^{\gamma/2d}} e^{C_0 t},$$

which proves that

$$\mathbb{P}(\omega) \leq \mathbb{P}\left(\forall t \in [0,T],\ W_1(f, f_N) \leq \frac{3e^{C_0 t}}{N^{\gamma/(2d)}}\right).$$

The bound 3.4 then gives Theorem 1.

3.3 From Theorem 4 to Theorem 2

In the cut-off case, one can derive Theorem 2 from Theorem 4 in the same manner. As we do not use the minimal distance in that case, the proof is simpler in the case $\alpha < d-1$ and we get a stronger convergence result.
We only have to consider $\omega = \omega_2 \cap \omega_3$, where ω_2 and ω_3 are defined according to (3.3). Then, the bound (3.4) is replaced for N larger than an explicit constant by

$$\mathbb{P}(\omega^c) \le C\, N^{\gamma}\, e^{-C\, N^{1-\gamma}} + e^{-CN^{1-\frac{1}{d}}} \le C\, N^{\gamma}\, e^{-C\, N^{-\lambda}} \tag{3.5}$$

Next, for any $Z^0 \in \omega$, we can apply the deterministic theorem 4 to get a time T^* that bound by below the time of existence of the strong solution f_N of the Vlasov equation (1.3) and obtain the stability estimate

$$W_1(f_N, \mu_N) \le W_\infty(f_N, \mu_N) \le \frac{e^{C_0 t}}{N^{\gamma/(2d)}}.$$

Moreover, Proposition 3 implies that the strong solution f with initial data f^0 is also defined at least on $[0, T^*)$. And from the condition (S^α_m) restated in (2.3) in term of ε, we get that

$$\|F - F_N\|_1 \le \varepsilon^{\bar{m}(d-\alpha)} \le \varepsilon,$$

since $\bar{m} \ge 1$ and $d - \alpha \ge 1$. So we can apply the stability estimate given by Proposition 4 with $F_1 = F$ and $F_2 = F_N$ and get that

$$W_1(f, f_N) \le [W_1(f^0, f^0_N) + \varepsilon]\, e^{C_0 t} \le \frac{3}{N^{\gamma/(2d)}}\, e^{C_0 t}.$$

From there, we obtain as before that for N large enough that

$$\mathbb{P}\left(\exists\, t \in [0,T],\ W_1(\mu_N(t), f(t)) \ge \frac{4 e^{C_0 t}}{N^{\gamma/(2d)}} \right) \le C N^{\gamma} e^{-C N^{\lambda}}.$$

4 Proof of Theorem 3 and 4

4.1 Definition of the transport

We try now to compare the the dynamics of μ_N and f_N, two distributions which have a compact support. For that, we choose an optimal transport $T^0 (= T^0_N)$ from f^0_N to μ^0_N for the infinite MKW distance The existence of such a transport is ensured by [CDPJ08]. T^0 is defined on the support of f^0_N, which is included in $B_{2d}(0, R^0)$ (the size of the support), and Proposition 2 implies that $W_\infty(f^0_N, \mu^0_N) \le \varepsilon$.

Thanks to the hypothesis of both theorems, the strong solution f_N to the Vlasov equation is well defined till a time T^*, infinite in the case of Theorem 3, that depends only on C_∞ and R_0 and not on N. Since we are dealing with strong solutions, there exists a well-defined underlying flow, that we will denote by $Z^f = (X^f, V^f)$: $Z^f(t,s,z)$ being the position-velocity at time t of a particle with position-velocity z at time s.
Moreover, by the hypothesis of Theorem 3 or because we use a cut-off in Theorem 4, the dynamic of the N particles is well defined, and we can also write in that case a flow $Z^\mu =$

(X^μ, V^μ), which is well defined at least at the position and velocity of the particles we are considering. A simple way to get a transport of $f_N(t)$ on $\mu_N(t)$ is to transport along the flows the map T^0, i.e. to define

$$T^t = Z^\mu(t,0) \circ T^0 \circ Z^f(0,t), \qquad \text{and} \quad T^t = (T^t_x, T^t_v)$$

We use the following notation, for a test-"particle" of the continuous system at the position $z_t = (x_t, v_t)$ at time t, $z_s = (x_s, v_s)$ will be its position at time s for $s \in [t-\tau, t]$. Precisely

$$z_s = Z^f(s,t,z_t)$$

Since f_N is the solution of a transport equation, we have $f_N(t, z_t) = f_N(s, z_s)$. And since the vector-field of that transport equation is divergence free, the flow Z^f is measure-preserving in the sense that for all smooth test function Φ

$$\int \Phi(z)\, f_N(s,z)\, dz = \int \Phi(Z^f(s,t,z))\, f_N(t,z)\, dz = \int \Phi(z_s)\, f_N(t,z_t)\, dz_t.$$

Finally let us remark that the f_N are solutions to the (continuous) Vlasov equations with an initial L^∞ norm and support that are uniformly bounded in N. Therefore the Proposition 3, and in particular the last assertion in it imply that this remains true uniformly in N for any finite time $T < T^*$. In particular the uniform bound on the support $R(T)$ implies since $/alpha < d-1$ the existence of a constant C independent of N such that for any $t \in [0,\ T]$

$$\begin{aligned}
&\|f_N(t,.,.)\|_\infty \le C, \qquad \|f_N(t,.,.)\|_{L^1} = 1,\\
&\operatorname{supp} f_N(t,.,.) \in B_{2d}(0,C),\\
&|E_{f_N}|_\infty(t) := \|E_{f_N}(t,\cdot)\|_\infty \le \sup_x \int |F(x-y)|\, f_N(t,y,w)\, dy\, dw \le C \qquad (4.1)\\
&|\nabla E_{f_N}(t,x)| \le \int |\nabla F(x-y)|\, f_N(t,y,w)\, dy\, dw \le C.
\end{aligned}$$

In what follows, the final time T is fixed and independent of N. For simplicity, C will denote a generic universal constant, which may actually depend on T, the size of the initial support, the infinite norms of the f_N... But those constants are always independent of N as in (4.1).

4.2 The quantities to control

We will not be able to control the infinite norm of the field (and its derivative) created by the empirical distribution, but only a small temporal average of this norm. For this, we introduce in the case without cut-off a small time step $\tau = \varepsilon^{r'}$ for some $r' > r$ and close to r (the precise condition will appear later).
In the case with cut-off where r and r' are useless, the time step will by $\tau = \varepsilon$.

Before going on, we define some important notations.

- **The MKW infinite distance between $\mu_N(t)$ and $f(t)$.**

  We wish to bound the infinite Wasserstein distance $W_\infty(\mu_N(t), f_N(t))$ between the empirical measure associated to the N particle system (1.2), and the solution of the

Vlasov equation (1.3) with blobs as initial condition. But for convenience we will work instead with the quantity

$$W_\infty(t) := \sup_{s\le t} \sup_{z_s\in \operatorname{supp} f_N(s)} |T^s(z_s) - (z_s)|, \tag{4.2}$$

where the sup on z_s should be understood precisely as a essential supremum with respect to the measure $f_N(s)$. This is not exactly the infinite Wasserstein distance between ${}_m u_N(t)$ and $f_N(t)$ (or its supremum in times smaller than $s \le t$). But, since for all s, the transport map T^s send the measure f_N onto μ_N by construction, we always have

$$W_\infty(\mu_N(t), f_N(t)) \le \sup_{s\le t} W_\infty(\mu_N(t), f_N(t)) \le W_\infty(t).$$

So that a control on $W_\infty(t)$ implies a control on $W_\infty(\mu_N(t), f_N(t))$. It is in fact a little stronger, since it means that rearrangement in the transport are not necessary to keep the MKW distance bounded. We introduce the supremum in time for technical reasons as it will be simpler to deal with a non decreasing quantity in the sequel.

- **The support of** μ_N

  We shall also need a uniform control on the support in position and velocity of the empirical distributions :

$$R^N(t) := \sup_{s\le t} \max_i |(X_i(t),\ V_i(t))|. \tag{4.3}$$

- **The infinite norm** $|\nabla^N E|_\infty$ **of the time averaged discrete derivative of the force field.**

  We also define a version of the infinite norm of its averaged derivative

$$|\nabla^N E|_\infty(t) := \sup_{i\ne j} \frac{1}{\tau}\int_{t-\tau}^{t} \frac{|E_N(X_i(s)) - E_N(X_j(s))|\, ds}{|X_i(s) - X_j(s)| + \varepsilon^{(1+r')}}\, ds. \tag{4.4}$$

  For $|\nabla^N E|_\infty$, we use the convention that when the interval of integration contains 0 (for $t < \tau$), the integrand is null on the right side for negative times. Remark that the control on $|\nabla^N E|_\infty$ is useless in the cut-off case.

- **The minimal distance in** $\mathbb{R}^{2d}$**,** d_N

  which has already be defined by the equation (2.1) in the Section 2.

