Quantitative particle approximation of nonlinear Fokker-Planck equations with singular kernel

Christian Olivera, Alexandre Richard, Milica Tomasevic

Quantitative particle approximation of nonlinear Fokker-Planck equations with singular kernel

Christian Olivera * Alexandre Richard† Milica Tomašević ‡

December 29, 2020

Abstract

In this work, we obtain quantitative convergence of moderately interacting particle systems towards solutions of nonlinear Fokker-Planck equations with singular kernels. In addition, we prove the well-posedness for the McKean-Vlasov SDEs involving these singular kernels and the trajectorial propagation of chaos for the associated moderately interacting particle systems.

Our results only require very weak regularity on the interaction kernel, including the Biot-Savart kernel, and attractive kernels such as Riesz and Keller-Segel kernels in arbitrary dimension. This seems to be the first time that such quantitative convergence results are obtained in Lebesgue and Sobolev norms for the aforementioned kernels. In particular, this convergence still holds (locally in time) for PDEs exhibiting a blow-up in finite time. The proofs are based on a semigroup approach combined with a fine analysis of the Sobolev regularity of infinite-dimensional stochastic convolution integrals, and we also exploit the regularity of the solutions of the limiting equation.

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

MSC2020 subject classification: 60K35, 60H30, 35K55, 35Q84.

*Departamento de Matemática, Universidade Estadual de Campinas, Brazil. colivera@ime.unicamp.br.
†Université Paris-Saclay, CentraleSupélec, MICS and CNRS FR-3487, France.
alexandre.richard@centralesupelec.fr.
‡CMAP, Ecole Polytechnique, CNRS, I.P. Paris, 91128 Palaiseau, France.
milica.tomasevic@polytechnique.edu.
C.O. is partially supported by CNPq through the grant 426747/2018-6. C.O. and A.R. acknowledge the support of the Math-Amsud program, through the grant 19-MATH-06. M.T. acknowledges the support of FMJH.

1 Introduction and main results

In this work, we are interested in the stochastic approximation of nonlinear Fokker-Planck Partial Differential Equations (PDEs) of the form

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

by means of moderately interacting particle systems. Here, K is a locally integrable kernel that may have a singular behaviour at 0.

Although we will not limit ourselves to kernels that derive from a potential, a typical family of singular kernels that we will consider in this work derives from Riesz potentials, defined in any dimension d as

$$V_s(x) := \begin{cases} |x|^{-s} & \text{if } s\in(0,d) \\ -\log|x| & \text{if } s=0 \end{cases}, \quad x\in\mathbb{R}^d. \tag{1.2}$$

The associated kernel is then $K_s := \pm\nabla V_s$, the sign deciding whether the interaction is attractive or repulsive. If $d\geq 2$ and $s=d-2$, this is the Coulomb potential that characterises electrostatic and gravitational forces (depending on the sign). Our results will also cover several classical models such as the $2d$ Navier-Stokes equation, which in vorticity form can be written as in (1.1) with the Biot-Savart kernel, and the parabolic-elliptic Keller-Segel PDE in any dimension $d\geq 1$, which models the phenomenon of chemotaxis. These kernels are presented and discussed in more details in Section 5.

The problem of deriving a macroscopic equation from a microscopic model of interacting particles can be traced back to the original inspiration of Kac [33], in the context of the Boltzmann equation. Since then, a huge literature has been devoted to interacting particle systems and their convergence to nonlinear Fokker-Planck equations such as (1.1). In the case of Lipschitz continuous interaction kernels, this problem is now well-understood. When the particles are diffusion processes, whether in mean-field or moderate interaction, one can refer e.g. to the works of Sznitman [54], Méléard [37] for a general account of the theory, and Oelschläger [42] and Méléard and Roelly-Coppoletta [39] on the moderate case. In this case, one can show that the propagation of chaos holds. That is to say, the empirical measure (on the space of trajectories) of the associated particle system converges in law towards the weak solution of a McKean-Vlasov SDE associated to (1.1). In particular, when one looks at the time marginal laws in this convergence, one recovers in the limit the Fokker-Planck equation.

There are fewer works when the interaction kernel is singular, despite the great importance it represents both theoretically and in applications. One can mention the early works of Marchioro and Pulvirenti [36] and Osada [45] on the $2d$ Navier-Stokes equation, and of Sznitman [53] and Bossy and Talay [5] on Burgers' equation. Cépa and Lépingle [10] studied one-dimensional electrical particles with repulsive interaction, and more recently Fournier and Hauray [21] 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. [14] (diffusive particles interacting through their hitting times), and several works on the Keller-Segel equation (Fournier and Jourdain [22], Cattiaux and Pédèches [8] and Jabir et al. [30] for instance).

When the interaction is attractive in addition to being singular (as it is the case in the Keller-Segel model), it may not be possible to define the particle system in mean-field interaction, and therefore to obtain propagation of chaos. Even when it is possible to define the particles, the propagation of chaos may not always hold. For example, in the tricky case of the $2d$ parabolic-elliptic Keller-Segel model, the mean-field particle system was shown to be well-defined [8, 22], but the convergence (on the level of measures on the space of trajectories) holds only for small values of the critical parameter of the equation (see [22]). For the d-dimensional Keller-Segel model with $d \geq 3$ and the attractive Riesz kernel with $s \in (d-2, d)$, we are not aware of any existence result for the mean-field particle system.

This work is motivated by the approximation of the PDE (1.1), including cases when the kernel is attractive and singular, and the associated mean-field interacting particle system is not known to be well-defined. We thus consider the following moderately interacting particle system:

$$dX_t^{i,N} = F_A\left(\frac{1}{N}\sum_{k=1}^{N}(K * V^N)(X_t^{i,N} - X_t^{k,N})\right) dt + \sqrt{2}\, dW_t^i, \quad t \leq T, \quad 1 \leq i \leq N, \tag{1.3}$$

where $(W^i)_{1\leq i\leq N}$ are independant standard Brownian motions, V^N is a mollifier and F_A is a smooth cut-off function. Hence, the existence of strong solutions for (1.3) is ensured.
However, the difficulty is to prove that at the limit $N \to \infty$, the system (1.3) behaves as the mean-field system would. In this work, we will show that:

(a) $\{\mu_t^N = \sum_{i=1}^N \delta_{X_t^{i,N}},\, t \in [0,T]\}$, the marginals of the empirical measure of (1.3), converge to the solution of the PDE (1.1) with a rate close to $\frac{1}{2(d+1)}$.

(b) The system (1.3) propagates chaos towards the following nonlinear equation (without the cut-off and the mollifier):

$$\begin{cases} dX_t = K * u_t(X_t)\, dt + \sqrt{2} dW_t, \quad t \leq T, \\ \mathcal{L}(X_t) = u_t,\ \mathcal{L}(X_0) = u_0. \end{cases} \tag{1.4}$$

The propagation of chaos implies that for N large, the particles in (1.3) behave approximately like independent copies of the process (1.4), which represents the evolution of a typical particle in an infinite system undergoing the dynamics prescribed by the Fokker-Planck equation.

Notice that it is not *a priori* clear that (1.4) is well-posed due to the singularity of K. Hence, in order to obtain this convergence, we obtain first a well-posedness result for (1.4).

As an application, we plan in a future work to use the rate in (a) and a time discretization of (1.3) to propose a numerical approximation of the PDE (1.1).

1.1 Notations and definitions.

- For any $\beta \in \mathbb{R}$ and $p \geq 1$, we denote by $H_p^\beta(\mathbb{R}^d)$ the *Bessel potential space*

$$H_p^\beta(\mathbb{R}^d) := \Big\{u \text{ tempered distribution; } \mathcal{F}^{-1}\Big(\big(1 + |\cdot|^2\big)^{\frac{\beta}{2}}\, \mathcal{F}u(\cdot)\Big) \in L^p(\mathbb{R}^d)\Big\},$$

where $\mathcal{F}u$ denotes the *Fourier transform* of u. This space is endowed with the norm

$$\|u\|_{\beta,p} = \left\|\mathcal{F}^{-1}\Big(\big(1+|\cdot|^2\big)^{\frac{\beta}{2}}\ \mathcal{F}u(\cdot)\Big)\right\|_{L^p(\mathbb{R}^d)}.$$

In particular, note that

$$\|u\|_{0,p} = \|u\|_{L^p(\mathbb{R}^d)} \quad \text{and for any } \beta \leq \gamma, \quad \|u\|_{\beta,p} \leq \|u\|_{\gamma,p}.$$

The space $H_p^\beta(\mathbb{R}^d)$ is associated to the fractional operator $(I-\Delta)^{\frac{\beta}{2}}$ defined as (see e.g. [57, p.180] for more details on this operator):

$$(I-\Delta)^{\frac{\beta}{2}} f := \mathcal{F}^{-1}\left((1+|\cdot|^2)^{\frac{\beta}{2}}\mathcal{F}f\right). \tag{1.5}$$

For $p=2$, these are Hilbert spaces when endowed with the scalar product

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

and the norm simply denoted by $\|u\|_\beta := \|u\|_{\beta,2} = \sqrt{\langle u,u\rangle_\beta}$.

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

$$\left(e^{t\Delta} f\right)(x) = \int_{\mathbb{R}^d} g_{2t}(x-y)\, f\,(y)\, dy,$$

where g denotes the usual d-dimensional Gaussian density function:

$$g_{\sigma^2}(x) = \frac{1}{(2\pi\sigma^2)^{\frac{d}{2}}} e^{-\frac{|x|^2}{2\sigma^2}}. \tag{1.6}$$

- For $\mathcal{X}$ some normed vector space, the space $\mathcal{C}(I;\mathcal{X})$ of continuous functions from the compact time interval I with values in $\mathcal{X}$ is classically endowed with the norm

$$\|f\|_{I,\mathcal{X}} = \sup_{s\in I} \|f_s\|_{\mathcal{X}}.$$

In case $I=[0,t]$ for some $t>0$, we will also use the notation $\|f\|_{t,\mathcal{X}} = \|f\|_{[0,t],\mathcal{X}}$.

- 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 $x \mapsto u(t,x)$.

1.2 Framework and assumptions

We start this section with a precise definition of the particle system (1.3).

Let $A>0$ and let F_A be defined as follows: let $f_A : \mathbb{R}\to\mathbb{R}$ 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 $x > A+1$ and $f_A(x) = -A$, 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

$$F_A : (x_1,\ldots,x_d)^T \mapsto (f_A(x_1),\ldots,f_A(x_d))^T. \tag{1.7}$$

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

$$V^N(x) := N^{d\alpha} V(N^\alpha x), \qquad \text{for some } \alpha\in[0,1]. \tag{1.8}$$

Below, α will be restricted to some interval $(0, \alpha_0)$, see Assumption $(\mathbf{A}_\alpha)$.

Let $T > 0$. For each $N \in \mathbb{N}$, the particle system (1.3) reads more precisely:

$$\begin{cases} dX_t^{i,N} = F_A\Big(\frac{1}{N}\sum_{k=1}^N (K * V^N)(X_t^{i,N} - X_t^{k,N})\Big)\, dt + \sqrt{2}\, dW_t^i, \quad t \le T,\ 1 \le i \le N, \\ X_0^{i,N},\ 1 \le i \le N, \quad \text{are independent of } \{W^i,\ 1 \le i \le N\}, \end{cases} \tag{1.9}$$

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

$$\mu_\cdot^N = \frac{1}{N}\sum_{i=1}^N \delta_{X_\cdot^{i,N}}\ , \tag{1.10}$$

and the mollified empirical measure by

$$u_\cdot^N := V^N * \mu_\cdot^N\ .$$

The following properties of the kernel will be assumed:

$(\mathbf{A}^K)$:

$(\mathbf{A}_i^K)$ $K \in L^1(\mathcal{B}_1)$;

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

$(\mathbf{A}_{iii}^K)$ There exists $\boldsymbol{r} \in (d \vee 2, +\infty)$, $\beta \in (\frac{d}{\boldsymbol{r}}, 1)$, $\boldsymbol{\zeta} \in (0, 1]$ and $C > 0$ such that for any $f \in L^1 \cap H_{\boldsymbol{r}}^\beta(\mathbb{R}^d)$, one has

$$\mathcal{N}_{\boldsymbol{\zeta}}(K * f) \le C\|f\|_{L^1\cap H_{\boldsymbol{r}}^\beta(\mathbb{R}^d)}.$$

The Assumption $(\mathbf{A}^K)$ 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 5, but as mentioned at the beginning of 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 $s < d - 1$, whether repulsive or attractive, satisfy $(\mathbf{A}^K)$.

Note that Assumption $(\mathbf{A}_{iii}^K)$ is rather mild, and we provide in Section 5.1 a sufficient condition which is easier to check in the examples.

Let us now state the assumptions on the initial conditions of the system:

$(\mathbf{A})$:

$(\mathbf{A}_i)$ For any $m \ge 1$,

$$\sup_{N\in\mathbb{N}} \mathbb{E}\left[\left\|\mu_0^N * V^N\right\|_{\beta,\boldsymbol{r}}^m\right] < \infty.$$

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

For instance, a sufficient condition for $(\mathbf{A})$ to hold is that particles are initially i.i.d. with a density that is smooth enough (see [16, Lemma 2.9] for a related result). The reader may also find interesting comments on this assumption in [20, Remark 1.2].

Finally, the restriction with respect to the key parameters in this setting is given by the following assumption:

($\mathbf{A}_\alpha$): The parameters α, β and $\boldsymbol{r}$ (which appear respectively in (1.8) and ($\mathbf{A}^K_{iii}$)) satisfy

$$0 < \alpha < \frac{1}{d + 2\beta + 2d(\frac{1}{2} - \frac{1}{\boldsymbol{r}})}.$$

We aim to prove the convergence of the mollified empirical measure to the PDE (1.1). As (1.1) preserves the total mass $M := \int_{\mathbb{R}^d} u_0(x)\, dx$, we will assume throughout the paper that $M = 1$.

Solutions to (1.1) will be understood in the following mild sense:

Definition 1.1. *Given $u_0 \in L^1 \cap L^\infty(\mathbb{R}^d)$ and $T > 0$, a function u on $[0, T] \times \mathbb{R}^d$ is said to be a mild solution to* (1.1) *on $[0, T]$ if*

(i) $u \in \mathcal{C}([0, T]; L^1 \cap L^\infty(\mathbb{R}^d))$;

(ii) u satisfies the integral equation

$$u_t = e^{t\Delta} u_0 - \int_0^t \nabla \cdot e^{(t-s)\Delta}(u_s\, K * u_s)\, ds, \quad 0 \le t \le T. \tag{1.11}$$

A function u on $[0, \infty) \times \mathbb{R}^d$ is said to be a global mild solution to (1.1) *if it is a mild solution to* (1.1) *on $[0, T]$ for all $T > 0$.*

In Section 2.2 the following result about the (local) well-posedness of the PDE (1.1) is established:

Proposition 1.2. *Assume that the kernel $K : \mathbb{R}^d \to \mathbb{R}^d$ satisfies ($\mathbf{A}^K_i$) and ($\mathbf{A}^K_{ii}$) and that the initial condition is $u_0 \in L^1 \cap L^\infty(\mathbb{R}^d)$. Then there exists $T > 0$ such that the PDE* (1.1) *admits a mild solution u in the sense of Definition 1.1 on $[0, T]$. In addition, this mild solution is unique.*

Remark 1.3. *In Corollary 2.3, we observe that if T is small enough, then $\|u\|_{T, L^1 \cap L^\infty(\mathbb{R}^d)}$ is controlled by $\|u_0\|_{L^1 \cap L^\infty(\mathbb{R}^d)}$, 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), it is possible to define a global cut-off that depends explicitly on the initial condition (see [44] for more details).*

From here we can explicit the cut-off A_T used to define the particle system. For a local mild solution u on $[0, T]$, we will use the following cut-off:

$$A_T := C_{K,d}\, \|u\|_{T, L^1 \cap L^\infty(\mathbb{R}^d)}, \tag{1.12}$$

where $C_{K,d}$ depends only on K and d and is given in Lemma 2.1.

Finally, as the kernel K may be singular, the PDEs we are interested in may only have local in time solutions (i.e. they may explode in finite time). For this reason, we will denote by T_{max} the maximal time of existence of a solution to (1.1) in the sense of Definition 1.1. This means that for any $T < T_{max}$, the PDE admits a mild solution on $[0, T]$. If there exists a global mild solution to our PDE, then $T_{max} = \infty$.

1.3 First main result: Rate of convergence to the PDE

Considering moderately interacting particles such as (1.3) dates back to Oelschläger [42], and was further developed by Méléard and Roelly-Coppoletta [39] who proved a propagation of chaos result for regular coefficients. Then let us mention Bossy and Talay [5] who studied Burgers equation, and Méléard [38] who studied the $2d$ Navier-Stokes equation. Jourdain and Méléard [32] studied the case where the nonlinearity may be as well in the diffusion coefficient, Jourdain [31] studied a nonlinear convection-diffusion equation with drift that is possibly unbounded, and Fournier and Hauray [21] obtained a rate of convergence for the intricate Landau equation, which is both singular and with interaction in the diffusion term (hence not of the form (1.3)).

