PROPAGATION OF CHAOS FOR THE VLASOV-POISSON-FOKKER-PLANCK SYSTEM IN 1D

MAXIME HAURAY
I2M
AMU, Centrale Marseille CNRS
Marseille, France

SAMIR SALEM
CEREMADE, Université Paris Dauphine
Paris, France

(Communicated by Francois Golse)

ABSTRACT. We consider a particle system in 1D, interacting via repulsive or attractive Coulomb forces. We prove the trajectorial propagation of molecular chaos towards a nonlinear SDE associated to the Vlasov-Poisson-Fokker-Planck equation. We obtain a quantitative estimate of convergence in the mean in MKW metric of order one, with an optimal convergence rate of order \( N^{-1/2} \). We also prove some exponential concentration inequalities of the associated empirical measures. A key argument is a weak-strong stability estimate on the (nonlinear) VPFP equation, that we are able to adapt for the particle system in some sense.

1. Introduction. We consider here a one dimensional system of \( N \) particles, with position \( X^N_i \in \mathbb{R} \) and velocity \( V^N_i \in \mathbb{R} \), interacting via the Poisson interaction, and submitted to independent Brownian noises and friction. The associated system of Stochastic Differential Equation (SDE) is the following:

\[
\begin{align*}
    dX^N_{i,t} &= V^N_{i,t} \, dt, \\
    dV^N_{i,t} &= \left( \frac{1}{N} \sum_{j=1}^{N} K(X^N_{i,t} - X^N_{j,t}) - V^N_{i,t} \right) \, dt + \sqrt{2} dB_{i,t},
\end{align*}
\]

where the \((B_{i,t})_{t \geq 0}\) are independent Brownian motions. The interaction kernel is defined (everywhere) by

\[
    K(x) := \pm \frac{1}{2} \operatorname{sign}(x) = \pm \frac{1}{2} \begin{cases} 
        1 & \text{if } x > 0, \\
        0 & \text{if } x = 0, \\
        -1 & \text{if } x < 0.
    \end{cases}
\]

The case \( K = \frac{1}{2} \operatorname{sign} \) corresponds to the repulsive case, while the interaction is attractive when \( K = -\frac{1}{2} \operatorname{sign} \). This two cases lead of course to very different dynamics, but concerning the propagation of chaos in finite time the sign of \( K \) is not very relevant, so we will handle both cases in the same way.

2010 Mathematics Subject Classification. Primary: 35Q83, 35Q84, 60K35; Secondary: 65C35.

Key words and phrases. Vlasov-Poisson equation, Propagation of chaos, concentration inequalities, quantitative estimates, weak-strong stability.
Existence and uniqueness in law for the particle system. We will state a result of existence and uniqueness in law of the solution of the process (1). The proof is a simple application of the Cameron-Martin-Girsanov Theorem, but to an appropriate system equivalent to (1), since the diffusion in that SDE is not full. This result in law will be enough for our aim, which is to state some result of propagation of chaos. The strong existence and uniqueness is a more difficult problem, which is to the best of our knowledge not covered by the actual literature [7].

Existence, and weak-strong stability of the limit nonlinear SDE. When \( N \) is large, and if the initial conditions \((X_{i,0}^N, V_{i,0}^N)_{1 \leq i \leq N}\) are independent (or asymptotically independent), then it is expected that the particles behave almost like \( N \) independent copies of the following non-linear limit process

\[
dY_t = W_t \, dt, \quad dW_t = \mathbb{E}_{Y_t} \left[ K(Y_t - \bar{Y}_t) \right] \, dt - W_t \, dt + \sqrt{2} \, dB_t, \tag{3}
\]

where \((B_t)_{t \geq 0}\) is a Brownian motion (independent of the rest) and \( \bar{Y}_t \) is an independent copy of \( Y_t \). The family of time marginals \( f_t = \mathcal{L}(Y_t, W_y) \) is a solution of the 1D Vlasov-Poisson-Fokker-Planck equation

\[
\partial_t f_t + v \, \partial_x f_t + (\rho_t \ast K) \partial_v f_t = \partial_v (\partial_v f_t + v \, f_t), \tag{4}
\]

where \( \rho_t = \int f_t \, dv \) is the law of \( Y_t \).

We will prove two results on that non-linear limit process or equation, both in Theorem 2.3 and Corollary 1:

- The existence of solution to the process (3) and to the VPFP equation (4), provided that there is a uniform control (exponential or polynomial) on the tail in velocity of the initial distribution;
- A weak-strong stability result, ensuring that when there exists a solution to (3) or (4) with a bounded density in position, i.e. \( \rho_t := \mathcal{L}(Y_t) \) bounded for all time \( t > 0 \), then it is the unique solution in a much larger class: the only condition being that the order one moment in position and velocity remains finite along time.

Quantitative propagation of chaos and concentration inequalities. Weak-strong uniqueness or stability results are very interesting for the propagation of chaos. In fact, a very standard procedure to prove propagation of chaos, is to prove that:

- the solutions of the particle systems form a tight sequence (compactness);
- the accumulation points are solution to the expected limit problem, possibly with some extra a priori estimates;
- uniqueness of the solutions holds, possibly under the extra a priori estimates.

That method goes back at least to [20], and is for instance used is recent works of the first authors and collaborators [12, 11]. When using the tightness-consistency-uniqueness method, the tightness of the particle system and the consistency are usually more simple to prove than the uniqueness of the limit process, as for instance in [12, 11]. Since we have here a very good uniqueness result on the limit problem, that strategy would be quite simple to use.

However, the situation is even better. In fact the proof of the weak-strong stability result on the limit equation could be adapted to get some stability between solutions of the particle system (1) and \( N \) independent copies of the limit system (3) and that will allow to state a (optimal) quantitative version of the propagation of chaos. We will prove two results:
• An estimate of propagation of chaos in the mean using Monge-Kantorovich-Wasserstein distance of order one (Theorem 2.5). Our proof rely on the fact that empirical measures associated to solutions of the particle system 1 are solution to the limit nonlinear SDE 3 plus a reminder which is roughly a martingale with distribution value. This allows to compare the interacting and non-interacting systems using standard arguments of martingale theory, namely Doob maximal inequality, in a quite subtle way.

• An exponential concentration inequality of the solution to the particle system (1) around the (unique) solution of the nonlinear SDE (Theorem 2.6). Our proof rely roughly on a perturbed version of the weak-strong stability result of Theorem 2.3, where the infinite norm of the distribution in position is replaced by some discrete infinite norm (See Definition 2.7). That perturbed version is in fact not stated for the nonlinear SDE, but directly applied in order to control the distance between the interacting particle system and the system of independent copies. And concentration inequality on the discrete infinite norm (in fact its supremum in time along independent copies of solution to a given SDE, obtained in Proposition 2) implies the requested concentration inequalities on the particle system.

Some related works. The literature on the convergence of particle systems towards non linear mean-field models is quite huge, so we will restrict ourselves to one dimensional models. The usual strategy, valid for smooth interaction, is well explained in Lecture notes by Sznitman [19]. In [3] Cépa and Lépingle prove the propagation of chaos for the Dyson model: a one dimensional first order model (i.e. without velocities) with a strongly singular interaction \( K(x) \sim |x|^{-1} \) modeling the behavior of eigenvalues of large hermitian matrices. Their proof relies on the use of maximal monotone operators. Recently, that convergence result was extended to similar systems with even stronger interaction \( K(x) \sim |x|^{-1-\alpha} \) with \( \alpha \in [0,1) \), by Berman and Önnheim [1] using Wasserstein gradient flows.

For second order models (involving positions and velocities), our result is to the best of our knowledge the two first result of propagation of chaos with the Poisson singularity in the stochastic case. The second one was obtained in the same time by P.-E. Jabin and Z. Wang in [15]. They precisely prove a much more general result: the propagation of chaos for interacting particle system (with or without independent noise) with bounded interaction force. That result is an average result, stated in relative entropy, with the optimal rate of convergence, but requires some regularity of the solution of the limit equation.

In the deterministic case, i.e. when the system under study is (1) without the Brownian motions (and the friction), then the mean field limit was proved originally by [21], and then by [5] as a special case of semi-geostrophic equations, and again by the first author [13]. Remark that the result of [15] also apply to that deterministic case.

2. Main results.

Well-posedness of the particle system. Since the force field is only of bounded variation and not Lipschitz, the standard theory for SDE does not apply. Neither does the theory of existence and uniqueness developed for uniformly elliptic diffusions (see for instance [22, 6, 2] among many others) since the diffusion act here only on the velocities. We also mention the very recent adaptation [7] of the previous works to the kinetic case, which still require too much regularity to be applied here (\( F \in W^{s,p} \), \( s > 2/3 \) and \( p > 6 \)). However, due to the particular geometry of the
problem, we can still get existence and uniqueness in law. The precise result is the following:

Theorem 2.1. For any $N \geq 2$, and any (deterministic) initial condition $(X_{i,0}^N, V_{i,0}^N)_{i \leq N} \in \mathbb{R}^{2N}$, weak existence and uniqueness in law hold for SDE (1).

Since (1) is linear, it also implies weak existence and uniqueness for any random initial condition. Theorem 2.1 is proved in Section A using the following strategy: we reformulate (1) in an SDE with memory involving the $(V_{i,t}^N)_{i \leq N,t \geq 0}$ only (simply because the $X_{i,t}^N$ are time integrals of the $V_{i,t}^N$). Then we apply to that new (non-markovian) SDE a standard technique relying on the Girsanov’s theorem. To the best of our knowledge, the strong existence and uniqueness of solution to that system are yet unknown, and we were not able to prove it.

A non-linear SDE related to the Vlasov-Poisson-Fokker-Planck equation. When the number of particles is large, and under the assumption that two particles are yet unknown, and we were not able to prove it.

For any $t \geq 0$, the $i$-th moment of the initial data with a uniform control on the velocity tails, and even a weak-strong stability estimate (here “strong solution” means that the law $\mathcal{L}(Y_t)$ of $Y_t$ remains uniformly bounded in time). Before stating the precise result, we introduce two useful norms:

Definition 2.2. For any $\lambda > 0$, and any $\gamma > 0$, we define for any $f \in L^\infty(\mathbb{R}^2)$ the two following norms:

\[
\| f \|_{e,\lambda} := \text{ess-sup}_{(x,v) \in \mathbb{R}^2} f(x,v)e^{\lambda |v|} \in [0, +\infty], \\
\| f \|_{p,\gamma} := \text{ess-sup}_{(x,v) \in \mathbb{R}^2} f(x,v)\langle v \rangle^\gamma \in [0, +\infty],
\]

where $\langle v \rangle = \sqrt{1 + v^2}$, so that equivalently $\| f \|_{p,\gamma} = \text{ess-sup}_{(x,v) \in \mathbb{R}^2} f(x,v) (1 + v^2)^{-\gamma/2}$, and the essential supremum are taken with respect to the Lebesgue measure.

Theorem 2.3. (i) Existence of strong solutions:
Assume that the law $f_0$ of the initial condition $(Y_0, W_0)$ satisfies $f_0 \in \mathcal{P}_1(\mathbb{R}^2) \cap L^1(\mathbb{R}^2)$ and also $\| f_0 \|_{e,\lambda} < +\infty$, for some $\lambda > 0$ or $\| f_0 \|_{p,\gamma} < +\infty$, for some $\gamma > 1$. Then, given any Brownian motion $(B_t)_{t \geq 0}$ (independent of the initial conditions) there exists a solution $(Y_t, W_t)$ to (3) with initial condition $(Y_0, W_0)$, and its law at time $t \geq 0$: $f_t = \mathcal{L}(Y_t, W_t)$ satisfies either

\[
\| f_t \|_{e,\lambda} \leq 2 e^{t+\lambda + \frac{\lambda^2}{2}} \| f_0 \|_{e,\lambda - t}, \quad \text{or} \quad \| f_t \|_{p,\gamma} \leq C_\gamma e^{\gamma t} \| f_0 \|_{p,\gamma},
\]

where $C_\gamma$ is a constant depending only on $\gamma$.

(ii) Weak/strong stability for solution with bounded density (in position).

If $(Y_1^t, W_1^t)$ and $(Y_2^t, W_2^t)$ are two solutions to (3) built on the same probability space with the same Brownian motion $(B_t)_{t \in \mathbb{R}}$, if the density $\rho_t^1 = \mathcal{L}(Y_t^1)$ is uniformly bounded at any time ($\forall t \geq 0$, $\| \rho_t \|_{\infty} < \infty$), then the following stability estimate holds:

\[
E[|Y_t^1 - Y_t^2| + |W_t^1 - W_t^2|] \leq e^{8(t + f_0^1 \| \rho_t^1 \|_{\infty} ds)} E[|Y_0^1 - Y_0^2| + |W_0^1 - W_0^2|].
\]
The proof of the weak-strong stability estimate relies on the crucial Lemma 3.1 that allows to control the singularity of the force, when comparing the evolution of two solutions, assuming that only one of them has a bounded density in position. That Lemma was already used in [13], to get similar results for the associated deterministic particle system: i.e. particles interacting via the same kernel (2) but without noise and friction. The proof of the propagation of the norm \(\|f\|_{e,\lambda}\) and \(\|f\|_{p,\gamma}\) is based on a probabilistic version of the method of characteristic, that is also a simple case of the Feymann-Kac formula. All of this is proved in Section 3. Once the propagation of these norms is known, the proof of the existence of solutions relies on an usual approximation procedure, and is postponed to the Appendix C.

The stability result on the Vlasov-Poisson-Fokker-Planck equation. The stability results on the process (3) simply translate on the associated Fokker-Planck or Kolmogorov forward equation, which is here the Vlasov-Poisson-Fokker-Planck equation (4). As the kernel \(K\) is bounded and defined everywhere by (2), remark that very few hypothesis are required to define solutions to (4) in the sense of distribution: if \(f \in L^1_{loc}(\mathbb{R}^+, \mathcal{P}(\mathbb{R}^2))\), where \(\mathcal{P}(\mathbb{R}^2)\) stand for the space of probability, then all the terms appearing in (4) define a distribution. Theorem 2.3 has the following consequence on VPFP equation:

**Corollary 1** (of Theorem 2.3). (i) Existence of strong solution. Let \(f_0 \in \mathcal{P}(\mathbb{R}^2) \cap L^1(\mathbb{R}^2)\) with a finite order one moment: \(\int (|x| + |v|) f_0(\,dx, dv)\), and satisfying either \(\|f_0\|_{e,\lambda} < +\infty\), for some \(\lambda > 0\) or \(\|f_0\|_{p,\gamma} < +\infty\), for some \(\gamma > 1\). Then, there exists a solution \(f_t\) to (4) with initial condition \(f_0\), and it satisfies (7).

(ii) Weak/strong uniqueness for solution with bounded density (in position).

If \(f^1_t\) and \(f^2_t\) are two solutions to (4) and if \(\rho^1_t = \int f^1_t(\,dv)\) is uniformly bounded for any time \(t \geq 0\), then the following stability estimate holds:

\[
W_1(f^1_t, f^2_t) \leq e^{\varepsilon \int_0^t \rho^1_s \, ds} W_1(f^1_0, f^2_0),
\]

where \(W_1\) stands for the order one Monge-Kantorovich-Wasserstein distance.

This corollary is proved in Appendix D. It is a direct consequence of Theorem 2.3, and of the fact that a weak solution \(f_t\) to the VPFP1D equation (4), can always be represented as the time marginals of a process \((Y_{i,t}, W_{i,t})_{i \geq 0}\) solution to (3), as it was first notice in [9, Theorem 2.6].

The quantitative propagation of chaos in the mean. When comparing a solution \((X^N_{i,t}, V^N_{i,t})_{i \leq N}\) of the particle system (1) with the limit process (3) and its associated Fokker-Planck equation (4), a very natural strategy (which goes back to McKean [17]) is to introduce \(N\) independent copies \((Y^N_{i,t}, W^N_{i,t})_{i \leq N}\) of the limit nonlinear SDE (3), constructed with the same Brownian motion as the \((X^N_{i,t}, V^N_{i,t})\), and with initial conditions coupled in an optimal way. In our case, we are able to prove a sharp estimate on the average distance between these two systems.

To state our result properly, we recall that by definition exchangeable random variables have a law that is invariant under permutation, and that chaotic sequences of r.v. are defined as follows:

**Definition 2.4.** Let \(f\) be a probability on \(\mathbb{R}^2\). A sequence \(\{(X^N_{i,t}, V^N_{i,t})_{i \leq N}\}_{N \in \mathbb{N}}\) of exchangeable random variables is said to be \(f\)-chaotic, if one of the equivalent conditions below is satisfied:

i) \(\forall k \in \mathbb{N}, \quad \mathcal{L}(\{(X^N_{i,t}, V^N_{i,t})_{i \leq k}\}_{N \in \mathbb{N}}) \overset{w}{\underset{N \to \infty}{\to}} f^{\otimes k}\),
ii) \( \mathcal{L}((X^N_1, V^N_1), (X^N_2, V^N_2)) \xrightarrow{w} f \otimes f \),

iii) \( \mu^N := \frac{1}{N} \sum_{i=1}^{N} \delta_{(X^N_i, V^N_i)} \xrightarrow{L} f \).

The subscript \( w \) on the arrow, refers to the weak convergence of measure, the subscript \( L \) to the convergence in law. If \( Y \) is a random variable of law \( f \), we will equivalently say that a sequence is \( f \)-chaotic or \( Y \)-chaotic. We refer to [19] for the equivalence of the three conditions above, and to [14] for a quantitative version of that equivalence.