- **Two useful integrals** $I_\alpha(t, \bar z_t, z_t)$ **and** $J_{\alpha+1}(t, \bar z_t, z_t)$

  Finally for any two test trajectories z_t and $\bar z_t$, we define

$$I_\alpha(t, \bar z_t, z_t) := \frac{1}{\tau}\int_{t-\tau}^{t} |F(T^s_x(\bar z_s) - T^s_x(z_s)) - F(\bar x_s - x_s)|\, ds, \tag{4.5}$$

which controls the difference of the two force fields at two point related by the "optimal" transport. We recall that we use here the convention $F(0) = 0$, in order to avoid self-interaction. It is important here since we have $T^s(z_s) = T^s(\bar z_s)$ for all $s \in [t-\tau, t]$, for a set of $(z_s, \bar{z}_s)$ of positive measure (those who are associated to the same particle (X_i, V_i)). Defining a second kernel as

$$K_\varepsilon := \min\left(\frac{1}{|x|^{1+\alpha}}, \frac{1}{\varepsilon^{1+r'}|x|^\alpha}\right) \quad \text{for } x \neq 0, \quad \text{and } K_\varepsilon(0) = 0, \tag{4.6}$$

we introduce a second useful quantity

$$\begin{aligned} J_{\alpha+1}(t, \bar{z}_t, z_t) &:= \frac{1}{\tau}\int_{t-\tau}^{t} K_\varepsilon(|T_x^s(\bar{z}_s) - T_x^s(z_s)|\, ds \\ &= \frac{1}{\tau}\int_{t-\tau}^{t} K_\varepsilon(|X_i(s) - X_j(s)|\, ds, \end{aligned} \tag{4.7}$$

if i and j is the indices such that $Z_i(t) = T^t(\bar{z}_t)$ and $Z_j(t) = T^t(z_t)$. $J_{\alpha+1}$ will be useful to control the discrete derivative of the field $|\nabla^N E|_\infty(t)$, and is thus useless in the cut-off case.

All previous quantities are relatively easily bounded by I_α and $J_{\alpha+1}$. Those last two will not be bounded by direct calculation on the discrete system, but we will compare them to similar ones for the continuous system, paying for that in terms of the distance between $\mu_N(t)$ and $f(t)$. That strategy is interesting because the integrals are easier to manipulate than the discrete sums.

Remark 10. *Before stating the next proposition, let us mentioned that we also define for $t < 0$ $W(t) = W(0)$ and $d_N(t) = d_N(0)$. This is just a helpful convention. With it the estimate of the next proposition are valid for any $t \geq 0$, and this will be very convenient in the conclusion of the proof of our main theorem. Remark also that $|E^N|_\infty(0) = 0$ and that $|\nabla^N E|_\infty(0) = 0$.*

We summarize the first easy bounds in the following

Proposition 7. *Under the assumptions of Theorem 3, one has for some constant C uniform in N, that for all $t \geq 0$*

$$\begin{aligned} &(i) \quad R_N(t) \leq W_\infty(t) + R(t) \leq W_\infty(t) + C, \\ &(ii) \quad W_\infty(t) \leq W_\infty(t-\tau) + \tau W_\infty(t) + C\,\tau \sup_{\bar{z}_t} \int_{|z_t| \leq R(t)} I_\alpha(t, \bar{z}_t, z_t)\, dz_t, \\ &(iii) \quad |\nabla^N E|_\infty(t) \leq C \sup_{\bar{z}_t} \int_{|z_t| \leq R(t)} J_{\alpha+1}(t, \bar{z}_t, z_t)\, dz_t, \\ &(iv) \quad d_N(t) + \varepsilon^{1+r'} \geq [d_N(t-\tau) + \varepsilon^{1+r'}] e^{-\tau(1+|\nabla^N E|_\infty(t))}. \end{aligned}$$

Note that the control on $R_N(t)$ is simple enough that it will actually be used implicitly in the rest many times, and that the *iv*) is a simple consequence of the *iii*). In fact, in that proposition the crucial estimates are the *ii*) and *iii*).
Remark also that in the case of very singular interaction force ($\alpha \geq 1$) with cut-off - in short (S_m^α) conditions - the control on minimal distance d_N and therefore the control on $|\nabla^N E|_\infty$ are useless, so that the only interesting inequality is the second one.

4.3 Proof of Prop. 7

Step 1. Let us start with (i). Simply write

$$R^N(t) = \sup_{s\le t}\sup_{z_s\in \operatorname{supp} f_N(s,\cdot)} |T^s(z_s)| \le \sup_{s\le t}\sup_{z_s\in \operatorname{supp} f_N(s,\cdot)} |T^s(z_s)-z_t| + \sup_{s\le t}\sup_{z_s\in \operatorname{supp} f_N(s,\cdot)} |z_s|,$$

So indeed by the bound (4.1) and the definition (4.2) of W_∞

$$R^N(t) \le W_\infty(t) + C.$$

Step 2. For (ii), for any time $t'\in[t-\tau,t]$ we have

$$\begin{aligned} |T_x^{t'}(\bar z_{t'}) - \bar x_{t'}| &\le |T_x^{t-\tau}(\bar z_{t-\tau}) - \bar x_{t-\tau}| + \int_{t-\tau}^{t'} |T_v^s(\bar z_s) - \bar v_s|\, ds \\ &\le |T_x^{t-\tau}(\bar z_{t-\tau}) - \bar x_{t-\tau}| + \tau W_\infty(t), \end{aligned} \tag{4.8}$$

and for the speeds

$$\begin{aligned} |T_v^{t'}(\bar z_{t'}) - \bar v_{t'}| \le{}& |T_v^{t-\tau}(\bar z_{t-\tau}) - \bar v_{t-\tau}| \\ &+ \int_{t-\tau}^{t'}\int |F(T_x^s(\bar z_s) - T^s(z_s)) - F(\bar x_s - x_s)| f_N(s,z_s)\, dz_s ds \\ \le{}& |T_v^{t-\tau}(\bar z_{t-\tau}) - \bar v_{t-\tau}| + \int_{t-\tau}^{t}\int |F(T_x^s(\bar z_s) - T^s(z_s)) - F(\bar x_s - x_s)| f_N(t,z_t)\, dz_t ds. \end{aligned}$$

where we used the fact that the change of variable $z_t\mapsto z_s$ preserves the measure. Since f_N is uniformly bounded in L^∞ and compactly supported in $B(0,R(t))$, one gets by the definition of I_α

$$|T_v^{t'}(\bar z_{t'}) - \bar v_{t'}| \le |T_v^{t-\tau}(\bar z_{t-\tau}) - \bar v_{t-\tau}| + C\tau \sup_{\bar z_t}\int_{|z_t|\le R(t)} I_\alpha(t,\bar z_t,z_t)\, dz_t. \tag{4.9}$$

Summing the two estimates (4.8) and (4.9), we get for the euclidian distance on $\mathbb{R}^{2d}$

$$|T^{t'}(\bar z_{t'}) - \bar z_{t'}| \le |T^{t-\tau}(\bar z_{t-\tau}) - \bar z_{t-\tau}| + C\tau\left(W_\infty(t) + \sup_{\bar z_t}\int_{|z_t|\le R(t)} I_\alpha(t,\bar z_t,z_t)\, dz_t\right).$$

Taking the supremum over all $\bar z_{t'}$ in the support of $f_N(t')$, and then the supremum over all $t'\in[t-\tau,t]$ we get

$$W_\infty(t) \le W_\infty(t-\tau) + \tau W_\infty(t) + C\tau \sup_{\bar z_t}\int_{|z_t|\le R(t)} I_\alpha(t,\bar z_t,z_t)\, dz_t$$

which is exactly (ii).

Step 3. Concerning $|\nabla^N E|_\infty(t)$ in (iii), noting that

$$\begin{aligned} \int_{t-\tau}^{t} \frac{|E_N(X_i(s)) - E_N(X_j(s))|}{|X_i(s)-X_j(s)| + \varepsilon^{1+r'}} ={}& \frac{1}{N}\sum_{k\neq i,j}\int_{t-\tau}^{t} \frac{|F(X_i(s)-X_k(s)) - F(X_j(s)-X_k(s))|}{|X_i(s)-X_j(s)| + \varepsilon^{1+r'}}\, ds. \\ &+ \frac{1}{N}\int_{t-\tau}^{t} \frac{|F(X_i(s)-X_j(s)) - F(X_j(s)-X_i(s))|}{|X_i(s)-X_j(s)| + \varepsilon^{1+r'}}\, ds \end{aligned}$$

By the assumption (1.8), one has that

$$|F(x)-F(y)| \le C\left(\frac{1}{|x|^{\alpha+1}}+\frac{1}{|y|^{\alpha+1}}\right)|x-y|.$$

So

$$\frac{|F(X_i(s)-X_k(s))-F(X_j(s)-X_k(s))|}{|X_i(s)-X_j(s)|+\varepsilon^{1+r'}} \le \frac{C}{|X_i(s)-X_k(s)|^{1+\alpha}}+\frac{C}{|X_j(s)-X_k(s)|^{1+\alpha}},$$

and that bound is also true for the remaining term where $k=i$ or j, if we delete the undefined term in the sum. One also obviously has, still by (1.8)

$$\begin{aligned}\frac{|F(X_i(s)-X_k(s))-F(X_j(s)-X_k(s))|}{|X_i(s)-X_j(s)|+\varepsilon^{1+r'}} \le& \frac{C}{\varepsilon^{1+r'}|X_i(s)-X_k(s)|^{\alpha}}\\ &+\frac{C}{\varepsilon^{1+r'}|X_j(s)-X_k(s)|^{\alpha}}.\end{aligned}$$

Therefore by the definition of K_ε

$$\frac{|F(X_i(s)-X_k(s))-F(X_j(s)-X_k(s))|}{|X_i(s)-X_j(s)|+\varepsilon^{1+r'}} \le C\big[K_\varepsilon(X_i(s)-X_k(s))+K_\varepsilon(X_j(s)-X_k(s))\big].$$

Summing up, this implies that

$$\begin{aligned}|\nabla^N E|_\infty(t) \le C\max_{i\neq j}\Big(&\frac{1}{\tau}\int_{t-\tau}^t \frac{1}{N}\sum_{k\neq i} K_\varepsilon(X_i(s)-X_k(s))\,ds\\ &+\frac{1}{\tau}\int_{t-\tau}^t \frac{1}{N}\sum_{k\neq j} K_\varepsilon(X_j(s)-X_k(s))ds\Big).\end{aligned}$$

Transforming the sum into integral thank to the transport, we get exactly the bound (*iii*) involving $J_{\alpha+1}$.

