



HAL
open science

Quantitative approximation of nonlinear Fokker-Planck equations with singular potential by interacting particle systems

Christian Olivera, Alexandre Richard, Milica Tomasevic

► **To cite this version:**

Christian Olivera, Alexandre Richard, Milica Tomasevic. Quantitative approximation of nonlinear Fokker-Planck equations with singular potential by interacting particle systems. 2020. hal-02914779v1

HAL Id: hal-02914779

<https://inria.hal.science/hal-02914779v1>

Preprint submitted on 12 Aug 2020 (v1), last revised 27 Jun 2022 (v4)

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

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

Quantitative approximation of nonlinear Fokker-Planck equations with singular potential by interacting particle systems

Christian Olivera ^{*} Alexandre Richard[†] Milica Tomašević [‡]

August 12, 2020

Abstract

We propose a new approach to show quantitative estimates for solutions of nonlinear Fokker-Planck equations with singular potential by stochastic interacting particle systems. Our result only requires very weak regularity on the interaction kernel, including the Biot-Savart kernel, the family of Keller-Segel kernels in arbitrary dimension, and more generally singular Riesz kernels. The proof is based on a semigroup approach, and we also exploit the regularity of the solutions of the limiting equation.

Key words and phrases: *Interacting particle systems, Nonlinear Fokker-Planck equation, Analytic semigroups, McKean-Vlasov SDEs, Propagation of chaos.*

1	Introduction	2
2	Main results	6
2.1	Framework and assumptions	6
2.2	Convergence of the particle system	8
2.3	Rate of convergence to the PDE solution	9
2.4	Propagation of chaos	10
2.5	Comparison with previous results	11
3	Properties of the PDE and of the PDE with cut-off	11
3.1	Some classical embeddings	12
3.2	Properties of mild solutions of the PDE	12
3.3	Properties of mild solutions of the PDE with cut-off	15
4	Convergence: Proof of Theorem 2.6	17
4.1	Strategy of the proof	17
4.2	Proofs of intermediate results	19
4.3	Proof of Corollary 2.7	24
4.4	Proof of Corollary 2.9	26
5	Rate of convergence: Proof of Theorem 2.11	27
6	Propagation of chaos	37
6.1	Proof of Proposition 2.14	37
6.2	Proof of Theorem 2.15	38

^{*}Departamento de Matemática, Universidade Estadual de Campinas, Brazil. colivera@ime.unicamp.br.

[†]Université Paris-Saclay, CentraleSupélec, MICS and CNRS FR-3487. alexandre.richard@centralesupelec.fr.

[‡]CMAP, Ecole Polytechnique, CNRS, I.P. Paris, 91128 Palaiseau, France. milica.tomasevic@polytechnique.edu.

7	Examples	40
7.1	General classes of kernels satisfying Assumption (\mathbf{H}^K)	40
7.2	Riesz and Coulomb potentials	41
7.3	Biot-Savart kernel and the 2d Navier-Stokes equation	41
7.4	Parabolic-elliptic Keller-Segel models	42
7.5	Attractive-repulsive kernels	42
	Appendix	43
A.1	Proofs of the tightness estimates	43
A.2	Regularisation property of the kernel	45
	References	48

1 Introduction

In this work, we are interested in the stochastic particle approximation of parabolic Partial Differential Equations (PDEs) of the form

$$\begin{cases} \partial_t u(t, x) = \Delta u(t, x) - \nabla \cdot (u(t, x) (\Phi(x) + K *_x u(t, x))), & t > 0, x \in \mathbb{R}^d, \\ u(0, x) = u_0(x), \end{cases} \quad (1.1)$$

where K is a locally integrable kernel that may have a singular behaviour at 0, and Φ is some bounded force applied to the system. A typical family of such singular kernels derives from Riesz potentials, defined in any dimension d as

$$V_s(x) := |x|^{-s}, \quad x \in \mathbb{R}^d, \quad \text{for } s \in (0, d). \quad (1.2)$$

The associated kernel is then $K_s := \nabla V_s$. If $d \geq 3$ and $s = d - 2$, this is the Coulomb potential that characterises electrostatic and gravitational forces (depending on the sign). Our motivation comes more particularly from several classical models. First, there is the 2d Navier-Stokes equation, which in vorticity form can be written as in (1.1) with the Biot-Savart kernel

$$K_{BS}(x) = \frac{1}{\pi} \frac{x^\perp}{|x|^2}.$$

(Note that $|K(x)| = |\nabla V_0(x)|$, where $V_0(x) = \log(|x|)$ is the fundamental solution of the Laplace equation in dimension 2).

We are also interested in the parabolic-elliptic Keller-Segel PDE in any dimension $d \geq 1$, for which the kernel is given by

$$K_{KS}(x) = -\chi \frac{x}{|x|^d}$$

for some $\chi > 0$. Equation (1.1) then describes the chemotaxy of biological cells subject to diffusive displacement and attracted by the concentration of nutrients in their environment. It entails a particularly tricky mathematical problem due to the fact that the dynamics is attractive and that a blow-up may occur in the PDE, depending on the value of χ .

Equation (1.1) also arise in many different contexts, as for instance swarming, aggregation phenomena, see e.g. [8, 43].

Note that by requiring enough regularity on Φ (typically bounded and Lipschitz continuous), the term $\Phi(x)$ does not bring any particular difficulty, and thus we assume without restriction that $\Phi \equiv 0$. Besides, Equation (1.1) preserves the total mass $M := \int_{\mathbb{R}^d} u_0(x) dx$. That is why, we may and will assume from now on that $M = 1$ in (1.1) and we will not change the notation for K .

The problem of deriving a macroscopic equation from a microscopic model of interacting particles can be traced back to the original inspiration of Kac [27], in the context of the Boltzmann

equation. Since then, a huge literature has been devoted to interacting particle systems and their convergence to (1.1), mostly in the case of Lipschitz continuous kernels of interaction: when the particles are interacting diffusions, this problem is now well-understood, see e.g. Sznitman [41] and Méléard [29] for a general account of the theory.

The case of singular kernels such as the aforementioned ones is more recent and there are fewer works on this topic, despite the great challenge it represents both theoretically and in applications. One can mention the early works of Bossy and Talay [5] and Sznitman [40] on Burgers' equation, and of Marchioro and Pulvirenti [28] and Osada [36] on the 2d Navier-Stokes equation. Cépa and Lépingle [11] studied one-dimensional electrical particles with repulsive interaction, and more recently Fournier and Hauray [19] studied a stochastic particle system approximating the Landau equation with moderately soft potentials. Outside the scope of physics, interesting biological models have arisen, for instance in neuroscience with the work of Delarue et al. [13] (diffusive particles interacting through their hitting times), and several works on the Keller-Segel equation (Fournier and Jourdain [20], Cattiaux and Pédèches [9] and Jabir et al. [25]). Let us also mention that similar singular non-diffusive problems (i.e. with deterministic particles) have seen significant progress recently, see e.g. the work of Serfaty [38] and Duerinckx [14].

Although this will not be our line of investigation in this work, let us briefly recall that the mean field particle system associated to (1.1) reads:

$$dX_t^{i,N} = \frac{1}{N} \sum_{j=1}^N K(X_t^{i,N} - X_t^{j,N}) dt + \sqrt{2} dW_t^i, \quad (1.3)$$

where $\{W_t^i, i \in \mathbb{N}\}$ is a family of independent d -dimensional Brownian motions (with the convention that $K(0) = 0$). Unfortunately, due to the singular interaction kernel, it is not obvious that this particle system is well-defined and that the propagation of chaos holds. The study of this system is usually done on a case-by-case basis. Another approach for tackling the convergence of particles interacting through singular kernels, that we adopt in the present paper, is to consider moderately interacting particles in the sense of Oelschläger [33]. In the moderate interaction setting, one introduces a smoothing of the interaction kernel. This has been successfully applied in the aforementioned works [5], [19], as well as by Méléard [30], and more recently with a new semigroup approach developed by Flandoli et al. [16, 18] and also in the companion paper [35].

Hence for $N \geq 1$, we consider the following particle system:

$$dX_t^{i,N} = F_A \left(\frac{1}{N} \sum_{k=1}^N (K * V^N)(X_t^{i,N} - X_t^{k,N}) \right) dt + \sqrt{2} dW_t^i, \quad t \leq T, \quad 1 \leq i \leq N, \quad (1.4)$$

where V^N is a mollifier, F_A is a smooth cut-off function that ensures that the drift driving each particle remains uniformly bounded in N , and $A > 0$ is the cut-off parameter. As such, the existence of strong solutions for (1.4) is ensured.

Our first objective is to prove the uniform convergence of its mollified empirical measure towards the solution of (1.1) when the number of particles goes to infinity, for a large class of interacting kernels, including those deriving from a singular potential that includes the Biot-Savart kernel, Coulomb potentials in any dimension (actually even more singular Riesz kernels with $s \in (0, d-1)$) and the attractive Keller-Segel potential.

For that purpose, we follow the new approach presented in Flandoli et al. [16], based on semigroup theory and developed first with application to the FKPP equation. This technique permits to approximate non-linear PDEs by smoothed empirical measures in strong functional topologies. It has already found many applications: see Flandoli and Leocata [15] for a PDE-ODE system related to aggregation phenomena; Simon and Olivera [39] for non-local conservation laws; and Flandoli et al. [18] for the 2d Navier-Stokes equation.

The main difficulty here is the singular nature of the kernel and then finding a suitable functional framework. In particular, it is crucial in our approach to get a solution theory for the class of PDEs (1.1). Hence in Section 3, we establish the (local) existence and uniqueness of mild solutions to (1.1), as well as an ansatz with cut-off corresponding to the limit of the particle system (1.4), in the

space $\mathcal{C}([0, T]; L^1 \cap W^{1,D}(\mathbb{R}^d))$ for some $D \geq d$. Then, we obtain the convergence in probability of the mollified empirical measure of the particle system (1.4) towards the solution of (1.1) in spaces of the form $L^2([0, T], H^\gamma(\mathbb{R}^d)) \cap L^\infty([0, T], L^1(\mathbb{R}^d))$, where H^γ is a nonhomogeneous Sobolev space (also known as Bessel potential space), see Theorem 2.6 and Corollary 2.7. Note that even for PDEs having a blow-up in finite time (such as the Keller-Segel model with $d \geq 2$ and χ large enough), this convergence happens, at least before the blow-up.

Beyond some fluctuation results as in [26, 34], the problem of quantifying the previous convergences is however very recent and the literature is very scarce. We are only aware of the following results: Méléard [30] obtained a non-explicit rate on the density of a particle for the 2d Navier-Stokes equation, Fournier and Mischler [21] obtained a rate for the so-called Nanbu particle system approaching the Boltzmann equation, Fournier and Hauray [19] approximated the Landau equation with moderately soft potentials, which is a step closer to the present work (see discussion in Section 2.5); and for a more systematic study, Jabin and Wang [24] also recently obtained a quantitative propagation of chaos result on the marginal densities with a $N^{-1/2}$ rate for some singular kernels including the Biot-Savart kernel (but not Riesz kernels). In [6], they also provide a slower rate of convergence for the tricky attractive case of the Keller-Segel model (in dimension 2). We shall discuss in details and compare these works to ours in Section 2.5.

Our second main result gives a rate of convergence of the mollified empirical measure of the particle system (1.4) towards the solution of (1.1). This rate is measured in the energy norm, provided some extra regularity on the law of the initial distribution of particles (see Theorem 2.11). This seems to be the first time that such quantitative convergence results are obtained in normed functional spaces, whether for the Biot-Savart, Riesz or Keller-Segel kernels. For instance, in the particular case of the parabolic-elliptic Keller-Segel PDE, our result permits to obtain for the first time an explicit quantitative result of convergence in any dimension. However the rates we provide are always below $\frac{1}{2}$, which is the rate one would classically hope to obtain for non-singular interactions of mean field type (see e.g. [10]) or even in the Biot-Savart case [24] (although the measured quantities are not the same). Nevertheless, our particles do not interact in a mean field way, but rather within a typical range of order $N^{-\alpha}$, for some $\alpha \in (0, 1)$. As it happens, we need $\alpha < \frac{1}{2(d+2)}$ for the convergence to hold, and we do obtain this rate in the end, which is in this sense optimal.

Finally, our last results concern the weak well-posedness of the McKean-Vlasov SDE associated to the particle system (1.4) (Proposition 2.14) and the propagation of chaos of the empirical measure of the particle system (1.4) towards the solution of the PDE with cut-off, at the level of probability measures on the space of trajectories (Theorem 2.15). This is done in the spirit of the work of Méléard and Roelly-Coppoletta [31], that we extend to singular kernels. As a consequence, and by choosing A large enough in (1.4), we obtain the weak well-posedness of the McKean-Vlasov SDE

$$\begin{cases} dX_t = (\Phi(X_t) + K * u_t(X_t)) dt + \sqrt{2}dW_t, & t \leq T, \\ \mathcal{L}(X_t) = u_t, \mathcal{L}(X_0) = u_0, \end{cases} \quad (1.5)$$

assuming that u_0 has enough regularity ($u_0 \in L^1 \cap W^{1,D}(\mathbb{R}^d)$, $D \geq \max(d, 2)$).

Although there have been some recent progress on the study of McKean-Vlasov SDEs with non-Lipschitz coefficients (usually measured in Wasserstein distance, see e.g. [2, 12, 23]), a drift of the form $K * \mu_t$ seems to be too singular to be continuous in Wasserstein distance for the aforementioned kernels of interest. Hence we provide a general well-posedness result for singular nonlinear SDEs. In some particular cases such as the Keller-Segel model, this extends previous results [20], but it requires more regularity on the initial data.

Plan of the paper. The framework, the assumptions on the kernel K , as well as the main results of this work are presented in Section 2. Since the well-posedness and regularity of the limit PDE play an important role in our argument, the existence and uniqueness of (1.1) and its

cut-off version are studied in Section 3 (actually the existence of (1.1) is obtained later in the paper as a limit point of the particle system). In Section 4, we detail the proof that gives the tightness of the mollified empirical measure (Theorem 2.6). This result is similar to the semigroup argument of [16, 18] but we give it more generality and extend it to some important functional spaces (Corollary 2.7). Then we prove that any limit point has enough regularity to be a mild solution and the uniqueness property of Section 3 permits to conclude for the convergence. Then the proof for the rate of convergence is developed in Section 5. In Section 6, the uniqueness of the martingale problem associated to the McKean-Vlasov SDE is proven (Proposition 2.14), as well as the propagation of chaos for the empirical measure of the interacting particle system (Theorem 2.15). Finally, we present some examples and applications of our results in Section 7. The proof of the tightness estimates and the regularisation property of the kernel are gathered in Appendix.

Notations and definitions. Sobolev spaces will be of frequent use in the paper: $W^{k,p}(\mathbb{R}^d)$, $k \in \mathbb{N}$, $p \geq 1$, denotes the usual Sobolev spaces of k times weakly differentiable elements of $L^p(\mathbb{R}^d)$ whose derivatives are also in $L^p(\mathbb{R}^d)$.

For any $\alpha \in \mathbb{R}$, we denote by $H^\alpha(\mathbb{R}^d)$ the *Bessel potential space*

$$H^\alpha(\mathbb{R}^d) := \left\{ u \in L^2(\mathbb{R}^d) ; \mathcal{F}^{-1} \left((1 + |\cdot|^2)^{\frac{\alpha}{2}} \mathcal{F}u(\cdot) \right) \in L^2(\mathbb{R}^d) \right\},$$

where $\mathcal{F}u$ denotes the *Fourier transform* of u . These are Hilbert spaces when endowed with the scalar product

$$\langle u, v \rangle_\alpha := \int_{\mathbb{R}^d} (1 + |\xi|^2)^{\frac{\alpha}{2}} \mathcal{F}u(\xi) \overline{\mathcal{F}v(\xi)} d\xi$$

and the norm denoted by $\|u\|_\alpha^2 = \langle u, u \rangle_\alpha$. In particular, note that

$$\|u\|_0 = \|u\|_{L^2(\mathbb{R}^d)} \quad \text{and, for any } \alpha \leq \beta, \quad \|u\|_\alpha \leq \|u\|_\beta.$$

The space $H^\alpha(\mathbb{R}^d)$ is in fact a Sobolev space of fractional order, associated to the fractional operator $(I - \Delta)^{\frac{\alpha}{2}}$ defined as (see e.g. [44, p.180] for more details on this operator):

$$(I - \Delta)^{\frac{\alpha}{2}} f := \mathcal{F}^{-1} \left((1 + |\cdot|^2)^{\frac{\alpha}{2}} \mathcal{F}f \right) \tag{1.6}$$

Depending on the context, the brackets $\langle \cdot, \cdot \rangle$ will denote either the scalar product in some L^2 space or the duality bracket between a measure and some functions it applies to.

For any $R > 0$, denote by \mathcal{B}_R the centred ball of \mathbb{R}^d of radius R . For positive α and any ball $\mathcal{B}_R \subset \mathbb{R}^d$, the space $H^\alpha(\mathcal{B}_R)$ is defined in Triebel [44, p.310], and corresponds roughly to distributions f on \mathcal{B}_R which are restrictions of $g \in H^\alpha(\mathbb{R}^d)$. Then $H_{\text{loc}}^\alpha(\mathbb{R}^d)$ is the space of distributions f on \mathbb{R}^d such that $f \in H^\alpha(\mathcal{B}_R)$ for any $R > 0$.

In this paper, $(e^{t\Delta})_{t \geq 0}$ is the heat semigroup. That is, for $f \in L^2(\mathbb{R}^d)$,

$$(e^{t\Delta} f)(x) = \int_{\mathbb{R}^d} \frac{1}{(4\pi t)^{\frac{d}{2}}} e^{-\frac{|x-y|^2}{4t}} f(y) dy.$$

Obviously, $\nabla e^{t\Delta} f = e^{t\Delta} \nabla f$. Applying the convolution inequality [7, Th. 4.15] for $p \geq 1$ and using the equality $\left\| \nabla \frac{1}{(4\pi t)^{\frac{d}{2}}} e^{-\frac{|\cdot|^2}{4t}} \right\|_{L^1(\mathbb{R}^d)} = \frac{C}{\sqrt{t}}$, it comes that

$$\left\| \nabla e^{t\Delta} \right\|_{L^p \rightarrow L^p} \leq \frac{C}{\sqrt{t}}. \tag{1.7}$$

The space $\mathcal{C}(I; L^1 \cap H^1(\mathbb{R}^d))$ of continuous functions from the time interval I with values in $L^1 \cap H^1(\mathbb{R}^d)$ is endowed with the norm

$$\|f\|_{\mathcal{C}(I; L^1 \cap H^1(\mathbb{R}^d))} = \sup_{s \in I} (\|f_s\|_{L^1(\mathbb{R}^d)} + \|f_s\|_{H^1(\mathbb{R}^d)}).$$

In case $I = [0, t]$ for some $t > 0$, we will also use the notation

$$\|f\|_{t, L^1 \cap H^1(\mathbb{R}^d)} = \|f\|_{[0, t], L^1 \cap H^1(\mathbb{R}^d)}.$$

For functions from \mathbb{R}^d to \mathbb{R} , we will encounter the space of n -times ($n \in \mathbb{N}$) differentiable functions, denoted by $\mathcal{C}^n(\mathbb{R}^d)$; the space of n -times ($n \in \mathbb{N}$) differentiable functions with bounded derivatives of any order between 0 and n , denoted by $\mathcal{C}_b^n(\mathbb{R}^d)$; and the space of n -times ($n \in \mathbb{N}$) differentiable functions with compact support, denoted by $\mathcal{C}_c^n(\mathbb{R}^d)$. For $n = 0$, we will denote the space of continuous (resp. bounded continuous, and continuous with compact support) by $\mathcal{C}(\mathbb{R}^d)$ (resp. $\mathcal{C}_b(\mathbb{R}^d)$ and $\mathcal{C}_c(\mathbb{R}^d)$).

Let also denote by \mathcal{N}_δ the Hölder seminorm of parameter $\delta \in (0, 1]$, that is, for any function f defined over \mathbb{R}^d :

$$\mathcal{N}_\delta(f) := \sup_{x \neq y \in \mathbb{R}^d} \frac{|f(x) - f(y)|}{|x - y|^\delta}.$$

The set of continuous and bounded functions on \mathbb{R}^d which have finite \mathcal{N}_δ seminorm is the Hölder space $\mathcal{C}^\delta(\mathbb{R}^d)$.

Finally, if u is a function or stochastic process defined on $[0, T] \times \mathbb{R}^d$, we will most of the time use the notation u_t to denote the mapping $x \mapsto u(t, x)$.

2 Main results

The aim of this section is to present and discuss our main results. First we give in Section 2.1 the framework to establish our theorems, then state some general assumptions. Then we state our results about the convergence of particle systems to the PDE (1.1) and give rates, as well as the propagation of chaos and convergence to the associated McKean-Vlasov SDE (1.5).

2.1 Framework and assumptions

Let us introduce a cut-off in the reaction term of Equation (1.1). Namely, for any $A > 0$, let F_A be defined as follows: let $f_A : \mathbb{R} \rightarrow \mathbb{R}$ be a $\mathcal{C}_b^2(\mathbb{R})$ function such that:

- (i) $f_A(x) = x$, for $x \in [-A, A]$,
- (ii) $f_A(x) = A$, for $|x| > A + 1$,
- (iii) $\|f'_A\|_\infty \leq 1$ and $\|f''_A\|_\infty < \infty$.

As a consequence, $\|f_A\|_\infty \leq A + 1$. Now F_A is given by

$$F_A : (x_1, \dots, x_d)^T \mapsto (f_A(x_1), \dots, f_A(x_d))^T. \quad (2.1)$$

The modified PDE with cut-off now reads:

$$\begin{cases} \partial_t \tilde{u}(t, x) = \Delta \tilde{u}(t, x) - \nabla \cdot (\tilde{u}(t, x) F_A(K * \tilde{u}(t, x))), & t > 0, x \in \mathbb{R}^d \\ \tilde{u}(0, x) = u_0(x). \end{cases} \quad (2.2)$$

Although this is implicit, \tilde{u} actually depends on A . Note that if F_A is replaced by the identity function, one recovers (1.1). Solutions to (1.1) and (2.2) will be understood in the following sense:

Definition 2.1. *Let $D \geq \max(d, 2)$. Given $u_0 \in L^1 \cap W^{1, D}(\mathbb{R}^d)$ and $A > 0$, a function u on $[0, T_0] \times \mathbb{R}^d$ is said to be a mild solution to (2.2) on $[0, T_0]$ if*

- (i) $u \in \mathcal{C}([0, T_0]; L^1 \cap W^{1, D}(\mathbb{R}^d))$;
- (ii) u satisfies the integral equation

$$u_t = e^{t\Delta} u_0 - \int_0^t \nabla \cdot e^{(t-s)\Delta} (u_s F_A(K * u_s)) ds, \quad 0 \leq t < T_0. \quad (2.3)$$

Given $T_0 > 0$, we say that this problem admits a unique solution if for any couple of mild solutions u^1, u^2 and for any $T \in (0, T_0)$, $\|u^1 - u^2\|_{T, L^1 \cap W^{1,D}(\mathbb{R}^d)} = 0$.

A function u on $[0, \infty) \times \mathbb{R}^d$ is said to be a global mild solution to (2.2) if it is a mild solution to (2.2) on $[0, T_0)$ for all $T_0 > 0$.

Remark 2.2. Similarly, a mild solution to the original PDE (1.1) satisfies Definition 2.1 i) and solves

$$u_t = e^{t\Delta} u_0 - \int_0^t \nabla \cdot e^{(t-s)\Delta} (u_s K * u_s) ds, \quad 0 < t < T. \quad (2.4)$$

Compared to the singular particle system (1.3), we introduce a mollifier that will be used both to regularise the particle system and its empirical measure. Let $V : \mathbb{R}^d \rightarrow \mathbb{R}_+$ be a smooth, rapidly decreasing probability density function, and assume further that V is even. For any $x \in \mathbb{R}^d$, define

$$V^N(x) := N^{d\alpha} V(N^\alpha x), \quad \text{for some } \alpha \in [0, 1]. \quad (2.5)$$

For each $N \in \mathbb{N}$, we consider the following interacting particle system:

$$\begin{cases} dX_t^{i,N} = F_A \left(\frac{1}{N} \sum_{k=1}^N (K * V^N)(X_t^{i,N} - X_t^{k,N}) \right) dt + \sqrt{2} dW_t^i, & t \leq T, \quad 1 \leq i \leq N, \\ X_0^{i,N}, \quad 1 \leq i \leq N, & \text{are independent of } \{W^i, \quad 1 \leq i \leq N\}, \end{cases} \quad (2.6)$$

where $\{(W_t^i)_{t \in [0, T]}, i \in \mathbb{N}\}$ is a family of independent standard \mathbb{R}^d -valued Brownian motions defined on a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$.

Let us denote the empirical measure of N particles by

$$\mu^N = \frac{1}{N} \sum_{i=1}^N \delta_{X^{i,N}}, \quad (2.7)$$

and the mollified empirical measure by

$$u^N := V^N * \mu^N.$$

The following hypotheses on the initial conditions of the system will be assumed:

(H):

(H_i) There exists $\beta > \frac{D}{2}$ such that for all $p \geq 2$, $\sup_{N \in \mathbb{N}} \mathbb{E} \left[\|u_0^N\|_\beta^p \right] < \infty$.

(H_{ii}) Let $u_0 \in L^1 \cap H^\beta(\mathbb{R}^d)$ such that $u_0 \geq 0$ and $\|u_0\|_{L^1(\mathbb{R}^d)} = 1$. Assume that $\langle u_0^N, \varphi \rangle \rightarrow \langle u_0, \varphi \rangle$ in probability, for any $\varphi \in \mathcal{C}_b(\mathbb{R}^d)$.

(H_{iii}) The parameters α and β (which appear respectively in (2.5) and **(H_i)**) satisfy

$$0 < \alpha < \frac{1}{d + 2\beta}.$$

Remark 2.3. The reason for choosing to work in $H^\beta(\mathbb{R}^d)$ at the level of the particles is twofold: first, it allows to work with L^2 spaces, which is more natural when it comes to the isometry property of the Itô integral. Second, it gives the extra regularity that is needed to get rates of convergence. Note that choosing $\beta > \frac{D}{2}$ is natural as it ensures the correct embedding in $L^1 \cap W^{1,D}(\mathbb{R}^d)$ (see Section 3.1).

Let us now state the assumptions on the kernel K :

(\mathbf{H}^K) :

(\mathbf{H}_i^K) There exists $\kappa_M > 1$ such that $K \in L^\kappa(\mathcal{B}_1)$ for all $\kappa \in [1, \kappa_M)$;

(\mathbf{H}_{ii}^K) $K \in L^q(\mathcal{B}_1^c)$, for some $q \in [1, +\infty]$;

(\mathbf{H}_{iii}^K) There exists $\mathfrak{p} \in [\max(q, D), +\infty]$ such that the matrix-valued kernel ∇K defines a convolution operator which is bounded component-wise from $L^1 \cap W^{1,D}(\mathbb{R}^d)$ to $L^{\mathfrak{p}}(\mathbb{R}^d)$.

The Assumption (\mathbf{H}^K) covers many interesting cases that are fundamental examples in physics and biology, some of which can be very singular. We will detail several examples in Section 7, but as mentioned in the introduction, we recall that the kernels of the 2d Navier-Stokes equation, of the parabolic-elliptic Keller-Segel PDE in any dimension, and the Riesz potentials (see (1.2)) up to $s < d - 1$ satisfy (\mathbf{H}^K) .

Note that Assumption (\mathbf{H}_{iii}^K) is rather mild, as all the examples we have in mind have stronger properties, as will be detailed in Section 7.