We also state the following proposition, that reformulate propagation of chaos in term of coupling. It is a consequence of [14, Theorem 1.2]

**Proposition 1.** Assume that \( ((X^N_i, V^N_i)_{i \leq N})_{N \in \mathbb{N}} \) is a sequence of exchangeable random variables with uniformly bounded order two moments: \( \sup_{N \in \mathbb{N}} E[|X^N_1|^2 + |V^N_1|^2] < +\infty \). Let \( f \) be a probability on \( \mathbb{R}^2 \). Then, the following statements are equivalent:

i) The sequence \( ((X^N_i, V^N_i)_{i \leq N})_{N \in \mathbb{N}} \) is \( f \)-chaotic;

ii) \( \mathbb{E}[W_t(\mu^N, f)] \xrightarrow{N \to \infty} 0 \), where \( \mu^N = \frac{1}{N} \sum_{i=1}^N \delta_{(X^N_i, V^N_i)} \) is the associated empirical measure;

iii) \( \min_{\text{coupling}} \mathbb{E} \left[ \frac{1}{N} \sum_i |X^N_i - Y_i| + |V^N_i - W_i| \right] \xrightarrow{N \to \infty} 0 \), where the minimum is taken on all the exchangeable coupling with \( (Y_i, V_i) \) i.i.d. with common law \( f \).

The above Proposition allows to state a precise result of propagation of molecular chaos. We emphasize that this result is a true result of propagation: it does not apply only to i.i.d. initial conditions, but to any chaotic initial conditions (with finite second order moment). However, the general case is somewhat more technical, so we warn the reader that the following theorem is simpler to understand if we only consider the case of i.i.d. initial conditions.

**Theorem 2.5.** Let \( f_0 \in \mathcal{P}(\mathbb{R}^2) \) with finite order two moment: \( \int (|x|^2 + |v|^2) f_0 (dx, dv) < \infty \), and such that there exists a (necessary unique by Corollary 1) solution \( f_t \) to (4) with initial condition \( f_0 \) satisfying \( \int_0^t \|\rho_s\|_\infty ds < +\infty \) for any time \( t \geq 0 \), where \( \rho_s \) stands for the density in position: \( \rho_s(x) := \int f_s(x, dv) \). We also denote by \( (Y_t, W_t)_{t \geq 0} \) the unique solution to (3) such that \( \mathcal{L}(Y_0, W_0) = f_0 \).

Let \( (X^N_{1,0}, V^N_{1,0})_{i \leq N} \) be a sequence of \( f_0 \)-chaotic random variable with uniformly bounded order two moment: \( \sup_{N \in \mathbb{N}} E[|X^N_{1,0}|^2 + |V^N_{1,0}|^2] < +\infty \). By Theorem 2.1, we may find a probability space together with a \( N \)-dimensional Brownian motion and a process \( (X^N_{i,t}, V^N_{i,t})_{i \leq N} \) solution to (1). Then, the sequence \( ((X^N_{i,t}, V^N_{i,t})_{i \leq N})_{n \in \mathbb{N}} \) is \( (Y_t, W_t)_{t \geq 0} \) chaotic.

More precisely, by standard arguments, we may also construct on that probability space \( N \) i.i.d. copies \( (Y^N_{i,t}, W^N_{i,t})_{i \leq N} \) of the solutions of (3) with the same Brownian motion, and with initial conditions of law \( f_0 \), coupled with \( (X^N_{1,0}, V^N_{1,0})_{i \leq N} \) in an exchangeable way. Then the following estimate holds:

\[
\mathbb{E} \left[ \sup_{s \in [0,t]} |X^N_{1,s} - Y^N_{1,s}| + |V^N_{1,s} - W^N_{1,s}| \right] \leq \left( \mathbb{E} \left[ |X^N_{1,0} - Y^N_{1,0}| + |V^N_{1,0} - W^N_{1,0}| \right] + \frac{9}{\sqrt{N}} \right) e^{(t+8 f_0^0 \|\rho_s\|_\infty ds)}.
\]
In particular, if \((X_{i,0}^N, V_{i,0}^N)_{i \leq N}\) are i.i.d. with law \(f_0\), then we choose initially \(X_{i,0}^N = Y_{i,0}^N, V_{i,0}^N = W_{i,0}^N\) and get:

\[
\mathbb{E} \left[ \sup_{s \leq t} \left| X_{i,s}^N - Y_{i,s}^N \right| + \left| V_{i,s}^N - W_{i,s}^N \right| \right] \leq \frac{9}{\sqrt{N}} e^{(t+8) \| \rho_s \|_{\infty}} ds.
\]

**Remark 1.** Thanks to Corollary 1, the last hypothesis on \(f_0\) in Theorem 2.5 is satisfied if \(\|f_0\|_{\varepsilon, \lambda} < \infty\) for some \(\lambda > 0\), or if \(\|f_0\|_{\rho, \gamma} < \infty\) for some \(\gamma > 1\).

That result is proved in Section 4: as for the stability part of Theorem 2.3, its proof strongly relies on Lemma 3.1 (the “rope argument”), that allows to write a kind of Grönwall inequality, which contains a fluctuation term. That fluctuation term is handled by “Poissonization” (replacement of the empirical measure by a very close Poisson Random Measure), which allows to use standard results on martingales, in a quite subtle way.

Using standard results on the convergence on empirical measures toward their mean [10, Theorem 1], we may deduce convergence results between the empirical measure of the particle system and the expected limit profile.

**Corollary 2 (of Theorem 2.5).** Under the same assumptions than in Theorem 2.5, and assuming moreover that \(\int (|x| + |v|)^q f_0(dx, dv) < \infty\) for some \(q > 2\) we obtain that for any time \(t \geq 0\),

\[
\mathbb{E}[W_1(\mu_t^N, f_t)] \leq C_t \left( \frac{\ln(1+N)}{\sqrt{N}} + \mathbb{E}[W_1(\mu_0^N, f_0)] \right),
\]

for some constant \(C_t > 0\) depending on \(t, q\) and the moments of \(f_0\).

**Exponential concentration inequalities for the particle system.** In the case where the initial conditions \((X_{i,0}^N, V_{i,0}^N)_{i \leq N}\) are i.i.d., we also prove concentration inequalities for the solutions of the particle system (1), precisely:

**Theorem 2.6.** Let \(f_0 \in \mathcal{P}(\mathbb{R}^2)\) with some finite exponential moment: \(\int e^{\lambda(|x|+|v|)} f_0(dx, dv) < \infty\) for some \(\lambda > 0\), and such that there exists a (necessary unique by Corollary 1) solution \(f_t\) to (4) with initial condition \(f_0\) satisfying \(\kappa_t := \sup_{s \leq t} \| \rho_s \|_{\infty}\) ds < \(t \infty\) for any time \(t \geq 0\), where \(\rho_s\) stands for the density in position: \(\rho_s(x) := \int f_s(x, dv)\). Let \((X_{i,0}^N, V_{i,0})_{i \leq N}\) be a sequence of \(N\) i.i.d random variables with common law \(f_0\).

By Theorem 2.1, we may find a probability space together with a \(N\)-dimensional Brownian motion and a process \((X_{i,t}^N, V_{i,t}^N)_{i \leq N}\) solution to (1). By standard arguments, we may also construct on that probability space \(N\) copies \((Y_{i,t}^N, W_{i,t}^N)_{i \leq N}\) of the solutions of (3) with the same initial conditions \((X_{i,0}^N, V_{i,0})_{i \leq N}\) and same Brownian motions.

Then the following concentration inequality holds for any \(\varepsilon\) satisfying \(\lambda N^{-1/2} \leq \varepsilon \leq (5\kappa_t \wedge 1) \min \{\frac{1}{16}, \frac{\lambda}{2}, \lambda^{-2}\}:

\[
\mathbb{P} \left( \frac{1}{N} \sum_{i=1}^N \sup_{s \leq t} \left| X_{i,s}^N - Y_{i,s}^N \right| + \left| V_{i,s}^N - W_{i,s}^N \right| \geq B_t \varepsilon \right) \leq (t + \varepsilon) \left( A_t + A'_t \sqrt{N} \varepsilon \right) N^2 e^{-2N \varepsilon^2},
\]

where the three constants \(A_t, A'_t, B_t\) depend on \(t, \lambda\) and the initial law \(f_0\) (See the end of Section 6 for more precise values).
Remark 2. Thanks to Corollary 1, the hypothesis on \( f_0 \) are satisfied if \( \|f_0\|_{c,\lambda'} < \infty \) for some \( \lambda' > \lambda \).

The proof of Theorem 2.6 relies on a different technique than the one of Theorem 2.5. Here, we rather use another perturbed version of the weak-strong stability estimate (8), applied to the N particle system and involving discrete infinite norms. Together with exponential concentration inequalities on discrete infinite norms of empirical measures, and on some fluctuation terms appearing naturally in the perturbed weak-strong estimate, it leads to above result.

Using deviation upper bounds for the approximation of probability measure by random empirical measures associated to i.i.d sample, as for instance in [10, Theorem 2], we could also obtain concentration inequalities between the empirical measures, and on some fluctuation terms appearing naturally in the perturbed weak-strong estimate, it leads to above result.

Corollary 3. Under the same assumption and notation as in Theorem 2.6, then for any \( T \geq 0 \), and any \( \alpha > 1 \), there exists two constants \( C_1, C_2 \) such that for any \( \varepsilon \geq 0 \) satisfying \( \lambda N^{-12} \leq \varepsilon \leq (5\kappa_4 \wedge 1) \min\left( \frac{1}{16}, \lambda, \lambda^{-2} \right) \), we have:

\[
\sup_{t \in [0,T]} \mathbb{P} \left[ \left| W_t^N (\mu_{X,t}^N, f_t) \right| \geq \varepsilon \right] \leq C_1 \left( N^4 e^{-C_2 N^2} + e^{-C_2 N^2} e^{-C_2 N^2 (1-\ln \varepsilon)^2} \right).
\]

Exponential concentration for discrete infinite norms. We start by defining the discrete infinite norm of a probability measure.

Definition 2.7. For any \( \varepsilon > 0 \), any \( f \in \mathcal{P}(\mathbb{R}) \), we define the infinite norm of \( f \) at scale \( \varepsilon \), denoted \( \|f\|_{\infty,\varepsilon} \) by:

\[
\|f\|_{\infty,\varepsilon} := \sup_{x \in \mathbb{R}} \frac{f(x - \varepsilon, x + \varepsilon)}{2\varepsilon} = \left\| f * \frac{1}{2\varepsilon} \mathbf{1}_{[-\varepsilon,\varepsilon]} \right\|_{\infty}.
\]

Remark 3. The quantity \( 2\varepsilon \|f\|_{\infty,\varepsilon} \) is usually referred as the concentration function of the measure \( f \) (a notion due to P. Levy). See for instance [18, Definition 3.1 on p. 165]. However, we use it for rather different purposes, and prefer to introduce a slightly different definition (with the factor \( (2\varepsilon)^{-1} \)), in order to emphasize on the relation with the infinite norm.

One of the key ingredients of the proof of Theorem 2.6 is a deviation inequality on time integral (or supremum) of discrete infinite norms for the \( \rho_t^N := N^{-1} \sum_i \delta_{Y_t^{N,i}} \), where the \( Y_t^{N,i} \) are \( N \) i.i.d. r.v. solutions to a given SDE. Such a result have a interest by itself so we state it below.

Proposition 2. Let \( (Y_t, W_t)_{t \in \mathbb{N}} \) be \( N \) i.i.d copies of a solution of (3) driven by independent Brownian motions \( (B_{i,t})_{t \in \mathbb{N}, t \geq 0} \), and assume that for \( t \geq 0 \):

- the common initial condition has a law \( f_0 \) which admits a finite exponential moment (in position and velocity) for some \( \lambda > 0 \): \( \mathbb{E} \left[ e^{\lambda |W_t|} + |W_t| \right] < +\infty \). In particular we denote \( c_\lambda := \frac{5}{2} + \frac{1}{\lambda} \ln \mathbb{E} \left[ e^{\lambda |W_0|} \right] \);
- \( \kappa_t := \sup_{0 \leq s \leq t} \|\rho_s\|_{\infty} < +\infty \) where \( \rho_s \) stands for the time marginal of \( Y_t \) at time \( s \).

Then provided that \( \lambda N^{-1/2} \leq \varepsilon \gamma \leq 5\kappa_t \min\left( \frac{1}{16}, \lambda, \lambda^{-2} \right) \), there exists a constant \( C_t \) depending on \( \kappa_t, \lambda \) and the moments of \( f_0 \) such that:

\[
\mathbb{P} \left( \sup_{s \in [0,t]} \|\rho_s^N\|_{\infty,\varepsilon} \geq \kappa_t + \gamma \right) \leq C_t \left( 1 + t \frac{2\kappa_t + \gamma}{\lambda} (e_\lambda + \sqrt{N} \varepsilon \gamma) \right) N^2 e^{-2N(\varepsilon \gamma)^2}.
\]
Essentially, the bound behave like $N^{3/2}e^{-2N(\epsilon \gamma)^2}$, in the most interesting range, when $\epsilon \gamma \sim N^{-1/2}$.

**Plan of the paper.** The paper is organized in the following way. In Section 3, we focus on the nonlinear limit SDE and prove the uniqueness part of Theorem 2.3 and the propagation of the norms defined in Definition 2.2. The remaining part of the proof of Theorem 2.3 and the proof of Corollary 1 are postponed respectively to Appendix C and D. Section 4 is devoted to the proof of Theorem 2.5 on the propagation of chaos in the mean, Section 5 to the proof of Proposition 2 and Section 6 to the proof of Theorem 2.6 on the exponential concentration. We prove Theorem 2.1 (weak existence and uniqueness of the particle system) in section A. An useful proposition about propagation of moments (Proposition 3) is proved in Appendix B. A useful regularity lemma is proved in the last Appendix E.

3. Proof of the uniqueness part of Theorem 2.3.

3.1. Weak-strong stability and uniqueness.

A key bound. Here we will use an argument that we call “rope argument”, which has already been used in [13] to treat the propagation of chaos for deterministic VP1D equation. We choose this name in order to illustrate the argument (which relies on the triangular inequality in a simple way) : if we imagine that two coupled particles are linked by a rope, then their relative displacement is limited. The argument consists in noticing that

$$K(x - \bar{x}) - K(y - \bar{y}) = 0,$$

as soon as $|y - \bar{y}| > |x - y| + |\bar{x} - \bar{y}|$. In fact, in that case $x - \bar{x}$ and $y - \bar{y}$ have necessarily the same sign. It is also interesting to replace the later condition by the stronger one $|y - \bar{y}| > 2 \max(|x - y|, |\bar{x} - \bar{y}|)$. Since $K$ is bounded by 1/2, it implies the following bound, that we will use many times in the sequel

$$|K(x - \bar{x}) - K(y - \bar{y})| \leq \mathbf{1}_{|y - \bar{y}| \leq 2|x - y| + \mathbf{1}_{|y - \bar{y}| \leq 2|\bar{x} - \bar{y}|}. \quad (10)$$

That bound is one-sided: the indicator functions in the r.h.s. take only $y - \bar{y}$ as argument, and this allows to prove a simple but important Lemma:

**Lemma 3.1.** Assume that $(X, Y)$ is a random couple of real numbers and that $(\bar{X}, \bar{Y})$ is an independent copy of that couple.

(i) Then,

$$\mathbb{E}[|K(X - \bar{X}) - K(Y - \bar{Y})|] \leq 8 \min(\|\rho_X\|_\infty, \|\rho_Y\|_\infty)\mathbb{E}[|X - Y|],$$

where $\rho_X, \rho_Y$ denote respectively the density (with respect to the Lebesgue measure) of the law of $X$ and $Y$.

(ii) For $\varepsilon > 0$, we also have a similar estimate involving the discrete infinite norms $\|\cdot\|_{\infty, \varepsilon}$ defined in Definition 2.7:

$$\mathbb{E}[|K(X - \bar{X}) - K(Y - \bar{Y})|] \leq 8 \min(\|\rho_X\|_{\infty, \varepsilon}, \|\rho_Y\|_{\infty, \varepsilon})\left(\mathbb{E}[|X - Y|] + \frac{\varepsilon}{2}\right).$$

(iii) In the case where $(X, Y)$ and $(\bar{X}, \bar{Y})$ are still independent but with possibly different distributions, we get the more general estimate: for $\varepsilon, \tilde{\varepsilon} \geq 0$ (in case set $\varepsilon = 0$ or $\tilde{\varepsilon} = 0$, set $\|\cdot\|_{\infty, 0} = \|\cdot\|_\infty$),

$$\mathbb{E}[|K(X - \bar{X}) - K(Y - \bar{Y})|] \leq 4 \|\rho_Y\|_{\infty, \varepsilon}\left(\mathbb{E}[|X - Y|] + \tilde{\varepsilon}/2\right) + 4 \|\rho_Y\|_{\infty, \tilde{\varepsilon}}\left(\mathbb{E}[|\bar{X} - \bar{Y}|] + \varepsilon/2\right).$$
Remark 4. Contrary to the result obtained for the Coulomb kernel in dimension larger than 2 (see [16]), the point (i) involves the $L^\infty$ norm of only one of the law of the random variables $(X,Y)$. This is key point in our works, and it enables to provide a weak-strong stability estimate.

We were not able to obtain a similar estimate in $W_p$ distance with $p \in (1, \infty)$. In that case, we need to bound quantities like

$$\mathbb{E} \left[ \left| K(X - \tilde{X}) - K(Y - \tilde{Y}) \right| |X - Y|^{p-1} \right],$$

in terms of $\mathbb{E}[|X - Y|^p]$. Unfortunately, we did not manage to do it with the help of the “rope argument”. But it may be only a technical issue, since a similar estimate could also be obtain for the $W_\infty$ distance. We refer to [4] for the use of that distance for a related system.