Step 4. Finally for $d_N(t)$, consider any $i\neq j$, differentiating the euclidian distance $|Z_i-Z_j|$, we get

$$\frac{d}{ds}|(X_i(s)-X_j(s),\ V_i(s)-V_j(s))| \ge -|V_i(s)-V_j(s)|-|E_N(X_i(s))-E_N(X_j(s))|.$$

Simply write

$$|E_N(X_i(s))-E_N(X_j(s))| \le \frac{|E_N(X_i(s))-E_N(X_j(s))|}{|X_i(s)-X_j(s)|+\varepsilon^{1+r'}}\,(|X_i(s)-X_j(s)|+\varepsilon^{1+r'})$$

to obtain that

$$\begin{aligned}\frac{d}{ds}|(X_i(s)-X_j(s),\ V_i(s)-V_j(s))| \ge& -\left(1+\frac{|E_N(X_i(s))-E_N(X_j(s))|}{|X_i(s)-X_j(s)|+\varepsilon^{1+r'}}\right)\\ &(|(X_i(s)-X_j(s),\ V_i(s)-V_j(s))|+\varepsilon^{1+r'}).\end{aligned}$$

Integrating this inequality and taking the minimum, we get

$$\begin{aligned}d_N(t)+\varepsilon^{1+r'} &\ge (d_N(t-\tau)+\varepsilon^{1+r'})\inf_{i\neq j}\exp\left(-\tau-\int_{t-\tau}^t \frac{|E_N(X_i(s))-E_N(X_j(s))|}{|X_i(s)-X_j(s)|+\varepsilon^{1+r'}}\,ds\right)\\ &\ge [d_N(t-\tau)+\varepsilon^{1+r'}]\exp^{-\tau(1+|\nabla^N E|_\infty(t))}.\end{aligned}$$

4.4 The bounds for I_α and $J_{\alpha+1}$

To close the the system of inequalities in Proposition 7, it remains to bound the two integrals involving I_α and J_α. It is done with the following lemmas

Lemma 2. *Assume that F satisfies an (S^α)-condition with $\alpha < 1$, and that τ is small enough such that for some constant C (precise in the proof)*

$$C\,\tau\,(1 + |\nabla^N E|_\infty(t))\,(W_\infty(t) + \tau) \leq d_N(t). \tag{4.10}$$

Then one has the following bounds, uniform in $\bar{z}_t$

$$\int_{|z_t|\leq R(t)} I_\alpha(t, \bar{z}_t, z_t)\,dz_t \leq C\left[W_\infty(t) + (W_\infty(t)+\tau)^d \tau^{-\alpha} + (W_\infty(t)+\tau)^{2d}(d_N(t))^{-\alpha}\tau^{-\alpha}\right].$$

$$\begin{aligned}\int_{|z_t\leq R(t)} J_{\alpha+1}(t, \bar{z}_t, z_t)\,dz_t \leq C\,\Big(1 + (W_\infty(t)+\tau)^d \varepsilon^{-(1+r')}\,\tau^{-\alpha} \\ + (W_\infty(t)+\tau)^{2d} \varepsilon^{-(1+r')}\,\tau^{-\alpha}\,(d_N(t))^{-\alpha}\Big).\end{aligned}$$

In the cut-off case where the interaction force satisfy a (S_m^α) condition (we recall that it means that the cut-off is of size $N^{-m} = \varepsilon^{\bar{m}}$ with $\bar{m} = \frac{2d}{\gamma} m$), we only need to bound the integral of I_α, with the result

Lemma 3. *Assume that $1 \leq \alpha < d-1$, and that F satisfies a (S_m^α) condition. Then one as the following bound, uniform in $\bar{z}_t$*

$$\int_{|z_t|\leq R(t)} I_\alpha(t, \bar{z}_t, z_t)\,dz_t \leq C\left(W_\infty(t) + (W_\infty(t)+\tau)^d \tau^{-1} \varepsilon^{\bar{m}(1-\alpha)} + (W_\infty(t)+\tau)^{2d} \varepsilon^{-\bar{m}\alpha}\right). \tag{4.11}$$

with the convention[1] *(if $\alpha = 1$) that $\varepsilon^0 = 1 + |\ln\varepsilon|$.*

The proofs with or without cut-off follow the same line and we will prove the above lemmas at the same time. We begin by an explanation of the sketch of the proof, and then perform the technical calculation.

4.4.1 Rough sketch of the proof

The point $\bar{z}_t = (\bar{x}_t, \bar{v}_t)$ is considered fixed through all this subsection (as the integration is carried over $z_t = (x_t, v_t)$). Accordingly we decompose the integration in z_t over several domains. First

$$A_t = \{z_t \mid |\bar{x}_t - x_t| \geq 4W_\infty(t) + 2\tau(|\bar{v}_t - v_t| + \tau|E|_\infty(t))\}.$$

This set consist of points z_t such that x_s and $T_x^s(z_s)$ are sufficiently far away from $\bar{x}_s$ on the whole interval $[t-\tau, t]$, so that they will not see the singularity of the force. The bound over this domain will be obtained using traditional estimates for convolutions.

[1] That convention may be justified by the fact that it implies a very simple algebra $(x^{1-\alpha})' \approx x^{-\alpha}$ even if $\alpha = 1$. It allows us to give an unique formula rather than three different cases.

One part of the integral can be estimated easily on A_t^c (the part corresponding to the flow of the regular solution f_N to the Vlasov equation). For the other part it is necessary to decompose further. The next domain is

$$B_t = A_t^c \bigcap \{z_t \mid |\bar{v}_t - v_t| \geq 4\, W_\infty(t) + 4\,\tau |E|_\infty|(t)\}.$$

This contains all particles z_t that are close to $\bar{z}_t$ in position (i.e. x_t close to $\bar{x}_t$), but with enough relative velocity not to interact too much. The small average in time will be useful in that part, as the two particles remains close only a small amount of time.
The last part is of course the remainder

$$C_t = (A_t \cup B_t)^c.$$

This is a small set, but where the particles remains close together a relatively long time. Here, we are forced to deal with the corresponding term at the discrete level of the particles. This is the only term which requires the minimal distance in $\mathbb{R}^{2d}$; and the only term for which we need a time step τ small enough as per the assumption in Lemma 2.

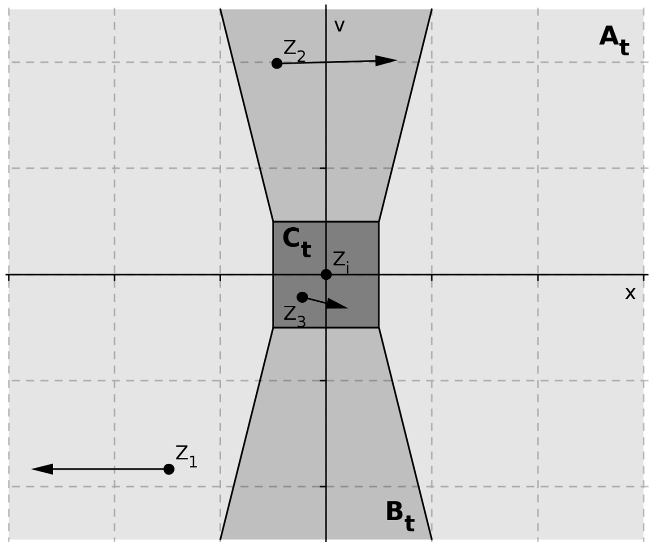

Figure 1: The partition of $\mathbb{R}^{2d}$.