2.2 Convergence of the particle system

The first result that we present establishes the (local) well-posedness of the PDE (1.1). This result is proven in Section 3.2.

Proposition 2.4. *Assume that the kernel $K : \mathbb{R}^d \rightarrow \mathbb{R}^d$ satisfies (\mathbf{H}^K) and that the initial condition satisfies $u_0 \in L^1 \cap W^{1,D}(\mathbb{R}^d)$. Then there exists $T_0 > 0$ such that the PDE (1.1) has a unique local mild solution u on $[0, T_0)$.*

For a local mild solution u on $[0, T_0)$, we will use the cut-off A_T defined for any $T \in [0, T_0)$ by

$$A_T := C_{K,d}(1 + \|u\|_{T, L^1 \cap W^{1,D}(\mathbb{R}^d)}), \quad (2.8)$$

where $C_{K,d}$ depends only on K and d and is given in Lemma 3.1.

Remark 2.5. *In Corollary 3.3, we observe that if T is small enough, then $\|u\|_{T, L^1 \cap W^{1,D}(\mathbb{R}^d)}$ is controlled by $\|u_0\|_{L^1 \cap W^{1,D}(\mathbb{R}^d)}$, so that the cut-off depends explicitly on the initial condition. If there exists a global solution, then in some cases (e.g. the parabolic-elliptic Keller-Segel PDE, see [35]), it is possible to define a global cut-off that depends explicitly on the initial condition.*

In the absence of information regarding the existence of a global solution, T will always denote a time smaller than the local existence time T_0 provided by Proposition 2.4.

The first main result of this paper is the following theorem.

Theorem 2.6. *Assume that the initial conditions $\{u_0^N\}_{N \in \mathbb{N}}$ satisfy (\mathbf{H}) and that the kernel K satisfies (\mathbf{H}^K) .*

Let T_0 be as in Proposition 2.4 and fix $T \in (0, T_0)$. Then let the dynamics of the particle system be given by (2.6) with A greater than A_T .

Then the sequence of mollified empirical measures $\{u_t^N, t \in [0, T]\}_{N \in \mathbb{N}}$ converges in probability, as $N \rightarrow \infty$, towards the unique mild solution u on $[0, T]$ of the PDE (1.1), in the following sense:

- $\forall \varphi \in L^2([0, T]; H^\beta(\mathbb{R}^d)), \int_0^T \langle u_t^N, \varphi_t \rangle_\beta dt \xrightarrow{\mathbb{P}} \int_0^T \langle u_t, \varphi_t \rangle_\beta dt;$
- *in the strong topology of $\mathcal{C}([0, T]; H_{loc}^\gamma(\mathbb{R}^d))$, for $\gamma < \beta$.*

Sections 4.1 and 4.2 are devoted to the proof of Theorem 2.6, following the approach presented in [18] using semi-group theory. In Section 4.3, we also obtain the following improvement with respect to the strong topologies considered:

Corollary 2.7. *Assume the hypothesis of Theorem 2.6 and assume further that the family of random variables $\{X_0^i, i \in \mathbb{N}\}$ is identically distributed. Then, the sequence of mollified empirical measures also converges in probability, as $N \rightarrow \infty$, towards the unique local mild solution u of the PDE (1.1) in the following strong topologies: for all $T \in (0, T_0)$,*

- $L^\infty([0, T]; L^1(\mathbb{R}^d))$,
- $L^2([0, T]; H^\gamma(\mathbb{R}^d))$, for $\gamma < \beta$.

Remark 2.8. *If there exists a global solution, then any $T_0 > 0$ works in the previous theorem and corollary, and T can be chosen equal to T_0 .*

As a consequence of the previous theorem, we obtain the propagation of chaos on the level of the one-dimensional time marginals of the empirical measure:

Corollary 2.9. *Assume the hypotheses of Theorem 2.6. Then for any $t \in (0, T_0)$, μ_t^N converges in probability as $N \rightarrow \infty$ to the constant random variable u_t .*

In addition, there exists $C > 0$ such that for any Lipschitz continuous and bounded function ϕ on \mathbb{R}^d , one has

$$|\langle u_t^N, \phi \rangle - \langle \mu_t^N, \phi \rangle| \leq \frac{C \|\phi\|_{Lip}}{N^\alpha} \quad a.s. \quad (2.9)$$

The proof of this result is written in Section 4.4.

Remark 2.10. *Observe that the previous result implies the convergence in law of μ_t^N to u_t , which is equivalent to saying that the law of (X_t^1, \dots, X_t^N) is u_t -chaotic (in the sense of [41, Def. 2.1]).*

2.3 Rate of convergence to the PDE solution

Now, we are interested in completing the above result by quantifying the obtained convergence.

Let us define, for $x \in \mathbb{R}^d$,

$$\psi_t^N(x) := (K * u_t^N)(x) \quad \text{and} \quad \psi_t(x) := (K * u_t)(x).$$

Our second main result is the following claim, whose proof is detailed in Section 5:

Theorem 2.11. *Let $\kappa = \beta_0 \alpha$ with $\beta_0 = (1 - \frac{d}{p}) \wedge (\beta - 1 - \frac{d}{2})$. Assume that the initial conditions $\{u_0^N\}_{N \in \mathbb{N}}$ satisfy **(H)** with $\beta > \frac{D}{2} + 1$ and that the kernel K satisfies **(H^K)**, with boundedness of ∇K from $W^{1,D}(\mathbb{R}^d)$ to $L^p(\mathbb{R}^d)$, with $p \geq D$ (instead of **(H_{iii}^K)**¹). Assume further that ∇K defines a convolution operator that is bounded on $L^2(\mathbb{R}^d)$, and that*

$$\forall N \in \mathbb{N}^*, \quad \mathbb{E} \|\psi_0 - \psi_0^N\|_{L^2(\mathbb{R}^d)}^2 + \mathbb{E} \|u_0 - u_0^N\|_{L^2(\mathbb{R}^d)}^2 \leq \frac{C_T}{N^\kappa} \quad (2.10)$$

and

$$\sup_{[0, T]} \|e^{t\Delta} \psi_0\|_{L^2(\mathbb{R}^d)} < \infty. \quad (2.11)$$

Then, there exists a constant $C_T > 0$ such that for all $N \in \mathbb{N}^*$,

$$(i) \quad \mathbb{E} \left(\sup_{[0, T]} \|u_t^N - u_t\|_{L^2(\mathbb{R}^d)}^2 + \int_0^T \|\nabla u_t^N - \nabla u_t\|_{L^2(\mathbb{R}^d)}^2 dt \right) \leq \frac{C_T}{N^\kappa};$$

$$(ii) \quad \mathbb{E} \left(\sup_{[0, T]} \|\psi_t^N - \psi_t\|_{L^2(\mathbb{R}^d)}^2 + \int_0^T \|\nabla \psi_t^N - \nabla \psi_t\|_{L^2(\mathbb{R}^d)}^2 dt \right) \leq \frac{C_T}{N^\kappa}.$$

¹this assumption implies obviously **(H_{iii}^K)**.

The approach used to prove Theorem 2.11 in Section 5 relies on energy estimates, which differs from the semigroup approach used in the proof of Theorem 2.6. However, note that Theorem 2.11 cannot stand alone without Theorem 2.6. Namely, in the proof of Theorem 2.6, we obtain certain properties of u^N that are useful in the proof of Theorem 2.11.

Remark 2.12. *Unlike weakly interacting particle systems ($\alpha = 0$, which corresponds to a mean field interaction), we cannot expect here a \sqrt{N} rate of convergence, which is usually the case for smoothly interacting particle systems [10], due to the short range of interaction of the particles which is of order $N^{-\alpha}$. At the macroscopic level, we also observe that the distance between a finite measure μ and its regularisation $V^N * \mu$ tested against a Lipschitz function ϕ is of order $N^{-\alpha}$ too. Hence it is reasonable to expect no better than an N^α rate of convergence. See [34] for a thorough discussion on this matter.*

In Theorem 2.11, provided that the kernel K allows to choose \mathfrak{p} as large as desired and that the initial conditions are chosen with $\beta \geq \frac{d}{2} + 2$, then the rate κ is almost α . So in view of the discussion of the previous paragraph, the rate we obtained is in some sense optimal. Note however that the condition (\mathbf{H}_{iii}) on α imposes that $\alpha < \frac{1}{d+2\beta}$, so choosing $\beta = \frac{d}{2} + 2$ yields $\kappa < \frac{1}{2(d+2)}$.

2.4 Propagation of chaos

Finally, our third objective is the propagation of chaos property of the particle system (2.6). We consider its expected limit, the following nonlinear stochastic differential equation:

$$\begin{cases} dX_t = F_A(K * u_t(X_t)) dt + \sqrt{2}dW_t, & t \leq T, \\ \mathcal{L}(X_t) = u_t, \quad \mathcal{L}(X_0) = u_0. \end{cases} \quad (2.12)$$

We will prove the convergence of the empirical measure μ^N towards the law of the weak solution to (2.12). First, we need to ensure (2.12) admits a unique weak solution. For that purpose, we will solve the associated nonlinear martingale problem. By classical arguments, one can then pass from a solution to this martingale problem to the existence of a weak solution to (2.12)

We consider the following nonlinear martingale problem related to (2.12):

Definition 2.13. *Consider the canonical space $\mathcal{C}([0, T]; \mathbb{R}^d)$ equipped with its canonical filtration. Let \mathbb{Q} be a probability measure on the canonical space and denote by \mathbb{Q}_t its one-dimensional time marginals. We say that \mathbb{Q} solves the nonlinear martingale problem (\mathcal{MP}) if:*

- (i) $\mathbb{Q}_0 = u_0$;
- (ii) For any $t \in (0, T]$, \mathbb{Q}_t has a density q_t w.r.t. Lebesgue measure on \mathbb{R}^d . In addition, it satisfies $q \in \mathcal{C}([0, T]; L^1 \cap W^{1,D}(\mathbb{R}^d))$;
- (iii) For any $f \in \mathcal{C}_c^2(\mathbb{R}^d)$, the process $(M_t)_{t \in [0, T]}$ defined as

$$M_t := f(w_t) - f(w_0) - \int_0^t \left[\Delta f(w_r) + \nabla f(w_r) \cdot F_A(K * q_r(w_r)) \right] dr$$

is a \mathbb{Q} -martingale, where $(w_t)_{t \in [0, T]}$ denotes the canonical process.

As the drift is bounded, $(M_t)_{t \geq 0}$ is well-defined. The following claim establishes the well-posedness of the martingale problem:

Proposition 2.14. *Assume that u_0 is a probability density function belonging to $W^{1,D}(\mathbb{R}^d)$ and that the kernel K satisfies (\mathbf{H}^K) . Then, for any $A > 0$ and F_A defined in (2.1), the martingale problem (\mathcal{MP}) admits a unique solution in the sense of Definition 2.13.*

This result, proven in Section 6.1 has important consequences. On an interval $[0, T]$ where a local solution to the PDE (1.1) exists, choosing $A > A_T$ as in (2.8), one has that the solution of (1.1) and the solution of (2.2) coincide, and $F_A(K * q_t) = K * q_t$. Thus, we can get rid of F_A in (2.12) and deduce the weak well-posedness to the McKean-Vlasov SDE (1.5) (see further comments on particular examples in Section 7).

Our third main result is the propagation of chaos of our particle system, proven in Section 6.2.

Theorem 2.15. *Let hypothesis of Theorem 2.6 hold and assume further that the family of random variables $\{X_0^i, i \in \mathbb{N}\}$ is identically distributed. Then, the empirical measure μ_t^N (defined in (2.7) on $\mathcal{C}([0, T]; \mathbb{R}^d)$) converges in law towards the unique weak solution of (2.12).*

2.5 Comparison with previous results

The approximation of nonlinear Fokker-Planck equations of the type (1.1) has been widely covered in the literature, but such convergence of a mollified empirical measure in strong topologies, for singular kernels, are quite new. In this sense, we extend previous results of [18]. The extension to several singular models, in particular to singular Riesz kernels and to attractive kernels (e.g. Keller-Segel) is completely new.

Concerning the rate, we consider the same class of PDE (1.1) as Jabin and Wang [24] and Bresch et al. [6]. In [6, 24], the authors were interested in the convergence, when $N \rightarrow \infty$, of the joint law of k fixed particles at a time t towards $u_t^{\otimes k}$, where u solves (1.1). They worked on a periodic domain $\Pi \subset \mathbb{R}^d$. Under the assumptions that the particles are well-defined and that $u \in L^\infty((0, T); W^{2, \infty}(\Pi))$, they proved using new techniques of relative entropy the above convergence in $L^\infty((0, T); L^1(\Pi^k))$.

Let us mention several improvements compared to their papers. More general kernel (e.g. Coulomb kernels, and even more singular Riesz kernels) can be dealt with here, a convergence rate in strong topology is obtained, milder assumptions on the PDE solution are imposed, and in particular it is not required to work on a bounded domain (although we work in \mathbb{R}^d , all the results can be easily adapted for the periodical domain case). More specifically for the Keller-Segel model, the rate obtained in [6] depends on $4\pi - \chi$ provided that the latter is positive, while we get a rate for any value of χ , even if $\chi \geq 8\pi$ and this rate is independent of χ . Moreover, we obtain the convergence rate without assuming that particles are initially independent with the same distribution, which is the usual assumption in the literature.

Nevertheless, the rate in [24] is $N^{-1/2}$ which is considered as optimal (recall from Remark 2.12 that we cannot get a better rate than $N^{-\alpha}$, and the condition (\mathbf{H}_{iii}) that we impose to get convergence leads to $\alpha < \frac{1}{2}$).

Recently, but for a slightly different class of PDEs, some quantitative results have also been obtained for the Landau equation by Fournier and Hauray [19]. Although this work has some specific difficulties that we do not encounter here (mostly related to non-constant, possibly vanishing, diffusion term), it is interesting to notice that the kernel used in the drift is of order $|\nabla b(v)| \approx |v|^\gamma$, with $\gamma \in (-2, 0)$ and $d = 3$. The authors obtain propagation of chaos for a mollified system when $\gamma \in (-2, 0)$ and a rate for $\gamma \in (-1, 0)$, which is the same range of parameter that we can get in Theorem 2.15 for the propagation of chaos ($s \in (-2, 0)$ if s is the parameter of a Riesz kernel, see Section 7.2) and for Theorem 2.11 ($s \in (-1, 0)$).

Finally, we obtained new well-posedness results for a large class of McKean-Vlasov SDEs given by (1.5). Let us emphasize the Keller-Segel model for which the kernel is given in Section 7.4. This model has attracted a lot of attention lately, for instance Fournier and Jourdain [20] proved the well-posedness of the McKean-Vlasov SDE for a value of the sensitivity parameter $\chi < 2\pi$ in dimension 2. Our result provides, assuming more regularity on the initial condition, global (in time) well-posedness whenever the PDE has a global solution, and local well-posedness whenever the solution of the PDE exhibits a blow-up. This behaviour depends on χ and the dimension. In particular in dimension 2, we get the global well-posedness of (1.5) whenever $\chi < 8\pi$, and local well-posedness of (1.5) when $\chi \geq 8\pi$.

3 Properties of the PDE and of the PDE with cut-off

We start this section with some classical embeddings that will be used throughout the article. Then in Subsection 3.2, we derive some general inequalities and prove Proposition 2.4. In Subsection 3.3, we prove that the PDE with cut-off (2.2) can have at most one mild solution.

3.1 Some classical embeddings

First, let us recall that $W^{1,d}(\mathbb{R}^d)$ is continuously embedded into $L^q(\mathbb{R}^d)$, for any $q \in [d, +\infty)$ (see e.g. [7, Corollary 9.11]), and that for $D > d$, $W^{1,D}(\mathbb{R}^d)$ is continuously embedded into $L^q(\mathbb{R}^d)$ for any $q \in [d, +\infty)$ (Morrey's Theorem, [7, Theorem 9.12]). Hence we have the following important consequence in our situation: by interpolation of Lebesgue spaces, $L^1 \cap W^{1,D}(\mathbb{R}^d)$ is continuously embedded into $L^q(\mathbb{R}^d)$, for any $q \in [1, +\infty)$ and also in $L^\infty(\mathbb{R}^d)$ if $D > d$: that is, for any $q \in [1, +\infty]$, there exists $C_q > 0$ such that

$$\begin{aligned} \text{for } q < \infty, \quad & \|f\|_{L^q(\mathbb{R}^d)} \leq C_q \|f\|_{L^1 \cap W^{1,D}(\mathbb{R}^d)}, \quad \forall f \in L^1 \cap W^{1,D}(\mathbb{R}^d) \\ \text{if } D > d, \quad & \|f\|_{L^q(\mathbb{R}^d)} \leq C_q \|f\|_{L^1 \cap W^{1,D}(\mathbb{R}^d)}, \quad \forall f \in L^1 \cap W^{1,D}(\mathbb{R}^d). \end{aligned} \quad (3.1)$$

Let us now have a look at the embedding for the Bessel potential spaces. If $\beta > \frac{D}{2}$, then we will prove that there exists $C_{d,\beta} > 0$ such that

$$\|f\|_{W^{1,D}(\mathbb{R}^d)} \leq C_{d,\beta} \|f\|_\beta, \quad \forall f \in W^{1,D}(\mathbb{R}^d). \quad (3.2)$$

Indeed, it is clear from the above that $\|f\|_{L^D(\mathbb{R}^d)} \lesssim \|f\|_\beta$. Now we distinguish two cases:

- if $\beta \in (\frac{D}{2}, \frac{D}{2} + 1)$, then the first assertion of Theorem 1.66 of [1] (a Sobolev embedding theorem for inhomogeneous Bessel spaces) yields that $\|\nabla f\|_{L^p(\mathbb{R}^d)} \lesssim \|\nabla f\|_{\beta-1} \lesssim \|f\|_\beta$, for $p \in [2, \frac{2d}{d-2(\beta-1)})$. The definition of D and β implies that $D \in [2, \frac{2d}{d-2(\beta-1)})$. The case $\beta = \frac{D}{2} + 1$ can be treated by considering $\beta' < \beta$.
- if $\beta > \frac{D}{2} + 1$, then by interpolation $\|\nabla f\|_{L^D(\mathbb{R}^d)} \lesssim \|\nabla f\|_{L^2 \cap L^\infty(\mathbb{R}^d)}$. The second assertion of Theorem 1.66 of [1] therefore gives $\|\nabla f\|_{L^2 \cap L^\infty(\mathbb{R}^d)} \lesssim \|\nabla f\|_{\beta-1} \lesssim \|f\|_\beta$.

Hence $\|\nabla f\|_{L^D(\mathbb{R}^d)} \lesssim \|f\|_\beta$ and we get the desired embedding.

As a consequence, $L^1 \cap H^\beta(\mathbb{R}^d)$ is continuously embedded into $L^1 \cap W^{1,D}(\mathbb{R}^d)$. That is, there exists $C_{d,\beta} > 0$ such that

$$\|f\|_{L^1 \cap W^{1,D}(\mathbb{R}^d)} \leq C_{d,\beta} \|f\|_{L^1 \cap H^\beta(\mathbb{R}^d)}, \quad \forall f \in L^1 \cap H^\beta(\mathbb{R}^d), \quad (3.3)$$

and $L^q(\mathbb{R}^d)$ embeds continuously into $L^1 \cap H^\beta(\mathbb{R}^d)$ for any $q \in [1, \infty]$.

3.2 Properties of mild solutions of the PDE

The following simple lemma is essential in the method presented here, as it will allow to control the reaction term, and connect the two PDEs (with and without cut-off).

Lemma 3.1. *Let K be satisfying (\mathbf{H}^K) . There exists $C_{K,d} > 0$ (which depends on K and d only) such that for any $f \in L^1 \cap W^{1,D}(\mathbb{R}^d)$*

$$\|K * f\|_{L^\infty(\mathbb{R}^d)} \leq C_{K,d} \|f\|_{L^1 \cap W^{1,D}(\mathbb{R}^d)}.$$

Proof. Let $\kappa \in (1, \kappa_M)$ and $\kappa' = \frac{\kappa}{\kappa-1}$ its conjugate exponent. Denote by \mathfrak{q}' the conjugate exponent of \mathfrak{q} . Then by Hölder's inequality, one gets

$$\begin{aligned} |K * f(x)| &\leq \int_{\mathcal{B}_1} |K(y)| |f(x-y)| dy + \int_{\mathcal{B}_1^c} |K(y)| |f(x-y)| dy \\ &\leq \|K\|_{L^\kappa(\mathcal{B}_1)} \|f\|_{L^{\kappa'}(\mathbb{R}^d)} + \|K\|_{L^{\mathfrak{q}}(\mathcal{B}_1^c)} \|f\|_{L^{\mathfrak{q}'}(\mathbb{R}^d)} \end{aligned}$$

Thus by the embedding (3.1), the desired result follows. \square

For each $T > 0$, let us now consider the space

$$\mathcal{X} := \mathcal{C}([0, T]; L^1 \cap W^{1,D}(\mathbb{R}^d)),$$

with the associated norm $\|\cdot\|_{T, L^1 \cap W^{1,D}(\mathbb{R}^d)}$, hereafter simply denoted by $\|\cdot\|_{\mathcal{X}}$. The proof of existence of local solutions relies on the continuity of the following bilinear mapping, defined on $\mathcal{X} \times \mathcal{X}$ as

$$B : (u, v) \mapsto \left(\int_0^t \nabla \cdot e^{(t-s)\Delta} (u_s K * v_s) ds \right)_{t \in [0, T]}.$$

Lemma 3.2. *The bilinear mapping B is continuous from $\mathcal{X} \times \mathcal{X}$ to \mathcal{X} .*

Proof. We shall prove that there exists $C > 0$ (independent of T) such that for any $u, v \in \mathcal{X}$ and any $t \in [0, T]$,

$$\|B(u, v)(t)\|_{L^1 \cap W^{1,D}(\mathbb{R}^d)} \leq C\sqrt{t} \|u\|_{\mathcal{X}} \|v\|_{\mathcal{X}}, \quad (3.4)$$

which suffices to prove the continuity of B .

First, using the property (1.7) of the Gaussian kernel with $p = 1$, observe that for any $t \in [0, T]$,

$$\|B(u, v)(t)\|_{L^1(\mathbb{R}^d)} \leq \int_0^t \frac{C}{\sqrt{t-s}} \|u_s K * v_s\|_{L^1(\mathbb{R}^d)} ds, \quad (3.5)$$

and since $v_s \in L^1 \cap W^{1,D}(\mathbb{R}^d)$, Lemma 3.1 yields

$$\|B(u, v)(t)\|_{L^1(\mathbb{R}^d)} \leq \int_0^t \frac{C}{\sqrt{t-s}} \|K * v_s\|_{L^\infty(\mathbb{R}^d)} \|u_s\|_{L^1(\mathbb{R}^d)} ds \quad (3.6)$$

$$\leq C\sqrt{t} \|u\|_{\mathcal{X}} \|v\|_{\mathcal{X}}. \quad (3.7)$$

As in (3.5)-(3.6), there is

$$\begin{aligned} \|B(u, v)(t)\|_{L^D(\mathbb{R}^d)} &\leq \int_0^t \frac{C}{\sqrt{t-s}} \|u_s K * v_s\|_{L^D(\mathbb{R}^d)} ds \\ &\leq C\sqrt{t} \|u\|_{\mathcal{X}} \|v\|_{\mathcal{X}}. \end{aligned} \quad (3.8)$$

The last term to consider is

$$\begin{aligned} \|\nabla B(u, v)(t)\|_{L^D(\mathbb{R}^d)} &\leq \int_0^t \frac{C}{\sqrt{t-s}} \|\nabla \cdot (u_s K * v_s)\|_{L^D(\mathbb{R}^d)} ds \\ &\leq \int_0^t \frac{C}{\sqrt{t-s}} (\|\nabla u_s \cdot K * v_s\|_{L^D(\mathbb{R}^d)} + \|u_s (\nabla \cdot K) * v_s\|_{L^D(\mathbb{R}^d)}) ds \\ &\leq \int_0^t \frac{C}{\sqrt{t-s}} \left(\|\nabla u_s\|_{L^D(\mathbb{R}^d)} \|K * v_s\|_{L^\infty(\mathbb{R}^d)} + \|u_s\|_{L^{\frac{pD}{p-D}}(\mathbb{R}^d)} \|(\nabla \cdot K) * v_s\|_{L^p(\mathbb{R}^d)} \right) ds, \end{aligned} \quad (3.9)$$

applying Hölder's inequality in the last line ($p \geq D$ from (\mathbf{H}_{iii}^K)), with the convention that $\frac{pD}{p-D} = D$ if $p = \infty$ and $\frac{pD}{p-D} = +\infty$ if $p = D$. Hence in view of Lemma 3.1 and Assumption (\mathbf{H}_{iii}^K) ,

$$\|\nabla B(u, v)(t)\|_{L^D(\mathbb{R}^d)} \leq C \left(\sqrt{t} \|u\|_{\mathcal{X}} \|v\|_{\mathcal{X}} + \int_0^t \frac{C}{\sqrt{t-s}} \|u_s\|_{L^{\frac{pD}{p-D}}(\mathbb{R}^d)} \|v_s\|_{L^1 \cap W^{1,D}(\mathbb{R}^d)} ds \right).$$

In view of the embedding (3.1) (note that if $p = D$, then this embedding still applies to $\|u_s\|_{L^\infty(\mathbb{R}^d)}$ since in that case $D > d$ because of $p \geq q > d$ in (\mathbf{H}_{ii}^K)), one gets

$$\|\nabla B(u, v)(t)\|_{L^D(\mathbb{R}^d)} \leq C\sqrt{t} \|u\|_{\mathcal{X}} \|v\|_{\mathcal{X}}. \quad (3.10)$$

Therefore, combining Equations (3.6), (3.8) and (3.10), one obtains (3.4). \square

The previous property of continuity of B now provides the existence of a local mild solution by a classical argument.

Proof of Proposition 2.4. For the existence part, note that for any $T > 0$, \mathcal{X} is a Banach space and that our aim is to find $T_0 > 0$ and $u \in \mathcal{X}$ such that

$$u_t = e^{t\Delta}u_0 - B(u, u)(t), \quad \forall t \in [0, T_0].$$

In view of Lemma 3.2, a standard contraction argument involving Banach fixed-point Theorem gives a local mild solution.