Proof. We first prove i). Starting from (10), we may bound

$$\mathbb{E} \left[ \left| K(X - \tilde{X}) - K(Y - \tilde{Y}) \right| \right] \leq \mathbb{E} \left[ \mathbb{I}_{|Y - \tilde{Y}| \leq 2|X - Y|} + \mathbb{I}_{|Y - \tilde{Y}| \leq 2|\tilde{X} - \tilde{Y}|} \right]
\leq 2 \mathbb{E} \left[ \mathbb{I}_{|Y - \tilde{Y}| \leq 2|X - Y|} \right]
\leq 2 \mathbb{E} \left[ \mathbb{E} \left[ \mathbb{I}_{|Y - \tilde{Y}| \leq 2|X - Y|} |(X,Y)\right] \right]
\leq 2 \mathbb{E} \left[ 4 \|\rho_Y\|_\infty |X - Y| \right] \leq 8 \|\rho_Y\|_\infty \mathbb{E}[|X - Y|].$$

In the second line, we used that the expectation remains unchanged if we permute $(X,Y)$ with $(\tilde{X}, \tilde{Y})$. In the third, we used that the law of $\tilde{Y}$ has density $\rho_Y$. If we apply the above bound to the couple $(Y,X)$ and $(\tilde{Y}, \tilde{X})$, we obtain a similar result, with $\|\rho_X\|_\infty$ in place of $\|\rho_Y\|_\infty$.

To obtain ii), remark that for any interval $[a,b] \subset \mathbb{R}$:

$$\mathbb{E} [ \mathbb{I}_A(X) ] = \int_a^b \rho_X(dx) \leq \|\rho_X\|_\infty,\varepsilon (|b - a| + 2\varepsilon).$$

Indeed, to use discrete infinite norm, we need to cover $[a,b]$ by a union of small intervals of length $2\varepsilon$. For this at most (the integer part of) $|b - a|/(2\varepsilon) + 1$ such intervals are requested.

For the third point, we cannot use the permutation $(X,Y) \leftrightarrow (\tilde{X}, \tilde{Y})$ in the previous calculation and have to estimate the two terms separately. The required adaptations are straightforward.

A simple Grönwall lemma. That bound allows us to prove the weak-strong stability part of Theorem 2.3. We introduce $(X_t, V_t)_{t \in \mathbb{R}}$ and $(Y_t, W_t)_{t \in \mathbb{R}}$ two solutions of the non-linear SDE (3) constructed on the same Brownian motion $(B_t)_{t \in \mathbb{R}}$, and also $(\tilde{X}_t, \tilde{V}_t, \tilde{Y}_t, \tilde{W}_t)_{t \in \mathbb{R}}$ an independent copy of the previous coupled processes. We also assume that $\rho_t := \mathcal{L}(X_t)$ has a bounded density for all $t \geq 0$. Then $(X_t - Y_t, V_t - W_t)_{t \in \mathbb{R}}$ solves the following ODE system (without noise):

$$\frac{d}{dt}(X_t - Y_t) = V_t - W_t,$$

$$\frac{d}{dt}(V_t - W_t) = -(V_t - W_t) + \mathbb{E}_{(\tilde{X}_t, \tilde{Y}_t)} \left[ K(X_t - \tilde{X}_t) - K(Y_t - \tilde{Y}_t) \right],$$

adaptations are straightforward.
which naturally leads to
\[
\sup_{s \in [0,t]} |X_s - Y_s| \leq |X_0 - Y_0| + \int_0^t |V_s - W_s| \, ds,
\]
\[
\sup_{s \in [0,t]} |V_s - W_s| \leq |V_0 - W_0| + \int_0^t |K(X_s - \tilde{X}_s) - K(Y_s - \tilde{Y}_s)| \, ds.
\]

If we take the expectation in the previous system, and apply the point (i) of Lemma 3.1 to the couple \((Y_t, X_t)\) and \((\tilde{Y}_t, \tilde{X}_t)\), we may write:
\[
E \left[ \sup_{s \in [0,t]} |X_s - Y_s| \right] \leq E[|X_0 - Y_0|] + \int_0^t E[|V_s - W_s|] \, ds,
\]
\[
E \left[ \sup_{s \in [0,t]} |V_s - W_s| \right] \leq E[|V_0 - W_0|] + 8 \int_0^t \|\rho_s\|_\infty E[|X_s - Y_s|] \, ds.
\]

Summing up the two inequalities and applying the Grönwall lemma lead to the requested estimate
\[
E \left[ \sup_{s \in [0,t]} (|X_s - Y_s| + |V_s - W_s|) \right] \leq E[|X_0 - Y_0| + |V_0 - W_0|] \exp \left( t + 8 \int_0^t \|\rho_s\|_\infty \, ds \right).
\]

**Proposition of moments.** We state here a useful proposition about propagation of moments. We emphasize that it apply to any kernel \(K\) satisfying \(\|K\|_\infty \leq \frac{1}{2}\).

**Proposition 3.** Let be \((Y_t, W_t)_{t \geq 0}\) be a weak solution to (3), with given (random) initial condition \((Y_0, W_0)\), and with an interaction kernel \(K\) which is not necessary given by (2) but satisfies \(\|K\|_\infty \leq \frac{1}{2}\). If \((Y_0, W_0)\) has an exponential moment of order \(\lambda > 0\), it holds for \(t \geq 0\):

\[
(i) \quad E \left[ e^{\lambda |W_t|} \right] \leq e^{2(3+\lambda)} E \left[ e^{\lambda |W_0|} \right],
\]
\[
(ii) \quad E \left[ e^{\lambda |Y_t|} \right] \leq 2 e^{\lambda (\frac{1}{4} + \lambda)} E \left[ e^{\lambda (|Y_0| + |W_0|)} \right].
\]

It also holds for any \(0 \leq s < t \leq s + \min \left( \frac{1}{\Pi}, \lambda^{-2} \right)\):

\[
(iii) \quad E \left[ e^{\lambda (t-s)^{-1} \sup_{s \leq \tau \leq t} |Y_\tau - Y_s|} \right] \leq e^{2(5+\lambda)} E \left[ e^{\lambda |W_0|} \right].
\]

Lastly, simpler estimates on the order one moments also hold: for any \(0 \leq s < t \leq s + \frac{1}{4}\),

\[
(iv) \quad E[|W_t|] \leq e^{-t} E[|W_0|] + 2, \quad E[|Y_t - Y_s|] \leq |s - t| \left( E[|W_0|] + 3 \right).
\]

The proof is performed in Appendix B.

3.2. **Propagation of uniform control on the velocity tails via Feymann-Kac estimates.** In this section, we prove the propagation of the norms \(\|f\|_{p, \lambda}\) and \(\|f\|_{p, \gamma}\) defined in Definition 2.2, along solutions of the non-linear SDE (3). To be fully rigorous, we will first prove that result of propagation on a mollified version of (3). This is a key point in the proof of existence of solution to the initial SDE, and the estimates pass to the limit and also apply to solutions of the non-mollified problem.
For this, we introduce a smoothing kernel \( \chi \in C^\infty(\mathbb{R}, \mathbb{R}^+) \) with support included in \([-1, 1]\) and satisfying \( \int_{\mathbb{R}} \chi(y) \, dy = 1 \). And then classically for \( \eta > 0 \), \( \chi_\eta := \eta^{-1} \chi(\frac{.}{\eta}) \), and the approximated kernel

\[
K_\eta := K * \chi_\eta, \quad \text{which satisfies} \quad \forall x \in \mathbb{R}, \quad |(K_\eta - K)(x)| \leq \mathbb{I}_{[-\eta, \eta]}(x). \tag{11}
\]

Given a stochastic basis, and a Brownian \((B_t)_{t \geq 0}\) motion on it, we consider the following non linear SDE:

\[
Y^\eta_t = Y^\eta_0 + \int_0^t W^\eta_s \, ds, \quad W^\eta_t = W^\eta_0 + \int_0^t \mathbb{E}_T[K_\eta(Y^\eta_s - \overline{Y})] \, ds - \int_0^t W^\eta_s \, ds + \sqrt{2} B_t,
\]

where \( \overline{Y} \) is an independent copy of \( Y^\eta \) and the initial condition \((Y^\eta_0, W^\eta_0)\) is defined as

\[
Y^\eta_0 := Y_0 + \eta U, \quad W^\eta_0 := W_0 + \eta V, \tag{13}
\]

where \((Y_0, W_0)\) has law \( f_0 \) and is independent of \((U, V)\) of law \( \chi \otimes \chi \) (and both are independent of the Brownian motion \((B_t)_{t \geq 0}\)). Then \((Y^\eta_t, W^\eta_t)\) has for law \( \mu^\eta_t := f_0 * \tilde{\chi}_\eta \) with \( \tilde{\chi}_\eta(y, w) := \chi_\eta(y) \chi_\eta(w) \), and is independent of the Brownian motion \((B_t)_{t \geq 0}\).

Since the kernel \( K_\eta \) is globally Lipschitz, \([19, \text{Thm 1.1}]\) implies the strong existence and uniqueness of the process \((Y^\eta_t, W^\eta_t)_{t \geq 0}\) solving (12). And by an application of Ito’s rule the family of the time marginals \((\mu^\eta_t)_{t \geq 0}\) of that process is a weak solution of the following regularized Vlasov-Poisson-Fokker-Planck equation:

\[
\frac{\partial}{\partial t} \mu^\eta_t + v \partial_x \mu^\eta_t + \overline{K}_\eta[\mu^\eta_t] \partial_x \mu^\eta_t = \partial_x(\partial_x \mu^\eta_t + v \mu^\eta_t), \tag{14}
\]

with the initial condition \( \mu^\eta_0 = \mathcal{L}(Y_0, W_0) = f_0 * \tilde{\chi}_\eta \) and the notation

\[
\overline{K}_\eta[\mu](x) = \int_{\mathbb{R}^2} K_\eta(x-y) \mu(\,dy, \, dw). \]

We begin by proving some \( \eta \) independent estimates on \( \mu^\eta_t \), for \( t \geq 0 \).

**Lemma 3.2.** Assume that the law of the initial condition of equation (3) \( f_0 \in \mathcal{P}_1 \cap L^1(\mathbb{R}^2) \) satisfies either \( \|f_0\|_{\infty, \lambda} < \infty \) for some \( \lambda > 0 \) or \( \|f_0\|_{p, \gamma} < \infty \) for some \( \gamma > 1 \). Then for all \( t > 0 \), the unique (measure) solution to the smoothed VPFP equation (14) with initial condition \( \mu^\eta_0 \) satisfies respectively

\[
\|\mu^\eta_t\|_{\infty, \lambda} \leq 2 e^{t + \lambda \eta + \frac{\lambda^2}{2}} \|f_0\|_{\infty, \lambda e^{-1}}, \quad \|\mu^\eta_t\|_{p, \gamma} \leq \|f_0\|_{p, \gamma} \langle \eta \rangle^\gamma C_\gamma e^{(1+\gamma)t}, \tag{15}
\]

where \( C_\gamma \) is a constant depending explicitly on \( \gamma \).

In particular, the associated spatial density \( \rho^\eta_t := \int_{\mathbb{R}} \mu^\eta_t(x, v) \, dv \) satisfies respectively

\[
\|\rho^\eta_t\|_{\infty} \leq \frac{4}{\lambda} e^{t + \frac{\lambda^2}{2}} e^{\lambda \eta} \|f_0\|_{\infty, \lambda e^{-1}}, \quad \|\rho^\eta_t\|_{\infty} \leq \frac{2 \gamma}{\gamma - 1} C_\gamma e^{(1+\gamma)t} \langle \eta \rangle^\gamma \|f_0\|_{p, \gamma}. \tag{16}
\]

**Proof.**

**Step 1. Regularization and Feynmann-Kac’s formula.** Fix \( t \geq 0 \) and consider the following “backward” SDE:

\[
Y^{x,v}_s = x - \int_0^s W^{x,v}_u \, du, \quad W^{x,v}_s = v - \int_0^s \overline{K}_\eta[\mu^\eta_t](Y^{x,v}_u) \, du + \int_0^s W^{x,v}_u \, du + \sqrt{2} B_s, \tag{17}
\]
First note that $\tilde{K}_n[\mu^n]$ is uniformly Lipschitz in position on $\mathbb{R}^+ \times \mathbb{R}^2$. So strong existence and uniqueness of solution to the (linear) SDE (17) are guaranteed by standard results. We set:

$$\theta_s = e^s \mu^n_{t-s}(Y^{x,v}_s, W^{x,v}_s).$$

Moreover the initial condition $\mu^n_0 = f_0 \ast \tilde{\chi}_n$ fulfills the hypothesis of Proposition 4 of the Appendix: $\partial^k_x \partial^l_t \mu^n_0 \in L^2(\mathbb{R}^2)$ for any $k, l \geq 0$. This implies that $\mu^n(t, x, v)$ possesses one continuous derivative in time, and two (continuous) derivatives in position and velocity. So, we may apply Ito's rule to $\theta$: we get

$$e^{-s} d\theta_s = \mu^n(t-s, Y^{x,v}_s, W^{x,v}_s) ds - \partial_t \mu^n(t-s, Y^{x,v}_s, W^{x,v}_s) ds$$

$$+ \partial_x \mu^n(t-s, Y^{x,v}_s, W^{x,v}_s) dY^{x,v}_s + \partial_v \mu^n(t-s, Y^{x,v}_s, W^{x,v}_s) dW^{x,v}_s$$

$$+ \Delta_v \mu^n(t-s, Y^{x,v}_s, W^{x,v}_s) \langle dW^{x,v}_s \rangle^2$$

$$= \left[ -\partial_t \mu^n - v \partial_x \mu^n - \tilde{K}_n[\mu^n] \partial_v \mu^n + \partial_x (v \mu^n) + \Delta_v \mu^n \right] (t-s, Y^{x,v}_s, W^{x,v}_s) ds$$

$$+ \partial_v \mu^n(t-s, Y^{x,v}_s, W^{x,v}_s) dB_s,$$

and since $\mu^n$ is a strong solution of (14), we get precisely that for any $0 \leq s \leq s' \leq t$:

$$\theta_{s'} - \theta_s = \int_s^{s'} e^{u-t} \partial_v \mu^n(t-u, Y^{x,v}_u, W^{x,v}_u) dB_u.$$

In particular, $(\theta_s)_{0 \leq s \leq t}$ is a martingale, so that

$$\mu^n(t, x, v) = \theta_0 = E[\theta_t] = e^t E[f_0 \ast \tilde{\chi}_n(Y^{x,v}_t, W^{x,v}_t)].$$

**Step 2. The case of uniform exponential tails.**

In this case, since $\tilde{\chi}$ has support in $[-1, 1]$, $f_0 \ast \tilde{\chi}(x, v) = \int f_0(x-y, v-w) \tilde{\chi}(y, w) dy dw \leq \|f_0\|_{e, \lambda} \int_{\mathbb{R}} e^{-\lambda|y-v'w'|} \chi(w') dw' \leq \|f_0\|_{e, \lambda} e^{-\lambda|v'|+\lambda\eta}.$

Moreover, the definition (17) of $W^n$ also implies that for $0 \leq s \leq t$:

$$W^{x,v}_s = e^s v - \int_0^s e^{s-u} \tilde{K}_n[\mu^n_u] du + M_s, \quad \text{with} \quad M_s := \sqrt{2} \int_0^s e^{s-u} dB_u.$$

Remark that $M_s$ is in fact a centered Gaussian variable with variance $e^{2s} - 1$. Since $\|\tilde{K}_n[\mu^n]\|_{\infty} \leq 1/2$, it leads to the following lower bound

$$|W^{x,v}_t| \geq e^t |v| - |M_t| - \frac{e^t}{2}.$$  

(20)

Using all of this in the representation formula (18) leads to

$$\mu^n_t(x, v) \leq \|f_0\|_{e, \lambda} e^{t+\lambda\eta} E[e^{-\lambda|W^{x,v}_t|}] \leq \|f_0\|_{e, \lambda} e^{t+\lambda(e^{t/2}+\eta)-\lambda\eta} E[e^{\lambda|M_t|}]$$

An application of (33) (in Appendix B) to $M_t$ leads to the following bound that is uniform in $\eta$ (for $\eta$ small):

$$\|\mu^n_t\|_{e, \lambda} e^t \leq 2 e^{t+\lambda\eta+\frac{\lambda}{2} e^{t/2} + \frac{\lambda^2}{2} |f_0|_{e, \lambda}}.$$  

(21)

The conclusion follows by replacing $\lambda$ by $e^{-t}$ in the above bound, and the estimate on $\|\mu^n_t\|_{\infty}$ follows by integration on $v$. 


Step 3. The case of uniform polynomial tails.

The simple inequality \(|v + w| \leq \sqrt{2}\langle v \rangle \langle w \rangle\) (recall that \(\langle v \rangle^2 := 1 + v^2\)) implies that \(\langle v - w \rangle^{-1} \leq \sqrt{2}\langle v \rangle^{-1} \langle w \rangle\), which allows to bound

\[
 f_0 \ast \tilde{\chi}_\eta(x, v) = \int f_0(x - y, v - w) \tilde{\chi}_\eta(y, w) \, dy \, dw \leq \|f_0\|_{p, \gamma} \int_{\mathbb{R}} \langle v - \eta w' \rangle^{-\gamma} \chi(w') \, dw',
\]

\[
 \leq 2^{\gamma/2} \|f_0\|_{p, \gamma} \langle v \rangle^{-\gamma} \int_{\mathbb{R}} \langle \eta w' \rangle^{-\gamma} \chi(w') \, dw' \leq 2^{\gamma/2} \|f_0\|_{p, \gamma} \langle v \rangle^{-\gamma}.
\]

Plugging it into the representation formula (18), using the lower bound (20) and the simple inequality, \(|v + \frac{1}{2} \rho t| \leq 2^{\gamma/2} (e^{\gamma t} + (2v)^{\gamma})\) (to prove it, simply separate the cases \(v \leq \frac{1}{2} e^{\rho t}\) and \(v \geq \frac{1}{2} e^{\rho t}\),

\[
 \mu^\eta(t, x, v) = e^{\rho t} \mathbb{E}\left[f_0 \ast \tilde{\chi}_\eta(Y_t^{x,v}, W_t^{x,v})\right] \leq 2^{\gamma/2} \|f_0\|_{p, \gamma} \langle \eta \rangle^{\gamma} \mathbb{E}\left[(W_t^{x,v})^{-\gamma}\right]
\]

\[
 \leq 2^{\gamma/2} \|f_0\|_{p, \gamma} \langle \eta \rangle^{\gamma} \mathbb{E}\left[(e^{|v|} - |M_t| - \frac{e^{\gamma t}}{2})^{-\gamma}\right]
\]

\[
 \leq 2^{\gamma} \|f_0\|_{p, \gamma} \langle \eta \rangle^{\gamma} \langle v \rangle^{-\gamma} \mathbb{E}\left[(|M_t| + \frac{e^{\gamma t}}{2})^{-\gamma}\right]
\]

\[
 \leq 2^{\gamma/2} \|f_0\|_{p, \gamma} \langle \eta \rangle^{\gamma} \langle v \rangle^{-\gamma} \left(e^{\gamma t} + 2^{\gamma} \mathbb{E}[|M_t|^{-\gamma}]\right).
\]

Since \(M_t \sim \mathcal{N}(0, e^{2t} - 1)\) we have \(\mathbb{E}[|M_t|^{-\gamma}] = (e^{2t} - 1)^{-\gamma} m_\gamma \leq e^{\gamma t} m_\gamma\), where \(m_\gamma\) stands for the moment of order \(\gamma\) of the law \(\mathcal{N}(0, 1)\). This implies the claimed bound on \(\|\mu_t^\eta\|_{p, \gamma}\) with \(C_\gamma := 2^{3\gamma/2}(1 + 2\gamma m_\gamma)\). The bound on \(\|\rho_t^\eta\|_\infty\) follows from the simple bound \(\langle v \rangle^{-\gamma} \leq \min(1, |v|^{-\gamma})\) and an integration in \(v\).