4.4.2 Step 1: Estimate over A_t

If $z_t \in A_t$, we have for $s \in [t - \tau, t]$

$$|\bar{x}_s - x_s| \;\geq\; |\bar{x}_t - x_t| - (t-s)|\bar{v}_t - v_t| - (t-s)^2 |E|_\infty(t) \geq \frac{|\bar{x}_t - x_t|}{2} \tag{4.12}$$

$$|T_x^s(\bar{z}_s) - T_x^s(z_s)| \;\geq\; |\bar{x}_s - x_s| - 2W_\infty(s) \geq \frac{|\bar{x}_t - x_t|}{2}. \tag{4.13}$$

For I_α, we use the direct bound for $z_t \in A_t$

$$\begin{aligned} |F(T_x^s(\bar{z}_s) - T_x^s(z_s)) - F(\bar{x}_s - x_s)| \;&\leq\; \frac{C}{|\bar{x}_t - x_t|^{1+\alpha}} (|T_x^s(\bar{z}_s) - \bar{x}_s| + |T_x^s(z_s) - x_s|) \\ &\leq\; \frac{C}{|\bar{x}_t - x_t|^{1+\alpha}} W_\infty(s) \leq \frac{C}{|\bar{x}_t - x_t|^{1+\alpha}} W_\infty(t), \end{aligned}$$

and obtain by integration on $[t-\tau,t]$

$$I_\alpha(t,\bar z_t,z_t)\le \frac{C}{|\bar x_t-x_t|^{1+\alpha}}W_\infty(t).$$

Then integrating in z_t we may get since $\alpha+1<2\le d$

$$\begin{aligned}\int_{A_t} I_\alpha(t,\bar z_t,z_t)\,dz_t &\le C W_\infty(t)\int_{A_t}\frac{dz_t}{|\bar x_t-x_t|^{1+\alpha}}\\ &\le C\,R(t)^{2d-1-\alpha}\,W_\infty(t)\le C\,W_\infty(t).\end{aligned}\tag{4.14}$$

For $J_{\alpha+1}$, we use (4.13) on the set A_t the bound

$$|K_\varepsilon(T^s_x(\bar z_s)-T^s_x(z_s))|\le \frac{C}{|\bar x_t-x_t|^{1+\alpha}}.$$

Integrating with respect to time and z_t we get since $1+\alpha<d$.

$$\begin{aligned}\int_{A_t} J_{\alpha+1}(t,\bar z_t,z_t)\,dz_t &\le C\int_{A_t}\frac{dz_t}{|\bar x_t-x_t|^{1+\alpha}}\\ &\le C\,R(t)^{2d-1-\alpha}\le C.\end{aligned}\tag{4.15}$$

For the cut-off case, the estimation on I_α for this step is unchanged.

4.4.3 Step 1' : Estimate over A_t^c for the "continuous" part of I_α .

For the remaining term in I_α, we use the rude bound

$$|F(T^s_x(\bar z_s)-T^s_x(z_s))-F(\bar x_s-x_s)|\le |F(T^s_x(\bar z_s)-T^s_x(z_s))|+|F(\bar x_s-x_s)|.$$

The term involving T^s is complicated and requires the additional decompositions. It will be treated in the next sections. The other term is simply bounded by

$$\begin{aligned}\int_{z_t\in A_t^c}\frac{1}{\tau}\int_{t-\tau}^t|F(\bar x_s-x_s)|ds\ dz_t &\le \frac{1}{\tau}\int_{t-\tau}^t\int_{z_t\in A_t^c}\frac{C\,dz_t}{|\bar x_s-x_s|^\alpha}\,ds\\ &\le \frac{1}{\tau}\int_{t-\tau}^t\int_{z_s\in Z^f(s,t,A_t^c)}\frac{C\,dz_s}{|\bar x_s-x_s|^\alpha}\,ds.\end{aligned}$$

From the bounds (4.1),we get that

$$|A_t^c|\le CR(t)^d[W_\infty(t)+\tau(1+|E|_\infty(t))]^d\le C(W_\infty(t)+\tau)^d,$$

where $|\cdot|$ denote the Lebesgue measure. Since the flow Z_f is measure preserving, the measure of the set $Z^f(s,t,A_t^c)$ satisfies the same bound. This set is also included in $B_{2d}(0,R)$. We use the above lemma which implies that above all the set $Z(s,t,A_t^c)$, the integral reaches is maximum when the set is a cylinder

Lemma 4. *Let $\Omega\subset B_{2d}(0,R)\subset\mathbb{R}^{2d}$. Then for any $a<d$, there exists a constant C_a depending on a and d such that*

$$\int_\Omega\frac{dz}{|x|^a}\le C_aR^a\,|\Omega|^{1-a/d}.$$

Proof of Lemma 4. We maximize the integral

$$\int_\omega |x|^{-a} dz$$

over all sets $\omega \subset \mathbb{R}^{2d}$ satisfying $\omega \subset B_{2d}(0,R)$ and $|\omega| = |\Omega|$. It is clear that the maximum is obtained by concentrating as much as possible ω near $x = 0$, *i.e.* with a cylinder of the form $B_d(0,r) \times B_d(0,R)$. Since $|\omega| = |\Omega|$ we have $(c_d)^2 r^d R^d = |\Omega|$, where c_d is the volume of the unit ball of dimension d. The integral over this cylinder can now be computed explicitly and gives the lemma. □

Applying the lemma, we get

$$\int_{z_t \in A_t^c} \frac{1}{\tau} \int_{t-\tau}^t |F(\bar{x}_s - x_s)|\, dz_t ds \le C[W_\infty(t) + \tau]^{d-\alpha}. \tag{4.16}$$

That term do not appear in Lemma 2 since it is strictly smaller than the bound of the remaining term (involving T), as we shall see in the next section.
For the cut-off case, the same bound is valid for I_α since $\alpha \le d-1 < d$ (The cut-off cannot in fact help to provide a better bound for this term).

At this point, the remaining term to bound in I_α is only

$$\int_{z_t \in A_t^c} \frac{1}{\tau} \int_{t-\tau}^t |F(T_x^s(\bar{z}_s) - T_x^s(z_s))|\, ds \tag{4.17}$$

and the remainder in $J_{\alpha+1}$ is (4.17)

$$\int_{A_t^c} J_{\alpha+1}(t, \bar{z}_t, z_t)\, dz_t = \frac{1}{\tau} \int_{A_t^c} \int_{t-\tau}^t K_\varepsilon(T_x^s(\bar{z}_s) - T_x^s(z_s))\, dz_t ds. \tag{4.18}$$

Therefore in the next sections we focus on giving a bound for (4.17) and (4.18).

4.4.4 Step 2: Estimate over B_t

We recall the definition of B_t

$$B_t = \left\{ z_t \;\text{ s.t. }\; \begin{array}{l} |\bar{x}_t - x_t| \le 4W_\infty(t) + 2\tau(|\bar{v}_t - v_t| + \tau|E|_\infty(t)) \\ |\bar{v}_t - v_t| \ge 4\,W_\infty(t) + 4\,\tau|E|_\infty|(t) \end{array} \right\}.$$

If $z_t \in B_t$, we have, as for A_t, for $s \in [t-\tau, t]$

$$\begin{aligned} |\bar{v}_s - v_s - \bar{v}_t + v_t| &\le 2\tau|E|_\infty(t) \le \frac{|\bar{v}_t - v_t|}{2}, & (4.19) \\ |T_v^s(\bar{z}_s) - T_v^s(z_s) - \bar{v}_t + v_t| &\le |\bar{v}_s - v_s - \bar{v}_t + v_t| + 2W_\infty(s) \le \frac{|\bar{v}_t - v_t|}{2}. & (4.20) \end{aligned}$$

This means that the particles involved are close to each others (in the positions variables), but with a sufficiently large relative velocity, so that they do not interact a lot on the interval $[t-\tau, t]$.

First we introduce a notation for the term of (4.17)

$$\int_{z_t\in B_t} I_{bc}(t,\bar z_t,z_t)\,dz_t, \qquad \text{with } I_{bc}(t,\bar z_t,z_t)=I_{bc}(t,i,j):=\frac{1}{\tau}\int_{t-\tau}^{t} F(T_x^s(\bar z_s)-T_x^s(z_s))\,ds, \tag{4.21}$$

where (i,j) are s.t. $T_x^s(\bar z_s)=X_i(s)$, $T_x^s(z_s)=X_j(s)$.
For $z_t\in B_t$, define for $s\in[t-\tau,\ t]$

$$\phi(s):=(T_x^s(\bar z_s)-T_x^s(z_s))\cdot\frac{\bar v_t-v_t}{|\bar v_t-v_t|}=(X_i(s)-X_j(s))\cdot\frac{\bar v_t-v_t}{|\bar v_t-v_t|}.$$

Note that $|\phi(s)|\le|T_x^s(\bar z_s)-T_x^s(z_s)|$ and that

$$\begin{aligned}\phi'(s)&=(T_v^s(\bar z_s)-T_v^s(z_s))\cdot\frac{\bar v_t-v_t}{|\bar v_t-v_t|}\\&=|\bar v_t-v_t|+(T_v^s(\bar z_s)-T_v^s(z_s)-(\bar v_t-v_t))\cdot\frac{\bar v_t-v_t}{|\bar v_t-v_t|}\ge\frac{|\bar v_t-v_t|}{2},\end{aligned}$$

where we have used (4.20). Therefore ϕ is an increasing function of the time on the interval $[t-\tau,t]$. If it vanishes at some time $s_0\in[t-\tau,t]$, then the previous bound by below on its derivative implies that

$$|T_x^s(\bar z_s)-T_x^s(z_s)|\ge|\phi(s)|\ge|t-s_0|\,\frac{|\bar v_t-v_t|}{2}. \tag{4.22}$$

If ϕ is always positive (resp. negative) on $[t-\tau,t]$, then the previous estimate is still true with the choice $s_0=t-\tau$ (resp. $s_0=t$). So in any case, estimate (4.22) holds true for some $s_0\in[t-\tau,t]$. Using this directly gives, as $\alpha<1$

$$|I_{bc}(t,\bar z_t,z_t)|\le\frac{C}{\tau}\,|\bar v_t-v_t|^{-\alpha}\int_{t-\tau}^{t}\frac{ds}{|s-s_0|^\alpha}\le C\,\tau^{-\alpha}\,|\bar v_t-v_t|^{-\alpha}. \tag{4.23}$$