More recently, the new semigroup approach developed by Flandoli et al. [17] allows one to approximate nonlinear PDEs by smoothed empirical measures in strong functional topologies. More precisely, the convergence of the mollified empirical measure of the moderately interacting particle system is obtained, and the approach was initially proposed for the FKPP equations. It has already found many applications: see Flandoli and Leocata [16] for a PDE-ODE system related to aggregation phenomena; Simon and Olivera [52] for non-local conservation laws; Flandoli et al. [20] for the $2d$ Navier-Stokes equation; and Olivera et al. [44] for the parabolic-elliptic Keller-Segel systems. However, in these recent works this convergence was not quantified and the propagation of chaos was not considered.

The general question of quantifying the convergence of interacting particle systems (whether in mean field or moderate interaction) towards the PDE in non-singular framework has been addressed thoroughly in the literature. See for example [32, 43, 54] or, more recently, Cortez and Fontbona [13] in the case of homogeneous Boltzmann equation for Maxwell molecules. However, in the singular case there are fewer results in the literature: Méléard [38] obtained a non-explicit rate on the density of one particle for the $2d$ Navier-Stokes equation, Bossy and Talay [5] got a rate for Burger's equation, Fournier and Mischler [23] obtained a rate for the so-called Nanbu particle system approaching the Boltzmann equation, and Fournier and Hauray [21] approximated the Landau equation with moderately soft potentials. Working at the level of the entropy solution of the Liouville equations associated to the mean field particle system, Jabin and Wang [29] obtained a quantitative convergence with a $N^{-1/2}$ rate for some kernels including the Biot-Savart kernel. Recently, Bresch et al. [6] also proved a rate of convergence between the entropy solutions of the Liouville equations and the PDE for the tricky case of the $2d$ Keller-Segel model. Let us finally mention the related problem of singular non-diffusive systems (i.e. with deterministic particles), which has seen significant progress recently with the work of Serfaty [51], who introduced the modulated energy method (see also [15] and [48]).

Our first main result is the following claim, whose proof is detailed in Section 3.2:

Theorem 1.4. *Assume that the initial conditions $\{u_0^N\}_{N\in\mathbb{N}}$ satisfy $(\mathbf{A})$ and that the kernel K satisfies $(\mathbf{A}^K)$. Moreover, let $(\mathbf{A}_\alpha)$ hold true. If $d=1$, assume further that $\beta \in (\frac{1}{2}+\frac{1}{\boldsymbol{r}}, 1)$.*
Let T_{max} be the maximal existence time for (1.1) and fix $T \in (0, T_{max})$. Then let the dynamics of the particle system be given by (1.9) with A greater than A_T.

Then, for any $\varepsilon > 0$ and any $m \geq 1$, there exists a constant $C > 0$ such that for all $N \in \mathbb{N}^*$,

$$\left\| \|u^N - u\|_{T, L^1\cap L^\infty(\mathbb{R}^d)} \right\|_{L^m(\Omega)} \leq \left\| \sup_{s\in[0,T]} \|e^{s\Delta}(u_0^N - u_0)\|_{L^1\cap L^\infty(\mathbb{R}^d)} \right\|_{L^m(\Omega)} + C N^{-\varrho+\varepsilon},$$

where

$$\varrho = \min\left(\alpha\boldsymbol{\zeta},\, \alpha(\beta - \frac{d}{\boldsymbol{r}}),\, \frac{1}{2}\left(1 - \alpha(d + [2\vee d])\right)\right).$$

Whether in mean-field or in moderate interaction, we are not aware of any comparable quantitative result on the empirical measure (or its mollified version) for the Biot-Savart, Riesz or Keller-Segel kernels. Note also that we do not necessarily require that our particles are initially i.i.d.

Several consequences of Theorem 1.4 are presented in Section 3.3: In Corollary 3.2, we obtain a similar rate of convergence in H_r^γ norm, with $\gamma < \beta$ and $r < \boldsymbol{r}$, and in Hölder norm; in Corollary 3.4, the rate in the previous results is shown to hold almost surely; and finally in Corollary 3.6, we obtain the same rate for the genuine empirical measure in a weak topology.

Moreover, we obtain the convergence in the limit case $H_{\boldsymbol{r}}^\beta$, but without a rate of convergence. Namely, we get the following corollary:

Corollary 1.5. *Let the same assumptions as in Theorem 1.4 hold. Then the sequence of mollified empirical measures $\{u_t^N,\ t\in[0,T]\}_{N\in\mathbb{N}}$ converges in probability, as $N\to\infty$, towards the unique* *mild solution u on $[0,T]$ of the PDE* (1.1)*, in the following sense:*

$$\forall \varphi \in L^2\left([0,T]; H^{-\beta}_{\boldsymbol{r'}}(\mathbb{R}^d)\right), \quad \int_0^T \langle u^N_t, \varphi_t\rangle_\beta \, dt \xrightarrow{\mathbb{P}} \int_0^T \langle u_t, \varphi_t\rangle_\beta \, dt,$$

where $\boldsymbol{r'} = \frac{\boldsymbol{r}}{\boldsymbol{r}-1}$ is the conjugate exponent of $\boldsymbol{r}$.

The proof is given in Section 3.6. Theorem 1.4 and Corollary 1.5 cover the convergence results for the mollified empirical measures obtained in a case by case basis in [20, 44]. We present here an alternative approach to the one presented initially in [17] to prove such convergence which, in addition, enables us to quantify it. We believe that this new approach could lead to rates of convergence in slightly different models such as [16, 17, 18].

The techniques used to prove Theorem 1.4 are a mix between analysis and stochastic calculus. On the analytical side, our approach relies on the regularisation property of the heat semigroup and of the convolution operator associated to K, which decides the functional framework of the whole study (i.e. the appropriate Bessel and fractional Sobolev spaces). On the probabilistic side, we write the SPDE satisfied by the mollified empirical measure u^N (see (3.3)), derive some *a priori* estimates on u^N and compare it to the mild solution of (1.1). This yields an expansion which involves several convolutions with singular terms, controlled with analytical techniques mentioned above, and a stochastic convolution integral. This integral is tricky to handle, as we need to get simultaneously its time regularity and its decay as $N \to \infty$, in all L^p norms (including $p = 1$) and Bessel norms with $\gamma \geq 0$ (see Appendix A.1). In particular, we use stochastic integration techniques in infinite-dimensional spaces [58] and Garsia-Rodemich-Rumsey's lemma to control the moments of $\sup_{t\in[0,T]} \|u^N_t - u_t\|$.

Remark 1.6. *Unlike particle systems in mean field interaction, we cannot expect here a $\sqrt{N}$ rate of convergence, which is the case when the interactions are smooth [9]. This is 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 $N^{-\alpha}$ rate of convergence. See [43] for a thorough discussion on this matter.*

In the light of the above remark, let us discuss the rate of convergence obtained in Theorem 1.4.

For the Biot-Savart, Riesz and Keller-Segel kernels, the parameters $\boldsymbol{r}, \beta$ and $\boldsymbol{\zeta}$ can be chosen in a way that ensures the rate in Theorem 1.4 is $\varrho = \alpha(1-\epsilon)$, for any $\epsilon > 0$ (this is detailed for each kernel in Section 5). In view of the previous remark, we cannot expect a better rate. However, the assumption ($\mathbf{A}_\alpha$) imposes a restriction on the possible values of α, and yields a rate

$$\varrho = \frac{1-\epsilon}{2(d+1)},$$

for $\epsilon > 0$ as small as desired, provided the initial condition is smooth enough (see Remark 3.8 for more details).

In particular, let us consider again the example of the $2d$ Keller-Segel model (see kernel (5.1)). We recall that the PDE has a global solution whenever the critical parameter χ satisfies $\chi < 8\pi$, and explodes in finite time otherwise (see [41]). In Theorem 1.4, we get a rate for any value of χ which is almost $\frac{1}{2(d+1)}$ if the law of the initial condition is smooth enough. This result holds even even if the PDE explodes ($\chi > 8\pi$) in finite time. In that case one works on $[0,T]$ for any $T < T_{\max}$.

1.4 Second main result: Propagation of chaos

We now tackle the question of propagation of chaos of (1.3) towards (1.4). Firstly, assuming that u_0 has enough regularity ($u_0 \in L^1 \cap L^\infty$), we prove the weak well-posedness of (1.4) in Proposition 1.8. Secondly, we prove the propagation of chaos of the empirical measure of the particle system (1.3) towards the law of this nonlinear SDE, at the level of probability measures on the space of trajectories (Theorem 1.9).

Recently, the well-posedness of McKean-Vlasov SDEs (or distribution-dependent SDEs) has gained much attention in the literature (see e.g. [2, 11, 12, 28, 35, 40, 47] and the references therein). The authors analyse well-posedness when the diffusion coefficient is also distribution-dependent and when the dependence on the law is not necessarily as in (1.4). The main difficulties there are to treat coefficients that may not be continuous (in the measure variable) w.r.t the Wasserstein distance and eventually to treat singular drift coefficients (in the space variable). As both of these difficulties appear in the specific distribution dependence in (1.4), our well-posedness result gives a new perspective on the matter. In some particular cases such as the Keller-Segel model, this extends previous results [22, 47], but it requires more regularity on the initial data.

To ensure that (1.4) admits a unique weak solution, 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 (1.4). Hence, consider the following nonlinear martingale problem related to (1.4):

Definition 1.7. *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 $q \in \mathcal{C}([0,T]; L^1 \cap L^\infty(\mathbb{R}^d))$;

(iii) For any $f \in \mathcal{C}^2_c(\mathbb{R}^d)$, the process $(M_t)_{t\in[0,T]}$ defined as

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

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 1.8. *Let $T < T_{max}$. Assume that u_0 is a probability density function belonging to $L^\infty(\mathbb{R}^d)$ and that the kernel K satisfies ($\mathbf{A}^K$). Then, the martingale problem ($\mathcal{MP}$) admits a unique solution in the sense of Definition 1.7.*

This result comes from the combination of the fact that the marginal laws of the process are uniquely determined as solutions of (1.1) with the fact that the linearised version of (1.4) admits a unique weak solution. This is one of the classical ways to prove the well-posedness of such distribution-dependent SDEs, along with the frequently used fixed point argument and convergence of the empirical measure argument. We choose this approach as we have *a priori* information about the PDE. We remark here that the weak solution is local in time if there is an explosion in finite time in the corresponding PDE, and global if the corresponding PDE is globally well-posed.

When it comes to our last result about the propagation of chaos, one important remark is in order. Usually, when dealing with singular nonlinear PDEs, the question of propagation of chaos of the related particle system is very demanding as it may happen (due to singular interaction) that it is not even possible to define the particle system. That is why, heavy techniques are used on a case by case basis (see e.g. [22]). Passing to the framework of a particle system in moderate interaction, we exhibit a powerful tool to approximate singular PDEs that circumvents the difficulty of well-posedness of the mean-field particle system and we prove the propagation of chaos of the moderately interacting particle system in very general and singular framework (even when there is an explosion in the associated PDE). This is particularly useful for numerical applications that we plan to tackle in a future work.

We present here our second main result about the propagation of chaos of the particle system. It will be proven in Section 4.2.

Theorem 1.9. *Let the hypotheses of Theorem 1.4 hold and assume further that the family of random variables* $\{X_0^i,\ i \in \mathbb{N}\}$ *is identically distributed. Then, the empirical measure* $\mu_\cdot^N$ *(defined in* (1.10) *as a measure on* $\mathcal{C}([0,T];\mathbb{R}^d)$*) converges in law towards the unique weak solution of* (1.4).

To obtain this result, we will first prove that the empirical measure on the space of trajectories converges to the law of the cut-off version of the SDE (1.4) (where the drift is $F_A(K * u_t)$). This is obtained in the spirit of the work of Méléard and Roelly-Coppoletta [39], that we extend to singular kernels. What allows us to do so, is the result about the convergence of the mollified empirical measure of Theorem 1.4 and its consequences. Then, choosing the parameter A conveniently, we are able to lift the cut-off and conclude the convergence of the empirical measure towards the law of (1.4).

We emphasize here that without the convergence of u^N in the convenient functional framework, it would not be possible to obtain the propagation of chaos in this singular setting. Hence, the result of Theorem 1.4 is very much related to the propagation of chaos and should be considered as the most important ingredient when proving the latter.

Remark 1.10. • *Let us focus on the Keller-Segel model, for which the kernel is given in* (5.1).

In Fournier and Jourdain [22], the well-posedness of the associated McKean-Vlasov SDE for a value of the sensitivity parameter $\chi < 2\pi$ *in dimension* 2 *is proven. The authors also proved tightness and consistency result for the associated particle system (one cannot properly speak about propagation of chaos as the PDE might not have a unique solution in their functional framework). Our result provides, assuming more regularity on the initial condition, global (in time) well-posedness of the McKean-Vlasov SDE whenever the PDE has a global solution, and local well-posedness whenever the solution of the PDE exhibits a blow-up. In particular in dimension* 2*, we get the global well-posedness of* (1.4) *whenever* $\chi < 8\pi$*, and local well-posedness of* (1.4) *when* $\chi \geq 8\pi$*. In both cases, we obtain the propagation of chaos for the moderately interacting particle system.*

• *Another singular drift given in the literature is the one in [47], where (roughly speaking) the interaction kernel is only integrable in the* $L^q([0,T];L^p(\mathbb{R}^d))$ *space where* $\frac{d}{p} + \frac{2}{q} < 1$*. The propagation of chaos in the mean-field case has been treated in [27] and in [56]. A typical example of such interaction given in [47] is of order* $\frac{1}{|x|^r}$ *where* $r \in [0,1)$*. A kernel with this kind of irregularity will satisfy* ($\mathbf{A}^K$)*. Actually, we can treat here the more singular cases of kernels* K *of the order* $\frac{1}{|x|^{s+1}}$ *for* $s \in (0, d-1)$*, in any dimension. However, in our framework one needs to be more flexible with the initial condition and our drift is not time dependent (though our techniques could support time dependence in the drift up to the order of magnitude that could allow us to use the Grönwall lemma when needed, we preferred not to introduce such a technicality in our computations).*

1.5 Organisation of the paper

The existence and uniqueness of the Fokker-Planck equation (1.1) and its cut-off version are studied in Section 2. Then in Section 3, we prove Theorem 1.4 and state its corollaries. In Section 4, the existence and uniqueness of the martingale problem associated to the cut-off McKean-Vlasov SDE are proven (Proposition 1.8), as well as the propagation of chaos for the empirical measure of (1.3) (Theorem 1.9). Finally, we present some examples and applications of our results in Section 5. In the Appendix, one may find the time and space estimates for some stochastic convolution integrals and the proof of a result about the boundedness of the mollified empirical measure in $L^2\left([0,T];H_{\boldsymbol{r}}^{\beta}(\mathbb{R}^d)\right)$.

1.6 Additional notations

- 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 a function.

- 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

$$\left\| \nabla e^{t\Delta} \right\|_{L^p \to L^p} \leq \frac{C}{\sqrt{t}}. \tag{1.13}$$

  By explicit computations in the Fourier space, we get that

$$\|g_{2t}\|_{\beta,1} = \left\| (I - \Delta)^{\frac{\beta}{2}} \frac{1}{(4\pi t)^{\frac{d}{2}}} e^{-\frac{|\cdot|^2}{4t}} \right\|_{L^1(\mathbb{R}^d)} \leq C\, t^{-\frac{\beta}{2}},$$

  hence the inequality (1.13) extends to

$$\left\| (I - \Delta)^{\frac{\beta}{2}} e^{t\Delta} \right\|_{L^p \to L^p} \leq C\, t^{-\frac{\beta}{2}}. \tag{1.14}$$

- 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 $\mathcal{C}^n_b(\mathbb{R}^d)$; and the space of n-times ($n \in \mathbb{N}$) differentiable functions with compact support, denoted by $\mathcal{C}^n_c(\mathbb{R}^d)$. 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 $\mathcal{N}_\delta$ the Hölder seminorm of parameter $\delta \in (0, 1]$, that is, for any function f defined over $\mathbb{R}^d$:

$$\mathcal{N}_\delta(f) := \sup_{x \neq y \in \mathbb{R}^d} \frac{|f(x) - f(y)|}{|x - y|^\delta}. \tag{1.15}$$

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

2 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 2.2, we derive some general inequalities and prove Proposition 1.2. In Subsection 2.3, we prove that the PDE with cut-off (2.8) can have at most one mild solution.