4. Proof of Theorem 2.5. We start with a useful lemma.

Lemma 4.1. Let \((X_1, \ldots, X_N)\) be \(N\) i.i.d random variables of law \(\rho \in \mathcal{P}(\mathbb{R}^d)\), and \(\rho_N = \frac{1}{N} \sum_1^N \delta_{X_i}\) be the associated empirical measure. Then, for all \(a \in \mathbb{R}^d\) we have:

\[
 \mathbb{E}\left[\sup_{u \in \mathbb{R}^d} \left|\int_{\mathbb{R}^d} 1_{|a - y| \leq u} (\rho_N - \rho)(dy)\right|\right] \leq \frac{3}{\sqrt{N}}
\]

Proof. In fact, we choose a sequence \((X_n)_{n \in \mathbb{N}}\) of i.i.d random variables of law \(\rho \in \mathcal{P}(\mathbb{R}^d)\) and \(L\) some Poisson random variable of parameter \(N\) independent of the \((X_n)_{n \in \mathbb{N}}\). We define two point processes by

\[
 M_N = \sum_{i=1}^L \delta_{X_i}, \quad \rho_N = \frac{1}{N} \sum_{i=1}^N \delta_{X_i}.
\]

\(M_N\) is in fact a Poisson Random Measure (PRM) with intensity measure \(N\rho\). Remark that \(\|M_N - N\rho_N\|_{TV} = |L - N|\). For all \(a \in \mathbb{R}^d\), we have

\[
 \left|\int 1_{|a - y| \leq u} (\rho_N - \rho)(dy)\right| \leq \int 1_{|a - y| \leq u} \left(\frac{1}{N} M_N - \rho\right)(dy)
\]

\[
 + \int 1_{|a - y| \leq u} \left(\frac{1}{N} M_N - \rho_N\right)(dy)
\]

\[
 \leq \frac{1}{N} \int 1_{|a - y| \leq u} (M_N - N\rho)(dy) + \frac{1}{N}\|M_N - N\rho_N\|_{TV}
\]
Given any initial conditions and Brownian paths: for all \(i\)

\[
\sup_{u \in \mathbb{R}_+} \left| \int [a-y] \leq u (\rho_N - \rho)(dy) \right| \leq \frac{1}{N} \sup_{u \in \mathbb{R}_+} |M_{u,i}^{N,a}| + \frac{|L - N|}{N},
\]

where \(M_{u,i}^{N,a} = \int [a-y] \leq u (M_N - \rho)(dy).

Since \(M_N\) is a PRM, \((M_{u,i}^{N,a})_{u \geq 0}\) is a martingale with respect to the filtration \((\mathcal{F}_u^a)_{u \geq 0} = (\sigma(M_N [B_{B(a,u)}]))_{u \geq 0}\), where \(B(a,u)\) the ball of center \(a\) and radius \(u\) in \(\mathbb{R}^d\). So using Doob’s inequality and the fact that \(M_{\infty,a}^{N} = L - N\) we find

\[
\mathbb{E} \left[ \sup_{u \geq 0} |M_{u,i}^{N,a}| \right] \leq \left( \mathbb{E} \left[ \sup_{u \geq 0} |M_{u,i}^{N,a}|^2 \right] \right)^{1/2} \leq 2 \left( \mathbb{E} \left[ |M_{\infty,a}^{N}|^2 \right] \right)^{1/2} \leq 2 \sqrt{\mathbb{V}[L]} = 2\sqrt{N}.
\]

Taking now the expectation in (22), using the above bound and \(\mathbb{E}[L - N] \leq \sqrt{\mathbb{V}(L)} = \sqrt{N}\), we conclude the proof.

We are now in position to prove Theorem 2.5, in two steps.

**Step 1. Coupling estimates.**

We begin by a calculation valid for any fixed realization of these processes (i.e. given any initial conditions and Brownian paths): for all \(i = 1, \cdots, N\) we have:

\[
\sup_{s \in [0,t]} |X_{i,s}^N - Y_{i,s}^N| \leq |X_{i,0}^N - Y_{i,0}^N| + \int_0^t |V_{i,s}^N - W_{i,s}^N| \, ds
\]

\[
\sup_{s \in [0,t]} |V_{i,s}^N - W_{i,s}^N| \leq |V_{i,0}^N - W_{i,0}^N| + \int_0^t \left| \frac{1}{N} \sum_{j=1}^N K(X_{i,s}^N - X_{j,s}^N) - \left( \int_{\mathbb{R}^2} K(Y_{i,s}^N - x) \mu_s(dx, dv) \right) \right| \, ds
\]

\[
\leq |V_{i,0}^N - W_{i,0}^N| + \int_0^t \left| \frac{1}{N} \sum_{j=1}^N K(X_{i,s}^N - X_{j,s}^N) - K(Y_{i,s}^N - Y_{j,s}^N) \right| \, ds
\]

\[
+ \frac{t}{N-1} + \int_0^t \Lambda_{i,s}^N \, ds
\]

with \(\Lambda_{i,s}^N := \left| \frac{1}{N - 1} \sum_{j \neq i}^N K(Y_{i,s}^N - Y_{j,s}^N) - \int_{\mathbb{R}^2} K(Y_{i,s}^N - x) \mu_s(dx, dv) \right|

We sum these inequalities over \(i = 1, \cdots, N\), divide by \(N \geq 2\) and then get:

\[
\frac{1}{N} \sum_{i=1}^N \sup_{s \in [0,t]} |X_{i,s}^N - Y_{i,s}^N| \leq \frac{1}{N} \sum_{i=1}^N |X_{i,0}^N - Y_{i,0}^N| + \int_0^t \frac{1}{N} \sum_{i=1}^N |V_{i,s}^N - W_{i,s}^N| \, ds
\]

\[
\frac{1}{N} \sum_{i=1}^N \sup_{s \in [0,t]} |V_{i,s}^N - W_{i,s}^N| \leq \frac{1}{N} \sum_{i=1}^N |V_{i,0}^N - W_{i,0}^N| + \frac{2t}{N} + \int_0^t \Lambda_s^N \, ds
\]

\[
+ \frac{1}{N^2} \sum_{i \neq j}^N \int_0^t \left| K(X_{i,s}^N - X_{j,s}^N) - K(Y_{i,s}^N - Y_{j,s}^N) \right| \, ds
\]

where \(\Lambda_s^N := \frac{1}{N} \sum_{i=1}^N \Lambda_{i,s}^N ds\).
Using equality (10) we find with the notation $\hat{\rho}_s^{i,N} = \frac{1}{N-1} \sum_{j \neq i} \delta_{Y_{i,s}^N}$,

\[
\frac{1}{N^2} \sum_{i \neq j}^N |K(X_{i,s}^N - X_{j,s}^N) - K(Y_{i,s}^N - Y_{j,s}^N)|
\]

\[
\leq \frac{2}{N^2} \sum_{i \neq j}^N \mathbb{I}[|Y_{i,s}^N - Y_{j,s}^N| \leq 2|X_{i,s}^N - Y_{i,s}^N|] = \frac{2(N - 1)}{N^2} \sum_{i = 1}^N \int \mathbb{I}[|Y_{i,s}^N - y| \leq 2|X_{i,s}^N - Y_{i,s}^N|] \hat{\rho}_s^{i,N}(dy)
\]

\[
\leq \frac{2}{N} \sum_{i = 1}^N \int \mathbb{I}[|Y_{i,s}^N - y| \leq 2|X_{i,s}^N - Y_{i,s}^N|] \rho_s(dy)
\]

\[
+ \frac{2}{N} \sum_{i = 1}^N \int \mathbb{I}[|Y_{i,s}^N - y| \leq 2|X_{i,s}^N - Y_{i,s}^N|](\hat{\rho}_s^{i,N} - \rho_s)(dy)
\]

\[
\leq \frac{8}{N} \|\rho_s\|_{\infty} \sum_{i = 1}^N |X_{i,s}^N - Y_{i,s}^N| + \Gamma_s^N,
\]

where

\[
\Gamma_s^N := \frac{2}{N} \sum_{i = 1}^N \Gamma_{i,s}^N, \quad \Gamma_{i,s}^N := \sup_{u \in \mathbb{R}_+} \left| \int \mathbb{I}[|Y_{i,s}^N - y| \leq u] (\hat{\rho}_s^{i,N} - \rho_s)(dy) \right|.
\] (24)

Then, $\beta(t) := \frac{1}{N} \sum_{i = 1}^N (\sup_{s \in [0,t]} |X_{i,s}^N - Y_{i,s}^N| + \sup_{s \in [0,t]} |V_{i,s}^N - W_{i,s}^N|)$ satisfies the integral inequality

\[
\beta(t) \leq \beta(0) + \int_0^t \left( (1 + 8\|\rho_s\|_{\infty}) \beta(s) + \Lambda_s^N + 2 \Gamma_s^N + \frac{2}{N} \right) ds.
\]

An application of Grönwall’s Lemma leads to

\[
\beta(t) \leq e^{(t+s)\int_0^s \|\rho_u\|_{\infty} ds} \left( \beta(0) + \int_0^t e^{-s} \left( \Lambda_s^N + 2 \Gamma_s^N + \frac{2}{N} \right) ds \right).
\]

Taking the expectation and using the symmetry of the laws of the of $(X_{i,t}^N, V_{i,t}^N)_{i=1,\ldots,N}$ and $(Y_{i,t}^N, W_{i,t}^N)_{i=1,\ldots,N}$ we find

\[
\mathbb{E} \left[ \sup_{s \in [0,t]} \left| X_{1,0}^N - Y_{1,0}^N \right| + \left| V_{1,0}^N - W_{1,0}^N \right| \right] \leq e^{(t+s)\int_0^s \|\rho_u\|_{\infty} ds}
\]

\[
\left( \mathbb{E} \left[ \left| X_{1,0}^N - Y_{1,0}^N \right| + \left| V_{1,0}^N - W_{1,0}^N \right| \right] + \int_0^t e^{-s} \left( \mathbb{E} [\Lambda_s^N + 2 \Gamma_s^N] + \frac{2}{N} \right) ds \right).
\] (25)

We will bound the expectation of stochastic terms appearing in the r.h.s. with the help of Lemma 4.1.

**Step 2. Control of the stochastic terms.**

We recall that $\hat{\rho}_t^{i,N} = \frac{1}{N-1} \sum_{j \geq 2} \delta_{Y_{i,t}^N}$ is the empirical measure associated to the $(Y_{i,t}^N)_{2 \leq i \leq N}$, and is then independent of $Y_{1,t}^N$. By the definition (24) of $\Gamma$ and Lemma 4.1, we have

\[
\mathbb{E} \left[ \Gamma_{1,t}^N \right] = \mathbb{E} \left[ \sup_{u \geq 0} \left| \int \mathbb{I}[|Y_{1,s}^N - y| \leq u] (\hat{\rho}_t^{1,N} - \rho_t)(dy) \right| \right],
\]

\[
= \mathbb{E} \left[ \sup_{u \geq 0} \left| \int \mathbb{I}[|Y_{1,s}^N - y| \leq u] (\hat{\rho}_t^{1,N} - \rho_t)(dy) \right| \left| Y_{1,t}^N \right| \right] \leq \frac{3}{\sqrt{N}},
\]
Moreover, using again the fact that the \((Y_{i,t}^N)_{1 \leq i \leq N}\) are i.i.d and that \(\|K\|_{\infty} \leq 1/2\), we find for \(N \geq 2\)

\[
E[A_{1,t}^N] = E\left[ E\left[ \left( \frac{1}{N-1} \sum_{j=2}^{N} K(Y_{i,t}^N - Y_{j,t}^N) - \int_{\mathbb{R}^2} K(Y_{i,t}^N - x)\mu_s(dx, dv) \right) \bigg| Y_{1,t}^N \right] \right],
\]

\[
\leq E\left[ \left( \left\| \frac{1}{N-1} \sum_{j=2}^{N} K(Y_{i,t}^N - Y_{j,t}^N) \right\|_{Y_{1,t}^N} \right)^{1/2} \right] \leq \frac{1}{\sqrt{2N-1}} \leq \frac{1}{\sqrt{2N}}.
\]

Plugging this bound into equation (25) leads to

\[
E\left[ \sup_{s \in [0,t]} \left| X_{1,s}^N - Y_{1,s}^N \right| + \left| V_{1,s}^N - W_{1,s}^N \right| \right] \leq e^{(1+8J_0} \rho_{\infty} \|ds\) \left( \frac{1}{\sqrt{N}}(1 + 2 \times 3) + \frac{2}{N} \right),
\]

and conclude the proof.

5. Proof of Proposition 2. We distinguish here the deviation upper bounds for \(\sup_{s \in [0,T]} \|\mu_{Y,L}^N\|_{\infty,\varepsilon} \) because such a result have an interest by itself, and because we will apply this result many times in the sequel. In the present proof, we apply the following strategy: we first find prove concentration inequalities for the quantity \(\|\mu_{Y,t}^N\|_{\infty,\varepsilon} \) at a fixed time \(t\), then for the variation of this quantity on small time intervals, and finally conclude by mixing both estimates in an optimal way. That proof could be extended to dimension larger than 1, but we prefer to restrict here to the case of interest.

5.1. Concentration inequalities at fixed time. We recall a special case of the Hoeffding’s inequality, a deviation upper bounds for Binomial variable.

**Lemma 5.1** (Hoeffding’s inequality for binomial variable). Let \(p \in [0,1]\), \(N \in \mathbb{N}^*\), \(X\) a Binomial variable of parameter \((N, p)\). Then for any \(\alpha > 0\):

\[
\mathbb{P}(|X - Np| \geq N\alpha) \leq 2e^{-2\alpha^2N}.
\]

Thanks to lemma 5.1 we are able to give some concentration inequalities for the empirical measure \(\rho^N = \frac{1}{N} \sum_{i=1}^{N} \delta_{Y_i}\), for i.i.d.r.v. \((Y_i)_{i \leq N}\).

**Lemma 5.2.** Let \(\alpha, \varepsilon > 0\). Assume that \((Y_1, \ldots, Y_N)\) are \(N\) independent random variables, all with law \(\rho \in L^\infty\) (we identify the law and its density). Assume also that \(\rho\) has an exponential moment of order \(\lambda > 0\): \(\mathcal{M}_\lambda(\rho) := \int e^{\lambda|y|}\rho(dy) < +\infty\). We denote by \(\rho^N := \frac{1}{N} \sum_{i=1}^{N} \delta_{Y_i}\) the associated empirical measure. Then for any \(\varepsilon, \alpha > 0\),

\[
\mathbb{P}(\|\rho^N\|_{\infty,\varepsilon} \geq \|\rho\|_{\infty} + \alpha) \leq \left( \frac{4\|\rho\|_{\infty}N(\varepsilon\alpha)}{\lambda} + 2 + N\mathcal{M}_\lambda(\rho) \right) e^{-2N(\varepsilon\alpha)^2}
\]

**Proof.** Step 1. A first bound valid on compact subset. For any \(0 < \delta < \varepsilon\) we set \(k = \left\lfloor \frac{R}{2\delta} \right\rfloor + 1\). It is clear that for all \(x \in [-R, R]\), there exists \(\ell \in \{-k, \ldots, k\}\) such that \(B(x, \varepsilon) \subset B(2\ell \delta, \varepsilon + \delta)\). It implies

\[
\mathbb{P}\left( \sup_{x \in [-R, R]} \frac{\rho^N[B(x, \varepsilon)]}{2\varepsilon} \geq \|\rho\|_{\infty} + \alpha \right) \leq \mathbb{P}\left( \sup_{\ell = -k, \ldots, k} \rho^N[B(2\ell \delta, \varepsilon + \delta)] \geq 2\varepsilon(\|\rho\|_{\infty} + \alpha) \right)
\]
\[
\leq \sum_{k=-k}^{k} \mathbb{P}\left( \rho^N \left[ B(2\ell\delta, \varepsilon + \delta) \right] \geq 2(\varepsilon + \delta)\|\rho\|_{\infty} + 2(\alpha\varepsilon - \delta\|\rho\|_{\infty}) \right).
\]

Since for any \( \ell \), \( N\rho^N (B(2\ell\delta, \varepsilon + \delta)) \) is a Binomial variable of parameter \( N \) and \( p_{\ell} = \int_{B(2\ell\delta, \varepsilon + \delta)} \rho(\mathbf{x}) \, d\mathbf{x} \leq 2(\varepsilon + \delta)\|\rho\|_{\infty} \), we may apply Lemma 5.1 and bound each term in the r.h.s. by \( \exp(-8N(\alpha\varepsilon - \delta\|\rho\|_{\infty})^2) \). By the definition of \( k \), and provided that \( \varepsilon\alpha \geq \delta\|\rho\|_{\infty} \) it leads to
\[
\mathbb{P}\left( \sup_{x \in [-R,R]} \frac{\rho^N [B(x, \varepsilon)]}{2\varepsilon} \geq \|\rho\|_{\infty} + \alpha \right) \leq \left( \frac{R}{\delta} + 2 \right) e^{-8N(\alpha\varepsilon - \delta\|\rho\|_{\infty})^2}.
\]