Now integrating

$$\begin{aligned}\int_{z_t\in B_t}|I_{bc}(t,\bar z_t,z_t)|\,dz_t&\le C\,\tau^{-\alpha}\int_{A_t^c}\frac{dz_t}{|\bar v_t-v_t|^\alpha}\\&\le C\,\tau^{-\alpha}\,[W_\infty(t)+\tau]^d\,[R(t)]^{d-\alpha},\end{aligned}$$

by using the fact that $B_t\subset B(0,C[W_\infty(t)+\tau])\times B(0,R(t))$. In conclusion

$$\int_{z_t\in B_t}|I_{bc}(t,\bar z_t,z_t)|\,dz_t\le C\,\tau^{-\alpha}\,[W_\infty(t)+\tau]^d. \tag{4.24}$$

With the cut-off where $\alpha>1$, the reasoning follows the same line up to the bound (4.23) which relies on the hypothesis $\alpha<1$. (4.23) is replaced by

$$\begin{aligned}|I_{bc}(t,\bar z_t,z_t)|&\le\frac{C}{\tau}\int_{t-\tau}^{t}\frac{ds}{(|s-s_0||\bar v_t-v_t|+4\varepsilon^{\bar m})^\alpha}\\&\le\frac{C}{\tau}\int_{t-\tau}^{s_0}\dots+\frac{C}{\tau}\int_{s_0}^{t}\dots\le\frac{2C}{\tau}\int_0^\tau\frac{ds}{(s|\bar v_t-v_t|+4\varepsilon^{\bar m})^\alpha}\\&\le\frac{C}{\tau|\bar v_t-v_t|}\int_0^{\tau|\bar v_t-v_t|}\frac{ds}{(s+4\varepsilon^{\bar m})^\alpha}\le\frac{C\varepsilon^{\bar m(1-\alpha)}}{\tau|\bar v_t-v_t|}.\end{aligned}$$

When $\alpha = 1$, the previous calculation leads to

$$|I_{bc}(t, \bar{z}_t, z_t)| \le \frac{C}{\tau|\bar{v}_t - v_t|} \ln\left(1 + C\varepsilon^{1-\bar{m}}\right) \le \frac{C}{\tau|\bar{v}_t - v_t|}\left(1 + \ln \varepsilon^{-\bar{m}}\right) \le \frac{C\varepsilon^0}{\tau}$$

For the first bound, we used the fact that the f_N are uniformly compactly supported , and the fact that $\tau = \varepsilon$ in the case with cut-off. The second bound follows from $\ln(1+x) \le 1 + \ln(x)$ if $x \ge 1$, an inequality we can apply if the constant C in the logarithm is larger than one. In the last bound, we use the convention $\varepsilon^0 = 1 + |\ln(\varepsilon)|$.
In both cases, the singular part in $1/|\bar{v}_t - v_t|$ is integrable on $\mathbb{R}^d$ and integrating that bound over B_t, we get the estimate

$$\begin{aligned} \int_{z_t \in B_t} |I_{bc}(t, \bar{z}_t, z_t)|\, dz_t &\le C\,\tau^{-1}\varepsilon^{\bar{m}(1-\alpha)} \int_{A_t^c} \frac{dz_t}{|\bar{v}_t - v_t|} \\ &\le C\,\tau^{-1}\,\varepsilon^{\bar{m}(1-\alpha)}[W_\infty(t) + \tau]^d\,[R(t)]^{d-1}, \\ &\le C\,\tau^{-1}\,\varepsilon^{\bar{m}(1-\alpha)}[W_\infty(t) + \tau]^d \end{aligned} \tag{4.25}$$

4.4.5 Step 3: Estimate over C_t

We recall the definition of C_t

$$C_t = \left\{ z_t \;\text{ s.t. }\; \begin{array}{l} |\bar{x}_t - x_t| \le 4W_\infty(t) + 2\tau(|\bar{v}_t - v_t| + \tau|E|_\infty(t)) \\ |\bar{v}_t - v_t| \le 4\,W_\infty(t) + 4\,\tau|E|_\infty|(t) \end{array} \right\}.$$

First remark that under the assumption (4.10) $C_t \subset \{|z_t - \bar{z}_t| \le C(W_\infty(t) + \tau)\}$, so that its volume is bounded by $C(W_\infty(t) + \tau)^{2d}$. From the previous steps, it only remains to bound

$$\int_{z_t \in C_t} I_{bc}(t, \bar{z}_t, z_t)\, dz_t.$$

We begin by the cut-off case, which is the simpler one. In that case, one simply bounded $I_{bc} \le C\,\varepsilon^{-\bar{m}\alpha}$ which implies

$$\int_{z_t \in C_t} I_{bc}(t, \bar{z}_t, z_t)\, dz_t \le C(W_\infty(t) + \tau)^{2d}\varepsilon^{-\bar{m}\alpha}. \tag{4.26}$$

It remains the case without cut-off. We denote $\tilde{C}_t = \{j \mid \exists z_t \in C_t,\ s.t.\ Z_j(t) = T^t(z_t)\}$, and transform the integral on C_t in a discrete sum

$$\int_{z_t \in C_t} I_{bc}(t, \bar{z}_t, z_t)\, dz_t = \sum_{j \in \tilde{C}_t} a_{ij} I_{Nc}(t, i, j) \quad \text{ with } I_{Nc}(t, i, j) = \frac{1}{\tau}\int_{t-\tau}^{t} \frac{dz_t}{|X_i(s) - X_j(s)|^\alpha}\, ds,$$

where i is the number of the particle associated to $\bar{z}_t$ ($T^t(\bar{z}_t) = Z_i(t)$) and

$$a_{ij} = |\{z_t \in C_t,\ T^t(z_t) = Z_j(t)\}|, \quad \text{so that } \sum_{j \in \tilde{C}_t} a_{ij} = |C_t|.$$

To bound I_{Nc} over $\tilde{C}_t$, we do another decomposition in j. Define

$$\begin{aligned} JX_t &= \left\{ j \in \tilde{C}_t\,,\ |X_j(t) - X_i(t)| \ge \frac{d_N(t)}{2} \right\}, \\ JV_t &= \left\{ j \in \tilde{C}_t\,,\ |X_j(t) - X_i(t)| \le |V_j(t) - V_i(t)| \text{ and } |V_j(t) - V_i(t)| \ge \frac{d_N(t)}{2} \right\}. \end{aligned}$$

By the definition of the minimal distance in $\mathbb{R}^{2d}$, $d_N(t)$, one has that $\tilde{C}_t = JX_t \cup JV_t$. Since

$$|T^t(z_t) - z_t| \le W_\infty(t),$$

one has by the definition of $\tilde{C}_t$ and C_t that for all $j \in \tilde{C}_t$, $|Z_j(t) - Z_i(t)| \le C\,(W_\infty(t) + \tau)$. Let us start with the bound over JX_t. If $j \in JX_t$, one has that

$$|X_j(s) - X_i(s)| \ge |X_j(t) - X_i(t)| - \int_s^t |V_j(u) - V_i(u)|\,du.$$

On the other hand, for $u \in [s,t]$,

$$|V_j(u) - V_i(u)| \le 2W_\infty(t) + |\bar{v}_u - v_u| \le 2(W_\infty(t) + \tau|E|_\infty) + |\bar{v}_t - v_t| \le C(W_\infty(t) + \tau).$$

Therefore assuming that with that constant C

$$C\,\tau(W_\infty(t) + \tau) \le d_N(t)/4, \tag{4.27}$$

we have that for any $s \in [t-\tau,\ t]$, $|X_j(s) - X_i(s)| \ge d_N(t)/4$. Consequently for any $j \in JX_t$

$$I_{Nc}(t,i,j) \le C\,[d_N(t)]^{-\alpha}. \tag{4.28}$$

For $j \in JV_t$, we write

$$|(V_j(s) - V_i(s)) - (V_j(t) - V_i(t))| \le \int_s^t |E_N(X_j(u)) - E_N(X_i(u))|\,du.$$

Note that

$$\begin{aligned} |X_j(s) - X_i(s)| &\le |X_j(t) - X_i(t)| + \int_s^t |V_j(u) - V_i(u)|\,du \\ &\le C(W_\infty(t) + \tau) + 2\int_s^t (W_\infty(u) + R(u))\,du \\ &\le C(W_\infty(t) + \tau). \end{aligned} \tag{4.29}$$

Hence we get for $s \in [t-\tau, t]$

$$\int_s^t |E_N(X_j(u)) - E_N(X_i(u))|\,du \le C\,\tau\,|\nabla^N E|_\infty\,(W_\infty(t) + \tau + \varepsilon^{1+r'}).$$

Note that the constant C still does not depend on $\tau = \varepsilon^{r'}$. Therefore provided that with the previous constant C

$$2C\,\tau\,|\nabla^N E|_\infty\,(W_\infty(t) + \tau) \le d_N(t)/4, \tag{4.30}$$

one has that

$$|V_j(s) - V_i(s) - (V_j(t) - V_i(t))| \le d_N(t)/4 \quad \text{and also} \quad |V_i(s) - V_j(s)| \ge \frac{d_N(t)}{4}.$$

As in the step for B_t (See equation (4.22)) this implies the dispersion estimate $|X_j(s) - X_i(s)| \ge |s - s_0|\,d_N(t)/4$ for some $s_0 \in [t-\tau,\ t]$. As a consequence for $j \in JV_t$,

$$I_{Nc}(t,i,j) \le \frac{C}{\tau}\,(d_N(t))^{-\alpha} \int_{t-\tau}^t \frac{ds}{|s - s_0|^\alpha} \le C\,\tau^{-\alpha}\,(d_N(t))^{-\alpha}. \tag{4.31}$$

Summing (4.28) and (4.31), one gets

$$\sum_{j\in\tilde{C}_t} a_{ij} I_{Nc}(t,i,j) \le C\,|C_t|\ \left((d_N(t))^{-\alpha} + \tau^{-\alpha}\ (d_N(t))^{-\alpha}\right).$$

Coming back to I_{bc}, using the bound on the volume of $|C_t|$ and keeping only the largest term of the sum

$$\int_{C_t} I_{bc}(t,\bar{z}_t,z_t)\,dz_t \le C\,(W_\infty(t)+\tau)^{2d}\tau^{-\alpha}\ (d_N(t))^{-\alpha}. \tag{4.32}$$

4.4.6 Conclusion of the proof of Lemmas 2, 3

Assumptions (4.27) and (4.30) are ensured by the hypothesis of the lemma. Summing up (4.14) for I_α or (4.15) for $J_{\alpha+1}$, with (4.16), (4.24) and (4.32), we indeed find the conclusion of the first lemma.
In the S_m^α case, no assumption is needed, and summing up the bounds (4.14), (4.16), (4.25), (4.26), we obtain the second lemma.