2.1 Some classical embeddings

If $\beta - \frac{d}{\boldsymbol{r}} > 0$, $H^\beta_{\boldsymbol{r}}(\mathbb{R}^d)$ is continuously embedded into $\mathcal{C}^{\beta - \frac{d}{\boldsymbol{r}}}(\mathbb{R}^d)$ (see [57, p.203]). In particular $H^\beta_{\boldsymbol{r}}$ is continuously embedded into $L^{\boldsymbol{r}} \cap L^\infty$. That is, there exists $C, C' > 0$ such that

$$\|f\|_{L^{\boldsymbol{r}} \cap L^\infty(\mathbb{R}^d)} \leq C \|f\|_{\beta,\boldsymbol{r}} \quad \text{and} \quad \|f\|_{\mathcal{C}^{\beta - \frac{d}{\boldsymbol{r}}}(\mathbb{R}^d)} \leq C' \|f\|_{\beta,\boldsymbol{r}}, \quad \forall f \in H^\beta_{\boldsymbol{r}}(\mathbb{R}^d). \tag{2.1}$$

Then by interpolation, $L^1 \cap H^\beta_{\boldsymbol{r}}(\mathbb{R}^d)$ is continuously embedded into $L^1 \cap L^\infty(\mathbb{R}^d)$. That is, there exists $C_{d,\beta,\boldsymbol{r}} > 0$ such that

$$\|f\|_{L^1 \cap L^\infty(\mathbb{R}^d)} \leq C_{d,\beta,\boldsymbol{r}} \|f\|_{L^1 \cap H^\beta_{\boldsymbol{r}}(\mathbb{R}^d)}, \quad \forall f \in L^1 \cap H^\beta_{\boldsymbol{r}}(\mathbb{R}^d). \tag{2.2}$$

Finally, we will need the continuous embedding $H^\alpha_p(\mathcal{B}_R) \subset W^{\alpha,p}(\mathcal{B}_R)$ for all $p \in [2, +\infty)$ and $\alpha \geq 0$ (see [57, p.327] and also [57, p.172] for the whole space).

2.2 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 2.1. *Let K be satisfying the Assumptions ($\mathbf{A}_i^K$) and ($\mathbf{A}_{ii}^K$). There exists $C_{K,d} > 0$ (which depends on K and d only) such that for any $f \in L^1 \cap L^\infty(\mathbb{R}^d)$,*

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

Proof. Recall that $\boldsymbol{q}'$ denotes the conjugate exponent of the paremeter $\boldsymbol{q}$ from ($\mathbf{A}_{ii}^K$). In view of Assumptions ($\mathbf{A}_i^K$) and ($\mathbf{A}_{ii}^K$), Hölder's inequality yields

$$\begin{aligned} |K * f(x)| &\leq \int_{\mathcal{B}_1} |K(y)||f(x-y)| \, dy + \int_{\mathcal{B}_1^c} |K(y)||f(x-y)| \, dy \\ &\leq \|K\|_{L^1(\mathcal{B}_1)} \|f\|_{L^\infty(\mathbb{R}^d)} + \|K\|_{L^{\boldsymbol{q}}(\mathcal{B}_1^c)} \|f\|_{L^{\boldsymbol{q}'}(\mathbb{R}^d)}. \end{aligned}$$

The conclusion follows from the interpolation inequality $\|f\|_{L^{\boldsymbol{q}'}(\mathbb{R}^d)} \lesssim \|f\|_{L^1 \cap L^\infty(\mathbb{R}^d)}$. □

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

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

with the associated norm $\|\cdot\|_{T, L^1 \cap L^\infty(\mathbb{R}^d)}$, 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

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

Lemma 2.2. *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 $t \in [0,T]$,

$$\|B(u,v)(t)\|_{L^1 \cap L^\infty(\mathbb{R}^d)} \leq C\sqrt{t} \ \|u\|_{\mathcal{X}} \|v\|_{\mathcal{X}}, \tag{2.3}$$

which suffices to prove the continuity of B.

First, using the property (1.13) of the Gaussian kernel with $p = \infty$, observe that for any $t \in [0,T]$,

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

and since $v_s \in L^1 \cap L^\infty(\mathbb{R}^d)$, Lemma 2.1 yields

$$\begin{aligned} \|B(u,v)(t)\|_{L^\infty(\mathbb{R}^d)} &\leq \int_0^t \frac{C}{\sqrt{t-s}} \|K * v_s\|_{L^\infty(\mathbb{R}^d)} \ \|u_s\|_{L^\infty(\mathbb{R}^d)} \, ds \\ &\leq C\sqrt{t} \ \|u\|_{\mathcal{X}} \|v\|_{\mathcal{X}}. \end{aligned} \tag{2.4}$$

On other hand we have, using similarly the property (1.13) of the Gaussian kernel with $p = 1$ and Lemma 2.1, that

$$\begin{aligned} \|B(u,v)(t)\|_{L^1(\mathbb{R}^d)} &\leq \int_0^t \frac{C}{\sqrt{t-s}} \|u_s \ K * v_s\|_{L^1(\mathbb{R}^d)} \, ds \\ &\leq \int_0^t \frac{C}{\sqrt{t-s}} \|u_s\|_{L^1(\mathbb{R}^d)} \|K * v_s\|_{L^\infty(\mathbb{R}^d)} \, ds \\ &\leq C\sqrt{t} \ \|u\|_{\mathcal{X}} \|v\|_{\mathcal{X}}. \end{aligned} \tag{2.5}$$

Therefore, combining Equations (2.4), (2.5) one obtains (2.3). □

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

Proof of Proposition 1.2. For the existence part, note that for any $T > 0$, $\mathcal{X}$ is a Banach space and that our aim is to find $T > 0$ and $u \in \mathcal{X}$ such that

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

In view of Lemma 2.2, such a local mild solution is obtained by a standard contraction argument (Banach fixed-point Theorem).

We will prove the uniqueness in the (slightly more complicated) case of the cut-off PDE (2.8), for any value of the cut-off A (see Proposition 2.6). 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 $[0,T]$, then Lemma 2.1 implies that for $i = 1,2$,

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

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

Corollary 2.3. *Assume that* $(\mathbf{A}_i^K)$ *and* $(\mathbf{A}_{ii}^K)$ *hold. Let* C *be the constant that appears in* (2.3)*. Then for* $T > 0$ *such that*

$$4C\sqrt{T}\ \|u_0\|_{L^1\cap L^\infty(\mathbb{R}^d)} < 1, \tag{2.6}$$

one can define a local mild solution u *to* (1.1) *up to time* T *and*

$$\|u\|_{T,L^1\cap L^\infty(\mathbb{R}^d)} \leq \frac{1 - \sqrt{1 - 4C\sqrt{T}\ \|u_0\|_{L^1\cap L^\infty(\mathbb{R}^d)}}}{2C\sqrt{T}}. \tag{2.7}$$

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

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

Proof. We rely on the bound (2.3), in order to get that for $t \leq T$,

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

Then by a standard argument (see e.g. Lemma 2.3 in [41]), choosing $T > 0$ which satisfies (2.6) ensures that (2.7) holds true. □

Remark 2.4. *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* $[0,T]$ *to* (1.1)*. Then in view of Definition 1.1-i) and by the inequality* $\|e^{(t-s)\Delta}\|_{L^p\to L^p} \leq C$*,*

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

using Lemma 2.1. Similarly, for any $s \in (0,t)$*,*

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

Hence by integration-by-parts

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

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 [41, Prop. 2.7] (although using $(\mathbf{A}^K_{iii})$ instead of the Poisson kernel), u is such that $u_t \geq 0$, for $(t,x) \in (0,T) \times \mathbb{R}^d$ (we can also deduce this fact later from the convergence of u^N to u). Hence the mass is preserved.

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

We aim to prove the convergence of the mollified empirical measure to the following PDE with cut-off:

$$\begin{cases} \partial_t \widetilde{u}(t,x) = \Delta \widetilde{u}(t,x) - \nabla \cdot \left(\widetilde{u}(t,x) F_A(K * \widetilde{u}(t,x))\right), \quad t>0,\ x \in \mathbb{R}^d \\ \widetilde{u}(0,x) = u_0(x). \end{cases} \tag{2.8}$$

Although this is implicit, $\widetilde{u}$ actually depends on A. Note that if F_A is replaced by the identity function, one recovers (1.1).

Remark 2.5. *Similarly to Definition 1.1, a mild solution to* (2.8) *satisfies Definition 1.1 i) and solves*

$$u_t = e^{t\Delta} u_0 - \int_0^t \nabla \cdot e^{(t-s)\Delta}(u_s F_A(K * u_s))\ ds, \quad 0 \leq t \leq T. \tag{2.9}$$

In this section, we consider the cut-off PDE (2.8) and its mild solution from Definition 1.1. Here, F_A is given in (1.7), but we denote it simply by F for the sake of readability.

Note that due to the boundedness of the reaction term in (2.8), any mild solution will always be global. This global solution will be rigorously obtained as the limit of the particle system (1.9). Thus we only consider global mild solutions when it comes to the PDE (2.8).

Proposition 2.6. *Assume that K satisfies Assumptions $(\mathbf{A}^K_i)$ and $(\mathbf{A}^K_{ii})$. Let $u_0 \in L^1 \cap L^\infty(\mathbb{R}^d)$. Then for any $A>0$ and F defined in* (1.7)*, there is at most one mild solution to the cut-off PDE* (2.8)*.*

Proof. Assume there are two mild solutions u^1 and u^2 to (2.8). Then,

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

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

$$\begin{aligned} \|u^1_t - u^2_t\|_{L^1(\mathbb{R}^d)} + \|u^1_t - u^2_t\|_{L^\infty(\mathbb{R}^d)} &\leq C \int_0^t \frac{1}{\sqrt{t-s}} \left(\|u^1_s - u^2_s\|_{L^1(\mathbb{R}^d)} + \|u^2_s\ K * (u^1_s - u^2_s)\|_{L^1(\mathbb{R}^d)} \right)\ ds \\ &\quad + C \int_0^t \frac{1}{\sqrt{t-s}} \left(\|u^1_s - u^2_s\|_{L^\infty(\mathbb{R}^d)} + \|u^2_s\ K * (u^1_s - u^2_s)\|_{L^\infty(\mathbb{R}^d)} \right)\ ds \\ &\leq C\sqrt{t}\ \|u^1 - u^2\|_{t, L^1 \cap L^\infty(\mathbb{R}^d)} \\ &\quad + C \int_0^t \frac{\|u^2_s\|_{L^\infty(\mathbb{R}^d)} + \|u^2_s\|_{L^1(\mathbb{R}^d)}}{\sqrt{t-s}}\ \|K * \left(u^1_s - u^2_s\right)\|_{L^\infty(\mathbb{R}^d)}\ ds. \end{aligned} \tag{2.10}$$

In view of Lemma 2.1, one has $\|K*(u^1_s-u^2_s)\|_{L^\infty(\mathbb{R}^d)}\le C\|u^1_s-u^2_s\|_{L^1\cap L^\infty(\mathbb{R}^d)}$, thus plugging this upper bound in (2.10) gives

$$\|u^1_t-u^2_t\|_{L^1(\mathbb{R}^d)}+\|u^1_t-u^2_t\|_{L^\infty(\mathbb{R}^d)}\le\|u^1-u^2\|_{t,L^1\cap L^\infty(\mathbb{R}^d)}C\sqrt{t}\left(1+\|u^2\|_{t,L^1\cap L^\infty(\mathbb{R}^d)}\right).$$

Hence for t small enough, we deduce that $\|u^1-u^2\|_{t,L^1\cap L^\infty(\mathbb{R}^d)}=0$. 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. □

Finally, we get the following stability estimate in $H^\beta_{\boldsymbol{r}}$ for the solution of the PDE.

Proposition 2.7. *Assume that K satisfies Assumptions $(\mathbf{A}^K_i)$ and $(\mathbf{A}^K_{ii})$. Let $u_0\in H^\beta_{\boldsymbol{r}}(\mathbb{R}^d)$ and $0<\beta<1$. Then for any $A>0$ and F defined in* (1.7)*, the unique solution of the cut-off PDE* (2.8) *verifies*

$$\sup_{t\in[0,T]}\|u_t\|_{\beta,\boldsymbol{r}}<\infty.$$

Proof. Let u be the unique solution of the cut-off PDE (2.8). Then using (1.5) and the fractional estimate (1.14), we have

$$\begin{aligned}
\|u_t\|_{\beta,\boldsymbol{r}}&\le\|e^{t\Delta}u_0\|_{\beta,\boldsymbol{r}}+\int_0^t\|\nabla\cdot e^{(t-s)\Delta}\left\{u_s\,F(K*u_s)\right\}\|_{\beta,\boldsymbol{r}}\,ds\\
&=\|e^{t\Delta}u_0\|_{\beta,\boldsymbol{r}}+\int_0^t\|\,(I-\Delta)^{\frac{\beta}{2}}\,\nabla\cdot e^{(t-s)\Delta}\left\{u_s\,F(K*u_s)\right\}\|_{L^{\boldsymbol{r}}(\mathbb{R}^d)}\,ds\\
&\le\|u_0\|_{\beta,\boldsymbol{r}}+C\int_0^t\frac{1}{(t-s)^{\frac{1+\beta}{2}}}\|u_s\,F(K*u_s)\|_{L^{\boldsymbol{r}}(\mathbb{R}^d)}\,ds\\
&\le\|u_0\|_{\beta,\boldsymbol{r}}+C\int_0^t\frac{1}{(t-s)^{\frac{1+\beta}{2}}}\|u_s\|_{L^{\boldsymbol{r}}(\mathbb{R}^d)}\,ds\\
&\le\|u_0\|_{\beta,\boldsymbol{r}}+C\int_0^t\frac{1}{(t-s)^{\frac{1+\beta}{2}}}\|u_s\|_{\beta,\boldsymbol{r}}\,ds.
\end{aligned}$$

Since $\beta<1$, we deduce the desired estimate from Grönwall's lemma. □

3 Rate of convergence

In this section, we establish our first main result (Theorem 1.4).

More precisely, we establish first the equation satisfied by the regularised empirical measure and derive its boundedness in a suitable space (Subsection 3.1). Then in Subsection 3.2, we prove Theorem 1.4. In Subsection 3.3, we state the corollaries of Theorem 1.4. The proofs of Corollaries 3.2, 3.6 and 1.5 are given respectively in Subsections 3.4, 3.5 and 3.6.

3.1 Properties of the regularised empirical measure

The proof of the rate of convergence relies on the following mild formulation of the mollified empirical measure:

$$\begin{aligned}
u^N_t(x)=e^{t\Delta}u^N_0(x)-\int_0^t\nabla\cdot e^{(t-s)\Delta}\langle\mu^N_s,V^N(\cdot-x)F\big(K*u^N_s(\cdot)\big)\rangle\,ds&\\
+\frac{1}{N}\sum_{i=1}^N\int_0^t e^{(t-s)\Delta}\nabla V^N(X^{i,N}_s-x)\cdot dW^i_s,&
\end{aligned}\tag{3.1}$$

and the following boundedness estimate:

Proposition 3.1. *Let the assumptions of Theorem 1.4 hold. Let $q \geq 1$. Then*

$$\sup_{N\in\mathbb{N}^*} \sup_{t\in[0,T]} \mathbb{E}\left[\left\|u_t^N\right\|_{\beta,\boldsymbol{r}}^q\right] < \infty.$$

This proposition will be proven in Appendix A.2. It is similar to Proposition 2.1 in [20] (the kernel plays no role here). Note that this is where the restriction $(\mathbf{A}_\alpha)$ on α appears.

Now, let us establish Equation (3.1). 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 (1.9) as

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

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

$$\begin{aligned} u_t^N(x) =& u_0^N(x) - \frac{1}{N}\sum_{i=1}^N \int_0^t \nabla V^N(x - X_s^{i,N}) \cdot F\big(K * u_s^N(X_s^{i,N})\big)\, ds \\ & - \frac{1}{N}\sum_{i=1}^N \int_0^t \nabla V^N(x - X_s^{i,N}) \cdot dW_s^i + \frac{1}{N}\sum_{i=1}^N \int_0^t \Delta V^N(x - X_s^{i,N})\, ds. \end{aligned} \tag{3.2}$$

Notice that

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

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 (3.2) and the fact that $\nabla V^N(-x) = -\nabla V^N(x)$ (because V^N is even) lead to

$$\begin{aligned} u_t^N(x) = u_0^N(x) + \int_0^t \langle \mu_s^N, \nabla V^N(\cdot - x)\cdot F\big(K * u_s^N(\cdot)\big)\rangle\, ds \\ + \frac{1}{N}\sum_{i=1}^N \int_0^t \nabla V^N(X_s^{i,N} - x)\cdot dW_s^i + \int_0^t \Delta u_s^N(x)\, ds. \end{aligned} \tag{3.3}$$

The following mild form is immediately derived:

$$\begin{aligned} u_t^N(x) = e^{t\Delta}u_0^N(x) + \int_0^t e^{(t-s)\Delta}\langle \mu_s^N, \nabla V^N(\cdot - x)\cdot F\big(K * u_s^N(\cdot)\big)\rangle\, ds \\ + \frac{1}{N}\sum_{i=1}^N \int_0^t e^{(t-s)\Delta}\nabla V^N(X_s^{i,N} - x)\cdot dW_s^i. \end{aligned} \tag{3.4}$$

Finally, developing the scalar product, one has

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

Combining the latter with the fact that $e^{t\Delta}\nabla\cdot f = \nabla\cdot e^{t\Delta} f$, we deduce the mild form (3.1) from (3.4).