We will prove the uniqueness in the (slightly more complicated) case of the cut-off PDE (2.2), for any value of the cut-off A (see Proposition 3.6). Admitting this result for now, let us observe that it implies uniqueness for the PDE without cut-off. Indeed, if u^1 and u^2 are mild solutions to (1.1) on some interval $[0, T_0)$, then Lemma 3.1 implies that for $i = 1, 2$, for any $T \in (0, T_0)$,

$$\|K * u^i\|_{T, L^\infty(\mathbb{R}^d)} \leq C_{K,d} \|u^i\|_{T, L^1 \cap W^{1,D}(\mathbb{R}^d)} < \infty.$$

Thus u^1 and u^2 are also mild solutions to the cut-off PDE with A larger than the maximum between $C_{K,d} \|u^1\|_{T, L^1 \cap W^{1,D}(\mathbb{R}^d)}$ and $C_{K,d} \|u^2\|_{T, L^1 \cap W^{1,D}(\mathbb{R}^d)}$. Hence the uniqueness result for the PDE with cut-off implies that u^1 and u^2 coincide. \square

Corollary 3.3. *Assume that (\mathbf{H}^K) holds. Let u be a local mild solution to (1.1) on some interval $[0, T)$. Let C be the constant that appears in (3.4). Then for $T > 0$ such that*

$$4C\sqrt{T} \|u_0\|_{L^1 \cap W^{1,D}(\mathbb{R}^d)} < 1, \quad (3.11)$$

one can define the solution for $T_0 > T$ and

$$\|u\|_{T, L^1 \cap W^{1,D}(\mathbb{R}^d)} \leq \frac{1 - \sqrt{1 - 4C\sqrt{T} \|u_0\|_{L^1 \cap W^{1,D}(\mathbb{R}^d)}}}{2C\sqrt{T}}. \quad (3.12)$$

For instance, assuming that $\|u_0\|_{L^1 \cap W^{1,D}(\mathbb{R}^d)} \neq 0$ and choosing T such that $4C\sqrt{T} \|u_0\|_{L^1 \cap W^{1,D}(\mathbb{R}^d)} = \frac{1}{2}$, one has

$$\|u\|_{T, L^1 \cap W^{1,D}(\mathbb{R}^d)} < 4 \|u_0\|_{L^1 \cap W^{1,D}(\mathbb{R}^d)}. \quad (3.13)$$

Proof. We rely on the bound (3.4), in order to get that for $t < T$,

$$\|u\|_{t, L^1 \cap W^{1,D}(\mathbb{R}^d)} \leq \|u_0\|_{L^1 \cap W^{1,D}(\mathbb{R}^d)} + C\sqrt{t} \|u\|_{t, L^1 \cap W^{1,D}(\mathbb{R}^d)}^2.$$

Then by a standard argument, choosing $T > 0$ which satisfies (3.11) ensures that (3.12) holds true. \square

Remark 3.4. *Since u^N is a probability density function, we expect that its limit, whenever it exists, stays nonnegative and has mass 1.*

Indeed, assume that u is a local mild solution on $(0, T_0)$ to (1.1). Then in view of Definition 2.1-i) and by the inequality $\|e^{(t-s)\Delta}\|_{L^p \rightarrow L^p} \leq C$,

$$\begin{aligned} \|e^{(t-s)\Delta} (u_s K * u_s)\|_{L^1(\mathbb{R}^d)} &\leq C \|u_s K * u_s\|_{L^1(\mathbb{R}^d)} \\ &\leq C \|K * u_s\|_{L^\infty(\mathbb{R}^d)} \|u_s\|_{L^1(\mathbb{R}^d)} \\ &< \infty, \end{aligned}$$

using Lemma 3.1. Similarly, for any $s \in (0, t)$,

$$\begin{aligned} \|\nabla \cdot e^{(t-s)\Delta} (u_s K * u_s)\|_{L^1(\mathbb{R}^d)} &\leq \frac{C}{\sqrt{t-s}} \|u_s K * u_s\|_{L^1(\mathbb{R}^d)} \\ &\leq \frac{C}{\sqrt{t-s}} \|K * u_s\|_{L^\infty(\mathbb{R}^d)} \|u_s\|_{L^1(\mathbb{R}^d)} \\ &< \infty. \end{aligned}$$

Hence by integration-by-parts

$$\int_{\mathbb{R}^d} \nabla \cdot e^{(t-s)\Delta} (u_s K * u_s)(x) dx = 0,$$

and it follows that

$$\int_{\mathbb{R}^d} u_t(x) dx = \int_{\mathbb{R}^d} u_0(x) dx.$$

Moreover, when the initial data is such that $u_0 \geq 0$ and $u_0 \not\equiv 0$, then by an argument similar to [32, Prop. 2.7] (although using the third property of (\mathbf{H}^K) instead of the Poisson equation), u is such that $u(t, x) > 0$, for $(t, x) \in (0, T_0) \times \mathbb{R}^d$.

Remark 3.5. We observe that it is possible to ask for less integrability in Assumption (\mathbf{H}^K) , at a cost that we explain below. The well-posedness of the PDEs (1.1) and (2.2) also holds in the space of mild solutions living in $\mathcal{C}([0, T_0], L^1 \cap W^{1, p_\circ}(\mathbb{R}^d))$, where $p_\circ = \frac{d}{2} \vee 1$, under the assumption (\mathbf{H}^K) , but assuming further that $\kappa_M > \max(\frac{d}{2}, 1)$.

Indeed, there are two ingredients in the above proofs that need to be preserved for the conclusions of Proposition 2.4 to hold. The main (and only) change comes from the embedding (3.1), which now reads: for any $q \in [1, p'_\circ]$ where p'_\circ is the conjugate exponent of p_\circ , there exists $C_q > 0$ such that

$$\|f\|_{L^q(\mathbb{R}^d)} \leq C_q \|f\|_{L^1 \cap W^{1, p_\circ}(\mathbb{R}^d)}, \quad f \in L^1 \cap W^{1, p_\circ}(\mathbb{R}^d).$$

The first ingredient is that in the proof of Lemma 3.1, we need $\|f\|_{L^{\kappa'}(\mathbb{R}^d)}$ to remain bounded, which imposes that $p_\circ < \kappa_M$. The second ingredient concerns the Cauchy-Schwarz inequality which is applied in (3.9). In this new framework, the quantities $\|u_s\|_{L^{2p_\circ}(\mathbb{R}^d)}$ and $\|v_s\|_{L^{2p_\circ}(\mathbb{R}^d)}$ appear, which means that we need now that $L^1 \cap W^{1, p_\circ}(\mathbb{R}^d)$ is embedded in $L^{2p_\circ}(\mathbb{R}^d)$. This imposes $p_\circ \geq \frac{d}{2}$, which in turns yields that κ_M has to be larger than $\frac{d}{2}$.

We do not discuss further this framework, but only observe that in dimension $d = 1, 2, 3$, this set of assumptions still permit to work with Coulomb potentials (see Section 7), while asking for less integrability of the solutions.

However, at the level of the particle systems, we will work in the H^β spaces with $\beta > \frac{D}{2}$, so with solutions that are $L^1 \cap L^\infty$ in space. In this regard, it is more convenient to work in the framework presented in Section 2 (i.e. in $L^1 \cap W^{1, D}(\mathbb{R}^d)$).

3.3 Properties of mild solutions of the PDE with cut-off

In this section, we consider the cut-off PDE (2.2) and its mild solution from Definition 2.1. Here, F_A is given in (2.1), but we denote it simply by F for the sake of readability.

Note that due to the boundedness of the reaction term in (2.2), any mild solution will always be global. This global solution will be rigorously obtained as the limit of the particle system (2.6). Thus we only consider global mild solutions when it comes to the PDE (2.2).

Proposition 3.6. Assume that K satisfies (\mathbf{H}^K) . Let $u_0 \in L^1 \cap W^{1, D}(\mathbb{R}^d)$. Then for any $A > 0$ and F defined in (2.1), there is at most one mild solution to the cut-off PDE (2.2).

Proof. Assume there are two mild solutions u^1 and u^2 to (2.2). Then,

$$\begin{aligned} u_t^1 - u_t^2 &= - \int_0^t \nabla \cdot e^{(t-s)\Delta} \{u_s^1 F(K * u_s^1) - u_s^2 F(K * u_s^2)\} ds \\ &= - \int_0^t \nabla \cdot e^{(t-s)\Delta} \{(u_s^1 - u_s^2) F(K * u_s^1) + u_s^2 (F(K * u_s^1) - F(K * u_s^2))\} ds. \end{aligned}$$

Hence there exists $C > 0$ (that depends on A) such that

$$\begin{aligned}
\|u_t^1 - u_t^2\|_{L^1(\mathbb{R}^d)} + \|u_t^1 - u_t^2\|_{L^D(\mathbb{R}^d)} &\leq C \int_0^t \frac{1}{\sqrt{t-s}} (\|u_s^1 - u_s^2\|_{L^1(\mathbb{R}^d)} + \|u_s^2 K * (u_s^1 - u_s^2)\|_{L^1(\mathbb{R}^d)}) ds \\
&\quad + C \int_0^t \frac{1}{\sqrt{t-s}} (\|u_s^1 - u_s^2\|_{L^D(\mathbb{R}^d)} + \|u_s^2 K * (u_s^1 - u_s^2)\|_{L^D(\mathbb{R}^d)}) ds \\
&\leq C\sqrt{t} \|u^1 - u^2\|_{t, L^1 \cap W^{1,D}(\mathbb{R}^d)} \\
&\quad + C \int_0^t \frac{\|u_s^2\|_{L^1(\mathbb{R}^d)} + \|u_s^2\|_{L^D(\mathbb{R}^d)}}{\sqrt{t-s}} \|K * (u_s^1 - u_s^2)\|_{L^\infty(\mathbb{R}^d)} ds.
\end{aligned} \tag{3.14}$$

In view of Lemma 3.1, one has $\|K * (u_s^1 - u_s^2)\|_{L^\infty(\mathbb{R}^d)} \leq C\|u_s^1 - u_s^2\|_{L^1 \cap W^{1,D}(\mathbb{R}^d)}$, thus plugging this upper bound in (3.14) gives

$$\|u_t^1 - u_t^2\|_{L^1(\mathbb{R}^d)} + \|u_t^1 - u_t^2\|_{L^D(\mathbb{R}^d)} \leq \|u^1 - u^2\|_{t, L^1 \cap W^{1,D}(\mathbb{R}^d)} C\sqrt{t} (1 + \|u^2\|_{t, L^1 \cap W^{1,D}(\mathbb{R}^d)}). \tag{3.15}$$

Consider now

$$\begin{aligned}
\nabla(u_t^1 - u_t^2) &= - \int_0^t \nabla e^{(t-s)\Delta} \{ \nabla \cdot \{ (u_s^1 - u_s^2) F(K * u_s^1) + u_s^2 (F(K * u_s^1) - F(K * u_s^2)) \} \} ds \\
&= - \int_0^t \nabla e^{(t-s)\Delta} (\nabla(u_s^1 - u_s^2) \cdot F(\nabla(G * u_s^1))) ds - \int_0^t \nabla e^{(t-s)\Delta} ((u_s^1 - u_s^2) \nabla \cdot F(K * u_s^1)) ds \\
&\quad - \int_0^t \nabla e^{(t-s)\Delta} (\nabla u_s^2 \cdot (F(K * u_s^1) - F(K * u_s^2))) ds \\
&\quad - \int_0^t \nabla e^{(t-s)\Delta} (u_s^2 (\nabla \cdot F(K * u_s^1) - \nabla \cdot F(K * u_s^2))) ds.
\end{aligned}$$

Then, in view of Inequality (1.7) and applying Hölder's inequality,

$$\begin{aligned}
\|\nabla(u_t^1 - u_t^2)\|_{L^D(\mathbb{R}^d)} &\leq \int_0^t \frac{C}{\sqrt{t-s}} \left(\|\nabla(u_s^1 - u_s^2)\|_{L^D(\mathbb{R}^d)} + \|(u_s^1 - u_s^2) \nabla \cdot F(K * u_s^1)\|_{L^D(\mathbb{R}^d)} \right. \\
&\quad \left. + \|\nabla u_s^2 \cdot K * (u_s^1 - u_s^2)\|_{L^D(\mathbb{R}^d)} + \|u_s^2 (\nabla \cdot F(K * u_s^1) - \nabla \cdot F(K * u_s^2))\|_{L^D(\mathbb{R}^d)} \right) ds \\
&\leq \int_0^t \frac{C}{\sqrt{t-s}} \left(\|u^1 - u^2\|_{t, L^1 \cap W^{1,D}(\mathbb{R}^d)} + \|u_s^1 - u_s^2\|_{L^{\frac{pD}{p-D}}(\mathbb{R}^d)} \|\nabla \cdot F(K * u_s^1)\|_{L^p(\mathbb{R}^d)} \right. \\
&\quad \left. + \|\nabla u_s^2\|_{L^D(\mathbb{R}^d)} \|K * (u_s^1 - u_s^2)\|_{L^\infty(\mathbb{R}^d)} \right. \\
&\quad \left. + \|u_s^2\|_{L^{\frac{pD}{p-D}}(\mathbb{R}^d)} \|\nabla \cdot F(K * u_s^1) - \nabla \cdot F(K * u_s^2)\|_{L^p(\mathbb{R}^d)} \right) ds \\
&=: C \int_0^t \frac{C}{\sqrt{t-s}} (\|u^1 - u^2\|_{t, L^1 \cap W^{1,D}(\mathbb{R}^d)} + I_1(s) + I_2(s) + I_3(s)) ds. \tag{3.16}
\end{aligned}$$

The boundedness of ∇F and the property (\mathbf{H}_{ii}^K) yield

$$I_1(s) \leq C \|u_s^1 - u_s^2\|_{L^{\frac{pD}{p-D}}(\mathbb{R}^d)} \|u_s^1\|_{L^1 \cap W^{1,D}(\mathbb{R}^d)},$$

and the embedding (3.1) now gives

$$I_1(s) \leq C \|u^1 - u^2\|_{t, L^1 \cap W^{1,D}(\mathbb{R}^d)} \|u^1\|_{t, L^1 \cap W^{1,D}(\mathbb{R}^d)}. \tag{3.17}$$

In view of Lemma 3.1, we immediately obtain

$$I_2(s) \leq C \|u^2\|_{t, L^1 \cap W^{1,D}(\mathbb{R}^d)} \|u^1 - u^2\|_{t, L^1 \cap W^{1,D}(\mathbb{R}^d)}. \tag{3.18}$$

To treat the last term we first notice that, for $i = 1, 2$,

$$\nabla \cdot F(K * u_s^i) = \sum_{j=1}^d f'_A(K_j * u_s^i) \partial_j(K_j * u_s^i).$$

Thus,

$$\begin{aligned} \|\nabla \cdot F(K * u_s^1) - \nabla \cdot F(K * u_s^2)\|_{L^p(\mathbb{R}^d)} &\leq \|f'_A\|_{L^\infty(\mathbb{R})} \sum_{j=1}^d \|\partial_j K_j * (u_s^1 - u_s^2)\|_{L^p(\mathbb{R}^d)} \\ &\quad + \sum_{j=1}^d \|\partial_j K_j * u_s^2\|_{L^p(\mathbb{R}^d)} \|f'_A(K_j * u_s^1) - f'_A(K_j * u_s^2)\|_{L^\infty(\mathbb{R}^d)}. \end{aligned}$$

Hence, using the boundedness of f'_A and property (\mathbf{H}_{ii}^K) , one gets

$$\begin{aligned} \|\nabla \cdot F(K * u_s^1) - \nabla \cdot F(K * u_s^2)\|_{L^p(\mathbb{R}^d)} &\leq C \|f'_A\|_{L^\infty(\mathbb{R})} \|u_s^1 - u_s^2\|_{L^1 \cap W^{1,D}(\mathbb{R}^d)} \\ &\quad + \|u_s^2\|_{L^1 \cap W^{1,D}(\mathbb{R}^d)} \|f''_A\|_{L^\infty(\mathbb{R}^d)} \|K * (u_s^1 - u_s^2)\|_{L^\infty(\mathbb{R}^d)}. \end{aligned}$$

It remains to use Lemma 3.1 to obtain

$$I_3(s) \leq C \|u^2\|_{t, L^1 \cap W^{1,D}(\mathbb{R}^d)} \|u^1 - u^2\|_{t, L^1 \cap W^{1,D}(\mathbb{R}^d)} (1 + \|u^2\|_{t, L^1 \cap W^{1,D}(\mathbb{R}^d)}). \quad (3.19)$$

Thus in view of (3.17), (3.18) and (3.19), we obtain

$$\begin{aligned} \|\nabla(u_t^1 - u_t^2)\|_{L^D(\mathbb{R}^d)} &\leq C\sqrt{t} \|u^1 - u^2\|_{t, L^1 \cap W^{1,D}(\mathbb{R}^d)} \\ &\quad \times \left(1 + \|u^1\|_{t, L^1 \cap W^{1,D}(\mathbb{R}^d)} + \|u^2\|_{t, L^1 \cap W^{1,D}(\mathbb{R}^d)} + \|u^2\|_{t, L^1 \cap W^{1,D}(\mathbb{R}^d)}^2\right). \end{aligned} \quad (3.20)$$

Therefore (3.15) and (3.20) yield

$$\begin{aligned} \|u^1 - u^2\|_{t, L^1 \cap W^{1,D}(\mathbb{R}^d)} &\leq C\sqrt{t} \|u^1 - u^2\|_{t, L^1 \cap W^{1,D}(\mathbb{R}^d)} \\ &\quad \times \left(1 + \|u^1\|_{t, L^1 \cap W^{1,D}(\mathbb{R}^d)} + \|u^2\|_{t, L^1 \cap W^{1,D}(\mathbb{R}^d)} + \|u^2\|_{t, L^1 \cap W^{1,D}(\mathbb{R}^d)}^2\right). \end{aligned}$$

Hence for t small enough, we deduce that $\|u^1 - u^2\|_{t, L^1 \cap W^{1,D}(\mathbb{R}^d)} = 0$. Therefore the uniqueness holds for mild solutions on $[0, t]$. Then by restarting the equation and using the same arguments as above, one gets uniqueness. \square

4 Convergence: Proof of Theorem 2.6

We now proceed to the proof of Theorem 2.6, which is split between Sections 4.1 and 4.2. In Section 4.1, we detail the steps of the proof and state the most important intermediate results that are used to prove the main theorem, and refer to Section 4.2 for the proof of these results. In Sections 4.3 and 4.4 we prove Corollaries 2.7 and 2.9, respectively.

4.1 Strategy of the proof

Step 1. First, it will be established in Subsection 4.2.1 that for any test function φ , u^N satisfies the following equation

$$\begin{aligned} \langle u_t^N, \varphi \rangle &= \langle u_0^N, \varphi \rangle + \int_0^t \langle \mu_s^N, \nabla(V^N * \varphi) \cdot F_A(K * u_s^N) \rangle ds \\ &\quad + \frac{1}{N} \sum_{i=1}^N \int_0^t \nabla(V^N * \varphi)(X_s^{i,N}) \cdot dW_s^i + \int_0^t \langle u_s^N, \Delta \varphi \rangle ds. \end{aligned} \quad (4.1)$$

In Subsection 4.2.2, we will then prove that $\{u^N\}$ is tight in the space

$$\mathcal{Y} := L_w^2([0, T]; H^\beta(\mathbb{R}^d)) \cap \mathcal{C}([0, T]; H_{\text{loc}}^\gamma(\mathbb{R}^d)), \quad \gamma < \beta, \quad (4.2)$$

where $L_w^2(\mathbb{R}^d)$ denotes the $L^2(\mathbb{R}^d)$ space endowed with the weak topology. It suffices for that to prove the boundedness of the sequence in a space that is compactly embedded in \mathcal{Y} . By Prokhorov's theorem, the tightness of $\{u^N\}$ implies that it is relatively compact in a sense that we precise now (because Prokhorov's theorem applies only in Polish spaces, and L_w^2 is not metrizable). Indeed, we will make a slight abuse of language in the following when we say that u^N converges in law (resp. in probability, or almost surely) in \mathcal{Y} : it will be understood that for any $\varphi \in L^2([0, T]; H^\beta(\mathbb{R}^d))$, $\langle u^N, \varphi \rangle$ converges in law (resp. in probability or a.s.), and of course that u^N converges in law (resp. in probability or a.s.) in $\mathcal{C}([0, T]; H_{\text{loc}}^\gamma(\mathbb{R}^d))$.

Hence there is a subsequence of u^N which converges in law in \mathcal{Y} , and we still denote this subsequence u^N by a slight abuse of notation. We deduce from the previous discussion and Skorokhod's representation theorem the following proposition.

Proposition 4.1. *There exists a probability space $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{P}})$ rich enough to support $\{u^N\}_{N \in \mathbb{N}}$ and there exists a \mathcal{Y} -valued random variable \tilde{u} defined on $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{P}})$ such that*

$$u^N \xrightarrow{\mathcal{Y}} \tilde{u} \quad \text{a.s.}$$

Remark 4.2. *For each N and $t \in [0, T]$, the definition of u_t^N yields that $u_t^N \in L^\infty(\mathbb{R}^d)$ a.s., and since u_t^N is a probability density function, it is also in $L^1(\mathbb{R}^d)$ a.s. Hence by interpolation, $u_t^N \in \bigcap_{p=1}^\infty L^p(\mathbb{R}^d)$ a.s.*

Now by Fatou's lemma, one gets that a.s., for almost all t , $\tilde{u}_t \in L^1(\mathbb{R}^d)$. Moreover, by Sobolev embedding, $\tilde{u}_t \in H^\beta(\mathbb{R}^d)$ with $\beta > d/2$ implies that $\tilde{u}_t \in L^\infty(\mathbb{R}^d)$ (see e.g. [1, Thm 1.66]). Hence by interpolation, $\tilde{u}_t \in \bigcap_{p=1}^\infty L^p(\mathbb{R}^d)$ a.s.

In Subsection 4.2.3, using (4.1) and the convergence result of Proposition 4.1, we will prove that the following equality holds for all $t > 0$ and all $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}^d)$,

$$\langle \tilde{u}_t, \varphi \rangle = \langle u_0, \varphi \rangle + \int_0^t \int_{\mathbb{R}^d} \tilde{u}_s \nabla \varphi(x) \cdot F_A(K * \tilde{u}_s)(x) \, dx \, ds + \int_0^t \langle \tilde{u}_s, \Delta \varphi \rangle \, ds. \quad (4.3)$$

Observing that $\tilde{u} \in \mathcal{Y}$, one deduces that $\int_0^t \nabla \cdot (\tilde{u}_s F_A(K * \tilde{u}_s)) \, ds \in L_{\text{loc}}^1(\mathbb{R}^d)$, hence the following mild formulation in distribution holds: for any $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}^d)$,

$$\langle \tilde{u}_t, \varphi \rangle = \langle e^{t\Delta} u_0, \varphi \rangle - \left\langle \int_0^t \nabla \cdot e^{(t-s)\Delta} (\tilde{u}_s F_A(K * \tilde{u}_s)) \, ds, \varphi \right\rangle.$$

Notice from the above equation that \tilde{u} is non-random and that, for any $t \in [0, T]$, it satisfies the following equation in \mathbb{R}^d :

$$\tilde{u}_t = e^{t\Delta} u_0 - \int_0^t \nabla \cdot e^{(t-s)\Delta} (\tilde{u}_s F_A(K * \tilde{u}_s)) \, ds.$$

The next proposition will be useful in identifying \tilde{u}_s as a mild solution to (2.2). Its proof is given in Subsection 4.2.4.

Proposition 4.3. *Let $T > 0$ chosen arbitrarily and let \tilde{u} be as in Proposition 4.1. Then, $\tilde{u} \in \mathcal{C}([0, T]; L^1 \cap W^{1,D}(\mathbb{R}^d))$. As a consequence, there exists a global mild solution to the cut-off PDE (2.2).*

It follows from this result and Proposition 3.6 that \tilde{u} is the unique global mild solution of (2.2).

Step 2. In Step 1 we have obtained that on the probability space $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{P}})$, u^N converges almost surely in \mathcal{Y} to \tilde{u} , which satisfies the mild formulation of (2.2).

Observe now that if $T < T_0$ (given in Proposition 2.4), and if the cut-off value A is chosen larger than A_T (which was defined in (2.8)), it comes that a mild solution \tilde{u} to (2.2) is also a mild solution to (1.1). Hence by the uniqueness in (1.1) (Proposition 2.4 again), \tilde{u} (hereafter denoted by u) is the unique mild solution to (1.1) on $[0, T]$, as claimed in Theorem 2.6.

Let us now come back to the original probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We have obtained that every subsequence of $\{u^N\}$ has a further subsequence that converges in law to u , the unique mild solution of (1.1), in \mathcal{Y} . Hence u^N converges in law to u , and since u is non-random, the convergence also happens in probability for the topology of \mathcal{Y} , which concludes the proof of Theorem 2.6.

4.2 Proofs of intermediate results

Recall that $V : \mathbb{R}^d \rightarrow \mathbb{R}_+$ is an even smooth probability density function and that V^N is defined by Equation (2.5), that $\{X^{i,N}\}$ is the particle system defined by (2.6) with cutoff F_A given in (2.1). In this section, we use again the notation F instead of F_A , for the sake of readability.

4.2.1 Equation satisfied by the regularised empirical measure: Proof of Equality (4.1)

Consider the mollified empirical measure

$$u_t^N := V^N * \mu_t^N : x \in \mathbb{R}^d \mapsto \int_{\mathbb{R}^d} V^N(x-y) d\mu_t^N(y) = \frac{1}{N} \sum_{k=1}^N V^N(x - X_t^{k,N}).$$

Using this definition, we rewrite the particle system in (2.6) as

$$dX_t^{i,N} = F(K * u_t^N(X_t^{i,N})) dt + \sqrt{2} dW_t^i, \quad t \in [0, T], \quad 1 \leq i \leq N. \quad (4.4)$$

Fix $x \in \mathbb{R}^d$ and $1 \leq i \leq N$. Apply Itô's formula to the function $V^N(x - \cdot)$ and the particle $X^{i,N}$. Then, sum for all $1 \leq i \leq N$ and divide by N . It comes

$$\begin{aligned} u_t^N(x) &= u_0^N(x) - \frac{1}{N} \sum_{i=1}^N \int_0^t \nabla V^N(x - X_s^{i,N}) \cdot F(K * u_s^N(X_s^{i,N})) ds \\ &\quad - \frac{1}{N} \sum_{i=1}^N \int_0^t \nabla V^N(x - X_s^{i,N}) \cdot dW_s^i + \frac{1}{N} \sum_{i=1}^N \int_0^t \Delta V^N(x - X_s^{i,N}) ds. \end{aligned} \quad (4.5)$$

Notice that

$$\frac{1}{N} \sum_{i=1}^N \int_0^t \nabla V^N(x - X_s^{i,N}) \cdot F(K * u_s^N(X_s^{i,N})) ds = \int_0^t \langle \mu_s^N, \nabla V^N(x - \cdot) \cdot F(K * u_s^N(\cdot)) \rangle ds$$

and

$$\frac{1}{N} \sum_{i=1}^N \int_0^t \Delta V^N(x - X_s^{i,N}) ds = \int_0^t \Delta u_s^N(x) ds.$$

The preceding equalities combined with (4.5) and the fact that $\nabla V^N(-x) = -\nabla V^N(x)$ (because V^N is even) lead to

$$\begin{aligned} u_t^N(x) &= u_0^N(x) + \int_0^t \langle \mu_s^N, \nabla V^N(\cdot - x) \cdot F(K * u_s^N(\cdot)) \rangle ds \\ &\quad + \frac{1}{N} \sum_{i=1}^N \int_0^t \nabla V^N(X_s^{i,N} - x) \cdot dW_s^i + \int_0^t \Delta u_s^N(x) ds. \end{aligned} \quad (4.6)$$

Now for $\varphi \in \mathcal{D}(\mathbb{R}^d)$, Equation (4.6) implies the desired Equality (4.1).