We now have to choose \( \delta \) in order to minimize the right hand side. The particular choice \( \delta\|\rho\|_{\infty} = (\varepsilon\alpha)/2 \) satisfies the previous restriction and already provides an interesting bound:
\[
\mathbb{P}\left( \sup_{x \in [-R,R]} \frac{\rho^N [B(x, \varepsilon)]}{2\varepsilon} \geq \|\rho\|_{\infty} + \alpha \right) \leq \left( \frac{2R\|\rho\|_{\infty}}{\varepsilon\alpha} + 2 \right) e^{-2N(\varepsilon\alpha)^2}.
\]

**Step 2. Extension to the whole space.** It is clear that
\[
\mathbb{P}\left( \|\rho^N\|_{\infty, \varepsilon} \geq \|\rho\|_{\infty} + \alpha \right) \leq \mathbb{P}\left( \sup_{x \in [-R,R]} \frac{\rho^N [B(x, \varepsilon)]}{2\varepsilon} \geq \|\rho\|_{\infty} + \alpha \right) + \mathbb{P}(\exists i \leq N, |Y_i| > R)
\]
\[
\leq \left( \frac{2R\|\rho\|_{\infty}}{\varepsilon\alpha} + 2 \right) e^{-2N(\varepsilon\alpha)^2} + N\mathcal{M}_\lambda(\rho) e^{-\lambda R},
\]
where we used (26) to bound the first term in the r.h.s, and a simple application of Chebyshev’s inequality to bound the second term: precisely that
\[
\mathbb{P}(\exists i \leq N, |Y_i| > R) \leq N\mathbb{P}(|Y_1| > R) \leq N\mathcal{M}_\lambda(\rho) e^{-\lambda R}.
\]
We choose \( R = \frac{2}{\lambda} N(\varepsilon\alpha)^2 \), and finally get:
\[
\mathbb{P}\left( \|\rho^N\|_{\infty, \varepsilon} \geq \|\rho\|_{\infty} + \alpha \right) \leq \left( \frac{4\|\rho\|_{\infty} N(\varepsilon\alpha)}{\lambda} + 2 \right) e^{-2N(\varepsilon\alpha)^2} + N\mathcal{M}_\lambda(\rho) e^{-2N(\varepsilon\alpha)^2},
\]
and the result is proved. \( \square \)

### 5.2. Conclusion of the proof of Proposition 2.

We begin with the following corollary of Proposition 3, which will be useful in the sequel

**Corollary 4.** Let be \((Y_t, W_t)_{t \in [0,T]}\) be a solution of (3) for some initial condition \((Y_0, W_0)\) of law \( f_0 \), having an exponential moment of order \( \lambda > 0 \). We define \( c_\lambda := \frac{5}{2} + \frac{1}{2} \ln \left( \mathbb{E}[e^{\lambda |W_0|}] \right) \) and denote by \( f_t \) the law of \((Y_t, W_t)\). For \( 0 \leq s < t \leq s + \min\left( \frac{1}{16}, \lambda^{-2} \right) \), and \( \beta > 0 \), it holds:
\[
\mathbb{P}\left( \sup_{s \leq u \leq t} |Y_u - Y_s| \geq (t - s)(c_\lambda + \beta) \right) \leq e^{-\frac{\beta}{2} \min(\beta, \lambda)}.
\]

And if the \((Y^N_t, W^N_t)\) are \( N \) independent copies of the previous process, with the same notation
\[
\mathbb{P}\left( \frac{1}{N} \sum_{i=1}^{N} \sup_{s \leq u \leq t} |Y^N_{i,u} - Y^N_{i,s}| \geq (t - s)(c_\lambda + \beta) \right) \leq e^{-\frac{\beta}{2} \min(\beta, \lambda) N}.
\]
Proof. Using point (iii) of the Proposition 3, and Chebyshev’s inequality, and the definition of $c_{\lambda}$ we get for any $0 < \lambda' \leq \lambda$:

\[
\mathbb{P}\left( \sup_{s \leq t} |Y_u - Y_s| \geq (t-s)(c_{\lambda} + \beta) \right) \leq e^{-\lambda'(c_{\lambda} + \beta)} \mathbb{E}\left[ e^{\lambda'(t-s) - 1} \sup_{s \leq t} |Y_u - Y_s| \right] \\
\leq e^{-\lambda'(c_{\lambda} + \beta)} e^{\frac{1}{2}\lambda'(5+\lambda')} \mathbb{E}\left[ e^{\lambda'|W_0|} \right] \\
\leq e^{-\lambda'\beta + \frac{1}{2}\lambda')^2} \mathbb{E}\left[ e^{\lambda'|W_0|} \right] \\
\leq e^{-\lambda'\beta + \frac{1}{2}\lambda')^2}.
\]

Optimization in $\lambda'$ leads to the particular choice $\lambda' = \beta$, when $\beta \leq \lambda$, and to the choice $\lambda' = \lambda$ otherwise. If we use $\beta - \frac{\lambda}{2} \geq \frac{\beta}{2}$ in the later case, we obtain the expected bound.

The proof of the second bound involving $N$ independent copies follows the same lines: by independence

\[
\mathbb{P}\left( \frac{1}{N} \sum_{i=1}^{N} \sup_{s \leq t} |Y_{i,u}^N - Y_{i,s}^N| \geq (t-s)(c_{\lambda} + \beta) \right) \\
\leq \left( e^{-\lambda'(c_{\lambda} + \beta)} \mathbb{E}\left[ e^{\lambda'(t-s) - 1} \sup_{s \leq t} |Y_u - Y_s| \right] \right)^N,
\]

and the conclusion follows with the same optimization on $\lambda'$.

\[\square\]

We are now in position to prove Proposition 2. We fix $\gamma > 0$, define $\alpha := \frac{\gamma}{2}$, and recall from Corollary 4 the notation $c_{\lambda} = \frac{\gamma}{2} + \frac{1}{\lambda} \ln \left( \mathbb{E}\left[ e^{\lambda'|W_0|} \right] \right)$ together with $\kappa_t := \sup_{0 \leq s \leq t} \|\rho_s\|_{\infty}$. We define

\[
\beta := \sqrt{N\varepsilon\gamma} \max\left(1, \frac{\sqrt{N\varepsilon\gamma}}{\lambda}\right), \quad \Delta t := \frac{\varepsilon\alpha}{(\kappa_t + \alpha)(c_{\lambda} + \beta)}.
\]

Remark that $\beta$ satisfies $\frac{\beta}{2} \min(\beta, \lambda) = 2N(\varepsilon\alpha)^2$ and that $\Delta t$ is always smaller than $\min\left(\frac{2}{\alpha}, \lambda^{-2}\right)$ by the assumptions in Proposition 2, so that we may apply Corollary 4.

We then choose $K = \left[\frac{t}{\Delta t}\right] + 1$, define $t_k = k\Delta t$ for all $0 \leq k \leq K$ (remark that $t_K \geq t$), and the two following events $\Omega^1$ and $\Omega^2$ as:

\[
\Omega^1 = \left\{ \exists 0 \leq k \leq K - 1, \|\rho_t^N\|_{\infty, \varepsilon+\Delta t(c_{\lambda} + \beta)} > \kappa_t + \alpha \right\},
\]

\[
\Omega^2 = \left\{ \exists 0 \leq k \leq K - 1, \exists i \leq N, \sup_{s \in [t_k, t_{k+1}]} |Y_{i,s}^N - Y_{i,t_k}^N| > \Delta t(c_{\lambda} + \beta) \right\}.
\]

If the events $\Omega^1$ and $\Omega^2$ are realized, then for any $0 \leq s \leq t$, we choose $k$ such that $s \in [t_k, t_{k+1})$ and get for any $x \in \mathbb{R}$

\[
\rho_s^N(B(x, \varepsilon)) \leq \rho_t^N(B(x, \varepsilon + \Delta t(c_{\lambda} + \beta)) \leq 2(\kappa_t + \alpha)(\varepsilon + \Delta t(c_{\lambda} + \beta)) \\
= 2\varepsilon(\kappa_t + \alpha)(1 + \frac{\alpha}{(\kappa_t + \alpha)}) = 2\varepsilon(\kappa_t + \gamma).
\]

The last equalities follow from the definition of $\alpha$ and $\Delta t$. It means that if $\Omega^1$ and $\Omega^2$ are realized, then

\[
\sup_{s \in [0,t]} \|\rho_s^N\|_{\infty, \varepsilon} \leq \kappa_t + \gamma.
\]
Next, we can bound \( P(\Omega_1) \) and \( P(\Omega_2) \) with the help of Lemma 5.2, Corollary 4, and point (ii) of Proposition 3 (on the control of the exponential moments of \( \rho_k \)): it leads precisely to

\[
P\left( \sup_{s \in [0,t]} \| \rho_s^N \|_{\infty,e} \geq \kappa_t + \gamma \right) \leq P(\Omega_1) + P(\Omega_2)
\]

\[
\leq \sum_{k=0}^{K} \left( \frac{4\|\rho_{t_k}\|_{\infty,E(\varepsilon a)}}{\lambda} + 2 + N\mathcal{M}_\lambda(\rho_{t_k}) \right) e^{-2N(\varepsilon a)^2} + 8NK e^{-\frac{2}{\varepsilon} \min(\beta,\lambda)}
\]

\[
\leq \left( \frac{t}{\Delta t} + 1 \right) \left( \frac{4\kappa_t N(\varepsilon a)}{\lambda} + 2 + 2N e^{\lambda(\frac{1}{2} + \lambda)t} \mathcal{M}_\lambda^{x,v}(f_0) + 8N \right) e^{-2N(\varepsilon a)^2}
\]

\[
= \left( \frac{(2\kappa_t + \gamma)(\varepsilon a + \beta)t}{\varepsilon \gamma} + 1 \right) \left( \frac{2\kappa_t(\varepsilon a)}{\lambda} + 2e^{\lambda(\frac{1}{2} + \lambda)t} \mathcal{M}_\lambda^{x,v}(f_0) + 10 \right) N e^{-\frac{1}{\varepsilon}N(\varepsilon a)^2}.
\]

Using that by assumption \( \varepsilon a \leq \frac{1}{2}\kappa_t \), and the expression of \( \beta \), leads to with

\[
P\left( \sup_{s \in [0,t]} \| \rho_s^N \|_{\infty,e} \geq \kappa_t + \gamma \right)
\]

\[
\leq C_1 \left( 1 + t \frac{2\kappa_t + \gamma}{\varepsilon \gamma} (\varepsilon a + \sqrt{N}\varepsilon \gamma \max(1, \sqrt{N}\varepsilon \gamma \lambda^{-1})) \right) N e^{-\frac{1}{\varepsilon}N(\varepsilon a)^2},
\]

where \( C_1 := 10 + \kappa_t^2 \lambda^{-1} + 2e^{\lambda(\frac{1}{2} + \lambda)t} \mathcal{M}_\lambda^{x,v}(f_0) \). The conclusion follows using that \( \lambda N^{-1/2} \leq \varepsilon a \) by assumption.

6. Proof of Theorem 2.6.

6.1. MKW estimates on deviation between particle and coupled systems.

The first step of the proof is similar to the first step of the proof of Theorem 2.5. Precisely we start with equation (23), which reads now

\[
\frac{1}{N} \sum_{i=1}^{N} \sup_{s \in [0,t]} |X_{i,s}^N - Y_{i,s}^N| \leq \int_0^t \frac{1}{N} \sum_{i=1}^{N} |V_{i,s}^N - W_{i,s}^N| ds
\]

\[
= \frac{1}{N} \sum_{i=1}^{N} \sup_{s \in [0,t]} |V_{i,s}^N - W_{i,s}^N| \leq \int_0^t \frac{1}{N^2} \sum_{i,j=1}^{N} |K(X_{i,s}^N - X_{j,s}^N) - K(Y_{i,s}^N - Y_{j,s}^N)| ds
\]

\[
+ \frac{t}{N-1} + \int_0^t \Lambda_s^N ds,
\]

where

\[
\Lambda_s^N := \frac{1}{N} \sum_{i=1}^{N} \Lambda_s^N ds = \frac{1}{N} \sum_{i=1}^{N} \left| \frac{1}{N} \sum_{j=1}^{N} K(Y_{i,s}^N - X_{j,s}^N) - \int_{\mathbb{R}^2} K(Y_{i,s}^N - x) \mu_s(dx, dv) \right|
\]

since the initial conditions are now equal. Remark that if we introduce \( \sigma, \tau \) two independent random variables with uniform law on \{1, 2, \ldots, N\}, then the sum involving \( K \) becomes

\[
\frac{1}{N^2} \sum_{i,j=1}^{N} |K(X_{i,s}^N - X_{j,s}^N) - K(Y_{i,s}^N - Y_{j,s}^N)| =
\]

\[
\mathbb{E}_{\sigma, \tau} \left[ |K(X_{\sigma,s}^N - X_{\tau,s}^N) - K(Y_{\sigma,s}^N - Y_{\tau,s}^N)| \right],
\]
where we emphasize that the expectation is taken only with respect to \((\sigma, \tau)\). So, we are in position to apply point (iii) Lemma 3.1: i.e. the part involving discrete uniform norms, and get:

\[
\frac{1}{N} \sum_{i=1}^{N} \sup_{s \in [0,t]} |Y_{i,s}^N - W_{i,s}^N| \leq 8 \int_0^t \|\rho_s^N\|_{\infty, \varepsilon} \left( \frac{1}{N} \sum_{i=1}^{N} |X_{i,s}^N - Y_{i,s}^N| + \frac{\varepsilon}{2} \right) ds \\
+ \frac{t}{N-1} + \int_0^t \Lambda_s^N ds \leq \int_0^t \left( 1 + 8\|\rho_s^N\|_{\infty, \varepsilon} \right) \left( \frac{1}{N} \sum_{i=1}^{N} |X_{i,s}^N - Y_{i,s}^N| + \frac{\varepsilon}{2} + \frac{1}{N-1} + \sup_{s \leq t} \Lambda_s^N \right) ds.
\]

where \(\rho_s^N\) is the empirical measure of the \((Y_i^N)_{1 \leq i \leq N}\): \(\rho_s^N := \frac{1}{N} \sum_{i=1}^{N} \delta_{Y_i^N}\). Applying finally Gronwall’s Lemma on the interval \([0, t]\) where the quantity \(\sup_{u \leq t} \Lambda_u^N\) is a constant, we get:

\[
\frac{1}{N} \sum_{i=1}^{N} \sup_{s \in [0,t]} \left( |X_{i,s}^N - Y_{i,s}^N| + |Y_{i,s}^N - W_{i,s}^N| \right) \leq \left( \frac{\varepsilon}{2} + \frac{1}{N-1} + \sup_{s \leq t} \Lambda_s^N \right) \exp \left( t + 8 \int_0^t \|\rho_s^N\|_{\infty, \varepsilon} ds \right),
\]

where

\[
\Lambda_s^N := \frac{1}{N} \sum_{i=1}^{N} \left| \frac{1}{N-1} \sum_{j \neq i} \int_{\mathbb{R} \times \mathbb{R}} K(Y_{i,s}^N - Y_{j,s}^N) - \int_{\mathbb{R} \times \mathbb{R}} K(Y_{i,s}^N - x)\mu_s(dx, dv) \right|.
\]

We now focus on finding some concentration inequalities for the random variable \(\sup_{t \in [0,T]} \Lambda_t^N\). In order to prove concentration inequalities for these supremum in time, we follow the same steps as in the proof of Proposition 2. Once it is be done, we will combine them with the concentration inequalities on \(\sup_{t \in [0,T]} \|\rho_t^N\|_{\infty, \varepsilon}\) given by Proposition 2, and we will obtain some deviation upper bounds for

\[
\frac{1}{N} \sum_{i=1}^{N} \sup_{s \in [0,t]} \left( |X_{i,s}^N - Y_{i,s}^N| + |Y_{i,s}^N - W_{i,s}^N| \right).
\]

6.2. Estimation of the fluctuation terms \(\Lambda_t^N\) at fixed time \(t\). We first establish the

**Lemma 6.1.** Let \((Y_i^N)_{1 \leq i \leq N}\) be i.i.d. random variables, with a diffuse common law (i.e. a law that does not charge any atom). Define the random variable \(\Lambda^N\) as follows:

\[
\Lambda^N = \frac{1}{N} \sum_{i=1}^{N} \left| \frac{1}{N-1} \sum_{j \neq i} K(Y_i^N - Y_j^N) - \mathbb{E}[K(Y_i^N - Y_j^N)|Y_i^N] \right|
\]

Then for all \(\alpha > 0\),

\[
\mathbb{P}(|\Lambda^N| \geq \alpha) \leq 2N e^{-2\alpha^2(N-1)}.
\]

**Proof.** **Step 1. A calculation with \(Y_i^N\) frozen.** To begin, we “freeze” \(Y_i^N = a\) and define

\[
\Lambda_i^N(a) := \frac{1}{N-1} \sum_{j \geq 2} K(a - Y_j^N) - \mathbb{E}[K(a - Y_j^N)].
\]
By the definition (2) of $K$, and since $\mathbb{P}(Y_j^N = a) = 0$ by assumption the random variable $\zeta_j^a$ defined by:

$$
\zeta_j^a = K(a - Y_j^N) + \frac{1}{2} = \begin{cases} 
0 & \text{if } a < Y_j^N \\
1/2 & \text{if } a = Y_j^N \\
1 & \text{if } a > Y_j^N
\end{cases}
$$

is a Bernoulli variable with parameter $p_a := \mathbb{P}(Y_j^N > a)$. But since $p_a - \frac{1}{2} = \mathbb{E}[K(a - Y_j^N)]$, we have $\sum_{j \geq 2} \zeta_j^a - (N - 1)p_a = \Lambda_1^N(a)$. So, by an application of Hoeffding’s inequality (Lemma 5.1) to the binomial variable $\sum_{j \geq 2} \zeta_j^a$:

$$
\mathbb{P}(|\Lambda_1^N(a)| \geq \alpha) = \mathbb{P}\left(\left|\sum_{j \geq 2} \zeta_j^a - (N - 1)p_a\right| \geq (N - 1)\alpha\right) \leq 2e^{-\alpha(N-1)\alpha^2},
$$