3.2 Rate in $L^1 \cap L^\infty$ norm: Proof of Theorem 1.4

Step 1 : A first upper bound on the $L^1 \cap L^\infty$ norm of $u^N - u$.

In view of the mild formulas (3.1) for u^N and (2.9) for u, it comes

$$\begin{aligned}
u_t^N(x) - u_t(x) &= e^{t\Delta}(u_0^N - u_0)(x) - \int_0^t \nabla \cdot e^{(t-s)\Delta} \left(\langle \mu_s^N, V^N(\cdot - x) F\big(K * u_s^N(\cdot)\big) \rangle - u_s F(K * u_s)(x) \right) ds \\
&\quad + \frac{1}{N} \sum_{i=1}^N \int_0^t e^{(t-s)\Delta} \nabla V^N(X_s^{i,N} - x) \cdot dW_s^i \\
&= e^{t\Delta}(u_0^N - u_0)(x) + \int_0^t \nabla \cdot e^{(t-s)\Delta} \left((u_s F(K * u_s) - u_s^N F(K * u_s^N)) * V^N \right)(x)\, ds \\
&\quad + E_t^{(1)}(x) + E_t^{(2)}(x) + M_t^N(x),
\end{aligned}$$

where we have set

$$\begin{aligned}
E_t^{(1)}(x) &:= \int_0^t \nabla \cdot e^{(t-s)\Delta} \left(u_s F(K * u_s) - (u_s F(K * u_s)) * V^N \right)(x)\, ds, \\
E_t^{(2)}(x) &:= \int_0^t \nabla \cdot e^{(t-s)\Delta} \langle u_s^N - \mu_s^N, V^N(\cdot - x) F\big(K * u_s^N(\cdot)\big) \rangle\, ds, \\
M_t^N(x) &:= \frac{1}{N} \sum_{i=1}^N \int_0^t e^{(t-s)\Delta} \nabla V^N(X_s^{i,N} - x) \cdot dW_s^i.
\end{aligned} \tag{3.5}$$

For any $p \in [1, +\infty]$, in view of the estimate (1.13), one has

$$\begin{aligned}
\|u_t^N - u_t\|_{L^p(\mathbb{R}^d)} &\le \|e^{t\Delta}(u_0^N - u_0)\|_{L^p(\mathbb{R}^d)} \\
&\quad + C \int_0^t \frac{1}{\sqrt{t-s}} \|(u_s F(K * u_s) - u_s^N F(K * u_s^N)) * V^N\|_{L^p(\mathbb{R}^d)} ds \\
&\quad + \|E_t^{(1)}\|_{L^p(\mathbb{R}^d)} + \|E_t^{(2)}\|_{L^p(\mathbb{R}^d)} + \|M_t^N\|_{L^p(\mathbb{R}^d)}
\end{aligned}$$

and it follows that

$$\begin{aligned}
\|u_t^N - u_t\|_{L^p(\mathbb{R}^d)} &\le \|e^{t\Delta}(u_0^N - u_0)\|_{L^p(\mathbb{R}^d)} + C \int_0^t \frac{1}{\sqrt{t-s}} \|F\|_{L^\infty(\mathbb{R}^d)} \|u_s^N - u_s\|_{L^p(\mathbb{R}^d)} ds \\
&\quad + C \int_0^t \frac{1}{\sqrt{t-s}} \|u_s\|_{L^p} \|F\|_{\mathrm{Lip}} \|K * (u_s - u_s^N)\|_{L^\infty(\mathbb{R}^d)} ds \\
&\quad + \|E_t^{(1)}\|_{L^p(\mathbb{R}^d)} + \|E_t^{(2)}\|_{L^p(\mathbb{R}^d)} + \|M_t^N\|_{L^p(\mathbb{R}^d)}.
\end{aligned}$$

From Proposition 2.7 and by interpolation, we have that

$$\|u\|_{T, L^p(\mathbb{R}^d)} < \infty, \quad \text{for any } p \ge 1.$$

Thus, for some $C > 0$ which depends on $\|u\|_{T, L^1 \cap L^\infty(\mathbb{R}^d)}$, $\|F\|_{L^\infty(\mathbb{R}^d)}$ and $\|F\|_{\mathrm{Lip}}$, it comes

$$\begin{aligned}
\|u_t^N - u_t\|_{L^p(\mathbb{R}^d)} &\le \|e^{t\Delta}(u_0^N - u_0)\|_{L^p(\mathbb{R}^d)} + C \int_0^t \frac{1}{\sqrt{t-s}} \|u_s^N - u_s\|_{L^p(\mathbb{R}^d)} ds \\
&\quad + C \int_0^t \frac{1}{\sqrt{t-s}} \|K * (u_s - u_s^N)\|_{L^\infty(\mathbb{R}^d)} ds \\
&\quad + \|E_t^{(1)}\|_{L^p(\mathbb{R}^d)} + \|E_t^{(2)}\|_{L^p(\mathbb{R}^d)} + \|M_t^N\|_{L^p(\mathbb{R}^d)}.
\end{aligned}$$

Finally we apply Lemma 2.1 and obtain

$$\begin{aligned} \|u_t^N - u_t\|_{L^p(\mathbb{R}^d)} \leq \|e^{t\Delta}(u_0^N - u_0)\|_{L^p(\mathbb{R}^d)} + C\int_0^t \frac{1}{\sqrt{t-s}}\|u_s^N - u_s\|_{L^p(\mathbb{R}^d)}ds \\ + C\int_0^t \frac{1}{\sqrt{t-s}}\|u_s - u_s^N\|_{L^1\cap L^\infty(\mathbb{R}^d)}ds \\ + \|E_t^{(1)}\|_{L^p(\mathbb{R}^d)} + \|E_t^{(2)}\|_{L^p(\mathbb{R}^d)} + \|M_t^N\|_{L^p(\mathbb{R}^d)}. \end{aligned} \tag{3.6}$$

Therefore, considering (3.6) for both $p = 1$ and $p = \infty$, we deduce that

$$\begin{aligned} \|u_t^N - u_t\|_{L^1\cap L^\infty(\mathbb{R}^d)} \leq \|e^{t\Delta}(u_0^N - u_0)\|_{L^1\cap L^\infty(\mathbb{R}^d)} + C\int_0^t \frac{1}{\sqrt{t-s}}\|u_s^N - u_s\|_{L^1\cap L^\infty(\mathbb{R}^d)}ds \\ + \|E_t^{(1)}\|_{L^1\cap L^\infty(\mathbb{R}^d)} + \|E_t^{(2)}\|_{L^1\cap L^\infty(\mathbb{R}^d)} + \|M_t^N\|_{L^1\cap L^\infty(\mathbb{R}^d)}, \end{aligned}$$

and for any $m \geq 1$, it comes

$$\begin{aligned} &\left\|\|u^N - u\|_{t,L^1\cap L^\infty(\mathbb{R}^d)}\right\|_{L^m(\Omega)} \\ &\leq \left\|\sup_{s\in[0,t]} \|e^{s\Delta}(u_0^N - u_0)\|_{L^1\cap L^\infty(\mathbb{R}^d)}\right\|_{L^m(\Omega)} \\ &+ C\int_0^t \frac{1}{(t-s)^{\frac{1}{2}}} \left\|\|u^N - u\|_{s,L^1\cap L^\infty(\mathbb{R}^d)}\right\|_{L^m(\Omega)} ds \\ &+ \left\|\|E^{(1)}\|_{t,L^1\cap L^\infty(\mathbb{R}^d)}\right\|_{L^m(\Omega)} + \left\|\|E^{(2)}\|_{t,L^1\cap L^\infty(\mathbb{R}^d)}\right\|_{L^m(\Omega)} + \left\|\|M^N\|_{t,L^1\cap L^\infty(\mathbb{R}^d)}\right\|_{L^m(\Omega)}. \end{aligned} \tag{3.7}$$

Step 2: L^1-estimates.

• *First, let us estimate* $\|E_t^{(1)}\|_{L^1(\mathbb{R}^d)}$. The property (1.13) on the derivative of the heat kernel gives

$$\sup_{s\in[0,t]} \|E_s^{(1)}\|_{L^1(\mathbb{R}^d)} \leq \int_0^t \frac{C}{(t-s)^{\frac{1}{2}}} \|e^{\frac{t-s}{2}\Delta}\left(u_s F(K*u_s) - (u_s F(K*u_s)) * V^N\right)\|_{L^1(\mathbb{R}^d)}ds. \tag{3.8}$$

Recall that the d-dimensional Gaussian probability density function g_{t-s} (defined in (1.6)) is the kernel associated to the operator $e^{\frac{t-s}{2}\Delta}$ and observe that

$$\begin{aligned} &\|e^{\frac{t-s}{2}\Delta}\left(u_s F(K*u_s) - (u_s F(K*u_s)) * V^N\right)\|_{L^1(\mathbb{R}^d)} \\ &= \int_{\mathbb{R}^d} \left|\int_{\mathbb{R}^d} u_s(y)F(K*u_s(y))\left(g_{t-s}(x-y) - g_{t-s}*V^N(x-y)\right)dy\right| dx \\ &\leq \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |u_s(y)F(K*u_s(y))| \left|g_{t-s}(x-y) - g_{t-s}*V^N(x-y)\right| dy\, dx. \end{aligned}$$

Hence using again that F is bounded and that $\|u_s\|_{L^1(\mathbb{R}^d)} = 1$,

$$\begin{aligned} \|e^{\frac{t-s}{2}\Delta}\left(u_s F(K*u_s) - (u_s F(K*u_s)) * V^N\right)\|_{L^1(\mathbb{R}^d)} &\leq C\int_{\mathbb{R}^d} \left|g_{t-s}(x) - g_{t-s}*V^N(x)\right| dx \\ &\leq C\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} V(y)\left|g_{t-s}(x) - g_{t-s}(x - \tfrac{y}{N^\alpha})\right| dy\, dx. \end{aligned}$$

Now for any $\theta \in (0, 1)$,

$$\begin{aligned} &\|e^{\frac{t-s}{2}\Delta}\left(u_s F(K*u_s) - (u_s F(K*u_s)) * V^N\right)\|_{L^1(\mathbb{R}^d)} \\ &\quad\leq C\int_{\mathbb{R}^d} V(y)\left(\int_{\mathbb{R}^d} g_{t-s}(x) + g_{t-s}(x - \tfrac{y}{N^\alpha})dx\right)^\theta \left(\int_{\mathbb{R}^d}\left|\frac{y}{N^\alpha}\cdot\int_0^1 \nabla g_{t-s}(x - r\frac{y}{N^\alpha})\,dr\right| dx\right)^{1-\theta} dy. \end{aligned}$$

Recall that g_{t-s} is a density, then by applying Hölder's inequality it comes

$$\begin{aligned}&\|e^{\frac{t-s}{2}\Delta}\left(u_sF(K*u_s)-(u_sF(K*u_s))*V^N\right)\|_{L^1(\mathbb{R}^d)}\\&\leq CN^{-(1-\theta)\alpha}\int_{\mathbb{R}^d}V(y)|y|^{1-\theta}\left(\int_0^1\int_{\mathbb{R}^d}\left|\nabla g_{t-s}(x-r\frac{y}{N^\alpha})\right|dx\,dr\right)^{1-\theta}dy\\&\leq CN^{-(1-\theta)\alpha}\int_{\mathbb{R}^d}V(y)|y|^{1-\theta}dy\ \left(\int_{\mathbb{R}^d}|\nabla g_{t-s}(x)|\,dx\right)^{1-\theta}.\end{aligned}$$

Now one recalls that $\int_{\mathbb{R}^d}|\nabla g_{t-s}(x)|\,dx\leq C(t-s)^{-\frac{1}{2}}$, and since one also has that $\int_{\mathbb{R}^d}V(y)|y|^{1-\theta}dy$ is finite (by assumption on V), it comes

$$\|e^{\frac{t-s}{2}\Delta}\left(u_sF(K*u_s)-(u_sF(K*u_s))*V^N\right)\|_{L^1(\mathbb{R}^d)}\leq C(t-s)^{-\frac{1}{2}(1-\theta)}N^{-(1-\theta)\alpha}.$$

Hence plugging this bound into (3.8) yields

$$\sup_{s\in[0,t]}\|E_s^{(1)}\|_{L^1(\mathbb{R}^d)}\leq CN^{-(1-\theta)\alpha}\int_0^t(t-s)^{-\frac{1}{2}-\frac{1}{2}(1-\theta)}ds,$$

where the integral is finite if $\theta>0$. Hence we have obtained that for any $\varepsilon\in(0,1)$ there exists C such that

$$\sup_{s\in[0,t]}\|E_s^{(1)}\|_{L^1(\mathbb{R}^d)}\leq C\,N^{-(1-\varepsilon)\alpha}. \tag{3.9}$$

• *We now search for a bound on* $\|\sup_{s\in[0,t]}\|E_s^{(2)}\|_{L^1(\mathbb{R}^d)}\|_{L^m(\Omega)}$. First, observe that due to the convolution inequality (1.13),

$$\begin{aligned}\sup_{s\in[0,t]}\|E_s^{(2)}\|_{L^1(\mathbb{R}^d)}&\leq C\int_0^t\frac{1}{(t-s)^{\frac{1}{2}}}\int_{\mathbb{R}^d}\left|\left\langle u_s^N-\mu_s^N,g_{t-s}*V^N(\cdot-x)\,F\left(K*u_s^N(\cdot)\right)\right\rangle\right|dx\,ds\\&=C\int_0^t\frac{1}{(t-s)^{\frac{1}{2}}}\left(\int_{\mathbb{R}^d}\left|\left\langle\mu_s^N,\int_{\mathbb{R}^d}V(y)\big\{g_{t-s}*V^N(\cdot-x)\,F\left(K*u_s^N(\cdot)\right)\right.\right.\right.\\&\qquad\left.\left.\left.-\,g_{t-s}*V^N(\tfrac{y}{N^\alpha}+\cdot-x)\,F\left(K*u_s^N(\tfrac{y}{N^\alpha}+\cdot)\right)\big\}dy\right\rangle\right|dx\right)ds.\end{aligned}$$

Hence by Fubini's theorem, one gets

$$\begin{aligned}&\sup_{s\in[0,t]}\|E_s^{(2)}\|_{L^1(\mathbb{R}^d)}\\&\leq C\int_0^t\frac{1}{(t-s)^{\frac{1}{2}}}\left(\left(\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}V(y)\left|g_{t-s}*V^N(z-x)\left(F(K*u_s^N(z))-F(K*u_s^N(z+\tfrac{y}{N^\alpha}))\right)\right|dy\,\mu_s^N(dz)\,dx\right)\right.\\&+\left.\left(\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}V(y)\left|F(K*u_s^N(z+\tfrac{y}{N^\alpha}))\left(g_{t-s}*V^N(z-x)-g_{t-s}*V^N(z-x+\tfrac{y}{N^\alpha})\right)\right|dy\,\mu_s^N(dz)\,dx\right)\right)ds\\&=:E^{(2,1)}+E^{(2,2)}.\end{aligned} \tag{3.10}$$

We shall now estimate $E^{(2,1)}$ and $E^{(2,2)}$.