4.5 Conclusion of the proof of Theorem 3 (without cut-off)

In this subsection, in order to make the argument clearer, we number explicitly the constants. Let us summarize the important information of Prop. 7 and Lemma 2. Let us also rescale the interested quantities s.t. all may be of order 1

$$\varepsilon\,\tilde{W}_\infty(t) = W_\infty(t), \quad \varepsilon^{1+r}\,\tilde{d}_N(t) = d_N(t).$$

Remark that by Proposition 2 $\tilde{W}_\infty(t) = c_\phi > 0$. By assumption (i) in Theorem 3, also note that $\tilde{d}_N(0) \ge 1$.
Recalling $\tau = \varepsilon^{r'}$ (with $r' > r > 1$), the condition of Lemma 2 after rescaling reads

$$C_1\,\varepsilon^{r'-r}\,(1+|\nabla^N E|_\infty(t))\,\tilde{W}_\infty(t) \le\ \tilde{d}_N(t). \tag{4.33}$$

In Lemma 2, we proved that there exist some constants C_0 and C_2 independent of N (and hence ε), such that if (4.33) is satisfied, then for any $t\in[0,\ T]$

$$\begin{aligned}
\tilde{W}_\infty(t) \qquad &\le \tilde{W}_\infty(t-\tau) + C_0\,\varepsilon^{r'}\ \left(\tilde{W}_\infty(t) + \varepsilon^{\lambda_1}\,\tilde{W}_\infty^d(t) + \varepsilon^{\lambda_2}\,\tilde{W}_\infty^{2d}(t)\,\tilde{d}_N^{-\alpha}(t)\right),\\
|\nabla^N E|_\infty(t) \quad &\le C_2\left(1 + \varepsilon^{\lambda_3}\,\tilde{W}_\infty^d(t) + \varepsilon^{\lambda_4}\,\tilde{W}_\infty^{2d}(t)\,\tilde{d}_N^{-\alpha}(t))\right)\\
\tilde{d}_N(t) + \varepsilon^{r'-r} \quad &\ge [\tilde{d}_N(t-\tau) + \varepsilon^{r'-r}]e^{-\tau(1+|\nabla^N E|_\infty(t))},
\end{aligned}$$

where ε appear four times with four different exponents $\lambda_i, i=1,\dots,4$ defined by

$$\begin{aligned}
\lambda_1 &= d-1-\alpha\,r', & \lambda_2 &= 2d-1-\ \alpha(1+r'+r),\\
\lambda_3 &= d-1-r'-\alpha\,r', & \lambda_4 &= 2d-1-r'-\ \alpha(1+r'+r).
\end{aligned}$$

To propagate uniform bounds as $\varepsilon\to 0$ and $N\to\infty$, we need all λ_i to be positive. As $r,r'>0$, it is clear that $\lambda_1>\lambda_3$ and $\lambda_2>\lambda_4$. Thus we need only check $\lambda_3>0$ and $\lambda_4>0$. As $r'>r$, it is sufficient to have

$$r' < \frac{d-1}{1+\alpha}, \quad \text{and} \quad r' < \frac{2d-1-\alpha}{1+2\,\alpha}.$$

Note that a straightforward calculation shows that

$$\frac{d-1}{1+\alpha} - \frac{2d-1-\alpha}{1+2\,\alpha} = \frac{\alpha^2-d}{(1+\alpha)(1+2\alpha)} < 0,$$

so that the first inequality is the stronger one. Thanks to the condition given in Theorem 3, $r < r^* := \frac{d-1}{1+\alpha}$, so that if we choose any $r' \in (r, r^*)$, the corresponding λ_i are all positive. We fix a r' as above and denote $\lambda = \min_i(\lambda_i)$. Then by a rough estimate,

$$\begin{aligned} \tilde{W}_\infty(t) &\leq \tilde{W}_\infty(t-\tau) + C_0\,\tau\left(\tilde{W}_\infty(t) + 2\,\varepsilon^\lambda\,\tilde{W}_\infty^{2d}(t)\,d_N^{-\alpha}(t)\right), \\ |\nabla^N E|_\infty(t) &\leq C_2\left(1\,+2\,\varepsilon^\lambda\,\tilde{W}_\infty^{2d}(t)\,\tilde{d}_N^{-\alpha}(t)\right), \\ \tilde{d}_N(t) &\geq [\,\tilde{d}_N(t-\tau) + \varepsilon^{r'-r}]e^{-(1+|\nabla^N E|_\infty(t))\tau} - \varepsilon^{r'-r}. \end{aligned} \tag{4.34}$$

If for some $t_0 > 0$ one has (4.33) on the whole time interval $[0, t_0]$and

$$\forall t \in [0, t_0], \quad 2\,\varepsilon^\lambda\,\tilde{W}_\infty^{2d}(t)\,\tilde{d}_N^{-\alpha}(t) \leq 1, \tag{4.35}$$

then we get $\tilde{W}_\infty(t) \leq \tilde{W}_\infty(t-\tau) + 2C_0\tau\tilde{W}_\infty(t)$ so that if $2C_0\tau < 1$

$$\begin{aligned} \tilde{W}_\infty(t) &\leq \tilde{W}_\infty(t-\tau)(1-2C_0\tau)^{-1}, \\ |\nabla^N E|_\infty(t) &\leq 2\,C_2, \\ \tilde{d}_N(t) &\geq\ e^{-(1+2C_2)\,t} - \varepsilon^{r'-r} \end{aligned} \tag{4.36}$$

for any $t \in [0, t_0]$. The last inequality implies $\tilde{d}_N(t) \geq \frac{1}{2}\,e^{-(1+2C_2)\,t}$ if $2\varepsilon^{r'-r}e^{(1+2C_2)\,T} < 1$. That condition is fulfilled for ε small enough, i.e. N large enough ($\ln N \geq CT$).
The first inequality in (4.36), iterated gives $W_\infty(t) \leq W_\infty(0)(1-2C_0\tau)^{-\frac{t}{\tau}}$. If $C_0\tau \leq \frac{1}{4}$, then we can use $-\ln(1-x) \leq 2x$ for $x \in [0, \frac{1}{2}]$, and get

$$\tilde{W_\infty}(t) \leq \tilde{W}_\infty(0)e^{4C_0 t}$$

To summarize, under the previous hypothesis it comes for all $t \in [0, t^0]$

$$\begin{aligned} \tilde{W}_\infty(t) &\leq e^{4C_0 t}, \\ |\nabla^N E|_\infty(t) &\leq 2\,C_2, \\ \tilde{d}_N(t) &\geq \frac{1}{2}\,e^{-(1+2C_2)\,t}. \end{aligned} \tag{4.37}$$

As we mention in the introduction, we only deals with continuous solutions to the N particles system (1.1) or (1.2). So $\tilde{W}_\infty(t)$ and $\tilde{d}_N(t)$ are continuous functions of the time, and $|\nabla^N E|_\infty(t)$ is also continuous in time thanks to the smoothing parameter that appears in its definition (4.4).
As we explain in Remark 10, $|\nabla^N E|_\infty(0) = 0$ and the conditions (4.33) and (4.35) are satisfied at time $t = 0$. In fact, at time 0 they maybe rewritten

$$C_1\varepsilon^{r'-r}\tilde{W}_\infty(0) \leq \tilde{d}_N(0), \qquad 2\varepsilon^\lambda\tilde{W}_\infty(0)^{2d}\tilde{d}_N(0)^{-\alpha} \leq 1$$

and this is true for N large enough because of our assumption on ε and $d_N(0)$. Then by continuity there exists a maximal time $t_0 \in]0,\ T]$ (possibly $t_0 = T$) such that they are satisfied on $[0,\ t_0]$.