For further use in Section 4.2.2, we also get the following mild form

$$\begin{aligned} u_t^N(x) &= e^{t\Delta} u_0^N(x) + \int_0^t e^{(t-s)\Delta} \langle \mu_s^N, \nabla V^N(\cdot - x) \cdot F(K * u_s^N(\cdot)) \rangle ds \\ &\quad + \frac{1}{N} \sum_{i=1}^N \int_0^t e^{(t-s)\Delta} \nabla V^N(X_s^{i,N} - x) \cdot dW_s^i. \end{aligned} \quad (4.7)$$

Finally, developing the scalar product, one has

$$\langle \mu_s^N, \nabla V^N(\cdot - x) \cdot F(K * u_s^N(\cdot)) \rangle = \nabla_x \cdot \langle \mu_s^N, V^N(\cdot - x) F(K * u_s^N(\cdot)) \rangle.$$

Combining the latter with the fact that $e^{t\Delta} \nabla \cdot f = \nabla \cdot e^{t\Delta} f$, (4.7) reads

$$\begin{aligned} u_t^N(x) &= e^{t\Delta} u_0^N(x) + \int_0^t \nabla \cdot e^{(t-s)\Delta} \langle \mu_s^N, V^N(\cdot - x) F(K * u_s^N(\cdot)) \rangle ds \\ &\quad + \frac{1}{N} \sum_{i=1}^N \int_0^t e^{(t-s)\Delta} \nabla V^N(X_s^{i,N} - x) \cdot dW_s^i. \end{aligned} \quad (4.8)$$

4.2.2 Tightness of u^N : Proof of Proposition 4.1

First, we point to [44, p.169, p.310] for the definition of the Sobolev spaces $W^{\eta,q}$. Here, we will need more particularly the space $W^{\eta,q}([0, T]; H^{-2}(\mathbb{R}^d))$, for $\eta \in (0, 1)$ and $q > 1$, whose norm is equivalent to the following norm (see [44, p.323]):

$$\|f\|_{W^{\eta,q}([0,T];H^{-2}(\mathbb{R}^d))}^p \sim \|f\|_{L^q([0,T];H^{-2}(\mathbb{R}^d))}^q + \int_0^T \int_0^T \frac{\|f_t - f_s\|_{-2}^q}{|t-s|^{1+q\eta}} ds dt.$$

Let us now prove the tightness of $\{u^N\}$ in the space \mathcal{Y} defined in (4.2). This will be achieved by proving boundedness in the following space

$$\mathcal{Y}_0 = L^p([0, T]; H^\beta(\mathbb{R}^d)) \cap W^{\eta,q}([0, T]; H^{-2}(\mathbb{R}^d)),$$

which is compactly embedded in \mathcal{Y} , as proven in [15] (see also Section 2.3 of [18]), if $\eta\vartheta \geq \frac{1-\vartheta}{p} + \frac{\vartheta}{q}$ where $\vartheta := \frac{\beta-\lambda}{2+\beta}$.

In the next two propositions, we compute the moments of u^N in \mathcal{Y}_0 .

Proposition 4.4. *Let Assumption (H) hold. Let $p \geq 2$. Then there exists a constant $C_{\beta,T,A,p} > 0$ such that, for all $t \in (0, T]$ and $N \in \mathbb{N}$, it holds:*

$$\mathbb{E} \left[\|u_t^N\|_\beta^p \right] \leq C_{\beta,T,A,p}.$$

Proposition 4.5. *Let Assumption (H) hold. Let $\eta \in (0, \frac{1}{2})$ and $q \geq 1$. There exists a constant $C_{\beta,T,A,q} > 0$ such that, for any $N \in \mathbb{N}$, it holds:*

$$\mathbb{E} \left[\int_0^T \int_0^T \frac{\|u_t^N - u_s^N\|_{-2}^q}{|t-s|^{1+q\eta}} ds dt \right] \leq C_{\eta,T,A,q}.$$

The proofs of these two results are similar to the proofs of Propositions 6 and 7 in [18] (the kernel plays no role here), but we reproduce them Appendix A.1 for the sake of completeness, and because this is where the restriction (\mathbf{H}_{iii}) on α appears.

The Chebyshev inequality then ensures that

$$\mathbb{P}(\|u^N\|_{\mathcal{Y}_0}^2 > R) \leq \frac{\mathbb{E}[\|u^N\|_{\mathcal{Y}_0}^2]}{R}, \quad \text{for any } R > 0.$$

Thus by Proposition 4.4 and Proposition 4.5, we obtain

$$\mathbb{P}(\|u^N\|_{\mathcal{Y}_0}^2 > R) \leq \frac{C}{R}, \quad \text{for any } R > 0, N \in \mathbb{N}.$$

Let \mathbb{P}_N be the law of u^N in \mathcal{Y} . The last inequality implies that there exists a bounded set $B_\epsilon \in \mathcal{Y}_0$ such that $\mathbb{P}_N(B_\epsilon) < 1 - \epsilon$ for all N , and therefore there exists a compact set $K_\epsilon \in \mathcal{Y}$ such that $\mathbb{P}_N(K_\epsilon) < 1 - \epsilon$. That is, the sequence of random variables $\{u^N\}$ is tight in \mathcal{Y} . Therefore we deduce that Proposition 4.1 holds.

4.2.3 Weak convergence to the PDE solution: Proof of Equality (4.3)

First it comes from Assumption **(H)** on the initial condition that

$$\langle u_0^N, \varphi \rangle \rightarrow \langle u_0, \varphi \rangle \quad \text{in probability.}$$

In view of Proposition 4.1, recall that $u^N \rightarrow \tilde{u}$ a.s. in the space \mathcal{Y} which was defined in Equation (4.2). First, it is clear that this result implies that we can pass to the limit in (4.1):

$$\int_0^t \langle u_s^N, \Delta \varphi \rangle ds \rightarrow \int_0^t \langle \tilde{u}_s, \Delta \varphi \rangle ds,$$

and

$$\mathbb{E} \left[\left(\frac{1}{N} \sum_{i=1}^N \int_0^t \nabla(V^N * \varphi)(X_s^{i,N}) \cdot dW_s^i \right)^2 \right] = \frac{1}{N^2} \sum_{i=1}^N \int_0^t \mathbb{E} \left[|\nabla(V^N * \varphi)(X_s^{i,N})|^2 \right] ds \rightarrow 0.$$

To conclude that \tilde{u} satisfies Equation (4.3), it remains to prove:

Lemma 4.6. *For any $t \in [0, T]$, the following convergence happens in the almost sure sense:*

$$\lim_{N \rightarrow \infty} \int_0^t \langle \mu_s^N, \nabla(V^N * \varphi) \cdot F(K * u_s^N) \rangle ds = \int_0^t \int_{\mathbb{R}^d} \tilde{u}_s(x) \nabla \varphi(x) \cdot F(K * \tilde{u}_s)(x) dx ds. \quad (4.9)$$

Proof. First, let $\epsilon > 0$ and let $R > 0$ be large enough to ensure, thanks to Assumption **(H^K)**, that

$$\int_{\mathbb{R}^d} \mathbb{1}_{\mathcal{B}_R^c}(y) |K(y)|^q dy \leq \epsilon^q, \quad (4.10)$$

where we recall that \mathcal{B}_R denotes the centred ball of \mathbb{R}^d with radius R . In view of Proposition 4.1, one has that for all $x \in \mathbb{R}^d$, there exists a random N large enough such that

$$\sup_{t \in [0, T], y \in \mathcal{B}_R} |u_t^N(x - y) - \tilde{u}_t(x - y)| \leq \epsilon \quad a.s.$$

It follows, using the Cauchy-Schwarz inequality in the second inequality and the bound (4.10) in the third, that

$$\begin{aligned} |K * (u_s^N - \tilde{u}_s)(x)| &= \left| \int_{\mathbb{R}^d} \mathbb{1}_{\mathcal{B}_R}(y) K(y) (u_s^N - \tilde{u}_s)(x - y) dy + \int_{\mathbb{R}^d} \mathbb{1}_{\mathcal{B}_R^c}(y) K(y) (u_s^N - \tilde{u}_s)(x - y) dy \right| \\ &\leq \epsilon \int_{\mathbb{R}^d} \mathbb{1}_{\mathcal{B}_R}(y) |K(y)| dy + \left(\int_{\mathbb{R}^d} \mathbb{1}_{\mathcal{B}_R^c}(y) |K(y)|^q dy \right)^{\frac{1}{q}} \left(\int_{\mathbb{R}^d} |(u_s^N - \tilde{u}_s)(x - y)|^{q'} dy \right)^{\frac{1}{q'}} \\ &\leq \epsilon \left(\int_{\mathcal{B}_R} |K(y)| dy + \|u_s^N - \tilde{u}_s\|_{L^{q'}(\mathbb{R}^d)} \right). \end{aligned}$$

Observe that by the Sobolev embedding Theorem, $L^2([0, T], L^1 \cap H^\beta(\mathbb{R}^d))$ is continuously embedded into $L^2([0, T], L^1 \cap W^{1,D}(\mathbb{R}^d))$ (see (3.3)), which is itself continuously embedded into

$L^2([0, T], L^q(\mathbb{R}^d))$, for any $q \in [1, +\infty)$ (see (3.1)). Hence, $s \mapsto \|u_s^N - \tilde{u}_s\|_{L^{q'}(\mathbb{R}^d)}$ is a.s. bounded in L^2 (see also Remark 4.2) and it follows from the previous inequality that almost surely,

$$\lim_{N \rightarrow \infty} \int_0^t \int_{\mathbb{R}^d} \tilde{u}_s(x) \nabla \varphi(x) \cdot F(K * u_s^N)(x) dx ds = \int_0^t \int_{\mathbb{R}^d} \tilde{u}_s(x) \nabla \varphi(x) \cdot F(K * \tilde{u}_s)(x) dx ds. \quad (4.11)$$

Next, observe that

$$\begin{aligned} & \left| \left\langle \mu_s^N, \nabla(V^N * \varphi) \cdot F(K * u_s^N) \right\rangle - \left\langle u_s^N, \nabla(V^N * \varphi) \cdot F(K * u_s^N) \right\rangle \right| \\ & \leq \sup_{x \in \mathbb{R}^d} \left| \nabla(V^N * \varphi)(x) \cdot F(K * u_s^N(x)) - (\nabla(V^N * \varphi) \cdot F(K * u_s^N)) * V^N(x) \right|. \end{aligned} \quad (4.12)$$

Using the fact that $\int_{\mathbb{R}^d} V = 1$ and $V \geq 0$, one first gets that

$$\begin{aligned} & \left| \nabla(V^N * \varphi)(x) \cdot F(K * u_s^N)(x) - (\nabla(V^N * \varphi) \cdot F(K * u_s^N)) * V^N(x) \right| \\ & \leq \int_{\mathbb{R}^d} V(y) \left| \nabla(V^N * \varphi)(x) \right| \left| F(K * u_s^N)(x) - F(K * u_s^N)(x - \frac{y}{N^\alpha}) \right| dy \\ & \quad + \int_{\mathbb{R}^d} V(y) \left| \nabla(V^N * \varphi)(x) - \nabla(V^N * \varphi)(x - \frac{y}{N^\alpha}) \right| \left| F(K * u_s^N)(x - \frac{y}{N^\alpha}) \right| dy \\ & \leq C \int_{\mathbb{R}^d} V(y) \left| \nabla(V^N * \varphi)(x) \right| \left| K * u_s^N(x) - K * u_s^N(x - \frac{y}{N^\alpha}) \right| dy \\ & \quad + \frac{C}{N^\alpha} \int_{\mathbb{R}^d} V(y) |y| dy, \end{aligned}$$

where the second inequality comes using the Lipschitz continuity and boundedness of F . Now in view of (A.8), for $\eta = 1 - \frac{d}{\mathfrak{p}}$ (with $\mathfrak{p} > d$ from Assumption (\mathbf{H}^K)), one has

$$\begin{aligned} \left| K * u_s^N(x) - K * u_s^N(x - \frac{y}{N^\alpha}) \right| & \leq \left(\frac{|y|}{N^\alpha} \right)^\eta \mathcal{N}_\eta(K * u_s^N) \\ & \leq C_{\mathfrak{p}, d} \|u_s^N\|_{L^1 \cap W^{1, D}(\mathbb{R}^d)} \frac{|y|^\eta}{N^{\eta\alpha}}. \end{aligned}$$

Therefore,

$$\begin{aligned} & \left| \nabla(V^N * \varphi)(x) \cdot F(K * u_s^N)(x) - (\nabla(V^N * \varphi) \cdot F(K * u_s^N)) * V^N(x) \right| \\ & \leq C \int_{\mathbb{R}^d} V(y) \left| \nabla(V^N * \varphi)(x) \right| C_{\mathfrak{p}, d} \|u_s^N\|_{L^1 \cap W^{1, D}(\mathbb{R}^d)} \frac{|y|^\eta}{N^{\eta\alpha}} dy + \frac{C}{N^\alpha} \int_{\mathbb{R}^d} V(y) |y| dy. \end{aligned}$$

Thus we have obtained

$$\left| \nabla(V^N * \varphi)(x) \cdot F(K * u_s^N)(x) - (\nabla(V^N * \varphi) \cdot F(K * u_s^N)) * V^N(x) \right| \leq C \left(\frac{1}{N^\alpha} + \frac{\|u_s^N\|_{L^1 \cap W^{1, D}(\mathbb{R}^d)}}{N^{\eta\alpha}} \right).$$

Recall that $\{u^N\}_{N \in \mathbb{N}}$ converges almost surely in $L^2([0, T], H^\beta(\mathbb{R}^d))$ for the weak topology, hence it is bounded in this space (by the uniform boundedness principle). Thus, $\sup_N \int_0^T \|u_s^N\|_{H^\beta(\mathbb{R}^d)}^2 ds < \infty$, and since $\|u_s^N\|_{L^1(\mathbb{R}^d)} = 1$, it follows from (3.3) that $\sup_N \int_0^T \|u_s^N\|_{L^1 \cap W^{1, D}(\mathbb{R}^d)}^2 ds < \infty$. Hence $\sup_N \int_0^T \|u_s^N\|_{L^1 \cap W^{1, D}(\mathbb{R}^d)} ds < \infty$ a.s. and therefore

$$\begin{aligned} \lim_{N \rightarrow \infty} \int_0^t \left\langle \mu_s^N, \nabla(V^N * \varphi) \cdot F(K * u_s^N) \right\rangle ds & = \lim_{N \rightarrow \infty} \int_0^t \left\langle u_s^N, \nabla(V^N * \varphi) \cdot F(K * u_s^N) \right\rangle ds \\ & = \lim_{N \rightarrow \infty} \int_0^t \int_{\mathbb{R}^d} u_s^N(x) \nabla(V^N * \varphi)(x) \cdot F(K * u_s^N)(x) dx ds \\ & = \int_0^t \int_{\mathbb{R}^d} \tilde{u}_s(x) \nabla \varphi(x) \cdot F(K * \tilde{u}_s)(x) dx ds, \end{aligned}$$

where in the last equality we used:

1. $u^N \xrightarrow{a.s.} \tilde{u}$ strongly in $L^2([0, T]; \mathcal{C}(D))$ for D the compact support of φ (recall that u^N converges a.s. in $\mathcal{C}([0, T], H_{\text{loc}}^\gamma(\mathbb{R}^d))$, hence by Sobolev embedding and dominated convergence, the convergence in $L^2([0, T]; \mathcal{C}(D))$ holds);
2. the convergence established in (4.11).

□

4.2.4 Time and space regularity of \tilde{u} : Proof of Proposition 4.3

As $\tilde{u} \in \mathcal{Y}$, we know that for almost all $t \in [0, T]$, $\tilde{u}_t \in L^1 \cap W^{1,D}(\mathbb{R}^d)$. Observe that for $p = 1, D$ and any $t \in [0, T]$, we have $\mathbb{E}\|u_t^N\|_{L^p(\mathbb{R}^d)} \leq C_{T,p}$: indeed, for $p = 1$, this is because u_t^N is a probability density function; for $p = D$, recall that when $\beta > \frac{D}{2}$, $L^1 \cap H^\beta(\mathbb{R}^d)$ is continuously embedded into $L^1 \cap W^{1,D}(\mathbb{R}^d)$ (see embedding (3.3)), hence it follows from Proposition 4.4 that $\mathbb{E}\|u_t^N\|_{W^{1,D}(\mathbb{R}^d)} \leq C_{T,p}$. Then Fatou's lemma implies that

$$\sup_{t \leq T} \|\tilde{u}_t\|_{L^1 \cap W^{1,D}(\mathbb{R}^d)} \leq C_T. \quad (4.13)$$

It only remains to prove that for any $t \in [0, T]$, one has

$$\lim_{\substack{s \rightarrow t \\ s \in (0, T)}} \|\tilde{u}_t - \tilde{u}_s\|_{L^1 \cap W^{1,D}(\mathbb{R}^d)} = 0. \quad (4.14)$$

This follows from the above properties of \tilde{u} and the mild form satisfied by \tilde{u} . Namely, almost everywhere in \mathbb{R}^d , one has

$$\tilde{u}_t = e^{(t-s)\Delta} \tilde{u}_s - \int_s^t \nabla \cdot e^{(t-r)\Delta} (\tilde{u}_r F(K * \tilde{u}_r)) \, dr.$$

To check (4.14), we need to ensure that

$$\lim_{s \rightarrow t} \int_s^t \|\nabla \cdot e^{(t-r)\Delta} (\tilde{u}_r F(K * \tilde{u}_r))\|_{L^1 \cap W^{1,D}(\mathbb{R}^d)} \, dr = 0. \quad (4.15)$$

This will follow from the continuity of the integral if the integral is well-defined. For $p \in \{1, D\}$, we have that

$$\int_s^t \|\nabla \cdot e^{(t-r)\Delta} (\tilde{u}_r F(K * \tilde{u}_r))\|_{L^p(\mathbb{R}^d)} \, dr \leq \int_s^t \frac{C_A}{\sqrt{t-r}} \|\tilde{u}_r\|_{L^p(\mathbb{R}^d)} \, dr.$$

In view of (4.13), the integral is well-defined. Now we turn to the norm of the gradient. Notice that

$$\begin{aligned} \int_s^t \left\| \nabla \left(\nabla \cdot e^{(t-r)\Delta} (\tilde{u}_r F(K * \tilde{u}_r)) \right) \right\|_{L^D(\mathbb{R}^d)} \, dr &\leq \int_s^t \|\nabla e^{(t-r)\Delta} (\nabla \tilde{u}_r \cdot F(K * \tilde{u}_r))\|_{L^D(\mathbb{R}^d)} \, dr \\ &\quad + \int_s^t \|\nabla e^{(t-r)\Delta} (\tilde{u}_r f'_A(K * \tilde{u}_r) \nabla \cdot K * \tilde{u}_r)\|_{L^D(\mathbb{R}^d)} \, dr \\ &=: I + II. \end{aligned}$$

Now let us use the boundedness of F , a convolution inequality and (4.13). Then it comes

$$I \leq C_A \int_s^t \frac{1}{\sqrt{t-r}} \|\nabla \tilde{u}_r\|_{L^D(\mathbb{R}^d)} \, dr \leq C_{T,A} \sqrt{t-s}.$$

For II , we use the boundedness of f'_A , a convolution inequality and Hölder's inequality (recall that by convention, $\frac{pD}{p-D} = D$ if $p = +\infty$ and $\frac{pD}{p-D} = +\infty$ if $p = D$) to get

$$\begin{aligned} II &\leq C_A \int_s^t \frac{1}{\sqrt{t-r}} \|\tilde{u}_r\|_{L^{\frac{pD}{p-D}}(\mathbb{R}^d)} \|\nabla K * \tilde{u}_r\|_{L^p(\mathbb{R}^d)} \, dr \\ &\leq C_A \int_s^t \frac{1}{\sqrt{t-r}} \|\tilde{u}_r\|_{L^{\frac{pD}{p-D}}(\mathbb{R}^d)} \|\tilde{u}_r\|_{L^1 \cap W^{1,D}(\mathbb{R}^d)} \, dr, \end{aligned}$$

using the boundedness of the convolution operator ∇K given in (\mathbf{H}_{iii}^K) . Thus, applying successively the embedding (3.1) and Inequality (4.13), it comes

$$II \leq C_A \int_s^t \frac{1}{\sqrt{t-r}} \|\tilde{u}_r\|_{L^1 \cap W^{1,D}(\mathbb{R}^d)}^2 dr \leq C\sqrt{t-s}.$$

Hence the proof is complete.

4.3 Proof of Corollary 2.7

For the first part of the corollary, we will prove the following stronger result:

Lemma 4.7. *Assume the same hypotheses as in Corollary 2.7. Then, one has*

$$\lim_{N \rightarrow \infty} \mathbb{E} \sup_{[0,T]} \|u_t^N - u_t\|_{L^1(\mathbb{R}^d)} = 0.$$

Proof. Let $R > 0$ to be precised afterwards. Observe that

$$\begin{aligned} \mathbb{E} \sup_{t \in [0,T]} \|u_t^N - u_t\|_{L^1(\mathbb{R}^d)} &\leq \mathbb{E} \sup_{t \in [0,T]} \|(u_t^N - u_t)\mathbb{1}_{\mathcal{B}_R}\|_{L^1(\mathbb{R}^d)} + \mathbb{E} \sup_{t \in [0,T]} \|(u_t^N - u_t)\mathbb{1}_{\mathcal{B}_R^c}\|_{L^1(\mathbb{R}^d)} \\ &=: I + II. \end{aligned} \tag{4.16}$$

Using the convergence in probability of u^N to u w.r.t. the strong topology of $\mathcal{C}([0,T]; L_{loc}^2(\mathbb{R}^d))$ (see Theorem 2.6), the uniform integrability of $\sup_{t \in [0,T]} \|(u_t^N - u_t)\mathbb{1}_{\mathcal{B}_R}\|_{L^1(\mathbb{R}^d)}$ in $L^1(\mathbb{P})$ (recall that $\|u_t^N\|_{L^1(\mathbb{R}^d)} = \|u_t\|_{L^1(\mathbb{R}^d)} = 1$), one deduces from the Vitali-Lebesgue Theorem [4, Th. 4.5.4] that I converges to zero as $N \rightarrow \infty$.

Now, we will show that, if R is large enough, II is arbitrary small. To do so, for a fixed $\varepsilon > 0$, we will bound both $\mathbb{E} \sup_{t \in [0,T]} \|(u_t^N)\mathbb{1}_{\mathcal{B}_R^c}\|_{L^1(\mathbb{R}^d)}$ and $\mathbb{E} \sup_{t \in [0,T]} \|u_t\mathbb{1}_{\mathcal{B}_R^c}\|_{L^1(\mathbb{R}^d)}$ by $\frac{\varepsilon}{2}$. As the computations are similar, we will only do them for u^N .

Consider a non-negative smooth cut-off function $0 \leq \zeta \leq 1$ supported on \mathcal{B}_1 and such that $\zeta = 1$ on $\mathcal{B}_{\frac{1}{2}}$. For any $R > 0$, introduce the rescaled function $\zeta_R(\cdot) = \zeta(\frac{\cdot}{R})$ which is supported on \mathcal{B}_R , and $\sigma_R = 1 - \zeta(\frac{\cdot}{R})$. We have that,

$$\langle u_t^N, \mathbb{1}_{\mathcal{B}_R^c} \rangle \leq \langle u_t^N, \sigma_R \rangle.$$

Since u_t^N and σ_R are nonnegative, and applying the Equality (4.7) for u_t^N , this leads us to

$$\begin{aligned} \mathbb{E} \sup_{t \in [0,T]} \|u_t^N \mathbb{1}_{\mathcal{B}_R^c}\|_{L^1(\mathbb{R}^d)} &= \mathbb{E} \sup_{t \in [0,T]} \int_{\mathbb{R}^d} u_t^N(x) \mathbb{1}_{\mathcal{B}_R^c}(x) dx \leq \mathbb{E} \sup_{t \in [0,T]} \int_{\mathbb{R}^d} u_t^N \sigma_R \\ &\leq \mathbb{E} \sup_{t \in [0,T]} \|\sigma_R e^{t\Delta} u_0^N\|_{L^1(\mathbb{R}^d)} \\ &\quad + \mathbb{E} \sup_{t \in [0,T]} \int_{\mathbb{R}^d} \sigma_R(x) \int_0^t e^{(t-s)\Delta} \langle \mu_s^N, \nabla V^N(\cdot - x) \cdot F((K * u_s^N)(\cdot)) \rangle ds dx \\ &\quad + \mathbb{E} \sup_{t \in [0,T]} \frac{1}{N} \sum_{i=1}^N \int_{\mathbb{R}^d} \sigma_R(x) \int_0^t e^{(t-s)\Delta} \nabla V^N(X_s^{i,N} - x) \cdot dW_s^i dx. \end{aligned}$$

Hence, applying the Fubini theorem on the first integral, and the stochastic Fubini Theorem from [45] on the stochastic integral, it comes, after noticing that $\int_{\mathbb{R}^d} \sigma_R(x) e^{(t-s)\Delta} \nabla V^N(\cdot - x) dx =$

$e^{(t-s)\Delta}V^N * \nabla\sigma_R(\cdot)$, that

$$\begin{aligned}
\mathbb{E} \sup_{t \in [0, T]} \|u_t^N \mathbb{1}_{\mathcal{B}_R^c}\|_{L^1(\mathbb{R}^d)} &= \mathbb{E} \sup_{t \in [0, T]} \int_{\mathbb{R}^d} u_t^N(x) \mathbb{1}_{\mathcal{B}_R^c}(x) dx \leq \mathbb{E} \sup_{t \in [0, T]} \int_{\mathbb{R}^d} u_t^N \sigma_R \\
&\leq \mathbb{E} \sup_{t \in [0, T]} \|\sigma_R e^{t\Delta} u_0^N\|_{L^1(\mathbb{R}^d)} \\
&\quad + \mathbb{E} \left(\int_0^T \left| \left\langle \mu^N, F(K * u^N(\cdot)) \cdot (e^{\Delta(t-s)}(V^N * \nabla\sigma_R)(\cdot)) \right\rangle \right| ds \right) \\
&\quad + \mathbb{E} \sup_{t \in [0, T]} \frac{1}{N} \sum_{i=1}^N \int_0^t (e^{(t-s)\Delta}(V^N * \nabla\sigma_R)(X_s^{i, N}) \cdot dW_s^i. \tag{4.17}
\end{aligned}$$

As $\nabla\sigma_R$ is bounded by $\frac{C}{R}$, we have that

$$\int_0^T \left| \left\langle \mu^N, F(K * u^N(\cdot)) \cdot e^{\Delta(t-s)}(V^N * \nabla\sigma_R)(\cdot) \right\rangle \right| ds \leq \frac{C_T}{R}. \tag{4.18}$$

Similarly, using the B urkholder-Davis-Gundy inequality, one gets

$$\mathbb{E} \sup_{[0, T]} \left| \frac{1}{N} \sum_{i=1}^N \int_0^t e^{(t-s)\Delta}(V^N * \nabla\sigma_R)(X_s^{i, N}) \cdot dW_s^i \right| \leq \frac{C_T}{R}. \tag{4.19}$$

Plugging (4.18) and (4.19) into (4.17), it comes

$$\mathbb{E} \sup_{[0, T]} \|u_t^N \mathbb{1}_{\mathcal{B}_R^c}\|_{L^1(\mathbb{R}^d)} \leq \mathbb{E} \sup_{[0, T]} \|\sigma_R e^{t\Delta} u_0^N\|_{L^1(\mathbb{R}^d)} + \frac{C_T}{R}.$$

Now we will show that $\mathbb{E} \sup_{[0, T]} \|\sigma_R e^{t\Delta} u_0^N\|_{L^1(\mathbb{R}^d)}$ is small for R large. We have

$$\begin{aligned}
\mathbb{E} \sup_{[0, T]} \|\sigma_R e^{t\Delta} u_0^N\|_{L^1(\mathbb{R}^d)} &= \mathbb{E} \sup_{[0, T]} \|u_0^N e^{t\Delta} \sigma_R\|_{L^1(\mathbb{R}^d)} \\
&= \mathbb{E} \sup_{[0, T]} \left\{ \|\mathbb{1}_{\mathcal{B}_{\frac{R}{4}}} u_0^N e^{t\Delta} \sigma_R\|_{L^1(\mathbb{R}^d)} + \mathbb{E} \sup_{[0, T]} \|\mathbb{1}_{\mathcal{B}_{\frac{R}{4}}^c} u_0^N e^{t\Delta} \sigma_R\|_{L^1(\mathbb{R}^d)} \right\}
\end{aligned}$$

Observe that if $x \in \mathcal{B}_{\frac{R}{4}}$, then

$$e^{t\Delta} \sigma_R(x) \leq \int_{\mathcal{B}_{\frac{R}{2}}^c} \frac{e^{-\frac{|x-y|^2}{4t}}}{4\pi t} dy \leq e^{-\frac{R^2}{128T}} \int_{\mathcal{B}_{\frac{R}{2}}^c} \frac{e^{-\frac{|x-y|^2}{8t}}}{4\pi t} dy \leq C e^{-\frac{R^2}{128T}},$$

and if $x \in \mathcal{B}_{\frac{R}{4}}^c$, then we might just bound $e^{t\Delta} \sigma_R(x)$ by a constant (independent of x and R). Hence

$$\mathbb{E} \sup_{[0, T]} \|\sigma_R e^{t\Delta} u_0^N\|_{L^1(\mathbb{R}^d)} \leq C \left(e^{-\frac{R^2}{128T}} + \mathbb{E} \int_{\mathcal{B}_{\frac{R}{4}}^c} u_0^N \right),$$

since $\int_{\mathbb{R}^d} u_0^N = 1$. To prove that $\mathbb{E} \int_{\mathcal{B}_{\frac{R}{4}}^c} u_0^N$ converges to 0 as $R \rightarrow +\infty$, uniformly in N , recall that X_0^1, \dots, X_0^N are identically distributed. Thus

$$\begin{aligned}
\sup_{N \in \mathbb{N}^*} \mathbb{E} \int_{\mathcal{B}_{\frac{R}{4}}^c} u_0^N &= \sup_{N \in \mathbb{N}^*} \mathbb{E} \int_{\mathcal{B}_{\frac{R}{4}}^c} V^N(X_0^1 - y) dy \\
&= \sup_{N \in \mathbb{N}^*} \mathbb{E} \int_{\mathbb{R}^d} \mathbb{1}_{\{|X_0^1 - N^{-\alpha}z| > \frac{R}{4}\}} V(z) dz.
\end{aligned}$$

Now it is clear that $\{|X_0^1 - N^{-\alpha}z| > \frac{R}{4}\} \subseteq \{|X_0^1| + |z| > \frac{R}{4}\} \subseteq \{|X_0^1| > \frac{R}{8}\} \cup \{|z| > \frac{R}{8}\}$. Therefore,

$$\begin{aligned} \sup_{N \in \mathbb{N}^*} \mathbb{E} \int_{\mathcal{B}_{\frac{R}{4}}^c} u_0^N &\leq \mathbb{P}(|X_0^1| > \frac{R}{8}) \int_{\mathbb{R}^2} V(z) dz + \int_{\mathcal{B}_{\frac{R}{8}}^c} V(z) dz \\ &\longrightarrow 0 \quad \text{as } R \rightarrow +\infty. \end{aligned}$$

It remains to choose R large enough to have

$$\mathbb{E} \sup_{t \in [0, T]} \|u_t^N \mathbb{1}_{\mathcal{B}_R^c}\|_{L^1(\mathbb{R}^d)} < \frac{\varepsilon}{2}.$$

This concludes the proof. \square

Now, to prove the second part of Corollary 2.7 we consider the space

$$\tilde{\mathcal{Y}} = \mathcal{Y} \cap L^\infty([0, T]; L^1(\mathbb{R}^d)). \quad (4.20)$$

The tightness of u^N in $\tilde{\mathcal{Y}}$ comes from Theorem 2.6 and the previous lemma. We will show that the convergence in $\tilde{\mathcal{Y}}$ implies strong convergence in $L^2([0, T], H^\gamma(\mathbb{R}^d))$ with $\gamma < \beta$. To do so, we argue as in Oelschläger [33] and Méléard and Roelly [31].