From ($\mathbf{A}_{iii}^K$) we deduce that $|F(K*u_s^N(z))-F(K*u_s^N(z+\frac{y}{N^\alpha}))|\leq\|F\|_{\text{Lip}}\left|\frac{y}{N^\alpha}\right|^{\boldsymbol{\zeta}}\|u_s^N\|_{L^1\cap H_{\boldsymbol{r}}^\beta(\mathbb{R}^d)}$. Hence

$$\begin{aligned}E^{(2,1)}&\leq\frac{C}{N^{\alpha\boldsymbol{\zeta}}}\int_0^t\frac{\|u_s^N\|_{L^1\cap H_{\boldsymbol{r}}^\beta(\mathbb{R}^d)}}{(t-s)^{\frac{1}{2}}}\left(\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}|y|^{\boldsymbol{\zeta}}\,V(y)\left|g_{t-s}*V^N(z-x)\right|dy\,\mu_s^N(dz)\,dx\right)ds\\&\leq\frac{C}{N^{\alpha\boldsymbol{\zeta}}}\int_0^t\frac{\|u_s^N\|_{L^1\cap H_{\boldsymbol{r}}^\beta(\mathbb{R}^d)}}{(t-s)^{\frac{1}{2}}}\left(\int_{\mathbb{R}^d}\left|g_{t-s}*V^N(x)\right|dx\right)ds,\end{aligned}$$

where in the first inequality, we used the fact that V is rapidly decreasing and therefore the integral with respect to y is finite. Then by the standard convolution inequality, $\|g_{t-s} * V^N\|_{L^1(\mathbb{R}^d)} \leq 1$. Hence it follows from Proposition 3.1 that

$$\begin{aligned}\|E^{(2,1)}\|_{L^m(\Omega)} &\leq \frac{C}{N^{\alpha\zeta}} \sup_{s\in[0,T]} \Big\| \|u_s^N\|_{L^1\cap H_r^\beta(\mathbb{R}^d)} \Big\|_{L^m(\Omega)} \int_0^t (t-s)^{-\frac{1}{2}} ds \\ &\leq \frac{C}{N^{\alpha\zeta}}. \end{aligned} \tag{3.11}$$

Consider now $E^{(2,2)}$. One has, using the boundedness of F,

$$\begin{aligned} &E^{(2,2)} \\ &\leq C \int_0^t \frac{1}{(t-s)^{\frac{1}{2}}} \left(\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} V(y) \left| \left(g_{t-s} * V^N(z-x) - g_{t-s} * V^N(z-x+\tfrac{y}{N^\alpha}) \right) \right| dy\, \mu_s^N(dz)\, dx \right) ds \\ &\leq \frac{C}{N^\alpha} \int_0^t \frac{1}{(t-s)^{\frac{1}{2}}} \left(\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} V(y)\, |y| \left(\int_0^1 |\nabla(g_{t-s} * V^N)(z-x+r\tfrac{y}{N^\alpha})|\, dr \right) dy\, \mu_s^N(dz)\, dx \right) ds. \end{aligned}$$

Applying Fubini's Theorem and a change of variables, it comes

$$E^{(2,2)} \leq \frac{C}{N^\alpha} \int_0^t \frac{1}{(t-s)^{\frac{1}{2}}} \|\nabla(g_{t-s} * V^N)\|_{L^1(\mathbb{R}^d)} ds. \tag{3.12}$$

The estimation of $\|\nabla(g_{t-s} * V^N)\|_{L^1(\mathbb{R}^d)}$ is twofold, depending on which term of the convolution we apply the gradient. First, by the convolution inequality (1.13),

$$\|\nabla(g_{t-s} * V^N)\|_{L^1(\mathbb{R}^d)} = \|\nabla g_{t-s} * V^N\|_{L^1(\mathbb{R}^d)} \leq \frac{C}{\sqrt{t-s}}. \tag{3.13}$$

Second, still by applying a convolution inequality, it comes

$$\begin{aligned}\|\nabla(g_{t-s} * V^N)\|_{L^1(\mathbb{R}^d)} = \|g_{t-s} * (\nabla V^N)\|_{L^1(\mathbb{R}^d)} &\leq \|g_{t-s}\|_{L^1(\mathbb{R}^d)} \|\nabla V^N\|_{L^1(\mathbb{R}^d)} \\ &\leq CN^\alpha. \end{aligned} \tag{3.14}$$

Hence, combining (3.13) and (3.14), we deduce that for any $\varepsilon \in [0,1]$,

$$\begin{aligned}\|\nabla(g_{t-s} * V^N)\|_{L^1(\mathbb{R}^d)} &= \|(\nabla g_{t-s}) * V^N\|_{L^1(\mathbb{R}^d)}^{1-\varepsilon} \|g_{t-s} * (\nabla V^N)\|_{L^1(\mathbb{R}^d)}^{\varepsilon} \\ &\leq C(t-s)^{-\frac{1}{2}(1-\varepsilon)} N^{\alpha\varepsilon}. \end{aligned}$$

Plugging the previous bound in (3.12) yields

$$E^{(2,2)} \leq CN^{-\alpha(1-\varepsilon)}. \tag{3.15}$$

In view of deterministic bounds obtained in Inequalities (3.11) and (3.15), we deduce from (3.10) that

$$\left\| \sup_{s\in[0,t]} \|E_s^{(2)}\|_{L^1(\mathbb{R}^d)} \right\|_{L^m(\Omega)} \leq C \left(N^{-\alpha\zeta} + N^{-\alpha(1-\varepsilon)} \right). \tag{3.16}$$

• *Finally, we turn to* $\|M_t^N\|_{L^1(\mathbb{R}^d)}$. One should be particularly careful when dealing with this term as $\left(\|M_s^N\|_{L^1(\mathbb{R}^d)}\right)_{s\geq 0}$ is not a martingale since the semigroup acts as a convolution in time within the stochastic integral (in particular Doob's maximal inequality does not hold). Besides, M_t^N is an $L^1 \cap L^\infty$-valued process, thus to control $\left\|\sup_{s\in[0,t]} \|M_s^N\|_{L^1(\mathbb{R}^d)}\right\|_{L^m(\Omega)}$ one cannot directly apply classical formulations of the Burkholder-Davis-Gundy (BDG) inequality. Instead, one should turn to generalizations of such inequalities in UMD Banach spaces (see van Neerven et al. [58]). There is a classical trick to apply BDG-type inequalities to stochastic convolution integrals, however it only leads to a bound on $\big\|\|M_t^N\|_{L^1(\mathbb{R}^d)}\big\|_{L^m(\Omega)}$ for a fixed $t>0$, instead of a bound on $\Big\|\sup_{s\in[0,t]}\|M_s^N\|_{L^1(\mathbb{R}^d)}\Big\|_{L^m(\Omega)}$. In order to keep the supremum in time inside the expectation, we will also use the lemma of Garsia, Rodemich and Rumsey [25]. Besides, there is an additional difficulty here which is that L^1 is not a UMD Banach space, hence the infinite-dimensional version of the BDG inequality cannot be applied directly.

As the computations are long and technical, we choose to do them in the Appendix A.1 and we give here the following result from Proposition A.1: for any $\varepsilon>0$ arbitrary small, there exists $C>0$ such that

$$\Big\|\sup_{s\in[0,t]}\|M_s^N\|_{L^1(\mathbb{R}^d)}\Big\|_{L^m(\Omega)} \le C\,N^{-\frac12(1-\alpha(d+2))+\varepsilon},\quad \forall N\in\mathbb{N}^*. \tag{3.17}$$

Step 3: L^∞-estimates.

- *We estimate the quantity* $\|E_t^{(1)}\|_\infty$. Applying (1.13), one has

$$\begin{aligned}
\sup_{s\in[0,t]}\|E_s^{(1)}\|_{L^\infty(\mathbb{R}^d)} &\le C\int_0^t \frac{1}{(t-s)^{\frac12}}\|u_sF(K*u_s)-(u_sF(K*u_s))*V^N\|_{L^\infty(\mathbb{R}^d)}ds\\
&\le C\int_0^t \frac{1}{(t-s)^{\frac12}}\int_{\mathbb{R}^d}V(y)\|(u_sF(K*u_s))(\cdot)-u_s(\cdot)F(K*u_s)(\cdot-\frac{y}{N^\alpha})\|_{L^\infty(\mathbb{R}^d)}dyds\\
&\quad + C\int_0^t \frac{1}{(t-s)^{\frac12}}\int_{\mathbb{R}^d}V(y)\|F(K*u_s)(\cdot-\frac{y}{N^\alpha})(u_s(\cdot)-u_s(\cdot-\frac{y}{N^\alpha}))\|_{L^\infty(\mathbb{R}^d)}dyds\\
&\le \frac{C}{N^{\alpha\boldsymbol{\zeta}}}\int_0^t \frac{1}{(t-s)^{\frac12}}\mathcal{N}_{\boldsymbol{\zeta}}(K*u_s)\|u_s\|_{L^\infty(\mathbb{R}^d)}\,ds + \frac{C}{N^{\alpha\delta}}\int_0^t\frac{1}{(t-s)^{\frac12}}\mathcal{N}_\delta(u_s)\,ds,
\end{aligned}$$

where $\mathcal{N}_{\boldsymbol{\zeta}}$ (resp. $\mathcal{N}_\delta$) is the Hölder seminorm of parameter $\boldsymbol{\zeta}$ (resp. δ) defined in (1.15). In view of the embedding (2.1) and Proposition 2.7, the Hölder regularity of u_s is $\delta=\beta-\frac{d}{\boldsymbol{r}}$ and

$$\mathcal{N}_\delta(u_s)\le C\|u_s\|_{H^\beta_{\boldsymbol{r}}(\mathbb{R}^d)}.$$

Thus, according to $(\mathbf{A}^K_{iii})$ and the embedding inequality $\|u_s\|_{L^\infty(\mathbb{R}^d)}\lesssim\|u_s\|_{\beta,\boldsymbol{r}}$, one has

$$\sup_{s\in[0,t]}\|E_s^{(1)}\|_{L^\infty(\mathbb{R}^d)}\le \frac{C}{N^{\alpha\boldsymbol{\zeta}}}\int_0^t\frac{1}{(t-s)^{\frac12}}\|u_s\|^2_{L^1\cap H^\beta_{\boldsymbol{r}}(\mathbb{R}^d)}ds+\frac{C}{N^{\alpha\delta}}\int_0^t\frac{1}{(t-s)^{\frac12}}\|u_s\|_{\beta,\boldsymbol{r}}\,ds.$$

Hence the boundedness of u in $L^\infty([0,T];L^1\cap H^\beta_{\boldsymbol{r}}(\mathbb{R}^d))$ (see again Proposition 2.7) yields

$$\sup_{s\in[0,t]}\|E_s^{(1)}\|_{L^\infty(\mathbb{R}^d)}\le C\left(N^{-\alpha\boldsymbol{\zeta}}+N^{-\alpha(\beta-\frac{d}{\boldsymbol{r}})}\right). \tag{3.18}$$

- *Now, we turn to* $\|E_t^{(2)}\|_\infty$. In view of (1.13), one has

$$\begin{aligned}
\|E_s^{(2)}\|_{L^\infty(\mathbb{R}^d)} &\le C\int_0^t\frac{1}{(t-s)^{\frac12}}\Big\|\int_{\mathbb{R}^d}(g_{t-s}*V^N)(z-\cdot)F\big(K*u_s^N(z)\big)(u_s^N-\mu_s^N)(dz)\Big\|_{L^\infty(\mathbb{R}^d)}ds\\
&\le C\int_0^t\frac{1}{(t-s)^{\frac12}}\Big\|\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}V(y)(g_{t-s}*V^N)(z-\frac{y}{N^\alpha}-\cdot)\\
&\qquad\times\Big(F\big(K*u_s^N(z-\frac{y}{N^\alpha})\big)-F\big(K*u_s^N(z)\big)\Big)\,dy\ \mu_s^N(dz)\Big\|_{L^\infty(\mathbb{R}^d)}ds\\
&\quad + C\int_0^t\frac{1}{(t-s)^{\frac12}}\Big\|\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}V(y)\Big((g_{t-s}*V^N)(z-\frac{y}{N^\alpha}-\cdot)-(g_{t-s}*V^N)(z-\cdot)\Big)\\
&\qquad\times F\big((K*u_s^N)(z)\big)dy\mu_s^N(dz)\Big\|_{L^\infty(\mathbb{R}^d)}ds\\
&=: E_t^{(2,1,\infty)}+E_t^{(2,2,\infty)}.
\end{aligned} \tag{3.19}$$

As above, using $(\mathbf{A}^K_{iii})$ yields

$$
\begin{aligned}
&E_t^{(2,1,\infty)}\\
&\leq \frac{C}{N^{\alpha\boldsymbol{\zeta}}}\int_0^t \frac{1}{(t-s)^{\frac12}}\|u_s^N\|_{L^1\cap H^\beta_{\boldsymbol{r}}(\mathbb{R}^d)}\,\Big\|\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} V(y)|y|^{\alpha\boldsymbol{\zeta}}(g_{t-s}*V^N)(z-\frac{y}{N^\alpha}-\cdot)dy\mu_s^N(dz)\Big\|_{L^\infty(\mathbb{R}^d)}ds\\
&\leq \frac{C}{N^{\alpha\boldsymbol{\zeta}}}\int_0^t \frac{1}{(t-s)^{\frac12}}\|u_s^N\|_{L^1\cap H^\beta_{\boldsymbol{r}}(\mathbb{R}^d)}\int_{\mathbb{R}^d} V(y)|y|^{\alpha\boldsymbol{\zeta}}\,\Big\|\int_{\mathbb{R}^d}(g_{t-s}*V^N)(z-\frac{y}{N^\alpha}-\cdot)\mu_s^N(dz)\Big\|_{L^\infty(\mathbb{R}^d)}dy\,ds\\
&\leq \frac{C}{N^{\alpha\boldsymbol{\zeta}}}\int_0^t \frac{1}{(t-s)^{\frac12}}\|u_s^N\|_{L^1\cap H^\beta_{\boldsymbol{r}}(\mathbb{R}^d)}\int_{\mathbb{R}^d} V(y)|y|^{\alpha\boldsymbol{\zeta}}\,\big\|g_{t-s}*u_s^N\big\|_{L^\infty(\mathbb{R}^d)}\,dy\,ds\\
&\leq \frac{C}{N^{\alpha\boldsymbol{\zeta}}}\int_0^t \frac{1}{(t-s)^{\frac12}}\|u_s^N\|^2_{L^1\cap H^\beta_{\boldsymbol{r}}(\mathbb{R}^d)}\,ds. \qquad (3.20)
\end{aligned}
$$

Observe that

$$
\begin{aligned}
E_t^{(2,2,\infty)} \leq C\int_0^t \frac{1}{(t-s)^{\frac12}}\int_{\mathbb{R}^d} V(y)\,\big\|&(F(K*u_s^N)\mu_s^N)*(g_{t-s}*V^N)(\frac{y}{N^\alpha}+\cdot)\\
&-(F(K*u_s^N)\mu_s^N)*(g_{t-s}*V^N)(\cdot)\big\|_{L^\infty(\mathbb{R}^d)}\,dy\,ds,
\end{aligned}
$$

where $F(K*u_s^N)\mu_s^N$ denotes the weighted empirical measure. Now, we shall prove that $(F(K*u_s^N)\mu_s^N)*(g_{t-s}*V^N)$ is bounded in $H^\beta_{\boldsymbol{r}}(\mathbb{R}^d)$. Recall the following representation for the $H^\beta_{\boldsymbol{r}}(\mathbb{R}^d)$ norm (see (1.5)):

$$
\begin{aligned}
\|(F(K*u_s^N)\mu_s^N)*(g_{t-s}*V^N)\|_{\beta,\boldsymbol{r}} &= \Big\|(I-\Delta)^{\frac\beta2}\big((F(K*u_s^N)\mu_s^N)*(g_{t-s}*V^N)\big)\Big\|_{L^{\boldsymbol{r}}(\mathbb{R}^d)}\\
&= \Big\|(I-\Delta)^{\frac\beta2}g_{t-s}*\big((F(K*u_s^N)\mu_s^N)*V^N\big)\Big\|_{L^{\boldsymbol{r}}(\mathbb{R}^d)}
\end{aligned}
$$

where the second equality holds because $(I-\Delta)^{\frac\beta2}$ acts as a convolution. Then from the inequality $\|(I-\Delta)^{\frac\beta2}g_{t-s}\|_{L^1(\mathbb{R}^d)}\leq C(t-s)^{-\frac\beta2}$ (see Equation (1.14)) and a convolution inequality, it comes

$$
\begin{aligned}
\|(F(K*u_s^N)\mu_s^N)*(g_{t-s}*V^N)\|_{\beta,\boldsymbol{r}} &\leq \frac{C}{(t-s)^{\frac\beta2}}\|(F(K*u_s^N)\mu_s^N)*V^N)\|_{L^{\boldsymbol{r}}(\mathbb{R}^d)}\\
&\leq \frac{C}{(t-s)^{\frac\beta2}}\|u_s^N\|_{L^{\boldsymbol{r}}(\mathbb{R}^d)}.
\end{aligned}
$$