We show that for N large enough, *i.e.* ε small enough, then one necessarily has $t_0 = T$. Then we will have (4.37) on $[0,\ T]$ which is the desired result. This is simple enough. By contradiction if $t_0 < T$ then

$$C_1\, \varepsilon^{(r-r')}\, (1 + |\nabla^N E|_\infty(t_0))\, \tilde{W}_\infty(t_0) = \tilde{d}_N(t_0), \quad \text{or} \quad 4\, \varepsilon^\lambda\, \tilde{W}_\infty^{2d}(t_0)\, \tilde{d}_N^{-\alpha}(t_0) = 1.$$

But until t_0, (4.37) holds. Therefore

$$\varepsilon^\lambda\, \tilde{W}_\infty^{2d}(t_0)\, \tilde{d}_N^{-\alpha}(t_0) \le \varepsilon^\lambda\, 2^\alpha\, e^{(\alpha + (4d+2\alpha)\max(C_0,C_2))\, t_0} < 1,$$

for ε small enough with respect to T and the C_i. This is the same for (4.33)

$$C_1\, \varepsilon^{(r-r')}\, (1 + |\nabla^N E|_\infty(t_0))\, \tilde{W}_\infty(t_0) \tilde{d}_N^{-1}(t_0) \le 2\varepsilon^{(r-r')} C_1 (1 + 2C_2) e^{(1+6\max(C_0,C_2))t_0} < 1.$$

Hence we obtain a contradiction and prove Theorem 3.

4.6 Conclusion of the proof of Theorem 4 (cut-off case)

In the cut-off case, using Lemma 3 together with the inequality *ii*) of the Proposition 7, we may obtain

$$W_\infty(t) \le W_\infty(t-\tau) + C_0 W_\infty(t) \left[1 + (W_\infty(t) + \tau)^{d-1} \tau^{-1} \varepsilon^{\bar{m}(1-\alpha)} + (W_\infty(t) + \tau)^{2d-1} \varepsilon^{-\bar{m}\alpha}\right].$$

We again rescale the quantity $W_\infty(t) = \varepsilon \tilde{W}_\infty(t)$. Choosing in that case $\tau = \varepsilon$, it comes for $1 \le \alpha < d-1$,

$$\tilde{W}_\infty(t) \le \tilde{W}_\infty(t-\tau) + C_0 \tilde{W}_\infty(t) \tau \left[1 + \varepsilon^{d-2-\bar{m}(\alpha-1)}\, \tilde{W}_\infty^{d-1}(t) + \varepsilon^{2d-1-\bar{m}\alpha}\, \tilde{W}_\infty^{2d-1}(t)\right].$$

As in the previous section, we will get a good bound provided that the power of ε appearing in parenthesis are positive. The two conditions read

$$\bar{m} < \bar{m}^* := \min\left(\frac{d-2}{\alpha-1}, \frac{2d-1}{\alpha}\right).$$

In that case, for N large enough (with respect to e^{Ct}), we get a control of the type

$$\frac{d}{dt} \tilde{W}_\infty(t) \le 4 C_0 \tilde{W}_\infty(t),$$

(but discrete in time) which gives the desired result.

Remark 11. *In the cut-off case (and also in the case without cut-off), it seems important to be able to say that the initial configurations Z we choose have a total energy close from the one of f^0. Because, if the empirical distribution μ_N^Z is close form f^0, but has a different total energy, we would not expect that they do not remains close a very long time. Fortunately, such a result is true and under the assumptions of Theorem 3 and 4, the total energy of the empirical distributions are close from the total energy of f^0.*
Unfortunately, we do not have a simple proof of this fact. But, it can be done using the argument for the proof of the deterministic theorems. First, the difference between the kinetic energy is easily controlled because our solutions are compactly supported and that there is no singularity there. Next, performing calculations very similar to the ones done in the proofs, we can control the difference between a small average in time of the potential energies, on the small interval of time $[0, \tau]$. Then, we control the average of the total energy, which is constant.

A Appendix

A.1 Large deviation on the infinite norm of f_N.

Proposition 8. *Assume that ρ is a probability on $\mathbb{R}^n$ with support included in $[-R^0,R^0]^n$ and and bounded density $f(x)\,dx$. Let ϕ be a bounded cut-off function, with support in $[-\frac{L}{2},\frac{L}{2}]^n$ and total mass one, and define the usual $\phi_\varepsilon := \frac{1}{\varepsilon^n}\phi(\frac{\cdot}{\varepsilon})$. For any configuration $Z_N=(Z_i)_{i\le N}$ we define*

$$f_N^Z := \mu_N^Z * \phi_\varepsilon(N).$$

If $\varepsilon(N)=N^{-\frac{\gamma}{n}}$ and the Z_N are distributed according to $f^{\otimes N}$, then we have the explicit "large deviations" bound with $c_\phi=(2L)^n\|\phi\|_\infty$ and $c_0=(2R^0+2)^nL^{-n}$

$$\forall\beta>1,\quad \mathbb{P}\left(\|f_N^Z\|_\infty\ge \beta c_\phi\|f\|_\infty\right)\le c_0N^\gamma e^{-(\beta\ln\beta-\beta+1)(2L)^n\|f\|_\infty N^{1-\gamma}}. \tag{A.1}$$

In particular, for $\phi=\mathbf{1}_{[-1/2,1/2]^n}$ and $\beta=2$, we get

$$\mathbb{P}\left(\|f_N^Z\|_\infty\ge 2^{1+n}\|f\|_\infty\right)\le (2R^0+2)^nN^\gamma e^{-(2\ln 2-1)2^n\|f\|_\infty N^{1-\gamma}}. \tag{A.2}$$

Proof. For any $Z\in\mathbb{R}^{nN}$ and $z\in\mathbb{R}^n$, we have

$$\begin{aligned}
f_N^Z(z) &= \frac{1}{N}\sum_{i=1}^N\phi_\varepsilon(z-Z_i)=\frac{1}{N\,\varepsilon^n}\sum_{i=1}^N\phi\left(\frac{z-Z_i}{\varepsilon}\right)\\
&\le \frac{\|\phi\|_\infty}{N\,\varepsilon^n}\#\{i \text{ s.t. } |z-Z_i|_\infty\le \tfrac{L\varepsilon}{2}\}\\
\|f_N^Z\|_\infty &\le \frac{\|\phi\|_\infty}{N\,\varepsilon^n}\sup_{z\in\mathbb{R}^n}\#\{i \text{ s.t. } |z-Z_i|_\infty\le \tfrac{L\varepsilon}{2}\},
\end{aligned}$$

where # stands for the cardinal (of a finite set). It remains to bound the supremum on all the cardinals. The first step will be to replace the sup on all the $z\in\mathbb{R}^n$ by a supremum on a finite number of point. For this, we cover $[-R^0,R^0]^n$ by M cubes C_k of size $L\varepsilon$, centered at the points $(c_k)_{k\le M}$. The number M of square needed depends on N via ε, and is bounded by

$$M\le\left[\frac{2(R^0+1)}{L\varepsilon}\right]^n.$$

Next, for any $z\in\mathbb{R}^d$, there exists a $k\le M$ such that $|z-c_k|\le\frac{L\varepsilon}{2}$. This implies that

$$\sup_{z\in\mathbb{R}^n}\#\{i \text{ s.t. } |z-Z_i|_\infty\le\tfrac{L\varepsilon}{2}\}\le\sup_{k\le M}\#\{i \text{ s.t. } |c_k-Z_i|_\infty\le L\varepsilon\}$$

Now we denote by $H_k^N := \#\{i \text{ s.t. } |c_k-Z_i|_\infty\le L\varepsilon\}$. H_k^N follows a binomial law $B(N,p_k)$ with $p_k=\int_{2C_k}f(z)\,dz$, where $2C_k$ denotes the square with center c_k, but size $2L\varepsilon$. Remark that

$$p_k\le\bar p:=(2L\varepsilon)^n\|f\|_\infty.$$

For any λ, the exponential moments of H_k^N are therefore given and bounded by

$$\begin{aligned}
\mathbb{E}(e^{\lambda H_k^N}) &= \left[1+(e^\lambda-1)p_k\right]^N\\
&\le \left[1+(e^\lambda-1)(2L\varepsilon)^n\|f\|_\infty\right]^N\\
&\le e^{(e^\lambda-1)N(2L\varepsilon)^n\|f\|_\infty}.
\end{aligned}$$

Now for the supremum of the H_k^N

$$\begin{aligned}\mathbb{E}(e^{\lambda \sup_k H_k^N}) &\leq \mathbb{E}(e^{\lambda H_1^N}) + \cdots + (e^{\lambda H_M^N}) \\ &\leq M e^{(e^\lambda - 1)N(2L\varepsilon)^n \|f\|_\infty} \\ &\leq \left[\frac{2(R^0+1)}{L\varepsilon}\right]^n e^{(e^\lambda - 1)N(2L\varepsilon)^n \|f\|_\infty}\end{aligned}$$

Using finally Chebyshev's inequality, we get for any $\beta > 0$

$$\begin{aligned}\mathbb{P}\left(\|f_N^Z\|_\infty \geq \beta(2L)^n \|\phi\|_\infty \|f\|_\infty\right) &\leq \mathbb{P}\left(\sup_k H_k^N \geq \beta \|f\|_\infty N(2L\varepsilon)^n\right) \\ &\leq \mathbb{E}(e^{\lambda \sup_k H_k^N}) e^{-\lambda\beta\|f\|_\infty N(2L\varepsilon)^n} \\ &\leq \left[\frac{2(R^0+1)}{L\varepsilon}\right]^n e^{\left(e^\lambda - 1 - \lambda\beta\right)N(2L\varepsilon)^n \|f\|_\infty}.\end{aligned}$$

For $\beta > 1$, the optimal λ is $\ln\beta$ and we get with $c_\phi = (2L)^n \|\phi\|_\infty$

$$\mathbb{P}\left(\|f_N^Z\|_\infty \geq \beta c_\phi \|f\|_\infty\right) \leq \left[\frac{2(R^0+1)}{L\varepsilon}\right]^n e^{-(\beta\ln\beta - \beta + 1)N(2L\varepsilon)^n \|f\|_\infty}.$$

With the scaling $\varepsilon(N) = N^{-\frac{\gamma}{n}}$, we get

$$\mathbb{P}\left(\|f_N^Z\|_\infty \geq \beta c_\phi \|f\|_\infty\right) \leq c_0 N^\gamma e^{-(\beta\ln\beta - \beta + 1)(2L)^n \|f\|_\infty N^{1-\gamma}}.$$

Remark finally that the choice of scale $\varepsilon(N) = (\ln N) N^{-\frac{1}{n}}$ is also sufficient to get a probability vanishing faster than any inverse power. □

A.2 Existence of strong solutions to Equation (1.3)

This subsection is devoted to the proof of lemma 1.