Step 2. Summing up on $N$. Using the notation introduced in the previous step, we can rewrite $\Lambda^N = \frac{1}{N} \sum_{i=1}^N |\Lambda_1^N(Y_i^N)|$, and then

$$
\mathbb{P}(\Lambda^N \geq \alpha) \leq \sum_{i=1}^N \mathbb{P}(\Lambda_1^N(Y_i^N) \geq \alpha) = \mathbb{P}\left(\sum_{i=1}^N |\Lambda_1^N(Y_i^N)| \geq \alpha\right) \leq N \mathbb{E}\left[\mathbb{P}(|\Lambda_1^N(Y_i^N)| \geq \alpha | Y_i^N)\right],
$$

where we have used the fact that the variables $(Y_i^N)_{1 \leq i \leq N}$ are exchangeable. But by independence of $Y_i^N$ and $(Y_i^N)_{i \geq 2}$, we obtain using the previous step that $\mathbb{P}(|\Lambda_1^N(Y_i^N)| \geq \alpha | Y_i^N) \leq 2e^{-2(N-1)\alpha^2}$ and the conclusion follows. $\square$

6.3. Estimation on the supremum in time of the fluctuation terms.

Lemma 6.2. Assume that $(Y_i^N, W_i^N)$ are $N$ independent copies of a solution $(Y, W)$ to (3) satisfying $\mathbb{E}[e^{\lambda(|Y_0| + |W_0|)}] < +\infty$ for some $\lambda > 0$, and $\Lambda_i^N$ is defined by (28). Then, with $\rho_s := \mathcal{L}(Y_s)$ $\kappa_t := \sup_{0 \leq s \leq t} \|\rho_s\|_\infty$ we have provided that $\varepsilon \leq \min\left(\frac{1}{16}, \frac{1}{2}, \lambda^{-2}\right)$:

$$
\mathbb{P}\left(\sup_{s \in [0, t]} |\Lambda_s^N| \geq C_t \varepsilon\right) \leq (\varepsilon + t) \left(\frac{5}{\varepsilon} + \frac{4\kappa_t N}{\lambda} + \frac{N}{\varepsilon} e^{\lambda(\frac{1}{2} + \lambda)} \mathbb{E}\left[e^{\lambda(|Y_0| + |W_0|)}\right]\right) e^{-2N\varepsilon^2},
$$

with $C_t := 36 + 80\kappa_t + (1 + 3\kappa_t)\frac{\varepsilon}{\lambda} \ln \mathbb{E}\left[e^{\lambda|W_0|}\right]$.

Proof. Step 1. Bounding the time fluctuations. Let $(Z_t)_{t \geq 0} = (Y_t, W_t)_{t \geq 0}$ be a solution to the non linear SDE (3). We show in that step that for $N \geq 2$:

$$
\sup_{s \in [t, t + \Delta t]} |\Lambda_s^N - \Lambda_i^N| \leq \left(16 \|\rho_i^N\|_{\infty, \varepsilon} + 4\|\rho_i\|_\infty\right) \sup_{s \in [t, t + \Delta t]} \frac{1}{N} \sum_{i=1}^N |Y_i^N - Y_i^N| + 4 \|\rho_i^N\|_{\infty, \varepsilon} |s - t| \left(\mathbb{E}[|W_0|] + 2\right) + 10 \|\rho_i^N\|_{\infty, \varepsilon} \varepsilon. \quad (29)
$$
Indeed, by the definition (28) of $\Lambda^N_s$: 

$$|\Lambda^N_s - \Lambda^N_t| \leq \frac{1}{N(N-1)} \sum_{i=1}^{N} \sum_{j=1}^{N} |K(Y^N_{i,s} - Y^N_{j,s}) - K(Y^N_{i,t} - Y^N_{j,t})|$$

$$+ \frac{1}{N} \sum_{i=1}^{N} \left| \int_\mathbb{R} K(Y^N_{i,s} - x) \rho_s(dx) - \int_\mathbb{R} K(Y^N_{i,t} - x) \rho_t(dx) \right|$$

Using the second point of Lemma 3.1 with two copies of a vector of joint law $\delta_{(Y^N_i,Y^N_j)}$, we may bound the first term in the r.h.s. by

$$\frac{1}{N(N-1)} \sum_{i=1}^{N} \sum_{j=1}^{N} |K(Y^N_{i,s} - Y^N_{j,s}) - K(Y^N_{i,t} - Y^N_{j,t})|$$

$$\leq 8 \frac{N}{N-1} \|\rho_N^t\|_{\infty,\varepsilon} \left( \frac{1}{N} \sum_{i=1}^{N} |Y^N_{i,s} - Y^N_{i,t}| + \frac{\varepsilon}{2} \right).$$

To estimate the second term in the r.h.s, we use the third point of Lemma 3.1, applied to independent couples: The first one with law $\frac{1}{N} \sum_i \delta_{(Y^N_i,Y^N_j)}$ and $(Y_t, Y_s)$. It leads to the estimate

$$\frac{1}{N} \sum_{i=1}^{N} \left| \int_\mathbb{R} K(Y^N_{i,s} - x) \rho_s(dx) - \int_\mathbb{R} K(Y^N_{i,t} - x) \rho_t(dx) \right|$$

$$\leq 4 \left( \|\rho_N^t\|_{\infty,\varepsilon} \frac{1}{N} \sum_{i=1}^{N} |Y^N_{i,s} - Y^N_{i,t}| + \|\rho_N^t\|_{\infty,\varepsilon} \left( \mathbb{E}[|Y_s - Y_t|] + \varepsilon/2 \right) \right).$$

Putting these two estimates together, using point (iv) of Proposition 3 in order to bound $\mathbb{E}[|Y_s - Y_t|]$ and taking the supremum in time leads to (29).

**Step 2. Controlling the deviation.** We define $t_k = k\varepsilon$ for $0 \leq k \leq k_M := \lfloor \varepsilon^{-1} \rfloor$ and recall that by assumption $\varepsilon \leq \min \left( \frac{1}{16}, \frac{1}{2}, \lambda^{-2} \right)$. Consider the events

$$\Omega_1 := \{ \|\rho_N^t\|_{\infty,\varepsilon} \leq \kappa_t + 1 \text{ for all } 0 \leq k \leq k_M \},$$

$$\Omega_2 := \{ \sup_{s \in [t_k, t_{k+\varepsilon}]} \frac{1}{N} \sum_{i=1}^{N} |Y^N_{i,s} - Y^N_{i,t_k}| \leq \varepsilon(c_\lambda + 2\varepsilon) \text{ for all } 0 \leq k \leq k_M \},$$

$$\Omega_3 := \{ |\Lambda^N_k| \leq 4\varepsilon \text{ for all } 0 \leq k \leq k_M \}.$$

By Lemma 5.2 and the point (ii) of Lemma 3,

$$\mathbb{P}(\Omega_1^c) \leq (k_M + 1) \left( 4\kappa_t N \varepsilon \lambda^{-1} + 2 + N e^{M(\frac{1}{2} + \lambda)} \mathbb{E}[e^{\lambda(|Y_0| + |W_0|)}] \right) e^{-2N\varepsilon^2}. $$

By Corollary 4, $\mathbb{P}(\Omega_2^c) \leq (k_M + 1) e^{-2N\varepsilon^2}$. By Lemma 6.1, $\mathbb{P}(\Omega_3^c) \leq 2(k_M + 1) e^{-8(N-1)\varepsilon^2} \leq 2(k_M + 1) e^{-2N\varepsilon^2}$ if $N \geq 2$. When the three above events are realized, we get with the help of the bound (29)

$$\sup_{s \in [0,t]} |\Lambda^N_s| \leq \sup_{k=0,\ldots,k_M} |\Lambda^N_k| + \sup_{k=0,\ldots,k_M, s \in [t_k, t_{k+\varepsilon}]} |\Lambda^N_k - \Lambda^N_k|$$

$$\leq 4\varepsilon + (20\kappa_t + 4)\varepsilon(c_\lambda + \varepsilon) + 4(\kappa_t + 1)\varepsilon \left( \mathbb{E}[|W_0|] + 2 \right) + 10(\kappa_t + 1)\varepsilon$$

$$\leq C_t^\varepsilon,$$
with $C'_t := 26 + 4c_\lambda + 4E[|W_0|] + \kappa_\lambda(20c_\lambda + 4E[|W_0|] + 30)$. Using the expression of $c_\lambda$ and the fact that $\lambda E[|W_0|] \leq \ln E[e^{\lambda|W_0|}]$, it is not difficult to show that $C'_t \leq C_t$, where $C_t$ is the constant introduced in Lemma 6.2. Moreover, gathering the three previous estimates on $P(\Omega_i^c)$ for $i = 1, 2, 3$, we get
\[
P(\Omega_1^c \cup \Omega_2^c \cup \Omega_3^c) \leq \frac{\varepsilon + t}{\varepsilon} \left( 5 + 4\kappa_t N\varepsilon \lambda^{-1} + N e^{\lambda N \varepsilon} \left( e^{\lambda|Y_0| + |W_0|} \right) \right) e^{-2N\varepsilon^2},
\]
which concludes the proof. □

6.4. Conclusion of the proof of Theorem 2.6. We now consider the two events
\[
\Omega_1 := \left\{ \sup_{s \in [0,t]} |A^N_s| \geq C_t \varepsilon \right\}, \quad \Omega_2 := \left\{ \sup_{s \in [0,t]} \|\rho^N_s\|_{\infty, \varepsilon} \geq \kappa_t + 1 \right\},
\]
where $C_t$ is the constant defined in 6.2. By Lemma 6.2, and since we also assume $\varepsilon \geq \lambda N^{-1/2}$,
\[
P(\Omega_1^c) \leq (\varepsilon + t) \left( \frac{5}{\varepsilon} + 4\kappa_t N\lambda^{-1} + \frac{N}{\varepsilon} e^{\lambda \left( \frac{1}{2} + \lambda \right)} \mathbb{E}[e^{\lambda(Y_0 + |W_0|)}] \right) e^{-2N\varepsilon^2},
\]
\[
\leq (\varepsilon + t) \lambda^{-1} N^2 \left( \frac{5}{\varepsilon} + 4\kappa_t + e^{\lambda \left( \frac{1}{2} + \lambda \right)} \mathbb{E}[e^{\lambda(Y_0 + |W_0|)}] \right) e^{-2N\varepsilon^2}.
\]
Now by Proposition 2, provided that $\varepsilon \leq 5\kappa_t \min\left( \frac{1}{16}, \frac{1}{2}, \lambda^{-2} \right)$, and since we also assume $\varepsilon \geq \lambda N^{-1/2}$, the following bound holds with
\[
D_t := 10 + \kappa_t^2 \lambda^{-1} + e^{\lambda \left( \frac{1}{2} + \lambda \right) t} \mathbb{E}[e^{\lambda(Y_0 + |W_0|)}] ;
\]
\[
P(\Omega_2^c) \leq D_t \left( 1 + t \frac{2\kappa_t + 1}{\varepsilon} \left( c_\lambda + \sqrt{N\varepsilon} \max(1, \sqrt{N\varepsilon}\lambda^{-1}) \right) \right) N e^{-\frac{1}{2}N\varepsilon^2}
\]
\[
\leq D_t \left( 1 + t \frac{2\kappa_t + 1}{\varepsilon} \left( c_\lambda + N\varepsilon^2 \lambda^{-1} \right) \right) N e^{-\frac{1}{2}N\varepsilon^2}
\]
\[
\leq D_t \lambda^{-1} \left( \varepsilon + t(2\kappa_t + 1) \left( c_\lambda + \sqrt{N\varepsilon} \right) \right) N^2 \varepsilon e^{-2N\varepsilon^2}.
\]
If the two events are satisfied, then thanks to (27) we get
\[
\frac{1}{N} \sum_{i=1}^{N} \sup_{s \in [0,t]} \left( |X^N_{i,s} - Y^N_{i,s}| + |V^N_{i,s} - W^N_{i,s}| \right)
\]
\[
\leq \left( \left( C_t + \frac{1}{2} \right) \varepsilon + \frac{1}{N - 1} \right) \exp \left( t + 8 \int_0^t \|\rho_s\|_{\infty} \, ds \right)
\]
\[
\leq \left( C_t + \frac{1}{2} + \frac{2}{\lambda} \right) \exp \left( t + 8 \int_0^t \|\rho_s\|_{\infty} \, ds \right) \varepsilon,
\]
where we have used that $\varepsilon \geq \lambda N^{-1/2} \geq \lambda(2(N - 1))^{-1}$ when $N \geq 2$, which concludes the proof since
\[
B_t := C_t + \frac{1}{2} + \frac{2}{\lambda}.
\]
Defining also
\[
A''_t := \lambda^{-1} \left( D_t \left( 1 + (2\kappa_t + 1) c_\lambda \right) + 5 + 4\kappa_t + e^{\lambda \left( \frac{1}{2} + \lambda \right)} \mathbb{E}[e^{\lambda(Y_0 + |W_0|)}] \right),
\]
\[
A'_t := \lambda^{-1} D_t (2\kappa_t + 1),
\]
we obtain from the previous bound \( P(\Omega_1^t \cup \Omega_2^t) \leq (t + \varepsilon)(A_t + A'_t \sqrt{N \varepsilon}) N^{2} e^{-2N\varepsilon^2} \), and that concludes the proof. In particular it can be checked that \( A''_t \leq A_t \) where \( A_t \) is defined by

\[
A_t := \lambda^{-1} \left[ 12 + \kappa_t^2 \lambda^{-1} + 2 e^{\lambda(\frac{1}{2} + \lambda) t} \mathbb{E} \left[ e^{\lambda(|V_0| + |W_0|)} \right] \right] (1 + (2\kappa_t + 1)c_\lambda),
\]

\[
A'_t := \lambda^{-1} \left[ 10 + \kappa_t^2 \lambda^{-1} + e^{\lambda(\frac{1}{2} + \lambda) t} \mathbb{E} \left[ e^{\lambda(|V_0| + |W_0|)} \right] \right] (2\kappa_t + 1),
\]

\[
B_t := 37 + \frac{2}{\lambda} + 80\kappa_t + (1 + 3\kappa_t) \frac{8}{\lambda} \ln \mathbb{E} \left[ e^{N|W_0|} \right].
\]

Appendix A. Proof of Theorem 2.1.

Step 1. A related SDE with memory. First note that a weak solution to equation (1) is some stochastic basis, together with a \( N \) dimensional Brownian motion \((B^N_t)_{t \geq 0}\) on it and a \( \mathbb{R}^{2N} \) valued processes \((X^N_t, V^N_t)_{t \geq 0}\) satisfying for all \( t \geq 0 \)

\[
X^N_t = X^N_0 + \int_0^t V^N_s ds,
\]

\[
V^N_t = V^N_0 + \int_0^t K^N(X^N_s)ds - \int_0^t V^N_s ds + \sqrt{2} B^N_t,
\]

where we denoted \( K^N(x_1, \ldots, x_N) \) the vector valued field which \( i \)-th component is \( \frac{1}{N} \sum_{j \neq i} K(x_i - x_j) \). But, from that system, we may write a SDE “with memory” involving \( V^N \)

\[
\forall t \geq 0, \quad V^N_t = V^N_0 + \int_0^t K^N \left( X^N_s + \int_0^s V^N_u du \right) ds - \int_0^t V^N_s ds + \sqrt{2} B^N_t. \quad (30)
\]

Conversely, given a solution to the SDE (30), it is not difficult to construct a solution to the original system. So it will be enough to prove weak existence and uniqueness in law for the delayed SDE (30).

Step 2. Weak existence. We follow a standard strategy in the case of SDE with constant diffusion, which relies on the Cameron-Martin-Girsanov (CMG) theorem. Roughly speaking, this Theorem ensures that adding a sufficiently regular drift to a diffusion does not change the shape (or regularity) of the trajectories, which still looks like the one of the Brownian motion: the support of the law of the Brownian motion, and the one of the trajectories of the drift-diffusion will be the same, with relative derivative given by the CMG theorem. With that in mind, the Girsanov’s strategy to construct a solution to an SDE is to start from a Brownian motion, to revert the SDE and define a new process, and finally to change the law of the trajectories so that the new process becomes a Brownian Motion, and the initial BM a solution of the required SDE.