Thus by the Sobolev embedding (2.1), we obtain that for $\delta=(\beta-\frac{d}{\boldsymbol{r}})$,

$$
\begin{aligned}
\|(F(K*u_s^N)\mu_s^N)*(g_{t-s}*V^N)(\frac{y}{N^\alpha}+\cdot)-(F(K*u_s^N)\mu_s^N)*(g_{t-s}*V^N)(\cdot)\|_{L^\infty(\mathbb{R}^d)}&\\
\leq N^{-\alpha\delta}\,\|(F(K*u_s^N)\mu_s^N)*(g_{t-s}*V^N)\|_{\mathcal{C}^\delta(\mathbb{R}^d)}&\\
\leq C\,N^{-\alpha\delta}\,\|(F(K*u_s^N)\mu_s^N)*(g_{t-s}*V^N)\|_{\beta,\boldsymbol{r}}&,
\end{aligned}
$$

which combined with the previous inequality yields

$$
E_t^{(2,2,\infty)} \leq \frac{C}{N^{\alpha(\beta-\frac{d}{\boldsymbol{r}})}}\int_0^t \frac{1}{(t-s)^{\frac{1+\beta}{2}}}\|u_s^N\|_{L^{\boldsymbol{r}}(\mathbb{R}^d)}\,ds. \qquad (3.21)
$$

From (3.19), (3.20) and (3.21), it comes that

$$
\begin{aligned}
\Big\|\sup_{s\in[0,t]}\|E_s^{(2)}\|_{L^\infty(\mathbb{R}^d)}\Big\|_{L^m(\Omega)} \leq \frac{C}{N^{\alpha(\beta-\frac{d}{\boldsymbol{r}})}}&\int_0^t \frac{1}{(t-s)^{\frac{1+\beta}{2}}}\,\big\|\|u_s^N\|_{L^{\boldsymbol{r}}(\mathbb{R}^d)}\big\|_{L^m(\Omega)}\,ds\\
&+\frac{C}{N^{\alpha\boldsymbol{\zeta}}}\int_0^t \frac{1}{(t-s)^{\frac12}}\,\Big\|\|u_s^N\|^2_{L^1\cap H^\beta_{\boldsymbol{r}}(\mathbb{R}^d)}\Big\|_{L^m(\Omega)}\,ds.
\end{aligned}
$$

Hence, in view of the fact that $\|u_s^N\|_{L^1(\mathbb{R}^d)} = 1$, of the uniform bound on $\left\| \|u_s^N\|^2_{\beta,\boldsymbol{r}} \right\|_{L^m(\Omega)}$ (Proposition 3.1) and the assumption $\beta < 1$ (see $(\mathbf{A}^K_{iii})$), we conclude that

$$\left\| \sup_{s\in[0,t]} \|E_s^{(2)}\|_{L^\infty(\mathbb{R}^d)} \right\|_{L^m(\Omega)} \leq C \left(N^{-\alpha(\beta-\frac{d}{\boldsymbol{r}})} + N^{-\alpha\boldsymbol{\zeta}} \right). \tag{3.22}$$

• *It remains to estimate M_t^N.* We proceed with the same care as when we got the bound in (3.17). The details may be found in Appendix A.1 and here we only apply Proposition A.1 for $p = \infty$ and conclude that for any $\varepsilon > 0$, there exists $C > 0$ such that

$$\left\| \sup_{s\in[0,t]} \|M_s^N\|_{L^\infty(\mathbb{R}^d)} \right\|_{L^m(\Omega)} \leq C\, N^{-\frac{1}{2}(1-\alpha(d+[2\vee d]))+\varepsilon}. \tag{3.23}$$

Step 4 : Conclusion.

From the Inequalities (3.7), (3.9), (3.16), (3.17), (3.18), (3.22) and (3.23), and using the Grönwall lemma, we conclude that for any $\varepsilon > 0$ small enough, there exists $C > 0$ such that for any $N \in \mathbb{N}^*$,

$$\begin{aligned} \left\| \|u^N - u\|_{t,L^1\cap L^\infty(\mathbb{R}^d)} \right\|_{L^m(\Omega)} &\leq \left\| \sup_{s\in[0,t]} \|e^{s\Delta}(u_0^N - u_0)\|_{L^1\cap L^\infty(\mathbb{R}^d)} \right\|_{L^m(\Omega)} \\ &\quad + C \left(N^{-\alpha\boldsymbol{\zeta}} + N^{-\alpha(\beta-\frac{d}{\boldsymbol{r}})} + N^{-\frac{1}{2}(1-\alpha(d+[2\vee d]))+\varepsilon} \right). \end{aligned}$$

3.3 Corollaries of Theorem 1.4

In Section 3.4, using an interpolation inequality between the results of Proposition 3.1 and Theorem 1.4, we obtain the following rate of convergence with respect to Sobolev norms:

Corollary 3.2. *Let the same assumptions as in Theorem 1.4 hold. Let ϱ be as in Theorem 1.4. Then, for any $\varepsilon > 0$ and any $m \geq 1$, there exists a constant $C > 0$ such that for all $N \in \mathbb{N}^*$,*

$$\left\| \sup_{t\in[0,T]} \|u_t^N - u_t\|_{\gamma,\boldsymbol{r}-\delta} \right\|_{L^m(\Omega)} \leq C \left(\left\| \sup_{s\in[0,T]} \|e^{s\Delta}(u_0^N - u_0)\|_{L^1\cap L^\infty(\mathbb{R}^d)} \right\|_{L^m(\Omega)} + N^{-\varrho+\varepsilon} \right)^{\frac{\gamma}{\beta}},$$

for $\delta \in (0,1)$ and $\gamma = \beta \frac{\boldsymbol{r}(\boldsymbol{r}-1-\delta)}{(\boldsymbol{r}-\delta)(\boldsymbol{r}-1)}$.

Remark 3.3. *It is clear that $\gamma < \beta$. It will also be important (in particular for the propagation of chaos in the next section) to ensure that $\gamma > \frac{d}{\boldsymbol{r}-\delta}$ so as to have an embedding in a space of Hölder continuous functions. This is indeed the case if δ is chosen small enough, see condition* (3.27)*.*

The following corollary is a direct consequence of Theorem 1.4, Corollary 3.2 and of Borel-Cantelli's lemma (see e.g. Lemma 2.1 in [34]):

Corollary 3.4. *Let the same assumptions as in Theorem 1.4 hold. Let ϱ be as in Theorem 1.4, and γ, δ be as in Corollary 3.2. Let $m \geq 1$ and assume further that*

$$\left\| \sup_{t\in[0,T]} \|e^{t\Delta}(u_0^N - u_0)\|_{L^1\cap L^\infty(\mathbb{R}^d)} \right\|_{L^m(\Omega)} \lesssim N^{-\varrho}.$$

Then for any $\varepsilon \in (0, \varrho)$, there exist random variables $X_1, X_2 \in L^m(\Omega)$ such that, almost surely,

$$\forall N \in \mathbb{N}^*, \quad \sup_{t\in[0,T]} \|u_t^N - u_t\|_{L^1\cap L^\infty(\mathbb{R}^d)} \leq \frac{X_1}{N^{\varrho-\varepsilon}} \quad \textit{and} \quad \sup_{t\in[0,T]} \|u_t^N - u_t\|_{\gamma,\boldsymbol{r}-\delta} \leq \frac{X_2}{N^{(\varrho-\varepsilon)\frac{\gamma}{\beta}}}.$$

Remark 3.5. *By a classical embedding recalled in Section 2.1, the results of Corollaries 3.2 and 3.4 imply the same rates in η-Hölder norm, with $\eta = \gamma - \frac{d}{\boldsymbol{r}-\delta}$, provided that this quantity is positive (see condition* (3.27)*).*

In view of the previous results, we also obtain a rate of convergence for the genuine empirical measure, which can be interpreted as propagation of chaos for the marginals of the empirical measure of the particle system. Following [4, Section 8.3], let us introduce the Kantorovich-Rubinstein metric which reads, for any two probability measures μ and ν on $\mathbb{R}^d$,

$$\|\mu - \nu\|_0 = \sup\left\{\int_{\mathbb{R}^d} \phi\, d(\mu - \nu)\,;\ \phi \text{ Lipschitz with } \|\phi\|_{L^\infty(\mathbb{R}^d)} \leq 1 \text{ and } \|\phi\|_{\text{Lip}} \leq 1\right\}. \tag{3.24}$$

Note that this distance metrizes the weak convergence of probability measures ([4, Theorem 8.3.2]).

Corollary 3.6. *Let ϱ be as in Theorem 1.4 and $m \geq 1$. Let the same assumptions as in Corollary 3.4 hold. Then for any $\varepsilon \in (0, \varrho)$, there exists $C > 0$ such that, for any $N \in \mathbb{N}^*$,*

$$\left\| \sup_{t\in[0,T]} \|\mu_t^N - u_t\|_0 \right\|_{L^m(\Omega)} \leq C\, N^{-\varrho+\varepsilon}.$$

As in Corollary 3.4, this gives an almost sure rate of convergence for $\sup_{t\in[0,T]} \|\mu_t^N - u_t\|_0$.

The proof is given in Section 3.5.

Remark 3.7. *Observe that the previous result implies the convergence in law of μ_t^N to u_t for a fixed $t \in (0, T_{max})$, which is equivalent to saying that the law of $(X_t^1, \dots, X_t^N)$ is u_t-chaotic (in the sense of [54, Def. 2.1]).*

In view of the three previous corollaries, let us make a remark on the best possible rate that one can hope for in this setting.

Remark 3.8. *Assuming that the kernel K is such that $\boldsymbol{\zeta}$ can be chosen equal or close to 1, the rate is really determined by $\alpha(\beta - \frac{d}{\boldsymbol{r}})$ and $\frac{1}{2}\left(1 - \alpha(d + [2 \vee d])\right)$. Then the supremum of $\alpha(\beta - \frac{d}{\boldsymbol{r}})$ under the constraint that $\alpha < (d + 2\beta + 2d(\frac{1}{2} - \frac{1}{\boldsymbol{r}}))^{-1}$ (see $(\mathbf{A}_\alpha)$) and $\beta < 1$ (see $(\mathbf{A}_{iii}^K)$) is equal to*

$$\frac{1}{2(d+1)},$$

which is approximated (but not attained) for β close to 1, $\boldsymbol{r}$ large enough and α close to $\frac{1}{d+2\beta+2d(\frac{1}{2}-\frac{1}{\boldsymbol{r}})}$ (which is then also close to $\frac{1}{2(d+1)}$). On the other hand, still assuming that β is close to 1 and that $d \geq 2$, one has $\frac{1}{2}\left(1 - \alpha(d + [2 \vee d])\right) \geq \frac{1}{2(d+1)}$. Hence the supremum for the rate (in the case $d \geq 2$) is $\frac{1}{2(d+1)}$ (not reached).

So when $d \geq 2$, the best possible rate is almost α, which in view of the discussion of Remark 1.6, is optimal (what might not be optimal is the constraint $(\mathbf{A}_\alpha)$ on α).

If $d = 1$, then by choosing again (whenever possible) $\beta \approx 1$ and $\boldsymbol{r}$ very large, then $\frac{1}{2}\left(1 - \alpha(d + [2 \vee d])\right)$ might be smaller than $\frac{1}{2(d+1)} = \frac{1}{4}$. Thus in that case, the best possible rate is attained when $\alpha = \frac{1}{2}\left(1 - \alpha(d + [2 \vee d])\right) = \frac{1}{2}(1 - 3\alpha)$. Hence the best possible rate when $d = 1$ is $\frac{1}{5}$ (not reached).

Finally, observing in Corollaries 3.2 and 3.4 that a rate is obtained in $H^{\gamma}_{\boldsymbol{r}-\delta}$, one could wonder if a convergence also happens in $H^{\beta}_{\boldsymbol{r}}$. Corollary 1.5 answers positively, thus extending the main convergence result in [20, 44] to general kernels.

3.4 Rate in Sobolev norm: Proof of Corollary 3.2

This result relies entirely on an interpolation inequality for Bessel potential spaces, and our previous results of convergence, and convergence with a rate.

Let us establish first the interpolation inequality that we shall use: let $\delta \in (0,1)$ and γ such that

$$\gamma = \beta \frac{\boldsymbol{r}(\boldsymbol{r}-\delta-1)}{(\boldsymbol{r}-\delta)(r-1)}. \tag{3.25}$$

The interpolation theorem for Bessel potential spaces, see [57, p.185], gives that for any $f \in H_1^0 \cap H_{\boldsymbol{r}}^\beta(\mathbb{R}^d)$ $(\equiv L^1 \cap H_{\boldsymbol{r}}^\beta(\mathbb{R}^d))$,

$$\|f\|_{\gamma,\boldsymbol{r}-\delta} \leq \|f\|_{0,1}^{\theta} \, \|f\|_{\beta,\boldsymbol{r}}^{1-\theta}, \tag{3.26}$$

where $\theta = \frac{\gamma}{\beta}$.

Hence it follows from (3.26) that for any $m \geq 1$,

$$\mathbb{E} \sup_{s\in[0,T]} \|u_s^N - u_s\|_{\gamma,\boldsymbol{r}-\delta}^m \leq \mathbb{E} \sup_{s\in[0,T]} \|u_s^N - u_s\|_{0,1}^{\theta m} \, \|u_s^N - u_s\|_{\beta,\boldsymbol{r}}^{(1-\theta)m},$$

and we deduce from Hölder's inequality that

$$\mathbb{E} \sup_{s\in[0,T]} \|u_s^N - u_s\|_{\gamma,\boldsymbol{r}-\delta}^m \leq \left(\mathbb{E} \sup_{s\in[0,T]} \|u_s^N - u_s\|_{L^1(\mathbb{R}^d)}^m \right)^{\theta} \left(\mathbb{E} \sup_{s\in[0,T]} \|u_s^N - u_s\|_{\beta,\boldsymbol{r}}^m \right)^{1-\theta}.$$

In view of the previous inequality and using Theorem 1.4 and Proposition 3.1, we deduce the rate of convergence in $L^m\left(\Omega; L^\infty([0,T], H_{\boldsymbol{r}-\delta}^\gamma(\mathbb{R}^d))\right)$.

Finally, note that it is always true that $\gamma < \beta$. Besides, it will be important to ensure that $\gamma > \frac{d}{\boldsymbol{r}-\delta}$ to have an embedding in the space of Hölder continuous functions (see (2.1)). For this, it suffices to choose δ which satisfies:

$$\frac{\boldsymbol{r}-\delta-1}{\boldsymbol{r}-1} > \frac{d/\boldsymbol{r}}{\beta}. \tag{3.27}$$

3.5 Propagation of chaos for the marginals: Proof of Corollary 3.6

Let $t \in (0, T_{max})$. Let us observe first that there exists $C > 0$ such that for any Lipschitz continuous function ϕ on $\mathbb{R}^d$, one has

$$\left| \langle u_t^N, \phi \rangle - \langle \mu_t^N, \phi \rangle \right| \leq \frac{C\|\phi\|_{\text{Lip}}}{N^\alpha} \quad a.s. \tag{3.28}$$

Indeed,

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

Recalling the definition (3.24) of the Kantorovich-Rubinstein distance, it comes

$$\begin{aligned} \left\| \sup_{t\in[0,T]} \|\mu_t^N - u_t\|_0 \right\|_{L^m(\Omega)} &\leq \left\| \sup_{t\in[0,T]} \|\mu_t^N - u_t^N\|_0 \right\|_{L^m(\Omega)} + \left\| \sup_{t\in[0,T]} \sup_{\|\phi\|_{L^\infty}\leq 1} \langle u_t^N - u_t, \phi \rangle \right\|_{L^m(\Omega)} \\ &\leq \left\| \sup_{t\in[0,T]} \|\mu_t^N - u_t^N\|_0 \right\|_{L^m(\Omega)} + \left\| \sup_{t\in[0,T]} \|u_t^N - u_t\|_{L^1(\mathbb{R}^d)} \right\|_{L^m(\Omega)}. \end{aligned}$$

Now applying Inequality (3.28) to the first term on the right-hand side of the above inequality, and Theorem 1.4 to the second term, we obtain the inequality of Corollary 3.6.

3.6 Convergence in $H^\beta_{\boldsymbol{r}}$: Proof of Corollary 1.5

Let us introduce the space

$$\mathcal{X} = L^2\left([0,T]; H^\beta_{\boldsymbol{r}}(\mathbb{R}^d)\right)$$

with the strong topology, and denote by $\mathcal{X}_w$ the same space endowed with the weak topology. Notice first that $L^2\big([0,T]; H^\beta_{\boldsymbol{r}}(\mathbb{R}^d)\big)$ is a reflexive Banach space for $1 < \boldsymbol{r} < \infty$ (see [57, p.198-199]). Hence by the Banach-Alaoglu theorem, it is compactly embedded in $\mathcal{X}_w$.