Let us fix $R > 0$ and $\varepsilon > 0$. By the definition of $H^\gamma(\mathbb{R}^d)$, one writes

$$\begin{aligned} \int_0^T \mathbb{E} \|u_s^N - u\|_\gamma^2 ds &= \int_0^T \mathbb{E} \int_{\mathbb{R}^d} (1 + |x|^2)^\gamma |\mathcal{F}u_s^N(x) - \mathcal{F}u_s(x)|^2 dx ds \\ &= \int_0^T \mathbb{E} \int_{\mathcal{B}_R} (1 + |x|^2)^\gamma |\mathcal{F}u_s^N(x) - \mathcal{F}u_s(x)|^2 dx ds \\ &\quad + \int_0^T \mathbb{E} \int_{\mathcal{B}_R^c} (1 + |x|^2)^\gamma |\mathcal{F}u_s^N(x) - \mathcal{F}u_s(x)|^2 dx ds \\ &=: I_N + II_N. \end{aligned}$$

In view of the previous lemma, $u^N \rightarrow u$ a.s. in $L^\infty([0, T]; L^1(\mathbb{R}^d))$, thus one has for N large enough,

$$I_N = \int_0^T \mathbb{E} \int_{\mathcal{B}_R} (1 + |x|^2)^\gamma |\langle u_s^N, e^{-ix \cdot} \rangle - \langle u_s, e^{-ix \cdot} \rangle|^2 dx ds \leq C \mathbb{E} \sup_{t \in [0, T]} \|u_t^N - u_t\|_{L^1(\mathbb{R}^d)} < \frac{\varepsilon}{2}.$$

Now, by the uniform boundedness principle (since u^N converges in weak topology) and choosing R large enough, one has

$$II_N \leq \frac{1}{(1 + |R|^2)^{\beta - \gamma}} \int_0^T \mathbb{E} \int_{\mathcal{B}_R^c} (1 + |x|^2)^\beta |\mathcal{F}u_s^N(x) - \mathcal{F}u_s(x)|^2 dx ds \leq \frac{C}{(1 + |R|^2)^{\beta - \gamma}} < \frac{\varepsilon}{2}.$$

Gathering the estimates on I_N and II_N , one deduces the desired result.

4.4 Proof of Corollary 2.9

We will show the following convergence in probability:

$$\langle \mu_t^N, \phi \rangle \rightarrow \langle u_t, \phi \rangle, \quad \text{as } N \rightarrow \infty, \quad (4.21)$$

for any $\phi \in \mathcal{C}_b(\mathbb{R}^d)$.

Recall first that weak convergence is generated by the complete metric

$$\|\mu - \nu\| = \sup_{\phi \in H} |\langle \mu, \phi \rangle - \langle \nu, \phi \rangle|,$$

where H is the set of all $\phi \in \mathcal{C}_b(\mathbb{R}^d)$, which are bounded by 1 and that are Lipschitz continuous with constant 1.

In view of Theorem 2.6, the following convergence holds in probability:

$$\langle u_t^N, \phi \rangle \rightarrow \langle u_t, \phi \rangle, \quad \text{as } N \rightarrow \infty, \quad (4.22)$$

for any $\phi \in \mathcal{C}_b(\mathbb{R}^d)$. Let us now control the distance between u^N and μ^N to complete the proof. There is

$$\begin{aligned} \|\mu_t^N - u_t^N\| &= \sup_{\phi \in H} |\langle \mu_t^N, \phi \rangle - \langle u_t^N, \phi \rangle| = \sup_{\phi \in H} |\langle \mu_t^N, (\phi - \phi * V^N) \rangle| \\ &\leq \sup_{\phi \in H} \left\langle \mu_t^N, \int_{\mathbb{R}^d} V(y) |\phi(\cdot) - \phi(\frac{y}{N^\alpha} - \cdot)| dy \right\rangle \leq \frac{C \|\phi\|_{\text{Lip}}}{N^\alpha}, \end{aligned}$$

which gives (2.9). Combining the previous inequality and the convergence (4.22), we deduce the convergence (4.21).

5 Rate of convergence: Proof of Theorem 2.11

We remind the reader the following notation:

$$\psi_t^N(x) = (K * u_t^N)(x) \quad \text{and} \quad \psi_t(x) = (K * u_t)(x).$$

Recall that \mathcal{N}_η denotes the Hölder seminorm of parameter $\eta \in (0, 1]$, defined at the beginning of the paper, and recall also that in this Section, it is assumed that ∇K defines a convolution operator that is bounded from $W^{1,D}(\mathbb{R}^d)$ to $L^p(\mathbb{R}^d)$. In particular, (A.8) reads here, for $\eta = 1 - \frac{d}{p}$:

$$\mathcal{N}_\eta(K * f) \leq C_{p,d,K} \|f\|_{W^{1,D}(\mathbb{R}^d)}. \quad (5.1)$$

Let us first state a technical result, which requires ∇K to define a bounded operator from $W^{1,D}(\mathbb{R}^d)$ to $L^p(\mathbb{R}^d)$ (instead of (\mathbf{H}_{iii}^K)).

Lemma 5.1. *Under the assumptions of Theorem 2.11, there exists $C > 0$ such that for any $u \in H^\beta(\mathbb{R}^d)$, one has $K * u \in W^{1,D}(\mathbb{R}^d)$, $\nabla K * u \in W^{1,D}(\mathbb{R}^d)$ and*

$$\|\nabla K * u\|_{W^{1,D}(\mathbb{R}^d)} \leq C \|u\|_\beta. \quad (5.2)$$

Proof. First, by interpolation (recall that $p \geq D \geq 2$), there exists $\theta \in [0, 1]$ such that

$$\|\nabla K * u\|_{L^D(\mathbb{R}^d)} \leq \|\nabla K * u\|_{L^p(\mathbb{R}^d)}^\theta \|\nabla K * u\|_{L^2(\mathbb{R}^d)}^{1-\theta}.$$

Thus by the boundedness of ∇K in L^2 and from $W^{1,D}(\mathbb{R}^d)$ to $L^p(\mathbb{R}^d)$, it comes

$$\|\nabla K * u\|_{L^D(\mathbb{R}^d)} \leq C \|u\|_{W^{1,D}(\mathbb{R}^d)}^\theta \|u\|_{L^2(\mathbb{R}^d)}^{1-\theta}.$$

The embeddings (3.2) and (3.1) yield

$$\|\nabla K * u\|_{L^D(\mathbb{R}^d)} \leq C \|u\|_\beta.$$

Hence $K * u \in W^{1,D}(\mathbb{R}^d)$.

Then, as before,

$$\begin{aligned} \|\nabla K * \nabla u\|_{L^D(\mathbb{R}^d)} &\leq C \|\nabla u\|_{W^{1,D}(\mathbb{R}^d)}^\theta \|\nabla u\|_{L^2(\mathbb{R}^d)}^{1-\theta} \\ &\leq C \|\nabla u\|_{\beta-1}^\theta \|\nabla u\|_{L^2(\mathbb{R}^d)}^{1-\theta}, \end{aligned}$$

since $\beta - 1 > \frac{D}{2}$, using embedding (3.2) and the fact that $\|\nabla u\|_{\beta-1} \lesssim \|u\|_\beta$. Thus $\nabla K * u \in W^{1,D}(\mathbb{R}^d)$ and the desired inequality (5.2) holds. \square

We have the following regularity properties.

Lemma 5.2. *Let $u^1, u^2 \in L^1 \cap H^\beta(\mathbb{R}^d)$ for $\beta > \frac{D}{2} + 1$. Denote $\psi^1 = K * u^1, \psi^2 = K * u^2$. Then we have the following inequalities:*

- For any $\eta \in (0, 1 - \frac{d}{p}]$, there exists $C > 0$ such that for $i = 1, 2$, one has

$$\mathcal{N}_\eta(\psi^i) \leq C \|u^i\|_\beta, \quad \mathcal{N}_\eta(\nabla \cdot K * \psi^i) \leq C \|u^i\|_\beta, \quad (5.3)$$

and

$$\mathcal{N}_\eta(F(\psi^1) \cdot (\nabla \cdot K * (\psi^2 - \psi^1) * V^N)) \leq C (\|u^2\|_\beta + \|u^1\|_\beta + \|u^2\|_\beta \|u^1\|_\beta + \|u^1\|_\beta^2);$$

- For $\beta_0 = (1 - \frac{d}{p}) \wedge (\beta - 1 - \frac{d}{2})$, there exists $C > 0$ such that for $i = 1, 2$, one has

$$\mathcal{N}_{\beta_0}(\nabla u^i) \leq C \|u^i\|_\beta, \quad (5.4)$$

and

$$\mathcal{N}_{\beta_0}(F(\psi^1) \cdot \nabla((u^2 - u^1) * V^N)) \leq C (\|u^2\|_\beta + \|u^1\|_\beta + \|u^2\|_\beta \|u^1\|_\beta + \|u^1\|_\beta^2).$$

Proof. • Recall that $p > d$ is defined in Assumption (\mathbf{H}^K) . In view of Inequality (5.1) applied with $\eta = 1 - \frac{d}{p}$ and then by the embeddings discussed in Section 3.1, one has

$$\mathcal{N}_\eta(\psi^i) \leq C_{p,d,K} \|u^i\|_{W^{1,D}(\mathbb{R}^d)} \leq C \|u^i\|_\beta.$$

Provided that ψ^i is a bounded function, the previous inequality holds for any $\eta \in (0, 1 - \frac{d}{p}]$. To see that $\psi^i \in L^\infty$, apply Lemma 3.1, so that $\|\psi^i\|_{L^\infty(\mathbb{R}^d)} \lesssim \|u^i\|_{L^1 \cap W^{1,D}(\mathbb{R}^d)} \lesssim \|u^i\|_{L^1 \cap H^\beta(\mathbb{R}^d)}$.

In order to get $\mathcal{N}_\eta(\nabla \cdot K * \psi^i)$, we will apply (5.1) to $f \equiv \nabla \cdot \psi^i$ and $\eta = 1 - \frac{d}{p}$. First we need to ensure that $\nabla \cdot K * \psi^i \in W^{1,p}(\mathbb{R}^d)$. In view of Lemma 5.1 and Equation (5.1), one has

$$\|\nabla \cdot K * \psi^i\|_{L^p(\mathbb{R}^d)} \leq C \|\psi^i\|_{W^{1,D}(\mathbb{R}^d)}$$

and

$$\|\nabla \cdot K * \nabla \psi^i\|_{L^p(\mathbb{R}^d)} \leq C \|\nabla \psi^i\|_{W^{1,D}(\mathbb{R}^d)}.$$

Now, we are ready to apply (5.1), the above calculations and the Inequality (5.2) to obtain

$$\mathcal{N}_\eta(\nabla \cdot K * \psi^i) \leq C \|\nabla \psi^i\|_{W^{1,D}(\mathbb{R}^d)} \leq C \|u^i\|_\beta.$$

We will see below that $\nabla \cdot K * \psi^i \in L^\infty(\mathbb{R}^d)$, so the above inequality is true for any $\eta \in (0, \frac{d}{p}]$.

For the next inequality, observe that

$$\begin{aligned} \mathcal{N}_\eta(F(\psi^1) \cdot (\nabla \cdot K * (\psi^2 - \psi^1) * V^N)) &\leq \|F(\psi^1)\|_{L^\infty(\mathbb{R}^d)} \mathcal{N}_\eta(\nabla \cdot K * (\psi^2 - \psi^1) * V^N) \\ &\quad + \mathcal{N}_\eta(F(\psi^1)) \|\nabla \cdot K * (\psi^2 - \psi^1) * V^N\|_{L^\infty(\mathbb{R}^d)} \\ &\leq C (\mathcal{N}_\eta(\nabla \cdot K * \psi^2) + \mathcal{N}_\eta(\nabla \cdot K * \psi^1)) + C \mathcal{N}_\eta(\psi^1) (\|\nabla \cdot K * \psi^2\|_{L^\infty(\mathbb{R}^d)} + \|\nabla \cdot K * \psi^1\|_{L^\infty(\mathbb{R}^d)}). \end{aligned}$$

Hence we deduce from (5.3) that

$$\begin{aligned} \mathcal{N}_\eta(F(\psi^1) \cdot (\nabla \cdot K * (\psi^2 - \psi^1) * V^N)) \\ \leq C_\eta (\|u^2\|_\beta + \|u^1\|_\beta + \|u^1\|_\beta (\|\nabla \cdot K * \psi^2\|_{L^\infty(\mathbb{R}^d)} + \|\nabla \cdot K * \psi^1\|_{L^\infty(\mathbb{R}^d)})). \end{aligned}$$

As $\nabla \cdot K * \psi^i \in W^{1,p}(\mathbb{R}^d)$ for $i = 1, 2$, by Morrey's inequality and the above computations, we have that

$$\|\nabla \cdot K * \psi^i\|_{L^\infty(\mathbb{R}^d)} \leq C \|\nabla \cdot K * \psi^i\|_{W^{1,p}(\mathbb{R}^d)} \leq C \|u^i\|_\beta.$$

• For the second assertion, we will apply a similar procedure as above starting from ∇u^i , $i = 1, 2$ that belong (component-wise) to $H^{\beta-1}(\mathbb{R}^d)$. If $\beta - 1 - \frac{d}{2} \in \mathbb{R} \setminus \mathbb{N}$, let $n = \lfloor \beta - 1 - \frac{d}{2} \rfloor$ be the integer part of $\beta - 1 - \frac{d}{2}$. Otherwise if $\beta - 1 - \frac{d}{2}$ is an integer, let $n = \beta - 2 - \frac{d}{2}$. Since β is assumed to be greater than $\frac{d}{2} + 1$ here, the embedding of [1, Th. 1.66] yields

$$\sup_{|\gamma| \leq n} \left(\|\partial^\gamma \nabla u^i\|_{L^\infty(\mathbb{R}^d)} + \mathcal{N}_{\beta-1-\frac{d}{2}-n}(\partial^\gamma \nabla u^i) \right) \lesssim \|\nabla u^i\|_{\beta-1}, \quad (5.5)$$

where $\gamma \in \mathbb{N}^d$ is a multi-index and $|\gamma|$ denotes the sum of its component. As $\|\nabla \cdot\|_{\beta-1} \lesssim \|\cdot\|_{\beta}$, the inequality of (5.4) follows.

Then, observe that

$$\begin{aligned} \mathcal{N}_{\beta_0}(F(\psi^1) \cdot \nabla((u^2 - u^1) * V^N)) &\leq \|F(\psi^1)\|_{L^\infty(\mathbb{R}^d)} \mathcal{N}_\eta(\nabla((u^2 - u^1) * V^N)) \\ &\quad + \mathcal{N}_\eta(F(\psi^1)) \|\nabla((u^2 - u^1) * V^N)\|_{L^\infty(\mathbb{R}^d)} \\ &\leq C(\mathcal{N}_\eta(\nabla u^2) + \mathcal{N}_\eta(\nabla u^1)) + C\mathcal{N}_\eta(\psi^1) (\|\nabla u^2\|_{L^\infty(\mathbb{R}^d)} + \|\nabla u^1\|_{L^\infty(\mathbb{R}^d)}). \end{aligned}$$

The embedding (5.5) also gives $\|\nabla u^i\|_{L^\infty(\mathbb{R}^d)} \lesssim \|u^i\|_{\beta}$. Hence, together with (5.3), one gets the last announced result. \square

We are now ready to prove Theorem 2.11.

Proof of Theorem 2.11. We split the proof into three parts. The first one is dedicated to obtain an upper bound on

$$\mathbb{E} \left(\sup_{[0, T]} \|u_t^N - u_t\|_{L^2(\mathbb{R}^d)}^2 + \int_0^T \|\nabla u_t^N - \nabla u_t\|_{L^2(\mathbb{R}^d)}^2 dt \right). \quad (5.6)$$

This bound will depend on the integral of $\mathbb{E} \left(\sup_{r \in [0, s]} \|\psi_r^N - \psi_r\|_{L^2(\mathbb{R}^d)}^2 \right)$.

Then, in the second part, we obtain an upper bound on

$$\mathbb{E} \left(\sup_{[0, T]} \|\psi_t^N - \psi_t\|_{L^2(\mathbb{R}^d)}^2 + \int_0^T \|\nabla \psi_t^N - \nabla \psi_t\|_{L^2(\mathbb{R}^d)}^2 dt \right)$$

in terms of the integrals of $\mathbb{E} \left(\sup_{r \in [0, s]} \|u_r^N - u_r\|_{L^2(\mathbb{R}^d)}^2 \right)$ and $\mathbb{E} \left(\sup_{r \in [0, s]} \|\psi_r^N - \psi_r\|_{L^2(\mathbb{R}^d)}^2 \right)$.

Finally, we conclude in the third part with Grönwall's Lemma.

1st Part. In order to obtain an upper bound on (5.6), one should apply Itô's formula to the random variable $|u_t(x) - u_t^N(x)|^2$ for a fixed $t \in [0, T]$ and $x \in \mathbb{R}^d$ and then integrate in space. The SDE satisfied by $u_t^N(x)$ is given in (4.6). However, as u_t is not a classical solution to our PDE, one needs to pass by a mollified version of it in order to perform the above operations.

Let $\phi \in C_c^\infty(\mathbb{R}^d)$ such that $\phi \geq 0$ and $\int \phi = 1$. For $\varepsilon > 0$, define the sequence of mollifiers $\phi^\varepsilon(x) := \frac{1}{\varepsilon^d} \phi(\frac{x}{\varepsilon})$. Denote $\phi_x^\varepsilon(\cdot) := \phi^\varepsilon(x - \cdot)$ and

$$u_t^\varepsilon(x) := u_t * \phi_x^\varepsilon(x) = \langle u_t, \phi_x^\varepsilon \rangle.$$

With such choice of mollifiers, one has $\lim_{\varepsilon \rightarrow 0} \|u_t - u_t^\varepsilon\|_{H^1(\mathbb{R}^d)} = 0$. In addition, a convolution inequality yields that $\|u_t^\varepsilon\|_{H^1(\mathbb{R}^d)} \leq \|u_t\|_{H^1(\mathbb{R}^d)}$. Thus, one has that u^ε converges to u in $L^2([0, T]; H^1(\mathbb{R}^d))$ as $\varepsilon \rightarrow 0$. From (4.3), one gets

$$u_t^\varepsilon(x) = u_0^\varepsilon(x) + \int_0^t \Delta u_s^\varepsilon(x) ds - \int_0^t \nabla \cdot \langle u_s F(\psi_s), \phi_x^\varepsilon \rangle ds.$$

Now, apply Itô's formula to the random variable $|u_t^\varepsilon(x) - u_t^N(x)|^2$. It comes,

$$\begin{aligned}
|u_t^\varepsilon(x) - u_t^N(x)|^2 &= |u_0^\varepsilon(x) - u_0^N(x)|^2 + 2 \int_0^t (u_s^\varepsilon(x) - u_s^N(x)) \Delta(u_s^\varepsilon(x) - u_s^N(x)) ds \\
&\quad - 2 \int_0^t (u_s^\varepsilon(x) - u_s^N(x)) \nabla \cdot \langle u_s F(\psi_s), \phi_x^\varepsilon \rangle ds - 2 \int_0^t (u_s^\varepsilon(x) - u_s^N(x)) \langle \mu_s^N, \nabla V^N(\cdot - x) \cdot F(K * u_s^N(\cdot)) \rangle ds \\
&\quad - \frac{2}{N} \sum_{i=1}^N \int_0^t (u_s^\varepsilon(x) - u_s^N(x)) \nabla V^N(X_s^{i,N} - x) \cdot dW_s^i + \frac{1}{N^2} \sum_{i=1}^N \int_0^t |\nabla V^N(X_s^{i,N} - x)|^2 ds.
\end{aligned}$$

Then integrate in space the above expression and perform the integration by parts where possible. One gets

$$\begin{aligned}
&\|u_t^\varepsilon - u_t^N\|_{L^2(\mathbb{R}^d)}^2 + 2 \int_0^t \|\nabla(u_s^\varepsilon - u_s^N)\|_{L^2(\mathbb{R}^d)}^2 ds = \|u_0^\varepsilon - u_0^N\|_{L^2(\mathbb{R}^d)}^2 \\
&\quad - 2 \int_0^t \int_{\mathbb{R}^d} \langle \mu_s^N, F(\psi_s^N) V^N(\cdot - x) \rangle \cdot \nabla(u_s^\varepsilon - u_s^N)(x) dx ds \\
&\quad + 2 \int_0^t \int_{\mathbb{R}^d} \langle u_s F(\psi_s), \phi_x^\varepsilon \rangle \cdot \nabla(u_s^\varepsilon - u_s^N)(x) dx ds \\
&\quad - \frac{2}{N} \sum_{i=1}^N \int_{\mathbb{R}^d} \int_0^t (u_s^\varepsilon - u_s^N)(x) \nabla V^N(x - X_s^{i,N}) \cdot dW_s^{i,N} dx + \frac{1}{N} \int_0^t \|\nabla V^N\|_{L^2(\mathbb{R}^d)}^2 ds. \tag{5.7}
\end{aligned}$$

For the stochastic integral, the stochastic Fubini Theorem of [45] can be applied since

$$\int_{\mathbb{R}^d} \left(\int_0^t |(u_s^\varepsilon - u_s^N)(x)|^2 |\nabla V^N(x - X_s^{i,N})|^2 ds \right)^{1/2} dx$$

is finite a.s. This comes from the fact that V is a rapidly decreasing probability density function and $u_s^\varepsilon(x)$ and $u_s^N(x)$ are uniformly bounded in time and space for fixed N and ε .