Precisely in our case, we let \((\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})\) be a stochastic basis and \((B^N_t)_{t \geq 0}\) be a \( N \)-dimensional Brownian motion on it, and define

\[
U^N_t := -\frac{1}{\sqrt{2}} \int_0^t K^N \left( X^N_0 + sV^N_0 + \sqrt{2} \int_0^s B^N_u du \right) ds + \int_0^t \left( B^N_s + \frac{1}{\sqrt{2}} V^N_0 \right) ds + B^N_t,
\]

\[
V^N_t := V^N_0 + \sqrt{2} B^N_t.
\]

The above definition of the two r.v. \((U^N_t, V^N_t)_{t \geq 0}\) implies that for any \( t \geq 0 \),

\[
V^N_t = V^N_0 + \int_0^t K^N \left( X^N_0 + \int_0^s V^N_u du \right) ds - \int_0^t V^N_s ds + \sqrt{2} U^N_t,
\]
which is exactly (30) with $B_t^N$ replaced by $U_t^N$. So, it remains to apply Cameron-
Martin-Girsanov (CMG) theorem: with an appropriate change of the reference
probability measure, $(U_t^N)_{t \geq 0}$ can be considered has a $N$-dimensional Brownian
motion. For this, remark that $dU_t^N = -H_t^N dt + dB_t^N$, where
\[
H_t^N := \frac{1}{\sqrt{2}} K^N \left( X_t^N + tV_0^N + \int_0^t \sqrt{2} B_u^N du \right) - B_t^N - \frac{1}{\sqrt{2}} V_0^N,
\]
is $F_t$-adapted and progressively measurable, and that for $0 < \gamma < \frac{1}{\sqrt{2}}$,
\[
E \left[ e^{\gamma |H_t^N|^2} \right] \leq e^{2\gamma \left( N \|K\|_2^2 + |V_0^N|^2 \right)} E \left[ e^{3\gamma (B_t^N)^2} \right] < +\infty.
\]
Therefore we deduce from classical results about exponential martingales that the
process $Z_t^N$ defined by
\[
Z_t^N = \exp \left( \int_0^t H_s^N \cdot dB_s^N - \frac{1}{2} \int_0^t |H_s^N|^2 ds \right)
\]
(where $\cdot$ stands for the scalar product) is a martingale, and due to CMG theorem
$(U_t^N)_{t \geq 0}$ is a $N$-dimensional Brownian motion under the probability $Q$ defined
for any $A$ in $\mathcal{F}_t$ by $Q(A) = \int_A Z_t^N d\mathbb{P}$. Therefore $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, Q, (U_t^N, V_t^N)_{t \geq 0})$ is a
weak solution to SDE (30).

**Step 3. Uniqueness in law.** Suppose that $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P}, (V_t^N, B_t^N)_{t \geq 0})$ is a
solution to equation (30) with initial condition $V_0^N$. We may use CMG theorem
again: in fact $2^{-1/2} dV_t^N = -\tilde{H}_t^N dt + d\tilde{B}_t^N$, where
\[
\tilde{H}_t^N := \frac{1}{\sqrt{2}} K^N \left( X_t^N + \int_0^t V_u^N du \right) + \frac{1}{\sqrt{2}} V_0^N.
\]
Moreover by (30), $V_t^N$ also satisfies
\[
V_t^N = e^{-t} V_0^N + \int_0^t e^{s-t} K^N \left( X_u^N + \int_0^u V_u^N du \right) ds + \sqrt{2} \int_0^t e^{s-t} dB_u^N,
\]
which implies that
\[
|\tilde{H}_t^N| \leq |V_t^N| + \frac{\sqrt{N}}{2} \leq |V_0^N| + \frac{3\sqrt{N}}{2} + |M_t^N|,
\]
where $M_t^N$ is a Gaussian r.v. with law $\mathcal{N}(0, (1-e^{2t})\text{Id})$. It follows that $E \left[ e^{\gamma |\tilde{H}_t^N|^2} \right] <
\infty$, for $t \geq 0$ and $\gamma < 1/2$. This implies that the process $\tilde{Z}_t^N$ defined by
\[
\tilde{Z}_t^N = \exp \left( \int_0^t \tilde{H}_s^N \cdot dB_s^N - \frac{1}{2} \int_0^t |\tilde{H}_s^N|^2 ds \right)
\]
\[
= \exp \left( \frac{1}{\sqrt{2}} \int_0^t \tilde{H}_s^N \cdot dB_s^N + \frac{1}{2} \int_0^t |\tilde{H}_s^N|^2 ds \right),
\]
is a martingale, and by Cameron-Martin-Girsanov theorem, $2^{-1/2}(V_t^N - V_0^N)_{t \geq 0}$ is a
$N$-dimensional Brownian motion on the filtered space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \tilde{Q})$ where $\tilde{Q}$
is defined for any $A$ in $\mathcal{F}_t$ by $\tilde{Q}(A) = \int_A \tilde{Z}_t^N d\tilde{P}$.  

Proof of Proposition 3. Using that \( \text{erf}(B_t) \) is a second solution. Since, solutions to (30) have continuous exponential moments are simple to obtain. In fact for a Gaussian variable \( Z \) defined as \( Z = \sqrt{\lambda} \sigma \sqrt{t} \), we finally get the following bound, that will be very useful in the sequel:

\[
E^2 \left( e^{\lambda |Z|} \right) \leq e^{2(\lambda + \lambda^2)} e^{\lambda^2 t/2},
\]

Together with the independence of \((B_t)_{t \geq 0}\) and the initial condition \( W_0 \), it leads to

\[
E^2 \left( e^{\lambda |W_1|} \right) \leq e^{2(\lambda + \lambda^2)} e^{\lambda^2 t/2}.
\]

Appendix B. Propagation of moments. This section is devoted to the proof of Proposition 3.

Proof of Point (i). First, introducing the notation \( f_t \) for the time marginal of \((Y_t, W_t)\) and \( F(t, x) := \int K(x - y) f_t(dy, dw) \) we have by (3)

\[
e^t W_t = W_0 + \int_0^t e^s F(s, Y_s) \, ds + \sqrt{2} \int_0^t e^s dB_s.
\]

Next, since \(|K|\) and also \(|F|\) are bounded by \(1/2\), we get a simple inequality:

\[
|W_t| \leq e^{-t}|W_0| + \frac{1}{2} + |M_t|, \quad \text{with} \quad M_t := \sqrt{2} \int_0^t e^{s-t} dB_s.
\]

The expression on the last line does not involve \((B_t^N)_{t \geq 0}\) anymore. Since the law of \((V_t^N)_{t \geq 0}\) under \( \hat{Q} \) is the law of the Brownian motion, and since \( \hat{Z}_t^N \) can be expressed in term of \( V_t^N \) only, that last expression does not depend on the specific solution we selected at the beginning of this paragraph: we will obtain exactly the same formula starting from a second solution. Since, solutions to (30) have continuous trajectories, it implies uniqueness in law of the solution to the SDE (30) and also to (1).
Proof of Point (ii). For the second point, we integrate the inequality (31) and get:

\[ Y_t = Y_0 + \int_0^t W_s \, ds = Y_0 + (1 - e^{-t})W_0 + \int_0^t (1 - e^{-s-t})F(s, Y_s) \, ds + N_t, \]

|Y_t| \leq |Y_0| + |W_0| + \frac{t}{2} + |N_t|, \quad \text{with} \quad N_t := \int_0^t e^{-s}M_s \, ds = \sqrt{2} \int_0^t (1 - e^{-u-t}) \, dB_u, \tag{34}

where we have used a stochastic version of Fubini’s Theorem in the last line. \(N_t\) is a centered random Gaussian variable with variance \(\sigma_n^2 := 4e^{-t} - e^{-2t} - 3 + 2t \leq 2t\). The bound (33) and the independence of the initial condition and \(N_t\) then lead to the claimed result:

\[
\mathbb{E}[e^{\lambda Y_t}] \leq e^{\frac{\lambda}{2} t} \mathbb{E}[e^{\lambda(Y_0+|W_0|)}] \mathbb{E}[e^{\lambda|N_t|}],
\]

\[
\leq e^{\frac{\lambda}{2} t} \mathbb{E}[e^{\lambda(Y_0+|W_0|)}] 2 e^{\frac{1}{2} \lambda^2 \sigma_n^2} \leq 2 e^{\frac{1}{2} \lambda^2 + \lambda^2 t} \mathbb{E}[e^{\lambda(|Y_0|+|W_0|)}].
\]

Proof of Point (iii). We first integrate equality (31) between \(s\) and \(u\), and take a supremum in time:

\[
Y_u = Y_s + (1 - e^{s-u})W_s + \int_s^u (1 - e^{v-u})F(v, Y_v) \, dv + N'_u,
\]

\[
\sup_{s \leq u \leq t} |Y_u - Y_s| \leq (t - s)|W_s| + \frac{t - s}{2} + \sup_{s \leq u \leq t} |N'_u|, \tag{36}
\]

with

\[
N'_u := \sqrt{2} \int_s^u (1 - e^{v-u}) \, dB_v,
\]

But, thanks to the properties of Brownian motion, by a change of time \(\tau = \sigma^2_{t-s} = 4e^{s-t} - e^{2(s-t)} - 3 + 2(t - s) \leq (t - s)^3\):

\[
\sup_{s \leq u \leq t} |N'_u| \overset{\mathcal{L}}{=} \sup_{0 \leq u \leq \tau} |B_u| \overset{\mathcal{L}}{=} \sqrt{\tau} \sup_{0 \leq u \leq 1} |B_u|.
\]

The law of supremum in absolute value of a 1D Brownian motion is explicitly known, see for instance [8, p. 342]. Here, we will use only simple estimates on the exponential moments:

\[
\mathbb{E}[e^{\lambda \sup_{0 \leq u \leq 1} |B_u|}] \leq \mathbb{E}[e^{\lambda \sup_{0 \leq u \leq 1} B_u}] + \mathbb{E}[e^{\lambda \sup_{0 \leq u \leq 1} (-B_u)}] \leq 2 \mathbb{E}[e^{\lambda |B_1|}] \leq 4e^{\frac{1}{2} \lambda^2}.
\]

In the second line, we use the well-known equality

\[
\sup_{0 \leq u \leq \tau} B_u \overset{\mathcal{L}}{=} \sup_{0 \leq u \leq \tau} (-B_u) \overset{\mathcal{L}}{=} |B_\tau|,
\]

and then the exponential moments given by (33). The constant 4 appearing above will raise some difficulties so we perform a little optimization to get rid of it. For any \(\theta \geq 1\), we may also bound

\[
\mathbb{E}[e^{\lambda \sup_{0 \leq u \leq 1} |B_u|}] \leq \mathbb{E}[e^{\lambda \theta \sup_{0 \leq u \leq 1} |B_u|}]^{\frac{1}{\theta}} \leq \left(4e^{\frac{1}{2} (\theta \lambda)^2}\right)^{\frac{1}{\theta}} = e^{\frac{1}{2} \ln 4 + \frac{1}{2} \lambda^2}.
\]

The optimal \(\theta\) seems to be \(\theta = \frac{2\sqrt{\ln 2}}{\lambda}\). It is admissible when \(\lambda \leq 2\sqrt{\ln 2}\). It leads to

\[
\mathbb{E}[e^{\lambda \sup_{0 \leq u \leq 1} |B_u|}] \leq e^{2\lambda \sqrt{\ln 2}} \leq e^{2\lambda}.
\]
Finally for $t-s \leq \lambda^{-2}$, the bound $\tau \leq (t-s)^3$ leads to $\lambda(t-s)^{-3/2} \leq 1$ and
\[ E \left[ e^{\lambda(t-s)^{-1} \sup_{s \leq u \leq t} |N_u|} \right] = E \left[ e^{\lambda(t-s)^{-1} \sqrt{\tau} \sup_{0 \leq u \leq 1} |B_u|} \right] \leq e^{2\lambda \sqrt{\tau} - s}. \]

Using the point (ii) of the Proposition and (36), we write
\[ E \left[ e^{\lambda(t-s)^{-1} \sup_{s \leq u \leq t} |Y_u - Y_s|} \right] \leq e^{\frac{1}{2} \lambda^2} E \left[ e^{\lambda |W_s|} \right] E \left[ e^{\lambda(t-s)^{-1} \sup_{s \leq u \leq t} |N_u|} \right] \leq e^{\frac{1}{2} \lambda^2} e^{(3+\lambda)} E \left[ e^{\lambda |W_0|} \right] e^{2\lambda \sqrt{\tau} - s}, \]
which concludes the proof, using that $\sqrt{\tau} - s \leq \frac{1}{2}$ by assumption.

**Proof of Point (iv).** Taking the expectation in (32),
\[ E[|W_t|] \leq e^{-\tau} E[|W_0|] + \frac{1}{2} + E[|M_t|]. \]
and using that $M_t$ is $N(0,1-e^{-2t})$ distributed, the expectation is simply bounded: $E[|M_t|] \leq \sqrt{2\pi} \leq 1$ and it implies the first bound of point (iv). The second bound uses (35) written with $s$ instead of 0, which leads in average to
\[ E[|Y_t - Y_s|] \leq (1 - e^{-\tau}) E[|W_s|] + \frac{t-s}{2} + E[|N_{t-s}|]. \]
Using, $1 - e^{-\tau} \leq t - s$ and the previous bound on $E[|W_s|]$, and the fact that $N_t$ is $N(0,\sigma^2_{t-s})$ distributed, with $\sigma^2_{t-s} \leq (t-s)^3$, it comes
\[ E[|Y_t - Y_0|] \leq (t-s) \left( E[|W_0|] + 2 + \frac{1}{2} + \sqrt{\frac{2(t-s)}{\pi}} \right), \]
and the conclusion follows when $t-s \leq \frac{1}{4}$.

**Appendix C. Proof of the existence part in Theorem 2.3.** This section is devoted to the conclusion of the proof of Theorem 2.3. The uniqueness part was proved in Section 3, together with a very useful propagation of uniform control on the velocity tails (Lemma 3.2).

**Step 1. Completeness estimates.** Thanks to the propagation of the uniform estimates on the tails in velocity and the crucial Lemma 3.1, we are in position to show that the family $(Y^n_t, W^n)_{n \geq 0}$ of solutions to (12) has the Cauchy property.

**Lemma C.1.** For $\eta, \eta' > 0$, let $(Y^n_t, W^n_t)_{t \geq 0}$ and $(Y'^n_t, W'^n_t)_{t \geq 0}$ be two (unique) solutions of the nonlinear SDE (12), constructed on a given probability basis, with a common Brownian motion $(B_t)_{t \geq 0}$ and initial condition chosen as (13). If the law of the initial condition satisfies either $\|f_0\|_{e,\lambda} < \infty$ for some $\lambda > 0$ or $\|f_0\|_{e,\gamma} < \infty$ for some $\gamma > 1$, then the following stability estimate holds for any $t \geq 0$:
\[ E \left[ \sup_{s \in [0,t]} \left( |Y^n_s - Y'^n_s| + |W^n_s - W'^n_s| \right) \right] \leq \frac{3}{2} \left( \eta + \eta' \right) \exp \left( (1 + 8K_{t,\eta+\eta'}) t \right), \tag{37} \]
where $K_{t,\eta+\eta'}$ is the constant appearing in the r.h.s of (16) with $\eta$ replaced by $\eta+\eta'$, i.e. respectively
\[ K_{t,\eta+\eta'} = \frac{4}{\lambda} e^{t+\frac{1}{2} + \frac{t^2}{2}} e^{\lambda(\eta+\eta')} \|f_0\|_{e,\lambda e^{-t}}, \text{ or } \quad K_{t,\eta+\eta'} = \frac{2\gamma}{\gamma - 1} C_\gamma e^{3t}(\eta+\eta')^\gamma \|f_0\|_{p,\gamma}. \]
Proof. Our strategy is the same as in the proof of the weak strong stability estimate in Subsection 3.1. For \( t \in [0,T] \) we have using (12) and (13):

\[
\sup_{s \in [0,t]} \left| Y_s^\eta - Y_s^{\eta'} \right| \leq \int_0^t \left| W_s^{\eta} - W_s^{\eta'} \right| \, ds + |\eta - \eta'| |U|
\]

\[
\sup_{s \in [0,t]} \left| W_s^{\eta} - W_s^{\eta'} \right| \leq \int_0^t \left| \bar{K}_\eta[\mu_s^{\eta}](Y_s^\eta) - \bar{K}_\eta[\mu_s^{\eta'}](Y_s^{\eta'}) \right| \, ds + |\eta - \eta'| |V|
\]

\[
\leq \int_0^t \left( \left| \bar{K}_\eta[\mu_s^{\eta}](Y_s^\eta) - \bar{K}_\eta[\mu_s^{\eta'}](Y_s^{\eta'}) \right| + \left| \bar{K}[\mu_s^{\eta}](Y_s^\eta) - \bar{K}[\mu_s^{\eta'}](Y_s^{\eta'}) \right| \right) \, ds + |\eta - \eta'|,
\]

since \(|V|\) and \(|U|\) are always bounded by 1 (we recall that \( \chi \) has its support included in \([-1,1]\)). But thanks to (11), for any \( y \in \mathbb{R} \):

\[
\left| \bar{K}_\eta[\mu_s^{\eta}](y) - \bar{K}[\mu_s^{\eta}](y) \right| = \left| \int (K_\eta - K)(y-y') \rho_s^{\eta}(dy') \right| \leq \int I_{[-\eta,\eta]}(y-y') \rho_s^{\eta}(dy') \leq 2\eta \|\rho_s^{\eta}\|_{\infty}.
\]

This allows to bound the first and third term, in the r.h.s. of the second inequality by \( 2 \int_0^t (|\eta|\|\rho_s^{\eta}\|_{\infty} + |\eta'|\|\rho_s^{\eta'}\|_{\infty}) \, ds \). And the second term is estimated in expectation with the help of Lemma 3.1:

\[
\mathbb{E}\left[ \left| \bar{K}[\mu_s^{\eta}](Y_s^\eta) - \bar{K}[\mu_s^{\eta'}](Y_s^{\eta'}) \right| \right] \leq 8 \min(\|\rho_s^{\eta}\|_{\infty}, \|\rho_s^{\eta'}\|_{\infty}) \mathbb{E}[|Y_s^\eta - Y_s^{\eta'}|] \leq 4(\|\rho_s^{\eta}\|_{\infty} + \|\rho_s^{\eta'}\|_{\infty}) \mathbb{E}[|Y_s^\eta - Y_s^{\eta'}|].
\]

Gathering all of this leads to

\[
\mathbb{E}\left[ \sup_{s \in [0,t]} \left( \left| Y_s^\eta - Y_s^{\eta'} \right| + \left| W_s^{\eta} - W_s^{\eta'} \right| \right) \right] \leq \int_0^t \alpha(s) \left( \mathbb{E}[|Y_s^\eta - Y_s^{\eta'}| + |W_s^{\eta} - W_s^{\eta'}|] + 3(\eta + \eta') \right) \, ds,
\]

with \( \alpha(s) := 1 + 4\|\rho_s^{\eta}\|_{\infty} + 4\|\rho_s^{\eta'}\|_{\infty} \). A simple application of the Gronwall’s lemma to \( \mathbb{E}\left[ \sup_{s \in [0,t]} (|Y_s^\eta - Y_s^{\eta'}| + |W_s^{\eta} - W_s^{\eta'}|) \right] + 3(\eta + \eta')/2 \) leads to

\[
\mathbb{E}\left[ \sup_{s \in [0,t]} \left( \left| Y_s^\eta - Y_s^{\eta'} \right| + \left| W_s^{\eta} - W_s^{\eta'} \right| \right) \right] \leq \frac{3}{2}(\eta + \eta') \exp\left( \int_0^t \alpha(s) \, ds \right).
\]

The conclusion follows from Lemma 3.2, which implies that \( \|\rho_s^{\eta}\|_{\infty} \leq K_{s,\eta} \leq K_{t,\eta+\eta'} \) for any \( s \in [0,t] \) (and a similar inequality with \( \eta' \) replacing \( \eta \)).