Now the Chebyshev inequality ensures that

$$\mathbb{P}\big(\|u^N\|^2_{\mathcal{X}} > R\big) \leq \frac{\mathbb{E}\big[\big\|u^N\big\|^2_{\mathcal{X}}\big]}{R} \leq \frac{\sup_{t\in[0,T]} \mathbb{E}\big[\big\|u^N_t\big\|^2_{\beta,\boldsymbol{r}}\big]}{R}, \quad \text{for any } R > 0.$$

Thus by Proposition 3.1, we obtain

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

Let $\mathbb{P}_N$ be the law of u^N in $\mathcal{X}$. The last inequality implies that for any $\epsilon > 0$, there exists a bounded set $B_\epsilon \in \mathcal{X}$ such that $\mathbb{P}_N(B_\epsilon) < 1-\epsilon$ for all N, and therefore by the preliminary remark, there exists a compact set $\mathcal{K}_\epsilon \in \mathcal{X}_w$ such that $\mathbb{P}_N(\mathcal{K}_\epsilon) < 1-\epsilon$. That is, $(\mathbb{P}_N)_{N\geq 1}$ is tight on $\mathcal{X}_w$.

Here, we cannot apply the usual Prokhorov Theorem, since $\mathcal{X}_w$ is not metrisable. However, $\mathcal{X}$ is reflexive and $\mathcal{X}'$ is separable, therefore any weak compact set in $\mathcal{X}_w$ is metrisable (see e.g. [49, Thm 3.16]). In addition, $\mathcal{X}_w$ is completely regular in the sense of [4, Def. 6.1.2] (as is any Hausdorff topological vector space). Hence, by Theorem 8.6.7 of [4], there exists a subsequence of probability measures $\mathbb{P}_{\epsilon_N}$ that converges weakly to some $\mathbb{P}_\infty$ on $\mathcal{X}_w$. In particular, if $F \in \mathcal{C}^0_b(\mathbb{R})$ and $\varphi \in L^2\left([0,T]; H^{-\beta}_{\boldsymbol{r}'}(\mathbb{R}^d)\right)$, we get that

$$\mathbb{E}F\left(\langle u^{\epsilon_N}, \varphi\rangle\right) \to \mathbb{E}F\left(\langle u^\infty, \varphi\rangle\right),$$

for some $\mathcal{X}_w$-valued random variable u^∞ with law $\mathbb{P}_\infty$.

On the other hand, from Theorem 1.4 we know that u^N converges almost surely to u in $L^\infty([0,T], L^1 \cap L^\infty(\mathbb{R}^d))$. By testing against any function $\varphi \in \mathcal{C}^\infty\left([0,T]\times\mathbb{R}^d\right)$, we deduce that $\mathbb{E}F\left(\langle u^N, \varphi\rangle\right) \to F\left(\langle u, \varphi\rangle\right)$ for any $F \in \mathcal{C}^0_b(\mathbb{R})$. Thus by uniqueness of the limit, we get that $\mathbb{E}F\left(\langle u^\infty, \varphi\rangle\right) = F\left(\langle u, \varphi\rangle\right)$. Since $\mathcal{C}^\infty\left([0,T]\times\mathbb{R}^d\right)$ is dense in $L^2\left([0,T]; H^{-\beta}_{\boldsymbol{r}'}(\mathbb{R}^d)\right)$, it follows that $\mathbb{P}_\infty$ is a Dirac measure at u.

We have thus obtained that any limit point of $(\mathbb{P}_N)_{N\geq 1}$ is the Dirac measure at u, therefore $(u^N)_{N\geq 1}$ converges in law to u in $\mathcal{X}_w$. Since the limit is deterministic, the convergence also holds in probability, in the sense of Corollary 1.5.

4 Propagation of chaos

In this section we study the well-posededness of the nonlinear SDE (1.4) and then the propagation of chaos of the particle system (1.9). More precisely, we prove Proposition 1.8 about the well-posedness of the martingale problem $(\mathcal{MP})$ related to (1.4). Then, we prove the convergence in law, when $N \to \infty$, of the empirical measure μ^N of the particle system (1.9) towards the unique solution of the martingale problem $(\mathcal{MP})$.

4.1 Proof of Proposition 1.8

Let $T < T_{max}$ and let u be the unique mild solution to (1.1) up to T.

The proof is organized as follows. For a solution 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 $(\mathcal{MP})$.

Let $\mathbb{Q}$ be a solution to $(\mathcal{MP})$. Notice first that as the family of marginal laws $(q_t)_{t\leq T}$ belongs to $q \in \mathcal{C}([0,T]; L^1 \cap L^\infty(\mathbb{R}^d))$, one has according to Lemma 2.1 that

$$\sup_{t\leq T} \|K * q_t\|_{L^\infty(\mathbb{R}^d)} \leq C_{K,d} \sup_{t\leq T} \|q_t\|_{L^1\cap L^\infty(\mathbb{R}^d)}. \tag{4.1}$$

To obtain the equation satisfied by $(q_t)_{t\leq T}$, 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. [55, Section 4]). One has

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

This equation is exactly (1.1) and we know it admits a unique mild solution in the sense of Definition 1.1 up to T. Meaning, as $q \in \mathcal{C}([0,T]; L^1 \cap L^\infty(\mathbb{R}^d))$, 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 u to (1.1) in the definition of the process $(M_t)_{t\leq T}$ from $(\mathcal{MP})$.

By Girsanov transformation, the equation

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

admits a weak solution. In addition, strong uniqueness holds (see [39, Remark 1.6]). Thus, 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 $(\widetilde{u}_t)_{0\leq t\leq T}$.

To do so, for $0 < t \leq T$, one derives the mild equation for $\mathbb{P}_t(dx) = p_t(x)dx$ (absolute continuity follows from bounded drift and Girsanov transformation). Following the same arguments as in [55, Section 4], as the drift is bounded, we have that for a.e. $x \in \mathbb{R}^d$,

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

Assume for a moment that $p \in \mathcal{C}([0,T]; L^1 \cap L^\infty(\mathbb{R}^d))$. The previous equation is a linearized version of Eq. (1.11) and with the same arguments as in Proposition 2.6, its cut off version admits a unique solution in $\mathcal{C}([0,T]; L^1 \cap L^\infty(\mathbb{R}^d))$. Then, by the same arguments as in Proposition 1.2, the above equation admits a unique solution in $\mathcal{C}([0,T]; L^1 \cap L^\infty(\mathbb{R}^d))$. Since both u and p solve this equation, they must coincide and we have the desired result : $(p_t)_{t\in[0,T]} = (u_t)_{t\in[0,T]}$.

It only remains to prove that $p \in \mathcal{C}([0,T]; L^1 \cap L^\infty(\mathbb{R}^d))$. Obviously, as we work with a family of probability density functions, we only need to prove that $p \in \mathcal{C}([0,T]; L^\infty(\mathbb{R}^d))$. Performing the same calculations as in the proof of Proposition 2.6, we get that

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

In view of Lemma 2.1, one has

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

One gets that

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

Grönwall's lemma implies that $p_t \in L^\infty(\mathbb{R}^d)$. Repeat the above computations for $p_t - p_s$ in place of p_t to conclude that $p \in \mathcal{C}([0,T]; L^\infty(\mathbb{R}^d))$. This concludes the proof.

4.2 Proof of Theorem 1.9

To prove Theorem 1.9, we will show that μ^N converges to the unique solution $\mathbb{Q}$ of the martingale problem $(\mathcal{MP})$. To do so, we will first prove the convergence towards an auxiliary martingale problem which is identical to $(\mathcal{MP})$ except that in the point *iii)* the process $(M_t)_{t\leq T}$ is the following:

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

Then, we will lift the cut-off F_A as A will be chosen large enough. Let us call this auxiliary martingale problem $(\mathcal{MP}_A)$ and denote its unique solution by $\mathbb{Q}$ by a slight abuse of notation.

A usual way to prove that μ^N converges to $\mathbb{Q}$ 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 $(M_t)_{t\leq T}$ 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

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

endowed with the weak topology of $\mathcal{P}(\mathcal{C}([0,T];\mathbb{R}^d))$ and the topology of $\mathcal{Y}$, where

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

(Recall from Section 3.6 that $\mathcal{X}_w$ is the space $L^2([0,T];H^\beta_{\boldsymbol{r}}(\mathbb{R}^d))$ endowed with the weak topology.)

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 $\mathcal{Y}$ (see Theorem 1.4 and Corollary 1.5).

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 4.1. *$\tilde{\Pi}^\infty$-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 Inequality (3.28). Take a test function $\varphi \in \mathcal{C}_c^\infty([0,T]\times\mathbb{R}^d)$ and define a functional $\phi(t,x) = \varphi(t,x_t)$, for $x \in \mathcal{C}([0,T];\mathbb{R}^d)$. Then,

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

where the last inequality comes from (3.28). 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 1.9.

Proposition 4.2. *Let $p \in \mathbb{N}$, $f \in \mathcal{C}_b^2(\mathbb{R}^d)$, $\Phi \in \mathcal{C}_b(\mathbb{R}^{dp})$ and $0 < s_1 < \cdots < s_p \leq s < t \leq T$. Define Γ as the following function on $\mathcal{H}$:*

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

Then

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

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

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

Step 2. We prove that Γ 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 $\lim_{n\to\infty} |\Gamma(\boldsymbol{\mu}^n, \mathbf{u}^n) - \Gamma(\boldsymbol{\mu}, \mathbf{u})| = 0$.

We decompose

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

Notice that

$$\begin{aligned} I_n &\leq \|\Phi\|_\infty \|\nabla f\|_\infty \langle \boldsymbol{\mu}^n, \int_s^t |F(K * \mathbf{u}_r^n(\cdot_r)) - F(K * \mathbf{u}_r(\cdot_r))| dr \rangle \\ &\leq C \int_s^t \langle \boldsymbol{\mu}_r^n, |K * (\mathbf{u}_r^n - \mathbf{u}_r)| \rangle \, dr \leq C \int_s^t \|K * (\mathbf{u}_r^n - \mathbf{u}_r)\|_\infty dr. \end{aligned} \tag{4.2}$$

In view of Lemma 2.1, one has

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

Now recall $\boldsymbol{r}$ and β are fixed in $(\mathbf{A}_{iii}^K)$, and let γ and δ satisfy (3.25) and (3.27), so that $\frac{d}{\boldsymbol{r}-\delta} < \gamma < \beta$. Then, use the Sobolev embedding $L^1 \cap H^\gamma_{\boldsymbol{r}-\varepsilon}(\mathbb{R}^d) \subset L^1 \cap L^\infty(\mathbb{R}^d)$ to get

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

Plug the latter in (4.2) to obtain

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

By the interpolation inequality (3.26),

$$\|\mathbf{u}_r^n - \mathbf{u}_r\|_{L^1 \cap H^\gamma_{\boldsymbol{r}-\varepsilon}(\mathbb{R}^d)} \leq \|\mathbf{u}_r^n - \mathbf{u}_r\|_{L^1(\mathbb{R}^d)} + \|\mathbf{u}_r^n - \mathbf{u}_r\|^\theta_{L^1(\mathbb{R}^d)} \|\mathbf{u}_r^n - \mathbf{u}_r\|^{1-\theta}_{H^\beta_{\boldsymbol{r}}(\mathbb{R}^d)},$$

for θ as in Section 3.4. Now since $\mathbf{u}^n$ converges in $\mathcal{Y}$, and converges in particular weakly in $L^2\left([0,T], H^{\beta}_{\boldsymbol{r}}(\mathbb{R}^d)\right)$, the uniform boundedness principle tells us that it is bounded in this space. Gathering this fact with the convergence in $L^\infty([0,T];L^1(\mathbb{R}^d))$ (by assumption), the previous inequality yields the convergence of $\mathbf{u}^n$ in $L^2\left([0,T], L^1\cap H^{\gamma}_{\boldsymbol{r}-\varepsilon}(\mathbb{R}^d)\right)$. Hence I_n converges to 0.

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 $G : C([0,T];\mathbb{R}^d) \to \mathbb{R}$ 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 $G(x^n)\to G(x)$ as $n\to\infty$, having in mind the properties of f and Φ, we should only concentrate on the term $\int_s^t F_A(K*\mathbf{u}_r(x^n_r))\cdot\nabla f(x^n_r)\,dr$. Here we use the continuity property $(\mathbf{A}^K_{iii})$ to deduce that $K*\mathbf{u}_r(x^n_r)$ converges to $K*\mathbf{u}_r(x_r)$ and by dominated convergence,

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

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

We have all the elements in hand to finish the proof of Theorem 1.9. By Lemma 4.1 and Proposition 4.2, we get that $\boldsymbol{\mu}\in\operatorname{supp}(\Pi^\infty)$ solves the nonlinear martingale problem $(\mathcal{MP}_A)$. Choose $A > A_T := C_{k,D}\|q\|_{T,L^1\cap L^\infty(\mathbb{R}^d)}$ and lift the cut-off (see (4.1)). Then, $\boldsymbol{\mu}$ 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 $\delta_{\mathbb{Q}}$.

5 Examples

5.1 A stronger, easier-to-check condition on the kernel

Assume that K satisfies $(\mathbf{A}^K_i)$ and $(\mathbf{A}^K_{ii})$. Assume further

> $(\widetilde{\mathbf{A}}^K_{iii})$ There exists $\boldsymbol{p}\in(d,+\infty]\cap[\boldsymbol{q},+\infty]$, $\boldsymbol{r}\in(d\vee 2,+\infty)$ and $\beta\in(\frac{d}{\boldsymbol{r}},1)$ such that the matrix-valued kernel ∇K defines a convolution operator which is bounded component-wise from $L^1(\mathbb{R}^d)\cap H^{\beta}_{\boldsymbol{r}}(\mathbb{R}^d)$ to $L^{\boldsymbol{p}}(\mathbb{R}^d)$.

We will show that if K satisfies $(\widetilde{\mathbf{A}}^K_{iii})$, then it satisfies $(\mathbf{A}^K_{iii})$. In the examples below, this new assumption $(\widetilde{\mathbf{A}}^K_{iii})$ will be easier to check.

First, we will make use of $(\mathbf{A}^K_i)$ and $(\mathbf{A}^K_{ii})$. Young's convolution inequality states that for $q_0 = (1+\frac{1}{\boldsymbol{p}}-\frac{1}{\boldsymbol{q}})^{-1}$ (in $(\widetilde{\mathbf{A}}^K_{iii})$, we assume that $\boldsymbol{p}\geq\boldsymbol{q}$, hence $q_0\geq 1$), there is for any $f\in L^{\boldsymbol{p}}\cap L^{q_0}(\mathbb{R}^d)$,

$$\begin{aligned}\|K*f\|_{L^{\boldsymbol{p}}(\mathbb{R}^d)} &\leq \|(\mathbb{1}_{\mathcal{B}_1}K)*f\|_{L^{\boldsymbol{p}}(\mathbb{R}^d)} + \left\|\left(\mathbb{1}_{\mathcal{B}_1^c}K\right)*f\right\|_{L^{\boldsymbol{p}}(\mathbb{R}^d)}\\ &\leq \|\mathbb{1}_{\mathcal{B}_1}K\|_{L^1(\mathbb{R}^d)}\|f\|_{L^{\boldsymbol{p}}(\mathbb{R}^d)} + \left\|\mathbb{1}_{\mathcal{B}_1^c}K\right\|_{L^{\boldsymbol{q}}(\mathbb{R}^d)}\|f\|_{L^{q_0}(\mathbb{R}^d)}\\ &\leq C_K\left(\|f\|_{L^{\boldsymbol{p}}(\mathbb{R}^d)} + \|f\|_{L^{q_0}(\mathbb{R}^d)}\right).\end{aligned}$$

In particular the previous inequality holds true if $f\in L^1\cap H^{\beta}_{\boldsymbol{r}}(\mathbb{R}^d)$, because by Sobolev embedding, f is in $L^1\cap L^\infty$ (see embedding (2.2)), and then the result holds by interpolation. Now in view of the previous fact and using the property $(\widetilde{\mathbf{A}}^K_{iii})$ of ∇K, one deduces that if $f\in L^1\cap H^{\beta}_{\boldsymbol{r}}(\mathbb{R}^d)$, then $K*f\in H^1_{\boldsymbol{p}}(\mathbb{R}^d)$. Hence it follows from Morrey's inequality [7, Th. 9.12] that there exists $C_{\boldsymbol{p},d}>0$ such that for any $f\in L^1\cap H^{\beta}_{\boldsymbol{r}}(\mathbb{R}^d)$,

$$\begin{aligned}\mathcal{N}_\eta\left(K*f\right) &\leq C\|K*f\|_{H^1_{\boldsymbol{p}}(\mathbb{R}^d)}\\ &\leq C_{\boldsymbol{p},d,K}\|f\|_{L^1\cap H^{\beta}_{\boldsymbol{r}}(\mathbb{R}^d)},\end{aligned}$$

where $\eta = 1 - \frac{d}{\boldsymbol{p}}$. Hence K satisfies ($\mathbf{A}^K_{iii}$) with $\boldsymbol{\zeta} = 1 - \frac{d}{\boldsymbol{p}}$.