Now, apply the Burkholder-Davis-Gundy inequality after the aforementioned stochastic Fubini theorem. It comes

$$\begin{aligned}
&\mathbb{E} \sup_{t \in [0, T]} \left| \frac{2}{N} \sum_{i=1}^N \int_{\mathbb{R}^d} \int_0^t (u_s^\varepsilon - u_s^N) \nabla V^N(x - X_s^{i,N}) \cdot dW_s^{i,N} dx \right| \\
&\leq \mathbb{E} \left[\left(\frac{4}{N^2} \sum_{i=1}^N \int_0^t \left| \int_{\mathbb{R}^d} (u_s^\varepsilon - u_s^N)(x) \nabla V^N(x - X_s^{i,N}) dx \right|^2 ds \right)^{\frac{1}{2}} \right] \\
&\leq \mathbb{E} \left[\left(\frac{4}{N^2} \sum_{i=1}^N \int_0^t \left| \int_{\mathbb{R}^d} \nabla(u_s^\varepsilon - u_s^N)(x) \cdot V^N(x - X_s^{i,N}) dx \right|^2 ds \right)^{\frac{1}{2}} \right] \\
&\leq \frac{2}{\sqrt{N}} \|V^N\|_{L^2(\mathbb{R}^d)} \mathbb{E} \left[\left(\int_0^t \|\nabla(u_s^\varepsilon - u_s^N)\|_{L^2(\mathbb{R}^d)}^2 ds \right)^{\frac{1}{2}} \right] \\
&\leq \frac{C_T}{N^{\frac{1}{2}(1-d\alpha)}}, \tag{5.8}
\end{aligned}$$

where we used the boundedness of the expectation (uniformly in ε and N). Observe also that

$\frac{1}{N} \int_0^t \|\nabla V^N\|_{L^2(\mathbb{R}^d)}^2 ds \leq C_T N^{-1+\alpha(d+2)}$. Hence, plugging this inequality and (5.8) into (5.7) gives

$$\begin{aligned} \mathbb{E} \sup_{t \in [0, T]} \|u_t^\varepsilon - u_t^N\|_{L^2(\mathbb{R}^d)}^2 + 2\mathbb{E} \int_0^T \|\nabla(u_s^\varepsilon - u_s^N)\|_{L^2(\mathbb{R}^d)}^2 ds &\leq \|u_0^\varepsilon - u_0^N\|_{L^2(\mathbb{R}^d)}^2 \\ &\quad \mathbb{E} \sup_{t \in [0, T]} \left(-2 \int_0^t \int_{\mathbb{R}^d} \langle \mu_s^N, F(\psi_s^N) V^N(\cdot - x) \rangle \cdot \nabla(u_s^\varepsilon - u_s^N)(x) dx ds \right. \\ &\quad \left. + 2 \int_0^t \int_{\mathbb{R}^d} \langle u_s F(\psi_s), \phi_x^\varepsilon \rangle \cdot \nabla(u_s^\varepsilon - u_s^N)(x) dx ds \right) \\ &\quad + C_T \left(\frac{1}{N^{\frac{1}{2}(1-d\alpha)}} + \frac{1}{N^{1-\alpha(d+2)}} \right). \end{aligned}$$

Here take the limit $\varepsilon \rightarrow 0$. One obtains

$$\begin{aligned} \mathbb{E} \sup_{t \in [0, T]} \|u_t - u_t^N\|_{L^2(\mathbb{R}^d)}^2 + 2\mathbb{E} \int_0^T \|\nabla(u_s - u_s^N)\|_{L^2(\mathbb{R}^d)}^2 ds &\leq \|u_0 - u_0^N\|_{L^2(\mathbb{R}^d)}^2 \\ &\quad \mathbb{E} \sup_{t \in [0, T]} \left(-2 \int_0^t \int_{\mathbb{R}^d} \langle \mu_s^N, F(\psi_s^N) V^N(\cdot - x) \rangle \cdot \nabla(u_s - u_s^N)(x) dx ds \right. \\ &\quad \left. + 2 \int_0^t \int_{\mathbb{R}^d} u_s(x) F(\psi_s(x)) \cdot \nabla(u_s - u_s^N)(x) dx ds \right) \\ &\quad + C_T \left(\frac{1}{N^{\frac{1}{2}(1-d\alpha)}} + \frac{1}{N^{1-\alpha(d+2)}} \right). \end{aligned}$$

Then observe that

$$\int_0^t \int_{\mathbb{R}^d} \langle \mu_s^N, F(\psi_s^N) V^N(\cdot - x) \rangle \cdot \nabla(u_s - u_s^N)(x) dx ds = \int_0^t \langle \mu_s^N, F(\psi_s^N) \cdot \nabla((u_s - u_s^N) * V^N) \rangle ds,$$

and that

$$\int_0^t \langle \mu_s^N, V^N * (F(\psi_s^N) \cdot \nabla((u_s - u_s^N) * V^N)) \rangle ds = \int_0^t \int_{\mathbb{R}^d} u_s^N F(\psi_s^N) * V^N(x) \cdot \nabla(u_s - u_s^N)(x) dx ds.$$

Hence using the two previous equalities, we have

$$\begin{aligned} \mathbb{E} \sup_{t \in [0, T]} \|u_t - u_t^N\|_{L^2(\mathbb{R}^d)}^2 + 2\mathbb{E} \int_0^T \|\nabla(u_s - u_s^N)\|_{L^2(\mathbb{R}^d)}^2 ds &\leq \|u_0 - u_0^N\|_{L^2(\mathbb{R}^d)}^2 \\ &\quad + \mathbb{E} \sup_{t \in [0, T]} \left(2 \int_0^t \int_{\mathbb{R}^d} \{u_s F(\psi_s) - u_s^N F(\psi_s^N)\} * V^N(x) \cdot \nabla(u_s - u_s^N)(x) dx ds \right. \\ &\quad + 2 \int_0^t \langle \mu_s^N, V^N * (F(\psi_s^N) \cdot \nabla((u_s - u_s^N) * V^N)) - F(\psi_s^N) \cdot \nabla((u_s - u_s^N) * V^N) \rangle ds \\ &\quad \left. + 2 \int_0^t \int_{\mathbb{R}^d} (u_s(x) F(\psi_s(x)) - \{u_s F(\psi_s)\} * V^N(x)) \cdot \nabla(u_s - u_s^N)(x) dx ds \right) \\ &\quad + C_T \left(\frac{1}{N^{\frac{1}{2}(1-d\alpha)}} + \frac{1}{N^{1-\alpha(d+2)}} \right) \\ &=: \|u_0 - u_0^N\|_{L^2(\mathbb{R}^d)}^2 + \mathbb{E} \sup_{t \in [0, T]} \sum_{i=1}^3 I_i + C_T \left(\frac{1}{N^{\frac{1}{2}(1-d\alpha)}} + \frac{1}{N^{1-\alpha(d+2)}} \right). \end{aligned} \tag{5.9}$$

Let us now analyse the terms I_1 to I_3 . First we have

$$\begin{aligned}
|I_1| &\leq 2 \int_0^t \left| \int_{\mathbb{R}^d} \{(u_s - u_s^N)F(\psi_s^N)\} * V^N(x) \cdot \nabla(u_s - u_s^N)(x) dx \right| ds \\
&\quad + 2 \int_0^t \left| \int_{\mathbb{R}^d} \{u_s(F(\psi_s) - F(\psi_s^N))\} * V^N(x) \cdot \nabla(u_s - u_s^N)(x) dx \right| ds \\
&=: I_{1,1} + I_{1,2}.
\end{aligned} \tag{5.10}$$

Applying the inequality $2ab \leq \frac{1}{2}a^2 + 2b^2$, the convolution inequality

$$\|f * V^N\|_{L^2(\mathbb{R}^d)} \leq \|f\|_{L^2(\mathbb{R}^d)} \|V^N\|_{L^1(\mathbb{R}^d)} = \|f\|_{L^2(\mathbb{R}^d)},$$

the boundedness of F , we deduce that

$$I_{1,1} \leq \frac{1}{2} \int_0^t \|\nabla(u_s - u_s^N)\|_{L^2(\mathbb{R}^d)}^2 ds + C \int_0^t \|u_s - u_s^N\|_{L^2(\mathbb{R}^d)}^2 ds, \tag{5.11}$$

and since F is Lipschitz continuous and $\|u_s\|_{L^\infty(\mathbb{R}^d)} \lesssim \|u_s\|_\beta$,

$$I_{1,2} \leq \frac{1}{2} \int_0^t \|\nabla(u_s - u_s^N)\|_{L^2(\mathbb{R}^d)}^2 ds + C \int_0^t \|u_s\|_\beta^2 \|\psi_s - \psi_s^N\|_{L^2(\mathbb{R}^d)}^2 ds.$$

Recall that from Proposition 4.4 we have that $\mathbb{E}\|u_t^N\|_\beta^2 \leq C_T$. Hence by Fatou's lemma, one has

$$\sup_{t \leq T} \|u_t\|_\beta^2 \leq C_T. \tag{5.12}$$

Therefore,

$$I_{1,2} \leq \frac{1}{2} \int_0^t \|\nabla(u_s - u_s^N)\|_{L^2(\mathbb{R}^d)}^2 ds + C_T \int_0^t \|\psi_s - \psi_s^N\|_{L^2(\mathbb{R}^d)}^2 ds. \tag{5.13}$$

Next, we have as in (4.12) that

$$\begin{aligned}
|I_2| &\leq \left| 2 \int_0^t \langle \mu_s^N, V^N * \{F(\psi_s^N) \cdot \nabla((u_s - u_s^N) * V^N)\} - F(\psi_s^N) \cdot \nabla((u_s - u_s^N) * V^N) \rangle ds \right| \\
&\leq 2 \int_0^t \sup_{x \in \mathbb{R}^d} |V^N * \{F(\psi_s^N) \cdot \nabla((u_s - u_s^N) * V^N)\}(x) - F(\psi_s^N(x)) \cdot \nabla((u_s - u_s^N) * V^N)(x)| ds \\
&\leq 2 \int_0^t \sup_{x \in \mathbb{R}^d} \int_{\mathbb{R}^d} V(y) \left| (F(\psi_s^N) \cdot \nabla((u_s - u_s^N) * V^N))(x - \frac{y}{N^\alpha}) - (F(\psi_s^N) \cdot \nabla((u_s - u_s^N) * V^N))(x) \right| dy ds \\
&\leq \frac{C}{N^{\beta_0 \alpha}} \int_0^t \mathcal{N}_{\beta_0}(F(\psi_s^N) \cdot \nabla((u_s - u_s^N) * V^N)) ds,
\end{aligned}$$

where we recall that \mathcal{N}_{β_0} denotes the Hölder seminorm of parameter $\beta_0 = (1 - \frac{d}{p}) \wedge (\beta - 1 - \frac{d}{2})$. In view of Lemma 5.2, it follows that

$$|I_2| \leq \frac{C}{N^{\beta_0 \alpha}} \int_0^t (\|u_s^N\|_\beta + \|u_s\|_\beta + \|u_s^N\|_\beta^2 + \|u_s^N\|_\beta \|u_s\|_\beta) ds. \tag{5.14}$$

Now for I_3 , observe first that

$$\int_{\mathbb{R}^d} \{u_s F(\psi_s)\} * V^N(x) \cdot \nabla(u_s - u_s^N)(x) dx = \int_{\mathbb{R}^d} u_s(x) F(\psi_s(x)) \cdot \nabla(u_s - u_s^N) * V^N(x) dx.$$

Hence we obtain, applying the previous equality and Lemma 5.2,

$$\begin{aligned}
|I_3| &= 2 \left| \int_0^t \int_{\mathbb{R}^d} (u_s(x)F(\psi_s(x)) - \{u_s F(\psi_s)\} * V^N(x)) \cdot \nabla(u_s - u_s^N)(x) dx ds \right| \\
&= 2 \left| \int_0^t \int_{\mathbb{R}^d} u_s(x)F(\psi_s(x)) \cdot (\nabla(u_s - u_s^N)(x) - \nabla(u_s - u_s^N) * V^N(x)) dx ds \right| \\
&\leq \frac{C}{N^{\alpha\eta}} \int_0^t \mathcal{N}_{\beta_0}(\nabla(u_s - u_s^N)) \int_{\mathbb{R}^d} |u_s(x)| dx ds \\
&\leq \frac{C}{N^{\alpha\beta_0}} \int_0^t (\|u_s\|_{\beta} + \|u_s^N\|_{\beta}) ds, \tag{5.15}
\end{aligned}$$

using that $\int_{\mathbb{R}^d} |u_s(x)| dx = 1$ in the last line.

Recall that $u \in \mathcal{Y}$ (defined in (4.2)), hence combined with Proposition 4.4, one deduces that $\mathbb{E} \int_0^t (\|u_s^N\|_{\beta} + \|u_s\|_{\beta} + \|u_s^N\|_{\beta}^2 + \|u_s^N\|_{\beta} \|u_s\|_{\beta}) ds \leq C_T$. Thus, from (5.9)-(5.15), one has

$$\begin{aligned}
\mathbb{E} \sup_{t \in [0, T]} \|u_t - u_t^N\|_{L^2(\mathbb{R}^d)}^2 &+ \int_0^T \mathbb{E} \|\nabla(u_s - u_s^N)\|_{L^2(\mathbb{R}^d)}^2 ds \leq \mathbb{E} \|u_0 - u_0^N\|_{L^2(\mathbb{R}^d)}^2 \\
&+ C \int_0^T \mathbb{E} \sup_{r \in [0, s]} \|u_r - u_r^N\|_{L^2(\mathbb{R}^d)}^2 ds + C_T \int_0^T \mathbb{E} \sup_{r \in [0, s]} \|\psi_r - \psi_r^N\|_{L^2(\mathbb{R}^d)}^2 ds \\
&+ C_T \left(\frac{1}{N^{\frac{1}{2}(1-d\alpha)}} + \frac{1}{N^{1-\alpha(d+2)}} \right) + \frac{C_T}{N^{\alpha\beta_0}}, \tag{5.16}
\end{aligned}$$

where we recall that $\beta_0 = (1 - \frac{d}{p}) \wedge (\beta - 1 - \frac{d}{2})$.

2nd Part. We will now estimate $\|\psi_t - \psi_t^N\|_{L^2(\mathbb{R}^d)}^2$.

Let us convolve in space the equality (4.5) with the kernel K , and use Fubini's Theorems to get the convolution inside the time integral. In particular for the stochastic integral, the stochastic Fubini Theorem of [45] can be applied since $\int_{\mathbb{R}^d} \left(\int_0^t |K(x-y)|^2 |\nabla V^N(y - X_s^{i,N})|^2 ds \right)^{1/2} dy$ is finite a.s due to the integrability properties of K and the fact that V is a rapidly decreasing function. Hence we have

$$\begin{aligned}
\psi_t^N(x) &= \psi_0^N(x) - \int_0^t \langle \mu_s^N, \nabla (K * V^N(x - \cdot) \cdot F(\psi_s^N)) \rangle ds \\
&\quad - \frac{1}{N} \sum_{i=1}^N \int_0^t \nabla (K * V^N)(x - X_s^{i,N}) dW_s^i + \int_0^t \Delta \psi_s^N(x) ds,
\end{aligned}$$

More classically, denote by $\psi_t^\varepsilon = K * u_t^\varepsilon$. It satisfies

$$\begin{aligned}
\psi_t^\varepsilon(x) &= \psi_0^\varepsilon(x) + \int_0^t \Delta \psi_s^\varepsilon(x) ds - \int_0^t K * \nabla \cdot \langle u_s F(\psi_s), \phi^\varepsilon \rangle(x) ds \\
&= \psi_0^\varepsilon(x) + \int_0^t \Delta \psi_s^\varepsilon(x) ds - \int_0^t \nabla K * \langle u_s F(\psi_s), \phi^\varepsilon \rangle(x) ds.
\end{aligned}$$

In order to show that the mollifier ψ^ε converges to ψ in $L^2([0, T], H^1(\mathbb{R}^d))$, we will show that ψ belongs to this space.

First, we show that $\psi \in L^\infty([0, T], L^2(\mathbb{R}^d))$. As u satisfies (2.3), we observe that ψ_t verifies almost surely

$$\psi_t = e^{t\Delta} \psi_0 - \int_0^t K * (\nabla e^{(t-s)\Delta} u_s F(\psi_s)) ds = e^{t\Delta} \psi_0 - \int_0^t \nabla K * (e^{(t-s)\Delta} u_s F(\psi_s)) ds.$$

Then, as ∇K defines a convolution operator that is bounded on $L^2(\mathbb{R}^d)$, we have

$$\begin{aligned}
\|\psi_t\|_{L^2(\mathbb{R}^d)} &\leq \|e^{t\Delta}\psi_0\|_{L^2(\mathbb{R}^d)} + \int_0^t \|\nabla K * (e^{(t-s)\Delta}u_s F(\psi_s))\|_{L^2(\mathbb{R}^d)} ds \\
&\leq \|e^{t\Delta}\psi_0\|_{L^2(\mathbb{R}^d)} + C \int_0^t \|u_s\|_{L^2(\mathbb{R}^d)} ds \leq C.
\end{aligned}$$

By hypothesis (2.11), we conclude that $\psi \in L^\infty([0, T], L^2(\mathbb{R}^d))$. Repeating the above computations for $\nabla\psi$, one obtains

$$\|\nabla\psi_t\|_{L^2(\mathbb{R}^d)} \leq \|\nabla e^{t\Delta}\psi_0\|_{L^2(\mathbb{R}^d)} + C \int_0^t \frac{\|u_s\|_{L^2(\mathbb{R}^d)}}{\sqrt{t-s}} ds.$$

By the Sobolev embedding (3.1), as u is a mild solution to the cut-off PDE, one obtains

$$\|\nabla\psi_t\|_{L^2(\mathbb{R}^d)} \leq \|\nabla e^{t\Delta}\psi_0\|_{L^2(\mathbb{R}^d)} + C_T \sup_{s \in [0, T]} \|u_s\|_{L^1 \cap W^{1, D}(\mathbb{R}^d)}.$$

Thus, by hypothesis (2.11), we conclude that $\psi \in L^\infty([0, T], H^1(\mathbb{R}^d))$.

From the above equations, Itô's formula and integration in space, we obtain

$$\begin{aligned}
\|\psi_t^\varepsilon - \psi_t^N\|_{L^2(\mathbb{R}^d)}^2 &= \|\psi_0^\varepsilon - \psi_0^N\|_{L^2(\mathbb{R}^d)}^2 \\
&\quad + 2 \int_0^t \int_{\mathbb{R}^d} \langle \mu_s^N, \nabla (K * V^N(x - \cdot) \cdot F(\psi_s^N)) \rangle \cdot (\psi_s^\varepsilon - \psi_s^N)(x) dx ds \\
&\quad - 2 \int_0^t \int_{\mathbb{R}^d} \nabla K * \{ \langle u_s F(\psi_s), \phi^\varepsilon \rangle \} (x) \cdot (\psi_s^\varepsilon - \psi_s^N)(x) dx ds \\
&\quad + 2 \int_0^t \int_{\mathbb{R}^d} \Delta(\psi_s^\varepsilon - \psi_s^N)(x) \cdot (\psi_s^\varepsilon - \psi_s^N)(x) dx ds \\
&\quad + \frac{2}{N} \sum_{i=1}^N \int_{\mathbb{R}^d} \int_0^t (\psi_s^\varepsilon - \psi_s^N)(x) \cdot \nabla (K * V^N)(x - X_s^{i, N}) dW_s^{i, N} dx \\
&\quad + \frac{1}{N} \int_0^t \|\nabla (K * V^N)\|_{L^2(\mathbb{R}^d)}^2 ds.
\end{aligned} \tag{5.17}$$

As ∇K defines a convolution operator that is bounded on $L^2(\mathbb{R}^d)$ and $\frac{t}{N} \|\nabla V^N\|_{L^2(\mathbb{R}^d)}^2 \leq C_T N^{-1+\alpha(d+2)}$, one has

$$\frac{t}{N} \|\nabla (K * V^N)\|_{L^2(\mathbb{R}^d)}^2 ds \leq C_T \frac{1}{N} \|V^N\|_{L^2(\mathbb{R}^d)}^2 = C_T \frac{1}{N^{1-\alpha(d+2)}}. \tag{5.18}$$

Now, as in the first part, we derive the first inequality below thanks to the stochastic Fubini lemma and the BDG inequality, then the second inequality by applying the Cauchy-Schwarz inequality and the fact that ∇K defines a convolution operator that is bounded on $L^2(\mathbb{R}^d)$, to get

$$\begin{aligned}
&\mathbb{E} \sup_{t \in [0, T]} \left| \frac{2}{N} \sum_{i=1}^N \int_{\mathbb{R}^d} \int_0^t (\psi_s^\varepsilon - \psi_s^N)(x) \cdot \nabla (K * V^N)(x - X_s^{i, N}) dW_s^{i, N} dx \right| \\
&\leq \frac{2}{N} \mathbb{E} \left[\left(\sum_{i=1}^N \int_0^T \left| \int_{\mathbb{R}^d} (\psi_s^\varepsilon - \psi_s^N)(x) \cdot (\nabla K * V^N)(x - X_s^{i, N}) dx \right|^2 ds \right)^{\frac{1}{2}} \right] \\
&\leq \frac{2}{N} \mathbb{E} \left[\left(\sum_{i=1}^N \int_0^T \|\psi_s^\varepsilon - \psi_s^N\|_{L^2(\mathbb{R}^d)}^2 \|V^N\|_{L^2(\mathbb{R}^d)}^2 ds \right)^{\frac{1}{2}} \right] \\
&\leq \frac{2}{\sqrt{N}} \|V^N\|_{L^2(\mathbb{R}^d)} \int_0^T \mathbb{E} \|\psi_s^\varepsilon - \psi_s^N\|_{L^2(\mathbb{R}^d)}^2 ds \leq C \int_0^T \mathbb{E} \|\psi_s^\varepsilon - \psi_s^N\|_{L^2(\mathbb{R}^d)}^2 ds.
\end{aligned} \tag{5.19}$$

Using similar arguments as when we derived (5.9), from (5.17), (5.18) and (5.19) we get

$$\begin{aligned}
& \mathbb{E} \sup_{t \in [0, T]} \|\psi_t - \psi_t^N\|_{L^2(\mathbb{R}^d)}^2 + 2\mathbb{E} \int_0^T \|\nabla(\psi_s - \psi_s^N)\|_{L^2(\mathbb{R}^d)}^2 ds \leq \mathbb{E} \|\psi_0 - \psi_0^N\|_{L^2(\mathbb{R}^d)}^2 \\
& + \mathbb{E} \sup_{t \in [0, T]} \left(-2 \int_0^t \int_{\mathbb{R}^d} \nabla K * V^N * \{u_s F(\psi_s) - u_s^N F(\psi_s^N)\}(x) \cdot (\psi_s - \psi_s^N)(x) dx ds \right. \\
& + 2 \int_0^t \langle S_s^N, V^N * (F(\psi_s^N) \cdot \{(\nabla \cdot (\psi_s - \psi_s^N)) * K * V^N\}) - F(\psi_s^N) \cdot \{(\nabla \cdot (\psi_s - \psi_s^N)) * K * V^N\} \rangle ds \\
& \left. - 2 \int_0^t \int_{\mathbb{R}^d} (\nabla K * \{u_s F(\psi_s)\}(x) - \nabla K * V^N * \{u_s F(\psi_s)\}(x)) \cdot (\psi_s - \psi_s^N)(x) dx ds \right) \\
& + C_T \frac{1}{N^{1-\alpha(d+2)}} + C \int_0^T \mathbb{E} \|\psi_s - \psi_s^N\|_{L^2(\mathbb{R}^d)}^2 ds \\
& =: \|\psi_0 - \psi_0^N\|_{L^2(\mathbb{R}^d)}^2 + \mathbb{E} \sup_{t \in [0, T]} \sum_{i=1}^3 J_i + C_T \frac{1}{N^{1-\alpha(d+2)}} + C \int_0^T \mathbb{E} \|\psi_s - \psi_s^N\|_{L^2(\mathbb{R}^d)}^2 ds. \tag{5.20}
\end{aligned}$$

We now treat the terms J_1 to J_3 . For J_1 , we proceed as for I_1 :

$$\begin{aligned}
|J_1| & \leq 2 \int_0^t \int_{\mathbb{R}^d} |\nabla K * V^N * \{(u_s - u_s^N)F(\psi_s^N)\}(x) \cdot (\psi_s - \psi_s^N)(x)| dx ds \\
& + 2 \int_0^t \int_{\mathbb{R}^d} |\nabla K * V^N * \{u_s(F(\psi_s) - F(\psi_s^N))\}(x) \cdot (\psi_s - \psi_s^N)(x)| dx ds \\
& \leq \int_0^t \|\nabla K * V^N * \{(u_s - u_s^N)F(\psi_s^N)\}\|_{L^2(\mathbb{R}^d)}^2 ds \\
& + \int_0^t \|\nabla K * V^N * \{u_s(F(\psi_s) - F(\psi_s^N))\}\|_{L^2(\mathbb{R}^d)}^2 ds + 4 \int_0^t \|\psi_s - \psi_s^N\|_{L^2(\mathbb{R}^d)}^2 ds,
\end{aligned}$$

where we used that $2ab \leq \frac{1}{2}a^2 + 2b^2$ in the last inequality. Now a convolution inequality, the fact that ∇K defines a convolution operator that is bounded on $L^2(\mathbb{R}^d)$, Inequality (5.12) and the boundedness and Lipschitz continuity of F yield

$$\begin{aligned}
|J_1| & \leq \int_0^t \|\nabla K * \{(u_s - u_s^N)F(\psi_s^N)\}\|_{L^2(\mathbb{R}^d)}^2 \|V^N\|_{L^1(\mathbb{R}^d)}^2 ds \\
& + \int_0^t \|\nabla K * \{u_s(F(\psi_s) - F(\psi_s^N))\}\|_{L^2(\mathbb{R}^d)}^2 \|V^N\|_{L^1(\mathbb{R}^d)}^2 ds + C \int_0^t \|\psi_s - \psi_s^N\|_{L^2(\mathbb{R}^d)}^2 ds \\
& \leq C \int_0^t \|u_s - u_s^N\|_{L^2(\mathbb{R}^d)}^2 ds + C \int_0^t \|u_s\|_{\beta}^2 \|\psi_s - \psi_s^N\|_{L^2(\mathbb{R}^d)}^2 ds + C \int_0^t \|\psi_s - \psi_s^N\|_{L^2(\mathbb{R}^d)}^2 ds \\
& \leq C \int_0^t \|u_s - u_s^N\|_{L^2(\mathbb{R}^d)}^2 ds + C_T \int_0^t \|\psi_s - \psi_s^N\|_{L^2(\mathbb{R}^d)}^2 ds. \tag{5.21}
\end{aligned}$$

For J_2 , we follow the computations for I_2 to get that for $\beta_0 = (1 - \frac{d}{p}) \wedge (\beta - 1 - \frac{d}{2})$,

$$\begin{aligned}
|J_2| & \leq 2 \int_0^t \sup_{x \in \mathbb{R}^d} \int_{\mathbb{R}^d} V(y) \left| (F(\psi_s^N) \cdot \{(\nabla \cdot (\psi_s - \psi_s^N)) * K * V^N\})(x - \frac{y}{N^\alpha}) \right. \\
& \quad \left. - (F(\psi_s^N) \cdot \{(\nabla \cdot (\psi_s - \psi_s^N)) * K * V^N\})(x) \right| dy ds \\
& \leq \frac{C}{N^{\beta_0 \alpha}} \int_0^t \mathcal{N}_{\beta_0}(F(\psi_s^N) \cdot \{(\nabla \cdot (\psi_s - \psi_s^N)) * K * V^N\}) ds.
\end{aligned}$$

Apply the Lemma 5.2 to get

$$|J_2| \leq \frac{C}{N^{\beta_0 \alpha}} \int_0^t (\|u_s^N\|_{\beta} + \|u_s\|_{\beta} + \|u_s^N\|_{\beta}^2 + \|u_s^N\|_{\beta} \|u_s\|_{\beta}) ds. \tag{5.22}$$

For J_3 , observe that

$$\begin{aligned}
|J_3| &= 2 \left| \int_0^t \int_{\mathbb{R}^d} \nabla K * \{u_s F(\psi_s)\} (x) \cdot ((\psi_s - \psi_s^N)(x) - (\psi_s - \psi_s^N) * V^N(x)) \, dx ds \right| \\
&= 2 \left| \int_0^t \int_{\mathbb{R}^d} (u_s F(\psi_s)) (x) \cdot K * \{ \nabla \cdot (\psi_s - \psi_s^N)(x) - \nabla \cdot (\psi_s - \psi_s^N) * V^N(x) \} \, dx ds \right| \\
&\leq \frac{C}{N^{\alpha\beta_0}} \int_0^t \mathcal{N}_{\beta_0}(K * (\nabla \cdot (\psi_s - \psi_s^N))) \int_{\mathbb{R}^d} u_s(x) |F(\psi_s(x))| \, dx ds
\end{aligned}$$

In view of Lemma 5.2 with $\beta_0 = (1 - \frac{d}{p}) \wedge (\beta - 1 - \frac{d}{2})$, it follows from the above computations that there is

$$|J_3| \leq \frac{C}{N^{\alpha\beta_0}} \int_0^t (\|u_s\|_{\beta} + \|u_s^N\|_{\beta}) \, ds. \quad (5.23)$$

Combining the expressions in (5.20) - (5.23), it comes

$$\begin{aligned}
&\mathbb{E} \sup_{t \in [0, T]} \|\psi_t - \psi_t^N\|_{L^2(\mathbb{R}^d)}^2 + 2\mathbb{E} \int_0^T \|\nabla(\psi_s - \psi_s^N)\|_{L^2(\mathbb{R}^d)}^2 \, ds \leq \mathbb{E} \|\psi_0 - \psi_0^N\|_{L^2(\mathbb{R}^d)}^2 \\
&+ C \int_0^T \mathbb{E} \|u_s - u_s^N\|_{L^2(\mathbb{R}^d)}^2 \, ds + C_T \int_0^T \mathbb{E} \|\psi_s - \psi_s^N\|_{L^2(\mathbb{R}^d)}^2 \, ds + \frac{C_T}{N^{\alpha\beta_0}} + C_T \frac{1}{N^{1-\alpha(d+2)}}. \quad (5.24)
\end{aligned}$$

Here, applying Grönwall's lemma just on $\mathbb{E} \sup_{t \in [0, T]} \|\psi_t - \psi_t^N\|_{L^2(\mathbb{R}^d)}^2$, one gets

$$\begin{aligned}
\mathbb{E} \sup_{t \in [0, T]} \|\psi_t - \psi_t^N\|_{L^2(\mathbb{R}^d)}^2 &\leq C_T \left[\mathbb{E} \|\psi_0 - \psi_0^N\|_{L^2(\mathbb{R}^d)}^2 + \int_0^T \mathbb{E} \sup_{r \in [0, s]} \|u_r - u_r^N\|_{L^2(\mathbb{R}^d)}^2 \, ds \right. \\
&\quad \left. + \frac{C_T}{N^{\alpha\beta_0}} + C_T \frac{1}{N^{1-\alpha(d+2)}} \right]. \quad (5.25)
\end{aligned}$$

3rd Part. First, plug (5.25) in (5.16) and apply Grönwall's lemma only on $\mathbb{E} \sup_{t \in [0, T]} \|u_t - u_t^N\|_{L^2(\mathbb{R}^d)}^2$. It comes

$$\begin{aligned}
\mathbb{E} \sup_{t \in [0, T]} \|u_t - u_t^N\|_{L^2(\mathbb{R}^d)}^2 &\leq C_T \left[\mathbb{E} \|u_0 - u_0^N\|_{L^2(\mathbb{R}^2)}^2 + \mathbb{E} \|\psi_0 - \psi_0^N\|_{L^2(\mathbb{R}^2)}^2 + \frac{1}{N^{\alpha\beta_0}} \right. \\
&\quad \left. + \frac{1}{N^{\frac{1}{2}(1-d\alpha)}} + \frac{1}{N^{1-\alpha(d+2)}} \right]. \quad (5.26)
\end{aligned}$$

In view of (5.26) and (5.25), one has

$$\begin{aligned}
&\mathbb{E} \sup_{t \in [0, T]} \|u_t - u_t^N\|_{L^2(\mathbb{R}^d)}^2 + \mathbb{E} \sup_{t \in [0, T]} \|\psi_t - \psi_t^N\|_{L^2(\mathbb{R}^d)}^2 \\
&\leq C_T \left[\mathbb{E} \|u_0 - u_0^N\|_{L^2(\mathbb{R}^2)}^2 + \mathbb{E} \|\psi_0 - \psi_0^N\|_{L^2(\mathbb{R}^2)}^2 + \frac{1}{N^{\alpha\beta_0}} + \frac{1}{N^{\frac{1}{2}(1-d\alpha)}} + \frac{1}{N^{1-\alpha(d+2)}} \right].
\end{aligned}$$

Using the Hypothesis (2.10), we deduce

$$\mathbb{E} \sup_{t \in [0, T]} \|u_t - u_t^N\|_{L^2(\mathbb{R}^d)}^2 + \mathbb{E} \sup_{t \in [0, T]} \|\psi_t - \psi_t^N\|_{L^2(\mathbb{R}^d)}^2 \leq \frac{C_T}{N^{\kappa}}, \quad (5.27)$$

where $\kappa = \min\{\beta_0\alpha, 1 - \alpha(d+2), \frac{1}{2}(1-d\alpha)\}$. Observe that with the constraint coming from (\mathbf{H}_{iii}) , one has $\alpha < \frac{1}{d+2\beta}$. Since $\beta > \frac{d}{2} + 1 \geq \max(\frac{d}{2}, 1) + 1$, we have that $\alpha < \frac{1}{d+4}$, which implies that $1 - \alpha(d+2) > \frac{1}{2}(1-d\alpha)$. This also implies that $\frac{1}{2}(1-d\alpha) > \beta_0\alpha$ as the latter is implied by $\alpha < \frac{1}{d+\beta_0\alpha}$ and $\beta_0 < 1$. Thus, $\kappa = \beta_0\alpha$.