**Step 2. The Cauchy property in the Wiener space.** We now consider the space \( \mathcal{A} \) of measurable applications (or random variables) from \( \Omega = C(\mathbb{R}^2;\mathbb{R}^2) \) (with the Wiener measure) into itself. We endow it with the topology of uniform convergence on compact (in time) subsets. An associated distance to that topology is for instance

\[
\forall Y, Z \in \mathcal{A}, \quad d(Y, Z) = \sum_{n \in \mathbb{N}^*} \frac{1}{2^n} \mathbb{E}\left( 1 \wedge \sup_{t \in [0,2^n]} \left( |Y_t^1 - Z_t^1| + |Y_t^2 - Z_t^2| \right) \right), \quad (38)
\]
for which \(\mathcal{A}\) is complete (the symbol \(\wedge\) stand for the minimum). Let us consider the “sequence” \((Y^n, W^n)_{\eta > 0}\) (the correct denomination is “net”). By Lemma C.1 it is a Cauchy “sequence” (or net) in \((\mathcal{A}, d)\), and then it converges towards a certain \((Y_t, W_t)_{t \geq 0}\) in \((\mathcal{A}, d)\).

At a fixed time \(t\), this implies the convergence in probability and then in law of \((Y^n_t, W^n_t)\) toward \((Y_t, W_t)\). In particular the time marginals \(\mu^n_t\) weakly converge (as measures) towards \(f_t\). Using a standard argument, we can pass in the limit in the uniform bound (15) obtained in Lemma 3.2. So that for any time \(t\), the density of the law \(f_t\) of \((Y_t, W_t)\) satisfies one of the bound in (7).

**Step 3. Identification of the limit.** In order to prove that \((Y_t, W_t)_{t \geq 0}\) is a solution to (3) we have to show that for any \(t \geq 0:\)

\[
\mathbb{E}\left[ \sup_{s \in [0,t]} |Y_s - Y_0 - \int_0^s W_u \, du| \right] = 0,
\]

\[
\mathbb{E}\left[ \sup_{s \in [0,t]} |W_s - W_0 + \int_0^s (W_u - \tilde{K}[\mu_u](Y_u)) \, du - B_s| \right] = 0.
\]

But this is something we know for the approximated process \((Y^n_t, W^n_t)_{t \geq 0}\). Precisely, by Definition of strong solution to (12):

\[
\mathbb{E}\left[ \sup_{s \in [0,t]} |Y^n_s - Y_0 - \eta U - \int_0^s W^n_u \, du| \right] = 0,
\]

\[
\mathbb{E}\left[ \sup_{s \in [0,t]} |W^n_s - W_0 - \eta V + \int_0^s (W^n_u - \tilde{K}[\mu^n_u](Y^n_u)) \, du - B_s| \right] = 0.
\]

But the convergence in \((\mathcal{A}, d)\) allows to pass to the limit in the first equality above, and we obtain the first equality in (39). In order to pass to the limit in the second inequality, and get the second part of (39), the main difficulty is to handle the non-linear term. Precisely, we will show in the rest of the proof that

\[
\sup_{s \in [0,t]} \mathbb{E}\left( |\tilde{K}_{\eta}[\mu^n_s](Y^n_s) - \tilde{K}[\mu_s](Y_s)| \right) \xrightarrow{\eta \to 0} 0.
\]

This will conclude the proof.

For this, we introduce a independent copy \((\tilde{Y}^n_t, \tilde{W}_t)\) of the couple \((Y^n_t, Y_t)\). We then rewrite the force with the help of that independent copy and estimate

\[
\mathbb{E}\left[ |\tilde{K}_{\eta}[\mu^n_s](Y^n_s) - \tilde{K}[\mu_s](Y_s)| \right] = \mathbb{E}\left[ |\tilde{K}_{\eta}(Y^n_s - \tilde{Y}^n_s) - K(Y_s - \tilde{Y}_s)| \right] \leq \mathbb{E}\left[ |K_{\eta} - K|(Y^n_s - \tilde{Y}^n_s) \right] + \mathbb{E}\left[ |K(Y^n_s - \tilde{Y}^n_s) - K(Y_s - \tilde{Y}_s)| \right].
\]

The first term in the r.h.s is bounded thanks to (11):

\[
\mathbb{E}\left[ |K_{\eta} - K|(Y^n_s - \tilde{Y}^n_s) \right] \leq \mathbb{E}\left[ |I[\eta > \eta]|(Y^n_s - \tilde{Y}^n_s) \right] \leq \mathbb{E}\left[ |I[\eta > \eta]|(Y^n_s - \tilde{Y}^n_s)|Y^n_s \right] \leq \mathbb{E}\left[ 2\|\rho^n_\eta\|_{\infty} \right] = 2\eta\|\rho^n_\eta\|_{\infty}.
\]

The second term in the r.h.s is bounded thanks to Lemma 3.1 and we get

\[
\mathbb{E}\left[ |K_{\eta}[\mu^n_s](Y^n_s) - \tilde{K}[\mu_s](Y_s)| \right] \leq 2\eta\|\rho^n_\eta\|_{\infty} + 8\|\rho_s\|_{\infty}\mathbb{E}[|Y^n_s - Y_s|].
\]

Since \(\|\rho^n_\eta\|_{\infty}\) and \(\|\rho_s\|_{\infty}\) are bounded uniformly in time on \([0,t]\) and for \(\eta \in (0,1)\), and thanks to the convergence of \((Y^n, W^n)\) towards \((Y, W)\) in \((\mathcal{A}, d)\), it is simple to conclude that the requested term goes to zero, as \(\eta\) goes to zero.
Appendix D. Proof of Corollary 1. The existence part is a simple consequence of the existence of the process solution to the nonlinear SDE (3), for a given initial condition with a polynomial or exponential decay of the velocity tails (use Ito’s rule). So we only prove the weak-strong stability estimate.

Let \((g_t)_{t \geq 0}\) be weak solution to (4) starting from \(g_0 \in \mathcal{P}_1(\mathbb{R}^2) \cap L^1(\mathbb{R}^2)\) and \((f_t)_{t \geq 0}\) be another weak solution to (4) starting from \(f_0 \in \mathcal{P}_1(\mathbb{R}^2) \cap L^1(\mathbb{R}^2)\) satisfying for any \(t > 0\)

\[
\sup_{s \leq t} \|\rho_s\|_{\infty} < \infty, \quad \text{where} \quad \rho_s(x) = \rho_s^f := \int_{\mathbb{R}} f_s(x, dv).
\]

By [9, Theorem 2.6] or [11, Proposition B.1], there exists a stochastic basis \((\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})\) and a Brownian motion \((B_t)_{t \geq 0}\) on this basis and a process \((X_t, V_t)_{t \geq 0}\) solution to (3), which has exactly the time marginal \(g_t\) at any time \(t \geq 0\). Extending the probability space, we may choose a r.v. \((Y_0, W_0) \sim f_0\) such that

\[
W_1(f_0, g_0) = \mathbb{E}\left[|Y_0 - X_0| + |W_0 - V_0|\right].
\]

Next, remark that the force field \(\tilde{K}\) created by \(f\), precisely \(\tilde{K}_s(x) = \int_{\mathbb{R}} K(x - y)\rho_s^f(dy)\), is Lipschitz in position. In fact,

\[
|\tilde{K}_s(x) - \tilde{K}_s(y)| \leq \int_{\mathbb{R}} |K(x - z) - K(y - z)|\rho_s(dz) \leq \int_{\mathbb{R}} \mathbb{I}_{[x,y]}(z)|\rho_s(dz) \leq \|\rho_s\|_{\infty}|x - y|.
\]

Then, standard results allow us to build on that probability space a stochastic process \((Y_t, W_t)_{t \geq 0}\) solution to

\[
Y_t = Y_0 + \int_0^t W_s \, ds, \quad W_t = W_0 + \int_0^t \tilde{K}_s(Y_s) \, ds - \int_0^t W_s \, ds + \sqrt{2} B_t.
\]

Note that the family of time marginals \((h_t)_{t \geq 0}\) of \((Y_t, W_t)\) is a solution to the following linear Vlasov-Fokker-Planck equation

\[
\partial_t h_t + v \partial_v h_t + \tilde{K}_t(x) \partial_v h_t = \partial_v (\partial_v h_t + vh_t), \quad \text{where} \quad \tilde{K}_s(x) = K * \rho_s^f(x), \quad (40)
\]

with initial condition \(h_{t=0} = f_0\). Of course, that equation is also satisfied by \(f\) by assumption. But since \(\tilde{K}\) is globally Lipschitz in the space variable, uniformly in the time variable, uniqueness holds for equation (40) in the class \(L^\infty(\mathcal{P}(\mathbb{R}^2))\) by standard results. So, we have \(h_t = f_t\) for all time \(t \geq 0\) and \((Y_t, W_t)_{t \geq 0}\) is actually a solution to (3) defined on the same probability space as \((X_t, V_t)_{t \geq 0}\). Then applying the point (ii) of Theorem 2.3 to those processes and recalling that

\[
W_1(f_t, g_t) \leq \mathbb{E}\left[|Y_t - X_t| + |W_t - V_t|\right],
\]

we get the expected estimate.

Appendix E. Regularity of solutions to the VPFP equation.

Proposition 4. Let \(K : (t,x) \in \mathbb{R}^+ \times \mathbb{R} \mapsto K_t(x) \in \mathbb{R}\) be a function belonging to \(L^\infty_{\text{loc}}(\mathbb{R}^+, C^k(\mathbb{R}))\) for all \(k \in \mathbb{N}\) and to \(C^1(\mathbb{R}^+ \times \mathbb{R}^2)\). Consider the unique (in the class of measures) solution to the following linear PDE

\[
\partial_t f_t + v \partial_v f_t + K_t(x) \partial_v f_t = \partial_v (\partial_v f_t + vf_t), \quad (41)
\]

for an initial condition \(f_0 \in \mathcal{P}(\mathbb{R}^2)\) satisfying \(\partial_v^k \partial_v^l f_0 \in L^2(\mathbb{R}^2)\) for all \(k,l \in \mathbb{N}\). Then \(f \in C^1(\mathbb{R}^+ \times \mathbb{R}^2)\) and is even two times continuously differentiable in \((x, v)\).
Proof. Differentiating equation (41) $k$ times in $x$ variable, and $l$ times in $v$ variable leads to
\[
\partial_t (\partial_x^k \partial_v^l f_t) + v \partial_x (\partial_x^k \partial_v^l f_t) + K_t(x) \partial_v (\partial_x^k \partial_v^l f_t) - \partial_v (v \partial_x^k \partial_v^l f_t) - \partial_x^2 (\partial_x^k \partial_v^l f_t)
\]
\[
= - \sum_{k'=0}^{k-1} \left( \begin{array}{c} k \\ k' \end{array} \right) \partial_x^{k-k'} K_t(x) (\partial_x^{k'} \partial_v^{l+1} f_t) - \partial_x^{k+1} \partial_x^{l-1} f_t + l \partial_x^2 \partial_v^l f_t, \quad (42)
\]
with the convention that a term containing a derivative with a negative power vanishes. Multiplying the above equation by $\partial_x^k \partial_v^l f_t$ and performing an integration by part, we obtain
\[
\frac{d}{dt} \| \partial_x^k \partial_v^l f_t \|_2^2 \leq \| \partial_x^k \partial_v^l f_t \|_2 \| \partial_x^{k+1} \partial_v^{l-1} f_t \|_2
\]
\[
+ C_{t,k} \sum_{k'=0}^{k-1} \left( \begin{array}{c} k \\ k' \end{array} \right) \| \partial_x^{k-k'} \partial_v^{l+1} f_t \|_2 \| \partial_x^k \partial_v^l f_t \|_2 + \left( l + \frac{1}{2} \right) \| \partial_x^k \partial_v^l f_t \|_2^2,
\]
where
\[
C_{t,k} = \sup_{k'=1, \ldots, k} \| \partial_x^{k'} K_t \|_{\infty}.
\]
Then summing these inequalities over the $k, l$ such that $k + l \leq m$, we find
\[
\frac{d}{dt} H_m(t) \leq \tilde{C}_{m,t} H_m(t), \quad \text{where} \quad H_m(t) = \sum_{k+l \leq m} \| \partial_x^k \partial_v^l f_t \|_2^2,
\]
and the constant $\tilde{C}_{m,t}$ is locally bounded in time. Since by assumption $H_m(0) < \infty$ for any $m \in \mathbb{N}$, we get that for any time $t \geq 0$, $\sup_{s \in [0,t]} H_m(s) < \infty$. Then by Morrey’s inequality for $m$ large enough ($m = 4$), $f \in L^\infty([0,t] \times C^2(\mathbb{R}^2))$, and so does $\partial_x^k \partial_v^l f$ for all $k, l \geq 0$. Using (42), we deduce that $\partial_t \partial_x^k \partial_v^l f \in L^\infty_{loc}(\mathbb{R}^+ \times \mathbb{R}^2)$, and in particular $\partial_x^k \partial_v^l f$ is continuous. So $f$ is two times continuously differentiable in $(x, v)$. And using finally (41), we see that $\partial_t f$ is itself continuous in all the variables, as a sum of continuous functions. This concludes the proof. \hfill \Box

REFERENCES

[1] R. J. Berman and M. Onnheim, Propagation of chaos, Wasserstein gradient flows and toric Kähler-Einstein metrics, Anal. PDE, 11 (2018), 1343–1380, arXiv:1501.07820, 2015.

[2] R. Catellier and M. Gubinelli, Averaging along irregular curves and regularisation of odes, Stochastic Process. Appl., 126 (2016), 2323–2366, arXiv:1205.1735.

[3] E. Cépa and D. Lépine, Diffusing particles with electrostatic repulsion, Probab. Theory Related Fields, 107 (1997), 429–449.

[4] Y.-P. Choi and S. Salem, Propagation of chaos for aggregation equations with no-flux boundary conditions and sharp sensing zones, M3AS, 28 (2018), 223–258.

[5] M. Cullen, W. Gangbo and G. Pisante, The semigeostrophic equations discretized in reference and dual variables, Arch. Ration. Mech. Anal., 185 (2007), 341–363.

[6] A. M. Davie, Uniqueness of solutions of stochastic differential equations, Int. Math. Res. Not., 2007 (2007), Art. ID rnm124, 26 pp.

[7] E. Fedrizzi, F. Flandoli, E. Priola and J. Vovelle, Regularity of stochastic kinetic equations, Electron. J. Probab., 22 (2017), Paper No. 48, 42 pp, arXiv:1606.01088.

[8] W. Feller, An introduction to probability theory and its applications, Second edition John Wiley & Sons, Inc., New York-London-Sydney, 1971.

[9] A. Figalli, Existence and uniqueness of martingale solutions for SDEs with rough or degenerate coefficients, J. Funct. Anal., 254 (2008), 109–153.

[10] N. Fournier and A. Guillin, On the rate of convergence in Wasserstein distance of the empirical measure, Probab. Theory Related Fields, 162 (2015), 707–738.

[11] N. Fournier and M. Hauray, Propagation of chaos for the Landau equation with moderately soft potentials, arXiv:1501.01802, Ann. Probab., 44 (2016), 3581–3660.
[12] N. Fournier, M. Hauray and S. Mischler, Propagation of chaos for the 2D viscous Vortex model, J. Eur. Math. Soc., 16 (2014), 1423–1466.
[13] M. Hauray, Mean field limit for the one dimensional Vlasov-Poisson equation, Séminaire Laurent Schwarz, Palaiseau, 2014, arXiv:1309.2531.
[14] M. Hauray and S. Mischler, On Kac’s chaos and related problems, J. Funct. Anal., 266 (2014), 6055–6157.
[15] P.-E. Jabin and Z. Wang, Mean field limit and propagation of chaos for vlasov systems with bounded forces, J. Funct. Anal., 271 (2016), 3588–3627, arXiv:1511.03769.
[16] G. Loeper, Uniqueness of the solution to the Vlasov-Poisson system with bounded density, In J. Math. Pures Appl., 86 (2006), 68–79.
[17] H. P. McKean, Jr. Propagation of chaos for a class of non-linear parabolic equations, In Stochastic Differential Equations (Lecture Series in Differential Equations, Session 7, Catholic Univ., 1967), pages 41–57.
[18] K. R. Parthasarathy, Probability measures on metric spaces, Academic Press, New York, 1967.
[19] A.-S. Sznitman, Topics in propagation of chaos, In École d’Été de Probabilités de Saint-Flour XIX—1989, volume 1464, chapter Lecture Notes in Math., pages 165–251. Springer, Berlin, 1991.
[20] A.-S. Sznitman, Équations de type de Boltzmann, spatialement homogènes, Z. Wahrsch. Verw. Gebiete, 66 (1984), 559–592.
[21] M. Trocheris, On the derivation of the one-dimensional Vlasov equation, Transport Theory Statist. Phys., 15 (1986), 597–628.
[22] A. Y. Veretennikov, On strong solutions and explicit formulas for solutions of stochastic integral equations, Math. USSR, Sb., 39 (1981), 387–403.

Received August 2016; revised March 2018.

E-mail address: maxime.hauray@univ-amu.fr
E-mail address: salem@ceremade.dauphine.fr