Proof of the lemma 1. Given the estimate on f, ρ also belongs to L^∞ with the bound

$$\|\rho(t,.)\|_{L^\infty(\mathbb{R}^d)} \leq C\, K(t)^d\, \|f(t,.,.)\|_{L^\infty(\mathbb{R}^{2d})}.$$

As we have (1.8) with $\alpha < d-1$, $E = F \star_x \rho$ is Lipschitz. Therefore the solution to (1.3) is given by the characteristics. Namely, we define X and V the unique solutions to

$$\begin{aligned}&\partial_t X(t,s,x,v) = V(t,s,x,v), \quad \partial_t V(t,s,x,v) = E(t, X(t,s,x,v)), \\ &X(s,s,x,v) = x, \quad V(s,s,x,v) = v.\end{aligned}$$

The solution f is now given by

$$f(t,x,v) = f(0, X(0,t,x,v), V(0,t,x,v)),$$

with the consequence that

$$R(t) \leq R(0) + \int_0^t K(s)\, ds, \quad K(t) \leq K(0) + \int_0^t \|E(s,.)\|_{L^\infty}\, ds.$$

Then

$$\|E\|_{L^\infty} \leq \|\rho\|_{L^1}^{1-\alpha/d}\, \|\rho\|_{L^\infty}^{\alpha/d},$$

which leads to the required inequality. To conclude it is enough to notice that the L^1 and L^∞ norms of f are preserved in this case. □

References

[Ars75] A. A. Arsen′ev. Existence in the large of a weak solution of Vlasov's system of equations. *Ž. Vyčisl. Mat. i Mat. Fiz.*, 15:136–147, 276, 1975.

[Bat01] Jürgen Batt. N-particle approximation to the nonlinear Vlasov-Poisson system. In *Proceedings of the Third World Congress of Nonlinear Analysts, Part 3 (Catania, 2000)*, volume 47, pages 1445–1456, 2001.

[BCC11] François Bolley, José A. Cañizo, and José A. Carrillo. Stochastic mean-field limit: non-Lipschitz forces and swarming. *Math. Models Methods Appl. Sci.*, 21(11):2179–2210, 2011.

[BGV07] François Bolley, Arnaud Guillin, and Cédric Villani. Quantitative concentration inequalities for empirical measures on non-compact spaces. *Probab. Theory Related Fields*, 137(3-4):541–593, 2007.

[BH77] W. Braun and K. Hepp. The Vlasov dynamics and its fluctuations in the $1/N$ limit of interacting classical particles. *Comm. Math. Phys.*, 56(2):101–113, 1977.

[BHJ10] J. Barré, M. Hauray, and P. E. Jabin. Stability of trajectories for N-particle dynamics with a singular potential. *Journal of Statistical Mechanics: Theory and Experiment*, 7:5–+, July 2010.

[BJ08] J. Barré and P. E. Jabin. Free transport limit for N-particles dynamics with singular and short range potential. *J. Stat. Phys.*, 131(6):1085–1101, 2008.

[Boi11] E. Boissard. *Problèmes d'interaction discret-continu et distances de Wasserstein.* PhD thesis, Université de Toulouse III, 2011.

[CDPJ08] Thierry Champion, Luigi De Pascale, and Petri Juutinen. The ∞-Wasserstein distance: local solutions and existence of optimal transport maps. *SIAM J. Math. Anal.*, 40(1):1–20, 2008.

[Deh00] W. Dehnen. A Very Fast and Momentum-conserving Tree Code. *The Astrophysical Journal*, 536:L39–L42, June 2000.

[DL89] Ronald J. DiPerna and Pierre-Louis Lions. Ordinary differential equations. *Invent. Math*, 98:511–547, 1989.

[Dob79] R. L. Dobrušin. Vlasov equations. *Funktsional. Anal. i Prilozhen.*, 13(2):48–58, 96, 1979.

[DY95] V. Dobrić and J. E. Yukich. Asymptotics for transportation cost in high dimensions. *J. Theoret. Probab.*, 8(1):97–118, 1995.

[Gao03] Fuqing Gao. Moderate deviations and large deviations for kernel density estimators. *J. Theoret. Probab.*, 16(2):401–418, 2003.

[GH91] Jonathan Goodman and Thomas Y. Hou. New stability estimates for the 2-D vortex method. *Comm. Pure Appl. Math.*, 44(8-9):1015–1031, 1991.

[GHL90] Jonathan Goodman, Thomas Y. Hou, and John Lowengrub. Convergence of the point vortex method for the 2-D Euler equations. *Comm. Pure Appl. Math.*, 43(3):415–430, 1990.

[GLV91] Keshab Ganguly, J. Todd Lee, and H. D. Victory, Jr. On simulation methods for Vlasov-Poisson systems with particles initially asymptotically distributed. *SIAM J. Numer. Anal.*, 28(6):1574–1609, 1991.

[GV89] Keshab Ganguly and H. D. Victory, Jr. On the convergence of particle methods for multidimensional Vlasov-Poisson systems. *SIAM J. Numer. Anal.*, 26(2):249–288, 1989.

[Hau04] M. Hauray. On Liouville transport equation with force field in BV_{loc}. *Comm. Partial Differential Equations*, 29(1-2):207–217, 2004.

[Hau09] Maxime Hauray. Wasserstein distances for vortices approximation of Euler-type equations. *Math. Models Methods Appl. Sci.*, 19(8):1357–1384, 2009.

[HJ07] Maxime Hauray and Pierre-Emmanuel Jabin. N-particles approximation of the Vlasov equations with singular potential. *Arch. Ration. Mech. Anal.*, 183(3):489–524, 2007.

[HM12] Maxime Hauray and Stéphane Mischler. On kac's chaos and related problems. To be published soon., 2012.

[Hor93] E. Horst. On the asymptotic growth of the solutions of the Vlasov-Poisson system. *Math. Methods Appl. Sci.*, 16(2):75–86, 1993.

[Kac56] M. Kac. Foundations of kinetic theory. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, vol. III*, pages 171–197, Berkeley and Los Angeles, 1956. University of California Press.

[Loe06] Grégoire Loeper. Uniqueness of the solution to the Vlasov-Poisson system with bounded density. *J. Math. Pures Appl. (9)*, 86(1):68–79, 2006.

[LP91] P.-L. Lions and B. Perthame. Propagation of moments and regularity for the 3-dimensional Vlasov-Poisson system. *Invent. Math.*, 105(2):415–430, 1991.

[MM11] S. Mischler and C. Mouhot. Kac's Program in Kinetic Theory. *ArXiv e-prints*, July 2011.

[NW80] H. Neunzert and J. Wick. The convergence of simulation methods in plasma physics. In *Mathematical methods of plasmaphysics (Oberwolfach, 1979)*, volume 20 of *Methoden Verfahren Math. Phys.*, pages 271–286. Lang, Frankfurt, 1980.

[Pfa92] K. Pfaffelmoser. Global classical solutions of the Vlasov-Poisson system in three dimensions for general initial data. *J. Differential Equations*, 95(2):281–303, 1992.

[Saa73] Donald G. Saari. Improbability of collisions in Newtonian gravitational systems. II. *Trans. Amer. Math. Soc.*, 181:351–368, 1973.

[Saa77] Donald G. Saari. A global existence theorem for the four-body problem of Newtonian mechanics. *J. Differential Equations*, 26(1):80–111, 1977.

[Sch91] Jack Schaeffer. Global existence of smooth solutions to the Vlasov-Poisson system in three dimensions. *Comm. Partial Differential Equations*, 16(8-9):1313–1335, 1991.

[Sch96] Steven Schochet. The point-vortex method for periodic weak solutions of the 2-D Euler equations. *Comm. Pure Appl. Math.*, 49(9):911–965, 1996.

[Spo91] Herbert Spohn. *Large scale dynamics of interacting particles*. Springer Verlag, New York, 1991.

[Szn91] Alain-Sol Sznitman. Topics in propagation of chaos. In *École d'Été de Probabilités de Saint-Flour XIX—1989*, volume 1464 of *Lecture Notes in Math.*, pages 165–251. Springer, Berlin, 1991.

[Var58] V. S. Varadarajan. On the convergence of sample probability distributions. *Sankhyā*, 19:23–26, 1958.

[Vil03] Cédric Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.

[Wol00] Stephen Wollman. On the approximation of the Vlasov-Poisson system by particle methods. *SIAM J. Numer. Anal.*, 37(4):1369–1398 (electronic), 2000.

[Xia92] Zhihong Xia. The existence of noncollision singularities in Newtonian systems. *Ann. of Math. (2)*, 135(3):411–468, 1992.