Finally, combine (5.27) with (5.16), resp. (5.24), to obtain the claim (i), resp. (ii), of the theorem. This concludes the proof. \square

6 Propagation of chaos

In this section we study the well-posedness of the nonlinear SDE (2.12) and then the propagation of chaos of the particle system (2.6). More precisely, we prove Proposition 2.14 about the well-posedness of the martingale problem (\mathcal{MP}) related to (2.12). Then, we prove the convergence in law, when $N \rightarrow \infty$, of the empirical measure μ^N of the particle system (2.6) towards the unique solution of the martingale problem (\mathcal{MP}) .

6.1 Proof of Proposition 2.14

First, for a solution \mathbb{Q} to the martingale problem, we study the mild equation for its time-marginals. We will see that this equation admits a unique solution in a suitable functional space. This will enable us to study the linear version of the martingale problem (\mathcal{MP}) . Analysing this linear martingale problem, we will get the uniqueness and existence for (\mathcal{MP}) .

To obtain the equation satisfied by $(q_t)_{t \geq 0}$, one derives the mild equation for the marginal distributions of the corresponding nonlinear process. This is done in the usual way as the drift component is bounded (see e.g. [42, Section 4]). One has

$$q_t = e^{t\Delta} u_0 - \int_0^t \nabla \cdot e^{(t-s)\Delta} (q_s F_A(K * q_s)) ds, \quad 0 \leq t \leq T.$$

This equation was already examined in Section 3.3 and in Proposition 4.3, and it was proved it admits a unique global solution in the sense of Definition 2.1. Meaning, as $q \in \mathcal{C}([0, T]; L^1 \cap W^{1,D}(\mathbb{R}^d))$, the one-dimensional time marginals of \mathbb{Q} are uniquely determined.

Define the corresponding linear martingale problem by fixing q to be the unique mild solution \tilde{u} to (2.2) in the definition of the process M_t from (\mathcal{MP}) .

As the SDE

$$Y_t = X_0 + \sqrt{2}W_t + \int_0^t F_A(K * \tilde{u}_s(Y_s)) ds$$

admits a unique strong solution (see [31, Remark 1.6]), we know that the associated linear martingale problem admits a unique solution. This immediately implies the uniqueness of solutions to (\mathcal{MP}) .

Now, a candidate for a solution to the (\mathcal{MP}) is the probability measure $\mathbb{P} := \mathcal{L}(Y)$. To prove the latter is the solution of (\mathcal{MP}) , we need to ensure that the family of marginal laws $(\mathbb{P}_t)_{0 \leq t \leq T}$ is exactly the family $(\tilde{u}_t)_{0 \leq t \leq T}$.

To do so, for a $0 < t \leq T$, one derives the mild equation for $\mathbb{P}_t = p_t(x)dx$ (absolute continuity follows from bounded drift and Girsanov transformation). Following the same arguments as in [42, Section 4], as the drift is bounded, we have that for a.e. $x \in \mathbb{R}^d$,

$$p_t = e^{t\Delta} u_0 - \int_0^t \nabla \cdot e^{(t-s)\Delta} (p_s F_A(K * \tilde{u}_s)) ds, \quad 0 \leq t \leq T.$$

Assume for a moment that $p \in \mathcal{C}([0, T]; L^1 \cap W^{1,D}(\mathbb{R}^d))$. The previous equation is a linearized version of the equation (2.3) and with the same arguments as in Proposition 3.6, it admits a unique solution in $\mathcal{C}([0, T]; L^1 \cap W^{1,D}(\mathbb{R}^d))$. Since both \tilde{u} and p solve this equation, they must coincide and we have the desired result : $(p_t)_{t \in [0, T]} = (\tilde{u}_t)_{t \in [0, T]}$.

It only remains to prove that $p \in \mathcal{C}([0, T]; L^1 \cap W^{1,D}(\mathbb{R}^d))$. Performing the same calculations as in the proof of Proposition 3.6, we get that

$$\|p_t\|_{L^D(\mathbb{R}^d)} \leq \|u_0\|_{L^D(\mathbb{R}^d)} + C \int_0^t \frac{\|p_s\|_{L^D(\mathbb{R}^d)}}{\sqrt{t-s}} \|K * \tilde{u}_s\|_{L^\infty(\mathbb{R}^d)} ds.$$

In view of Lemma 3.1, one has

$$\|p_t\|_{L^1(\mathbb{R}^d)} + \|p_t\|_{L^D(\mathbb{R}^d)} \leq \|u_0\|_{L^1(\mathbb{R}^d)} + \|u_0\|_{L^D(\mathbb{R}^d)} + C \|\tilde{u}\|_{T, L^1 \cap W^{1,D}(\mathbb{R}^d)} \int_0^t \frac{\|p_s\|_{L^1 \cap W^{1,D}(\mathbb{R}^d)}}{\sqrt{t-s}} ds. \quad (6.1)$$

Now,

$$\|\nabla p_t\|_{L^D(\mathbb{R}^d)} \leq C \int_0^t \frac{1}{\sqrt{t-s}} (\|\nabla p_s\|_{L^D(\mathbb{R}^d)} + \|p_s \nabla \cdot F(K * \tilde{u}_s)\|_{L^D(\mathbb{R}^d)}) ds.$$

As before, using the Hölder inequality, one has

$$\|p_s \nabla \cdot F(K * \tilde{u}_s)\|_{L^D(\mathbb{R}^d)} \leq \|p_s\|_{L^{\frac{pD}{p-D}}(\mathbb{R}^d)} \|\nabla \cdot F(K * \tilde{u}_s)\|_{L^p(\mathbb{R}^d)}.$$

The boundedness of ∇F and the property (\mathbf{H}_{iii}^K) yield

$$\|\nabla \cdot F(K * \tilde{u}_s)\|_{L^p(\mathbb{R}^d)} \leq C \|\tilde{u}_s\|_{L^1 \cap W^{1,D}(\mathbb{R}^d)},$$

and the embedding (3.1) now gives

$$\|p_s \nabla \cdot F(K * \tilde{u}_s)\|_{L^d(\mathbb{R}^d)} \leq \|p_s\|_{L^1 \cap W^{1,D}(\mathbb{R}^d)} \|\tilde{u}_s\|_{L^1 \cap W^{1,D}(\mathbb{R}^d)}.$$

It comes

$$\|\nabla p_t\|_{L^D(\mathbb{R}^d)} \leq C \int_0^t \frac{1}{\sqrt{t-s}} (\|p_s\|_{L^1 \cap W^{1,D}(\mathbb{R}^d)} + \|p_s\|_{L^1 \cap W^{1,D}(\mathbb{R}^d)} \|\tilde{u}\|_{T, L^1 \cap W^{1,D}(\mathbb{R}^d)}) ds. \quad (6.2)$$

Combining (6.1) and (6.2), one gets that

$$\|p_t\|_{L^1 \cap W^{1,D}(\mathbb{R}^d)} \leq \|u_0\|_{L^1 \cap W^{1,D}(\mathbb{R}^d)} + C \|\tilde{u}\|_{T, L^1 \cap W^{1,D}(\mathbb{R}^d)} \int_0^t \frac{\|p_s\|_{L^1 \cap W^{1,D}(\mathbb{R}^d)}}{\sqrt{t-s}} ds.$$

Grönwall's lemma implies that $p_t \in L^1 \cap W^{1,D}(\mathbb{R}^d)$ and it comes that $p \in \mathcal{C}([0, T]; L^1 \cap W^{1,D}(\mathbb{R}^d))$. This concludes the proof.

6.2 Proof of Theorem 2.15

To prove Theorem 2.15, we will prove that μ^N converges to the unique solution \mathbb{Q} of the martingale problem $(\mathcal{M}\mathcal{P})$.

A usual way to do so consists in proving the tightness of the family $\Pi^N := \mathcal{L}(\mu^N)$ in the space $\mathcal{P}(\mathcal{P}(\mathcal{C}([0, T]; \mathbb{R}^d)))$ and, then, in proving that any limit point Π^∞ of Π^N is $\delta_{\mathbb{Q}}$. The latter is done by showing that under Π^∞ a certain quadratic function of the canonical measure in $\mathcal{P}(\mathcal{C}([0, T]; \mathbb{R}^d))$ is zero. The form of this function depends on the form of the process $(M_t)_{t \geq 0}$ specified in the definition of the martingale problem. Moreover, one must analyse this function under Π^N and use the convergence of Π^N to Π^∞ to get the desired result. This is where μ^N and the particle system appear.

However, here the situation is slightly modified. Namely, at the level of Π^N , we need to keep track not just of μ^N , but also of the mollified empirical measure u^N that appears in the definition of the particle system. That is why we will need to use the convergence of u^N towards u proved before and keep track of the couple (μ^N, u^N) . This random variable lives in the product space

$$\mathcal{H} := \mathcal{P}(\mathcal{C}([0, T]; \mathbb{R}^d)) \times \tilde{\mathcal{Y}}.$$

endowed with the weak topology of $\mathcal{P}(\mathcal{C}([0, T]; \mathbb{R}^d))$ and the topology of $\tilde{\mathcal{Y}}$, where we recall that $\tilde{\mathcal{Y}}$ was defined in (4.20) as $\tilde{\mathcal{Y}} = \mathcal{Y} \cap L^\infty([0, T]; L^1(\mathbb{R}^d))$. We will denote by $(\boldsymbol{\mu}, \mathbf{u})$ the canonical projections in \mathcal{H} .

Now, for $N \geq 1$ we denote by $\tilde{\Pi}^N$ the law of the random variables (μ^N, u^N) that take values in \mathcal{H} . The sequence $(\tilde{\Pi}^N, N \geq 1)$ is tight if and only if $(\tilde{\Pi}^N \circ \boldsymbol{\mu}, N \geq 1)$ and $(\tilde{\Pi}^N \circ \mathbf{u}, N \geq 1)$ are tight. The tightness of $(\tilde{\Pi}^N \circ \boldsymbol{\mu}, N \geq 1)$ is classical, as the drift of the particle system is bounded. As for $(\tilde{\Pi}^N \circ \mathbf{u}, N \geq 1)$, we have already proven the convergence of $(u^N, N \geq 1)$ in $\tilde{\mathcal{Y}}$ (see Theorem 2.6 and Corollary 2.7).

Once we have the tightness of $(\tilde{\Pi}^N, N \geq 1)$, let $\tilde{\Pi}^\infty$ be a limit point of $(\tilde{\Pi}^N, N \geq 1)$. By a slight abuse of notation, we denote the subsequence converging to it by $(\tilde{\Pi}^N, N \geq 1)$ as well. We will study the support of $\tilde{\Pi}^\infty$ in order to describe the support of $\Pi^\infty := \tilde{\Pi}^\infty \circ \boldsymbol{\mu}$.

The following lemma shows that the marginals of $\boldsymbol{\mu}$ and \mathbf{u} coincide under the limit probability measure. This will be extremely useful to obtain that the support of Π^∞ is concentrated around \mathbb{Q} .

Lemma 6.1. $\tilde{\Pi}^\infty$ -almost surely, $\boldsymbol{\mu}_t$ is absolutely continuous w.r.t. Lebesgue measure and its density is $\boldsymbol{\mu}_t(dx) = \mathbf{u}_t(x)dx (= u_t(x)dx)$.

Proof. This is a consequence of Corollary 2.9. Take a test function $\varphi \in \mathcal{C}_b^1([0, T] \times \mathbb{R}^d) \cap L^2([0, T] \times \mathbb{R}^d)$ and define a functional $\phi(t, x) = \varphi(t, x_t)$, for $x \in \mathcal{C}([0, T]; \mathbb{R}^d)$. Then,

$$\begin{aligned} \mathbb{E}_{\tilde{\Pi}^\infty} |\langle \mathbf{u}, \varphi \rangle - \langle dt \otimes \boldsymbol{\mu}, \phi \rangle| &= \lim_{N \rightarrow \infty} \mathbb{E}_{\tilde{\Pi}^N} |\langle \mathbf{u}, \varphi \rangle - \langle dt \otimes \boldsymbol{\mu}, \phi \rangle| = \lim_{N \rightarrow \infty} \mathbb{E} |\langle u^N, \varphi \rangle - \langle dt \otimes \mu_t^N, \varphi \rangle| \\ &\leq \lim_{N \rightarrow \infty} \mathbb{E} \int_0^T |\langle u_t^N, \varphi(t, \cdot) \rangle - \langle \mu_t^N, \varphi(t, \cdot) \rangle| dt \\ &\leq C_T \sup_{t \in [0, T]} \|\varphi(t, \cdot)\|_{\text{Lip}} \times \lim_{N \rightarrow \infty} \frac{1}{N^\alpha}, \end{aligned}$$

where the last inequality comes from Corollary 2.9. Thus, we obtain that $\tilde{\Pi}^\infty$ -a.s. the following measures on $\mathbb{R}^d \times [0, T]$ are equal:

$$\mathbf{u}_t(x)dx dt = \boldsymbol{\mu}_t(dx)dt,$$

hence $\tilde{\Pi}^\infty$ -a.s., for almost all $t \in [0, T]$,

$$u_t(x)dx = \mathbf{u}_t(x)dx = \boldsymbol{\mu}_t(dx).$$

□

The following proposition will be the last ingredient needed for the proof of Theorem 2.15.

Proposition 6.2. Let $p \in \mathbb{N}$, $f \in \mathcal{C}_b^2(\mathbb{R}^d)$, $\Phi \in \mathcal{C}_b(\mathbb{R}^{dp})$ and $0 < s_1 < \dots < s_p \leq s < t \leq T$. Define H as the following function on \mathcal{H} :

$$\begin{aligned} H(\boldsymbol{\mu}, \mathbf{u}) &= \int_{\mathcal{C}([0, T]; \mathbb{R}^d)} \Phi(x_{s_1}, \dots, x_{s_p}) \left[f(x_t) - f(x_s) - \int_s^t \Delta f(x_r) dr \right. \\ &\quad \left. + \int_s^t F_A(K * \mathbf{u}_r(x_r)) \cdot \nabla f(x_r) dr \right] d\boldsymbol{\mu}(x). \end{aligned} \quad (6.3)$$

Then

$$\mathbb{E}_{\tilde{\Pi}^\infty}(H^2) = 0.$$

Proof. Step 1. Notice that $\lim_{N \rightarrow \infty} \mathbb{E}_{\tilde{\Pi}^N}(H^2) = 0$. Indeed, by Itô's formula applied on $\frac{1}{N} \sum_{i=1}^N (f(X_t^i) - f(X_s^i))$, one has

$$\mathbb{E}_{\tilde{\Pi}^N}(H^2) = \mathbb{E}(H(\mu^N, u^N)^2) = \mathbb{E} \left(\frac{1}{N} \sum_{i=1}^N \int_s^t \nabla f(X_r^i) \cdot dW_r^i \right)^2 = \frac{1}{N^2} \sum_{i=1}^N \mathbb{E} \left(\int_s^t \nabla f(X_r^i) \cdot dW_r^i \right)^2 \leq \frac{C}{N}.$$

Step 2. We prove that H is continuous on \mathcal{H} . Let $(\boldsymbol{\mu}^n, \mathbf{u}^n)$ be a sequence converging in \mathcal{H} to $(\boldsymbol{\mu}, \mathbf{u})$. Let us prove $\lim_{n \rightarrow \infty} |H(\boldsymbol{\mu}^n, \mathbf{u}^n) - H(\boldsymbol{\mu}, \mathbf{u})| = 0$.

We decompose

$$|H(\boldsymbol{\mu}^n, \mathbf{u}^n) - H(\boldsymbol{\mu}, \mathbf{u})| \leq |H(\boldsymbol{\mu}^n, \mathbf{u}^n) - H(\boldsymbol{\mu}^n, \mathbf{u})| + |H(\boldsymbol{\mu}^n, \mathbf{u}) - H(\boldsymbol{\mu}, \mathbf{u})| =: I_n + II_n.$$

Notice that

$$\begin{aligned} I_n &\leq \|\Phi\|_\infty \|\nabla f\|_\infty \langle \boldsymbol{\mu}^n, \int_s^t |F(K * \mathbf{u}_r^{\cdot}(\cdot)) - F(K * \mathbf{u}_r(\cdot))| dr \rangle \\ &\leq C \int_s^t \langle \boldsymbol{\mu}_r^n, |K * (\mathbf{u}_r^n - \mathbf{u}_r)| \rangle dr \leq C \int_s^t \|K * (\mathbf{u}_r^n - \mathbf{u}_r)\|_\infty dr. \end{aligned} \quad (6.4)$$

By Lemma 3.1, one has

$$\|K * (\mathbf{u}_r^n - \mathbf{u}_r)\|_{L^\infty(\mathbb{R}^d)} \leq C_{K,d} \|\mathbf{u}_r^n - \mathbf{u}_r\|_{L^1 \cap W^{1,D}(\mathbb{R}^d)}.$$

Now choose γ such that $\frac{D}{2} < \gamma < \beta$ and use the Sobolev embedding $L^1 \cap H^\gamma(\mathbb{R}^d) \subset L^1 \cap W^{1,D}(\mathbb{R}^d)$ (see (3.3)). Then, one gets

$$\|K * (\mathbf{u}_r^n - \mathbf{u}_r)\|_{L^\infty(\mathbb{R}^d)} \leq C_{K,d} \|\mathbf{u}_r^n - \mathbf{u}_r\|_{L^1 \cap H^\gamma(\mathbb{R}^d)}.$$

Plug the latter in (6.4) to obtain

$$I_n \leq C_{K,d} \int_s^t \|\mathbf{u}_r^n - \mathbf{u}_r\|_{L^1 \cap H^\gamma(\mathbb{R}^d)} dr.$$

The right-hand side converges to zero as the convergence of \mathbf{u}^n to \mathbf{u} in $\tilde{\mathcal{Y}}$ implies the a.s. convergence in $L^2([0, T]; H^\gamma(\mathbb{R}^d))$, for $\gamma < \beta$ and in $L^\infty([0, T]; L^1(\mathbb{R}^d))$.

To prove that II_n converges to zero, as $\boldsymbol{\mu}^n$ converges weakly to $\boldsymbol{\mu}$, we should prove the continuity of the functional $F : C([0, T]; \mathbb{R}^d) \rightarrow \mathbb{R}$ defined by

$$G(x) = \Phi(x_{s_1}, \dots, x_{s_p}) [f(x_t) - f(x_s) - \int_s^t \Delta f(x_r) dr - \int_s^t F_A(\nabla G * \mathbf{u}_r(x_r)) \cdot \nabla f(x_r) dr].$$

Let $(x^n)_{n \geq 1}$ a sequence converging in $\mathcal{C}([0, T]; \mathbb{R}^d)$ to x . To prove $G(x^n) \rightarrow G(x)$ as $n \rightarrow \infty$, having in mind the properties of f and Φ , we should only concentrate on the term $\int_s^t F_A(K * \mathbf{u}_r(x_r^n)) \cdot \nabla f(x_r^n) dr$. Here is where we use the property (A.8) applied to $K * \mathbf{u}_r$. According to it, $K * \mathbf{u}_r(x_r^n)$ converges to $K * \mathbf{u}_r(x_r)$ and by dominated convergence,

$$\int_s^t F_A(K * \mathbf{u}_r(x_r^n)) \cdot \nabla f(x_r^n) dr \rightarrow \int_s^t F_A(K * \mathbf{u}_r(x_r)) \cdot \nabla f(x_r) dr, \quad \text{as } n \rightarrow \infty.$$

Conclusion. Combine Step 1 and Step 2 to finish the proof. □

We have all the elements in hand to finish the proof of Theorem 2.15. By Lemma 6.1 and Proposition 6.2, we get that $\boldsymbol{\mu} \in \text{supp}(\Pi^\infty)$ solves the nonlinear martingale problem (\mathcal{MP}) . As we have the uniqueness for (\mathcal{MP}) , we get that there is only one limit value of the sequence Π^N which is $\delta_{\mathbb{Q}}$.

7 Examples

7.1 General classes of kernels satisfying Assumption (\mathbf{H}^K)

The first two points of Assumption (\mathbf{H}^K) are simple technical conditions and may not require specific comments. The third assumption is much more interesting. Let us start with the simple example of a kernel K such that ∇K is integrable. Then it is clear that ∇K defines a convolution operator and by a convolution inequality, this operator is bounded in any $L^p(\mathbb{R}^d)$, $p \in [1, +\infty)$. As a consequence of the embedding (3.1), ∇K satisfies (\mathbf{H}_{iii}^K) for any $p \in [1, \infty)$.

Let us now look at a more singular example. We will encounter in the next paragraph the Coulomb potential, defined as $V_C(x) := |x|^{-(d-2)}$, $x \in \mathbb{R}^d$ whenever $d \geq 3$, and $V_C(x) := -\log|x|$, $x \in \mathbb{R}^d$ for $d = 1, 2$. The associated kernel $K_C = \nabla V_C$ is a generalisation in any dimension of the classical Coulomb force, and ∇K_C is not integrable. Nevertheless, it is possible to define the convolution operator of kernel ∇K_C as the Principal Value integral acting on the space of smooth, rapidly decaying functions (i.e. the Schwarz space), thus defining a tempered distribution.

The previous example is a special case of operator defined as a singular integral, which under certain assumptions on the kernel (see the three conditions of [22, Chapter 4.4]) extends to a bounded operator in $L^p(\mathbb{R}^d)$, for any $p \in (1, +\infty)$. In particular, ∇K_C verifies these conditions and therefore K_C satisfies (\mathbf{H}^K) .

7.2 Riesz and Coulomb potentials

The Coulomb potential belongs to a more general class of interaction potentials, called Riesz potentials. These are defined in any dimension as

$$V_s(x) := |x|^{-s}, \quad x \in \mathbb{R}^d, \quad \text{for } s \in (0, d).$$

If $d \geq 3$ and $s = d - 2$, this is the Coulomb potential presented in the previous subsection. We denote the associated kernel by $K_s := \nabla V_s$. K_s satisfies Assumption (\mathbf{H}^K) , provided that $s < d - 1$ (in dimension 1 and 2, one often chooses $V_s(x) = -\log|x|$ instead):

- The parameter κ_M can then be chosen equal to $\frac{d}{s+1}$, which is greater than 1.
- K_s is integrable outside the unit ball for any $q > \frac{d}{s+1}$.
- – If $s < d - 2$, then ∇K_s is not bounded in any L^p but it is however bounded from L^p to L^q whenever $p > 1$ and $\frac{1}{q} = \frac{1}{p} + \frac{s+2}{d} - 1$ (see [37, Theorem 25.2]). This is still enough, thanks to the embedding (3.1), to ensure property (\mathbf{H}_{iii}^K) . In particular, one can choose $D = \max(d, 2)$ for these kernels. Nevertheless we cannot apply Theorem 2.11, which crucially relies on the L^2 boundedness of ∇K . Note however that ∇K_s is locally integrable, hence one could consider a slightly modified kernel \tilde{K}_s by taming K_s at infinity so that $\nabla \tilde{K}_s \in L^1(\mathbb{R}^d)$.
- If $s = d - 2$, then ∇K_s is a typical kernel satisfying the conditions of [22, Chapter 4.4], and therefore it defines a bounded operator in any $L^p(\mathbb{R}^d)$, $p \in (1, \infty)$. Hence all our results apply to this particular kernel. Again, $D = \max(d, 2)$ is a natural choice in this case.

It also follows from the L^p boundedness of ∇K that in Theorem 2.11, κ can be chosen as close as desired to $\alpha((\beta - 1 - \frac{d}{2}) \wedge 1)$. Given condition (\mathbf{H}_{iii}) and if one chooses $\beta = 2 + \frac{d}{2}$, the convergence happens at a rate κ , for κ as close as desired to $\frac{1}{2(d+2)}$.

- If $s \in (d - 2, d - 1)$, then one can verify (see e.g. [14, Lemma 2.5]) that ∇K_s defines a convolution operator from $L^1 \cap \mathcal{C}^\sigma(\mathbb{R}^d)$ to $L^\infty(\mathbb{R}^d)$, with $\sigma \in (2 - d + s, 1)$:

$$\begin{aligned} \|\nabla K_s * f\|_{L^\infty(\mathbb{R}^d)} &\lesssim \|f\|_{L^1(\mathbb{R}^d)} + \|f\|_{L^\infty(\mathbb{R}^d)} + \mathcal{N}_\sigma(f) \\ &\lesssim \|f\|_{L^1 \cap W^{1,D}(\mathbb{R}^d)} + \|\nabla f\|_{L^p(\mathbb{R}^d)}, \end{aligned}$$

where these inequalities follows from Morrey's inequality [7, Theorem 9.12] with $\sigma = 1 - \frac{d}{p}$ and the embedding (3.1) with $D > d$. Hence, if one chooses $p = D$, it follows that ∇K_s satisfies (\mathbf{H}_{iii}^K) . Note that the previous conditions on p and σ imply that one must choose $D > \frac{d}{d-(1+s)}$. We are not able to give a rate of convergence in this case, as ∇K_s is not bounded in L^2 .

- Riesz kernels with $s \geq d - 1$ are out of the scope of the present paper, as K_s is not locally integrable.

Besides obtaining rates of convergence, Proposition 2.14 proves the well-posedness of the McKean-Vlasov SDE (1.5) for all Riesz kernels with $s \in (0, d - 1)$, which seems to be a new result for all the values of the parameter s , and most notably for the largest values $s \geq d - 2$. The uniform propagation of chaos (Theorem 2.15) is also new for this whole class of particle systems.

7.3 Biot-Savart kernel and the 2d Navier-Stokes equation

By considering the vorticity field ξ associated to the incompressible two-dimensional Navier-Stokes solution u , one gets equation (1.1) with the Biot-Savart kernel $K_{BS}(x) = \frac{1}{\pi} \frac{x^\perp}{|x|^2}$, where $x^\perp = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}$. The original Navier-Stokes solution is then recovered thanks to the formula $u_t = K_{BS} * \xi_t$.

The Biot-Savart kernel is an example of repulsive kernel, in the sense that

$$x \cdot K(x) \geq 0, \quad \text{on the domain of definition of } K. \quad (7.1)$$

In this case, the Biot-Savart kernel is merely repulsive since $x \cdot K(x) = 0$.

It is well-known that with such kernel, Eq. (1.1) has a unique global solution, and that $\|K * \xi_t\|_{L^\infty(\mathbb{R}^2)}$ can be bounded by $C(1 + \|\xi_0\|_{L^\infty(\mathbb{R}^2)})$ (see [18] and references therein). This permits to choose the cut-off value A_T from (2.8) independently on T .

The kernel K_{BS} is covered by our assumption (\mathbf{H}^K) , for the same reason as the Coulomb kernel with $d = 2$ ($K_0(x) = \frac{1}{|x|}$). In that case, we recover the Theorem 2 of Flandoli et al. [18] within our Theorem 2.6.

All the other results of this paper apply, and in particular, if the initial condition is smooth enough ($\beta \geq 2 + \frac{d}{2}$), the rate of convergence κ given by Theorem 2.11 is almost $\frac{1}{8}$.

7.4 Parabolic-elliptic Keller-Segel models

Another important example covered by this paper is the Keller-Segel PDE, which takes the form (1.1) with $K_{KS}(x) = -\chi \frac{x}{|x|^d}$, for some $\chi > 0$. This is a model of chemotaxis, but we refer to the companion paper [35] for more information on this equation.

The additional difficulty of this model compared to the aforementioned ones is that the kernel is attractive, in the following sense:

$$x \cdot K(x) < 0, \quad \text{on the domain of definition of } K. \quad (7.2)$$

This leads to important issues that we discuss in more details in [35], but let us just mention as an example that in dimension 2, the PDE admits a global solution if and only if $\chi < 8\pi$ (see e.g. Biler [3] for a recent review). Note that in that case ($d = 2$ and $\chi < 8\pi$) it is again possible to choose a value of the cut-off A_T independently of T (see [35]).

Theorem 2.11 gives a rate of convergence of the particle system to the Keller-Segel PDE, whenever a local or global solution exists, which is, as far as we know, the first time that such quantitative results appear for this class of PDEs (see [6] for related results and Section 2.5 to discuss how this compares to the present work).

As a consequence of Theorem 2.14, we also deduce the weak well-posedness of the McKean-Vlasov (1.5) for all values of the concentration parameter $\chi < 8\pi$, globally in time, which is a new result (see [35] for a thorough discussion and comparison with previous results).

7.5 Attractive-repulsive kernels

There is at least another very interesting class of kernels that enters our framework. The attractive-repulsive kernels are attractive in a region of space, i.e. they satisfy (7.2) on a subdomain D of the domain of definition of K , and repulsive (i.e. satisfying (7.1)) on the complement of D .

The most famous example of such attractive-repulsive kernels might be the Lennard-Jones potential in molecular dynamics: this isotropic potential (i.e. $V(x) \equiv V(|x|)$) reads

$$V(r) = V_0 (r^{-12} - r^{-6}), \quad r > 0,$$

for some $V_0 > 0$. Then $K(x) = \nabla V(x)$ satisfies the first condition of (\mathbf{H}^K) (local integrability) only if the dimension is greater or equal to 14, which may not be of the greatest physical relevance.

A similar, but less singular potential is proposed by Flandoli et al. [17] to model the adhesion of cells in biology. It can be expressed in general as

$$V(r) = V_\alpha r^{-\alpha} - V_\beta r^{-\beta},$$

with $\alpha, \beta > 0$ and $V_\alpha, V_\beta > 0$. One can now refer to the discussion on Riesz kernels in Section 7.2 to determine the values of α and β that ensure the applicability of our results.

Appendix

A.1 Proofs of the tightness estimates

Proof of Proposition 4.4. **Step 1.** Recall that the operators $(I - \Delta)^\beta$, $\beta \in \mathbb{R}$ were defined in the Notations section, see Equation (1.6), with a clear link the Sobolev norm $\|\cdot\|_\beta$. Let $H = L^2(\mathbb{R}^d)$ and let F stand for the function F_A defined in (2.1). From (4.8) after applying $(I - \Delta)^{\beta/2}$ and by the triangular inequality we have

$$\left\| (I - \Delta)^{\beta/2} u_t^N \right\|_{L^p(\Omega; H)} \leq \left\| (I - \Delta)^{\beta/2} e^{t\Delta} u_0^N \right\|_{L^p(\Omega; H)} \quad (\text{A.1})$$

$$+ \int_0^t \left\| (I - \Delta)^{\beta/2} \nabla \cdot e^{(t-s)\Delta} (V^N * (F(K * u_s^N) \mu_s^N)) \right\|_{L^p(\Omega; H)} ds \quad (\text{A.2})$$

$$+ \left\| \frac{1}{N} \sum_{i=1}^N \int_0^t (I - \Delta)^{\beta/2} \nabla e^{(t-s)\Delta} (V^N (X_s^{i,N} - \cdot)) \cdot dW_s^i \right\|_{L^p(\Omega; H)}. \quad (\text{A.3})$$

Step 2. Noticing that by a convolution inequality

$$\left\| (I - \Delta)^{\beta/2} e^{t\Delta} u_0^N \right\|_{L^2(\mathbb{R}^d)} \leq \|e^{t\Delta}\|_{L^2 \rightarrow L^2} \left\| (I - \Delta)^{\beta/2} u_0^N \right\|_{L^2(\mathbb{R}^d)},$$

one gets that the first term (A.1) can be estimated by

$$\left\| (I - \Delta)^{\beta/2} e^{t\Delta} u_0^N \right\|_{L^p(\Omega; H)} \leq \left\| (I - \Delta)^{\beta/2} u_0^N \right\|_{L^p(\Omega; H)} \leq C_\beta,$$

with $C_\beta > 0$, where the boundedness of the norm of u_0^N comes from Assumption (\mathbf{H}_i) .

Step 3. Let us come to the second term (A.2):

$$\begin{aligned} & \int_0^t \left\| (I - \Delta)^{\beta/2} \nabla \cdot e^{(t-s)\Delta} (V^N * (F(K * u_s^N) \mu_s^N)) \right\|_{L^p(\Omega; H)} ds \\ & \leq C \int_0^t \left\| \nabla \cdot e^{\frac{t-s}{2}\Delta} \right\|_{L^2 \rightarrow L^2} \left\| (I - \Delta)^{\beta/2} e^{\frac{t-s}{2}\Delta} (V^N * (F(K * u_s^N) \mu_s^N)) \right\|_{L^p(\Omega; H)} ds. \end{aligned}$$

Observe that for any $x \in \mathbb{R}^d$,

$$|V^N * (F(K * u_s^N) \mu_s^N)(x)| \leq \|F(K * u_s^N)\|_\infty |V^N * \mu_s^N(x)| \leq A |u_s^N(x)|.$$

Now by the positivity of the operator $(I - \Delta)^{\beta/2} e^{\frac{t-s}{2}\Delta}$ (see Lemma 16 in [16]), we have

$$\begin{aligned} & \left\| (I - \Delta)^{\beta/2} e^{\frac{t-s}{2}\Delta} (V^N * (F(K * u_s^N) \mu_s^N)) \right\|_{L^p(\Omega; H)} \\ & \leq C_A \|e^{\frac{t-s}{2}\Delta} (I - \Delta)^{\beta/2} u_s^N\|_{L^p(\Omega; H)} \leq C_A \left\| (I - \Delta)^{\beta/2} u_s^N \right\|_{L^p(\Omega; H)}. \end{aligned}$$

By the combination of the two previous inequalities and (1.7), one gets

$$\begin{aligned} & \int_0^t \left\| (I - \Delta)^{\beta/2} \nabla \cdot e^{(t-s)\Delta} (V^N * (F(K * u_s^N) \mu_s^N)) \right\|_{L^p(\Omega; H)} ds \\ & \leq C_{\beta, A} \int_0^t \frac{1}{\sqrt{t-s}} \left\| (I - \Delta)^{\beta/2} u_s^N \right\|_{L^p(\Omega; H)} ds. \end{aligned}$$

This bounds the second term.

Step 4. For the third term (A.3), we follow the proof of Lemma 11 in [18] which gives, for any $\delta > 0$,

$$\left\| \frac{1}{N} \sum_{i=1}^N \int_0^t (I - \Delta)^{\beta/2} \nabla e^{(t-s)\Delta} (V^N (X_s^{i,N} - \cdot)) \cdot dW_s^i \right\|_{L^p(\Omega; H)}^p \leq C_p \left(\frac{1}{N} \int_0^t \frac{1}{(t-s)^{1-\delta}} \|V^N\|_{\beta+\delta} ds \right)^{\frac{p}{2}}.$$

One can easily verify that there exists a constant $C > 0$ independent of N such that

$$\|V^N\|_{\beta+\delta} \leq CN^{\alpha(2\delta+2\beta+d)}.$$

Hence there exists $C_{\beta,T,p,\delta} > 0$ such that

$$\left\| \frac{1}{N} \sum_{i=1}^N \int_0^t (\mathbf{I} - \Delta)^{\beta/2} \nabla e^{(t-s)\Delta} (V^N (X_s^{i,N} - \cdot)) \cdot dW_s^i \right\|_{L^p(\Omega;H)} \leq C_{\beta,T,p,\delta} N^{\frac{1}{2}(\alpha(d+2\delta+2\beta)-1)}.$$

Therefore, taking $\alpha < \frac{1}{d+2\beta}$ and δ small enough, the last quantity is bounded by some $C_{\beta,T,p}$.

Collecting the three bounds of Steps 2 to 4, we get

$$\left\| (\mathbf{I} - \Delta)^{\beta/2} u_t^N \right\|_{L^p(\Omega;H)} \leq C_{\beta,T,p} + C_{\beta,A} \int_0^t \frac{1}{\sqrt{t-s}} \left\| (\mathbf{I} - \Delta)^{\beta/2} u_s^N \right\|_{L^p(\Omega;H)} ds.$$

We can now apply Grönwall's Lemma to deduce

$$\left\| (\mathbf{I} - \Delta)^{\beta/2} u_t^N \right\|_{L^p(\Omega;H)} \leq C_{\beta,T,A,p}.$$

□

Proof of Proposition 4.5 . Let us now prove the second estimate on u^N given in Proposition 4.5. In this proof we use the fact that $L^2(\mathbb{R}^d) \subset H^{-2}(\mathbb{R}^d)$ with continuous embedding, and that the linear operator Δ is bounded from $L^2(\mathbb{R}^d)$ to $H^{-2}(\mathbb{R}^d)$.

In view of (4.6), we first observe that

$$\begin{aligned} u_t^N(x) - u_s^N(x) &= \int_s^t \langle \mu_r^N, \nabla V^N(\cdot - x) \cdot F(K * u_r^N) \rangle dr \\ &\quad + \int_s^t \Delta u_r^N(x) dr \\ &\quad + \frac{1}{N} \sum_{i=1}^N \int_s^t \nabla V^N(X_r^{i,N} - x) \cdot dW_r^i. \end{aligned}$$

Thus we obtain (below the symbol ‘ \cdot ’ denotes the variable of integration with respect to the measure μ_r^N)

$$\mathbb{E} \left[\|u_t^N - u_s^N\|_{-2}^q \right] \leq (t-s)^{q-1} \int_s^t \mathbb{E} \left[\|\langle \mu_r^N, \nabla V^N(\cdot - \cdot) \cdot F(K * u_r^N) \rangle\|_{-2}^q \right] dr \quad (\text{A.4})$$

$$+ (t-s)^{q-1} \frac{1}{2} \int_s^t \mathbb{E} \left[\|\Delta u_r^N\|_{-2}^q \right] dr \quad (\text{A.5})$$

$$+ \mathbb{E} \left[\left\| \frac{1}{N} \sum_{i=1}^N \int_s^t \nabla V^N(X_r^{i,N} - \cdot) \cdot dW_r^i \right\|_{-2}^q \right]. \quad (\text{A.6})$$

To estimate the first term (A.4), we observe first that

$$\begin{aligned} \mathbb{E} \left[\|\langle \mu_r^N, \nabla V^N(\cdot - \cdot) \cdot F(K * u_r^N) \rangle\|_{-2}^q \right] &= \mathbb{E} \left[\|\nabla(\mu_r^N F(K * u_r^N) * V^N)\|_{-2}^q \right] \\ &\leq \mathbb{E} \left[\|\mu_r^N F(K * u_r^N) * V^N\|_{-1}^q \right] \\ &\leq C_A \mathbb{E} \left[\|u_t^N\|_{L^2(\mathbb{R}^d)}^q \right] \leq C_A. \end{aligned}$$

Moreover, for the second term (A.5) we have

$$\mathbb{E} \left[\|\Delta u_r^N\|_{-2}^q \right] \leq C \mathbb{E} \left[\|u_r^N\|_{L^2(\mathbb{R}^d)}^q \right] \leq C. \quad (\text{A.7})$$

Now, we bound the last term (A.6):

$$\mathbb{E} \left[\left\| \frac{1}{N} \sum_{i=1}^N \int_s^t \nabla V^N (X_r^{i,N} - \cdot) \cdot dW_r^i \right\|_{-2}^q \right] \leq C_q \mathbb{E} \left[\frac{1}{N^2} \sum_{i=1}^N \int_s^t \left\| \nabla V^N (X_r^{i,N} - \cdot) \right\|_{-2}^q dr \right]^{q/2}.$$

Then we have

$$\begin{aligned} & \frac{1}{N^2} \int_{\mathbb{R}^d} \sum_{i=1}^N \int_s^t \left| (\mathbb{I} - \Delta)^{-1} \nabla V^N (X_r^{i,N} - x) \right|^2 dr dx \\ &= (t-s) \frac{1}{N} \|V^N\|_{-1}^2 \leq (t-s) \frac{1}{N} \|V^N\|_0^2 \leq CN^{d\alpha-1}(t-s) \leq C(t-s), \end{aligned}$$

since Assumption (\mathbf{H}_{iii}) implies that $d\alpha - 1 < 0$. In order to conclude the lemma, we need to divide (A.4)–(A.6) by $|t-s|^{1+q\eta}$. From the previous estimates, we always get a term of the form $|t-s|^\rho$ with $\rho < 1$ (using the assumption $\eta < \frac{1}{2}$). \square

A.2 Regularisation property of the kernel

Recall that K satisfies (\mathbf{H}^K) . First, we will make use of the first two points of this assumption. Young's convolution inequality states that for any $q_1 \in \left(\max(1, (1 + \frac{1}{p} - \frac{1}{\kappa_M})^{-1}), p \right]$, $q_2 = (1 + \frac{1}{p} - \frac{1}{q})^{-1}$, and any $f \in L^{q_1} \cap L^{q_2}(\mathbb{R}^d)$,

$$\begin{aligned} \|K * f\|_{L^p(\mathbb{R}^d)} &\leq \|(\mathbb{1}_{\mathcal{B}_1} K) * f\|_{L^p(\mathbb{R}^d)} + \|(\mathbb{1}_{\mathcal{B}_1^c} K) * f\|_{L^p(\mathbb{R}^d)} \\ &\leq \|\mathbb{1}_{\mathcal{B}_1} K\|_{L^{r_1}(\mathbb{R}^d)} \|f\|_{L^{q_1}(\mathbb{R}^d)} + \|\mathbb{1}_{\mathcal{B}_1^c} K\|_{L^q(\mathbb{R}^d)} \|f\|_{L^{q_2}(\mathbb{R}^d)} \\ &\leq C_K (\|f\|_{L^{q_1}(\mathbb{R}^d)} + \|f\|_{L^{q_2}(\mathbb{R}^d)}), \end{aligned}$$

where $r_1 = (1 + \frac{1}{p} - \frac{1}{q_1})^{-1} \in [1, \kappa_M)$, which ensures that C_K is finite.

Hence for q_1, q_2 as above and recalling the property (\mathbf{H}_{iii}^K) of ∇K , one deduces that if $f \in L^1 \cap L^{q_1} \cap L^{q_2} \cap W^{1,D}(\mathbb{R}^d)$, then $K * f \in W^{1,p}(\mathbb{R}^d)$. Hence it follows from Morrey's inequality [7, Th. 9.12] that there exists $C_{p,d} > 0$ such that for any $f \in L^1 \cap L^{q_1} \cap L^{q_2} \cap W^{1,D}(\mathbb{R}^d)$,

$$\begin{aligned} \mathcal{N}_\eta(K * f) &\leq C_{p,d} \|\nabla K * f\|_{L^p(\mathbb{R}^d)} \\ &\leq C_{p,d,K} \|f\|_{L^1 \cap W^{1,D}(\mathbb{R}^d)}, \end{aligned} \tag{A.8}$$

where $\eta = 1 - \frac{d}{p}$.

References

- [1] H. Bahouri, J.-Y. Chemin, and R. Danchin. *Fourier Analysis and Nonlinear Partial Differential Equations*, volume 343. Grundlehren der Mathematischen Wissenschaften, Springer Verlag, 2011.
- [2] M. Bauer, T. Meyer-Brandis, and F. Proske. Strong solutions of mean-field stochastic differential equations with irregular drift. *Electron. J. Probab.*, 23:Paper No. 132, 35, 2018.
- [3] P. Biler. Mathematical challenges in the theory of chemotaxis. *Ann. Math. Sil.*, 32(1):43–63, 2018.
- [4] V. I. Bogachev. *Measure theory. Vol. I*. Springer-Verlag, Berlin, 2007.
- [5] M. Bossy and D. Talay. A stochastic particle method for the McKean-Vlasov and the Burgers equation. *Math. Comp.*, 66(217):157–192, 1997.

- [6] D. Bresch, P.-E. Jabin, and Z. Wang. On mean-field limits and quantitative estimates with a large class of singular kernels: Application to the Patlak–Keller–Segel model. *C. R. Math. Acad. Sci. Paris*, 357(9):708–720, 2019.
- [7] H. Brezis. *Functional analysis, Sobolev Spaces and Partial Differential Equations*. Springer, 2010.
- [8] J. A. Carrillo, M. Fornasier, G. Toscani, and F. Vecil. Particle, kinetic, and hydrodynamic models of swarming. In *Mathematical modeling of collective behavior in socio-economic and life sciences*, Model. Simul. Sci. Eng. Technol., pages 297–336. Birkhäuser Boston, Boston, MA, 2010.
- [9] P. Cattiaux and L. Pédèches. The 2-D stochastic Keller–Segel particle model: existence and uniqueness. *ALEA*, 13:447–463, 2016.
- [10] P. Cattiaux, A. Guillin, and F. Malrieu. Probabilistic approach for granular media equations in the non-uniformly convex case. *Probab. Theory Related Fields*, 140(1-2):19–40, 2008.
- [11] E. Cépa and D. Lépingle. Diffusing particles with electrostatic repulsion. *Probab. Theory Related Fields*, 107(4):429–449, 1997.
- [12] P. E. Chaudru de Raynal. Strong well posedness of McKean–Vlasov stochastic differential equations with Hölder drift. *Stochastic Process. Appl.*, 130(1):79–107, 2020.
- [13] F. Delarue, J. Inglis, S. Rubenthaler, and E. Tanré. Particle systems with a singular mean-field self-excitation. Application to neuronal networks. *Stochastic Process. Appl.*, 125(6):2451–2492, 2015.
- [14] M. Duerinckx. Mean-field limits for some Riesz interaction gradient flows. *SIAM J. Math. Anal.*, 48(3):2269–2300, 2016.
- [15] F. Flandoli and M. Leocata. A particle system approach to aggregation phenomena. *J. Appl. Probab.*, 56(1):282–306, 2019.
- [16] F. Flandoli, M. Leimbach, and C. Olivera. Uniform convergence of proliferating particles to the FKPP equation. *J. Math. Anal. Appl.*, 473(1):27–52, 2019.
- [17] F. Flandoli, M. Leocata, and C. Ricci. On the macroscopic limit of Brownian particles with local interaction. *Preprint [arXiv:2006.01781](https://arxiv.org/abs/2006.01781)*, 2020.
- [18] F. Flandoli, C. Olivera, and M. Simon. Uniform approximation of 2d Navier–Stokes equation by stochastic interacting particle systems. *Preprint [arXiv:2004.00458](https://arxiv.org/abs/2004.00458)*, 2020.
- [19] N. Fournier and M. Hauray. Propagation of chaos for the Landau equation with moderately soft potentials. *Ann. Probab.*, 44(6):3581–3660, 2016.
- [20] N. Fournier and B. Jourdain. Stochastic particle approximation of the Keller–Segel equation and two-dimensional generalization of Bessel processes. *Ann. Appl. Probab.*, 27(5):2807–2861, 2017.
- [21] N. Fournier and S. Mischler. Rate of convergence of the Nanbu particle system for hard potentials and Maxwell molecules. *Ann. Probab.*, 44(1):589–627, 2016.
- [22] L. Grafakos. *Classical Fourier analysis*, volume 249 of *Graduate Texts in Mathematics*. Springer, New York, third edition, 2014.
- [23] X. Huang and F.-Y. Wang. McKean–Vlasov SDEs with Drifts Discontinuous under Wasserstein Distance. *Preprint, [arXiv:2002.06877](https://arxiv.org/abs/2002.06877)*, 2020.
- [24] P.-E. Jabin and Z. Wang. Quantitative estimates of propagation of chaos for stochastic systems with $W^{-1,\infty}$ kernels. *Invent. Math.*, 214(1):523–591, 2018.

- [25] J.-F. Jabir, D. Talay, and M. Tomašević. Mean-field limit of a particle approximation of the one-dimensional parabolic-parabolic Keller-Segel model without smoothing. *Electron. Commun. Probab.*, 23, 2018.
- [26] B. Jourdain and S. Méléard. Propagation of chaos and fluctuations for a moderate model with smooth initial data. *Ann. Inst. H. Poincaré Probab. Statist.*, 34(6):727–766, 1998.
- [27] M. Kac. Foundations of kinetic theory. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, vol. III*, pages 171–197. University of California Press, Berkeley and Los Angeles, 1956.
- [28] C. Marchioro and M. Pulvirenti. Hydrodynamics in two dimensions and vortex theory. *Comm. Math. Phys.*, 84(4):483–503, 1982.
- [29] S. Méléard. Asymptotic behaviour of some interacting particle systems; McKean-Vlasov and Boltzmann models. In *Probabilistic models for nonlinear partial differential equations (Montecatini Terme, 1995)*, volume 1627 of *Lecture Notes in Math.*, pages 42–95. Springer, Berlin, 1996.
- [30] S. Méléard. A trajectorial proof of the vortex method for the two-dimensional Navier-Stokes equation. *Ann. Appl. Probab.*, 10(4):1197–1211, 2000.
- [31] S. Méléard and S. Roelly-Coppoletta. A propagation of chaos result for a system of particles with moderate interaction. *Stochastic Process. Appl.*, 26(2):317–332, 1987.
- [32] T. Nagai. Global existence and decay estimates of solutions to a parabolic-elliptic system of drift-diffusion type in \mathbb{R}^2 . *Differential Integral Equations*, 24(1–2):29–68, 2011.
- [33] K. Oelschläger. A law of large numbers for moderately interacting diffusion processes. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 69(2):279–322, 1985.
- [34] K. Oelschläger. A fluctuation theorem for moderately interacting diffusion processes. *Probab. Theory Related Fields*, 74(4):591–616, 1987.
- [35] C. Olivera, A. Richard, and M. Tomasevic. Particle approximation of the 2-d parabolic-elliptic Keller-Segel system in the subcritical regime. *Preprint arXiv:2004.03177*, 2020.
- [36] H. Osada. Propagation of chaos for the two-dimensional Navier-Stokes equation. *Proc. Japan Acad. Ser. A Math. Sci.*, 62(1):8–11, 1986.
- [37] S. G. Samko, A. A. Kilbas, and O. I. Marichev. *Fractional integrals and derivatives*. Gordon and Breach Science Publishers, Yverdon, 1993. Theory and applications, Edited and with a foreword by S. M. Nikolskii, Translated from the 1987 Russian original, Revised by the authors.
- [38] S. Serfaty. Mean field limit for Coulomb-type flows. *To appear in Duke Math. J.*, *arXiv:1803.08345*, 2020.
- [39] M. Simon and C. Olivera. Non-local conservation law from stochastic particle systems. *J. Dynam. Differential Equations*, 30(4):1661–1682, 2018.
- [40] A.-S. Sznitman. A propagation of chaos result for Burgers’ equation. *Probab. Theory Relat. Fields*, 71(4):581–613, 1986.
- [41] A.-S. Sznitman. Topics in propagation of chaos. In *École d’Été de Probabilités de Saint-Flour XIX—1989*, volume 1464 of *Lecture Notes in Math.*, pages 165–251. Springer, Berlin, 1991.
- [42] D. Talay and M. Tomašević. A new McKean–Vlasov stochastic interpretation of the parabolic-parabolic Keller–Segel model: The one-dimensional case. *Bernoulli*, 26(2):1323–1353, 2020.

- [43] C. M. Topaz, A. L. Bertozzi, and M. A. Lewis. A nonlocal continuum model for biological aggregation. *Bull. Math. Biol.*, 68(7):1601–1623, 2006.
- [44] H. Triebel. *Interpolation theory, function spaces, differential operators*, volume 18 of *North-Holland Mathematical Library*. North-Holland Publishing Co., Amsterdam-New York, 1978.
- [45] M. Veraar. The stochastic Fubini theorem revisited. *Stochastics*, 84(4):543–551, 2012.