5.2 General classes of kernels satisfying Assumption ($\mathbf{A}^K$)

The first two points of Assumption ($\mathbf{A}^K$) are simple technical conditions and may not require specific comments, except that it would be interesting to lift the first integrability condition in order to be able to consider more singular kernels. The third assumption is much more interesting. Let us start with the simple example of a kernel K such that ∇K is integrable. Then ∇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 (2.2), ∇K satisfies ($\widetilde{\mathbf{A}}^K_{iii}$) for any $\boldsymbol{p} \in [1, \infty]$ and any β and $\boldsymbol{r}$ such that $\beta - \frac{d}{\boldsymbol{r}} > 0$. Hence it satisfies ($\mathbf{A}^K_{iii}$) with $\boldsymbol{\zeta} = 1$ (see Section 5.1).

Let us now look at a more singular example. We will discuss further in the next paragraph the Coulomb potential, defined as

$$V_C(x) := \begin{cases} |x|^{-(d-2)} & \text{if } d \geq 3 \\ -\log|x| & \text{if } d = 1, 2 \end{cases}, \quad x \in \mathbb{R}^d.$$

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 Schwartz 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 [26, 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 ($\mathbf{A}^K$) with $\boldsymbol{\zeta} = 1 - \frac{d}{\boldsymbol{p}}$ (see Section 5.1), for any $\boldsymbol{p} \in (1, \infty)$ and any β and $\boldsymbol{r}$ such that $\beta - \frac{d}{\boldsymbol{r}} > 0$.

5.3 Riesz and Coulomb potentials

The Coulomb potential belongs to a more general class of interaction potentials, called Riesz potentials, which were defined in (1.2). 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 ($\mathbf{A}^K$), provided that $d \geq 2$ and $s \in [0, d-1)$: Indeed,

- These kernels are locally integrable if and only if $0 \leq s < d - 1$.
- K_s is integrable outside the unit ball for any $q > \frac{d}{s+1}$.
-
  - If $d \geq 3$ and $s < d - 2$, then ∇K_s is not bounded in any L^p but it is however bounded from L^p to L^q whenever $p \in (1, \frac{d}{d-(s+2)})$ and $\frac{1}{q} = \frac{1}{p} + \frac{s+2}{d} - 1$ (see [50, Theorem 25.2]). This is still enough, thanks to the embedding (2.2), to ensure property ($\widetilde{\mathbf{A}}^K_{iii}$). In particular, one can choose any $\boldsymbol{p} \in (1, \frac{d}{d-(s+2)})$ and any β and $\boldsymbol{r}$ such that $\beta - \frac{d}{\boldsymbol{r}} > 0$ for these kernels. In particular ($\mathbf{A}^K_{iii}$) holds for $\boldsymbol{\zeta} = 1 - \frac{d}{\boldsymbol{p}}$ (see Section 5.1). Hence all our results can be applied to this kernel.
  - If $s = d - 2$, then ∇K_s is a typical kernel satisfying the conditions of [26, 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, one can choose $\boldsymbol{\zeta} = 1 - \frac{d}{\boldsymbol{p}}$ for any $\boldsymbol{p} \in (1, \infty)$, and any β and $\boldsymbol{r}$ such that $\beta - \frac{d}{\boldsymbol{r}} > 0$.

    In particular, Theorem 1.4 and its corollaries 3.2, 3.4 and 3.6 are applicable, and choosing $\boldsymbol{q}$ and $\boldsymbol{p}$ very large, one can obtain with an appropriate choice of α and β (see Remark 3.8) a rate which is as closed as desired to $\frac{1}{2(d+1)}$.

– If $s \in [d-2, d-1)$, then one can verify (see e.g. [15, Lemma 2.5]) that ∇K_s defines a convolution operator from $L^1 \cap L^\infty \cap \mathcal{C}^\sigma(\mathbb{R}^d)$ to $L^\infty(\mathbb{R}^d)$, with $\sigma \in (2-d+s, 1)$:

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

for some β and $\boldsymbol{r}$ such that $\sigma = \beta - \frac{d}{\boldsymbol{r}}$, thanks to the embedding (2.1). Hence, one must choose $\boldsymbol{p} = \infty$, and $\beta < 1$ and $\boldsymbol{r}$ such that $\beta - \frac{d}{\boldsymbol{r}} > 2 - d + s$.

Theorem 1.4 and its corollaries 3.2, 3.4 and 3.6 are applicable, choosing $\boldsymbol{q} = \boldsymbol{p} = \infty$ is allowed, so with an appropriate choice of α and β (see Remark 3.8), one obtains a rate which is as close as desired to $\frac{1}{2(d+1)}$.

Besides obtaining rates of convergence, Proposition 1.8 proves the well-posedness of the McKean-Vlasov SDE (1.4) for all Riesz kernels with $s \in (0, d-1)$, which is new for the values of the parameter s, and most notably for the largest values $s \geq d-2$. The trajectorial propagation of chaos (Theorem 1.9) is also new for this whole class of particle systems.

5.4 Parabolic-elliptic Keller-Segel models

An important and tricky example covered by this paper is the Keller-Segel PDE, which takes the form (1.1) with the kernel defined by

$$K_{KS}(x) = -\chi \frac{x}{|x|^d}, \tag{5.1}$$

for some $\chi > 0$.

The difficulty in this model comes from the fact that the kernel is attractive, in the following sense:

$$x \cdot K(x) < 0, \quad \text{on the domain of definition of } K. \tag{5.2}$$

This leads to important issues that we discuss in more details in [44], 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 [44]).

Theorem 1.4, Corollaries 3.2, 3.4 and 3.6 give rates of convergence of the particle system to the Keller-Segel PDE, even for solutions that exhibit a blow-up.

As a consequence of Theorem 1.8, we also deduce the local-in-time weak well-posedness of the McKean-Vlasov (1.4) for all values of the concentration parameter χ, and global-in-time weak well-posedness for $\chi < 8\pi$, which is a new result (see [44] for a thorough discussion and comparison with previous results).

5.5 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

$$x \cdot K(x) \geq 0, \quad \text{on the domain of definition of } K. \tag{5.3}$$

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 $C\left(1 + \|\xi_0\|_{L^\infty(\mathbb{R}^2)}\right)$ (see [20] and references therein). This permits to choose the cut-off value A_T from (1.12) independently of T.

The kernel K_{BS} is covered by our assumption ($\mathbf{A}^K$), for the same reason as the Coulomb kernel with $d = 2$ ($K_0(x) = \frac{1}{|x|}$). In that case, we recover and extend with a rate the Theorem 1.3 of Flandoli et al. [20] within Theorem 1.4 and Corollary 1.5.

All the other results of this paper apply, and in particular, if the initial condition is smooth enough (i.e. $\beta \approx 1$ for Theorem 1.4), we obtain a rate ρ in $L^1 \cap L^\infty$ norm which is almost $\frac{1}{6}$.

5.6 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 (5.2) on a subdomain D of the domain of definition of K, and repulsive (i.e. satisfying (5.3)) 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

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

for some $V_0 > 0$. Then $K(x) = \nabla V(x)$ satisfies the first condition of ($\mathbf{A}^K$) (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. [19] to model the adhesion of cells in biology. It can be expressed in general as

$$V(r) = V_a\, r^{-a} - V_b\, r^{-b},$$

with $a, b > 0$ and $V_a, V_b > 0$. One can now refer to the discussion on Riesz kernels in Section 5.3 to determine the values of a and b that ensure the applicability of our results.

Appendix

A.1 Time and space estimates of the stochastic convolution integrals

In this section we study the moments of the supremum in time of $\|M_t^N\|_{L^p(\mathbb{R}^d)}$, where the stochastic convolution integral M_t^N was defined in (3.5). Such estimates will be used in the proof of Theorem 1.4 for $p = 1$ and $p = \infty$ only, but the result is established for any p without any additional effort.

Proposition A.1. *Let the assumption of Theorem 1.4 hold. Let $m_0 \geq 1$, $p \in [1, +\infty]$ and $\varepsilon > 0$. Then there exists $C > 0$ such that for any $t \in [0, T]$ and $N \in \mathbb{N}^*$,*

$$\left\| \sup_{s \in [0,t]} \|M_s^N\|_{L^p(\mathbb{R}^d)} \right\|_{L^{m_0}(\Omega)} \leq C\, N^{-\frac{1}{2}(1 - \alpha(d + \varkappa)) + \varepsilon},$$

where $\varkappa = \max\left(2, d(1 - \frac{2}{p})\right)$.

The proof of Proposition A.1 relies on the following proposition, which we prove at the end of this section:

Proposition A.2. *For any $p \in [1, \infty]$, any $m \geq 1$ and any $\delta \in (0, 1]$, there exists $C > 0$ such that*

$$(i) \quad \left\| \|M_t^N\|_{L^p(\mathbb{R}^d)} \right\|_{L^m(\Omega)} \leq C\, N^{-\frac{1}{2}(1 - \alpha(d + 3\delta + \varkappa))}, \quad \forall t \in [0, T],\ \forall N \in \mathbb{N}^*,$$

$$(ii) \quad \left\| \|M_t^N - M_s^N\|_{L^p(\mathbb{R}^d)} \right\|_{L^m(\Omega)} \leq C\, (t - s)^{\frac{\delta}{2}}\, N^{-\frac{1}{2}(1 - \alpha(d + 5\delta + \varkappa))}, \quad \forall s \leq t \in [0, T],\ \forall N \in \mathbb{N}^*,$$

where $\varkappa$ was defined in Proposition A.1.

When $p \geq 2$, Proposition A.2 relies itself on the following result.

Proposition A.3. *Let $\gamma \geq 0$ and $m \geq 1$. For any $\delta \in (0,1]$, there exists $C>0$ such that*

$$(i) \qquad \big\| \|M_t^N\|_\gamma \big\|_{L^m(\Omega)} \leq C\, N^{-\frac{1}{2}(1-\alpha(d+2\delta+2\gamma))}, \quad \forall t \in [0,T],\ \forall N \in \mathbb{N}^*,$$

$$(ii) \qquad \big\| \|M_t^N - M_s^N\|_\gamma \big\|_{L^m(\Omega)} \leq C\,(t-s)^{\frac{\delta}{2}}\, N^{-\frac{1}{2}(1-\alpha(d+4\delta+2\gamma))}, \quad \forall s \leq t \in [0,T],\ \forall N \in \mathbb{N}^*.$$

The final ingredient in the proof of Proposition A.1 is a consequence of Garsia-Rodemich-Rumsey's Lemma [25], given in the following lemma (for $\mathbb{R}$-valued processes, it already appears in [46, Corollary 4.4], and the extension to Banach spaces is consistent with Garsia-Rodemich-Rumsey's Lemma with no additional difficulty, see e.g. [24, Theorem A.1]).

Lemma A.4. *Let E be a Banach space and $(Y^n)_{n\geq 1}$ be a sequence of E-valued continuous processes on $[0,T]$. Let $m \geq 1$ and $\eta > 0$ such that $m\eta > 1$ and assume that there exists constants $\rho > 0$, $C>0$, and a sequence $(\delta_n)_{n\geq 1}$ of positive real numbers such that*

$$\Big(\mathbb{E}\Big[\big\|Y_s^n - Y_t^n\big\|_E^m\Big]\Big)^{\frac{1}{m}} \leq C|s-t|^\eta\, \delta_n^\rho, \quad \forall s,t \in [0,T],\ \forall n \geq 1.$$

Then for any $m_0 \in (0,m]$, there exists a constant $C_{m,m_0,\eta,T} > 0$, depending only on m, m_0, η and T, such that $\forall n \geq 1$,

$$\left(\mathbb{E}\Big[\sup_{t\in[0,T]} \big\|Y_t^n - Y_0^n\big\|_E^{m_0}\Big]\right)^{\frac{1}{m_0}} \leq C_{m,m_0,\eta,T}\ \delta_n^\rho.$$

We are now ready to prove the main result of this section.

Proof of Proposition A.1. We aim to apply Lemma A.4 to M^N in the Banach space $L^p(\mathbb{R}^d)$, for some $p \in [1,+\infty]$. It comes

$$\Big\| \big\|M_s^N - M_t^N\big\|_{L^p(\mathbb{R}^d)} \Big\|_{L^m(\Omega)} \leq \Big\| \big\|M_s^N - M_t^N\big\|_{L^p(\mathbb{R}^d)} \Big\|_{L^m(\Omega)}^{\frac{1}{2}} \times \left(\Big\| \big\|M_s^N\big\|_{L^p(\mathbb{R}^d)} \Big\|_{L^m(\Omega)} + \Big\| \big\|M_t^N\big\|_{L^p(\mathbb{R}^d)} \Big\|_{L^m(\Omega)} \right)^{\frac{1}{2}}.$$

Then, applying Proposition A.2(ii) to the first term on the right-hand side of the previous inequality, and Proposition A.2(i) to the two others, it follows that, for some constant C independent of N, s and t,

$$\Big\| \big\|M_s^N - M_t^N\big\|_{L^p(\mathbb{R}^d)} \Big\|_{L^m(\Omega)} \ \leq\ C|s-t|^{\frac{\delta}{4}}\ N^{-\frac{1}{2}(1-\alpha(d+\varkappa))+2\alpha\delta}. \tag{A.1}$$

Let now $\varepsilon > 0$ and $m_0 > 0$, and choose δ such that $2\alpha\delta = \varepsilon$. Set $\eta = \frac{\delta}{4}$ and $\rho = -\frac{1}{2}(1-\alpha(d+\varkappa)) + 2\alpha\delta = -\frac{1}{2}(1-\alpha(d+\varkappa)) + \varepsilon$. Hence, choosing $m \geq 1 \vee m_0$ large enough so that $m\eta > 1$, the estimation in (A.1) satisfies the conditions in Lemma A.4, and it follows that, for some constant $C>0$,

$$\Big\| \sup_{t\in[0,T]} \big\|M_t^N\big\|_{L^p(\mathbb{R}^d)} \Big\|_{L^{m_0}(\Omega)} \leq C\ N^{-\frac{1}{2}(1-\alpha(d+\varkappa))+\varepsilon},$$

which is the desired result. □

It remains to prove Propositions A.3 and A.2.

Proof of Proposition A.3. Let us formulate some preliminary remarks. As the semigroup acts as a convolution in time within the stochastic integral, $\big(M_t^N\big)_{t\geq 0}$ is not a martingale. Thus, we define the process $\widetilde{M}^N$ in the following way: For any fixed $x \in \mathbb{R}^d$ and fixed $t > 0$, set for any $r \leq t$:

$$\widetilde{M}_r^N(x) := \frac{1}{N}\sum_{i=1}^N \int_0^r e^{(t-s)\Delta} \nabla V^N(X_s^{i,N} - x) \cdot dW_s^i.$$

Now, $\widetilde{M}^N$ is a martingale that takes values in an inifinite-dimensional space, and $\widetilde{M}^N_t = M^N_t$. Recall that the operators $(I-\Delta)^{\frac{\gamma}{2}}$, $\gamma \in \mathbb{R}$ were defined in the Notations section, see Equation (1.5), with the relation $\|(I-\Delta)^{\frac{\gamma}{2}} f\|_{L^2(\mathbb{R}^d)} = \|f\|_\gamma$.

(i) We aim at evaluating the $L^2(\mathbb{R}^d)$ norm of $(I-\Delta)^{\frac{\gamma}{2}}\widetilde{M}^N_t$. To apply a BDG-type inequality on it, we turn to the generalization of such inequality to UMD Banach spaces given in [58]. We are in a position to apply [58, Cor. 3.11] to $(I-\Delta)^{\frac{\gamma}{2}}\widetilde{M}^N$ in $L^2(\mathbb{R}^d)$ (which is UMD), and since $(I-\Delta)^{\frac{\gamma}{2}}\widetilde{M}^N_t(x) = (I-\Delta)^{\frac{\gamma}{2}} M^N_t(x)$, it comes

$$\big\|\|M^N_t\|_\gamma\big\|_{L^m(\Omega)} \leq C\left(\mathbb{E}\left[\left\|\left(\frac{1}{N^2}\sum_{i=1}^N \int_0^t |(I-\Delta)^{\frac{\gamma}{2}} e^{(t-s)\Delta}\nabla V^N(X^{i,N}_s - \cdot)|^2 ds\right)^{\frac{1}{2}}\right\|^m_{L^2(\mathbb{R}^d)}\right]\right)^{1/m}. \tag{A.2}$$

As in the proof of Lemma 3.1 in [20], one gets that for any $\delta > 0$, there exists $C > 0$ such that

$$\left(\mathbb{E}\left[\left\|\left(\frac{1}{N^2}\sum_{i=1}^N \int_0^t |(I-\Delta)^{\frac{\gamma}{2}} e^{(t-s)\Delta}\nabla V^N(X^{i,N}_s - \cdot)|^2 ds\right)^{\frac{1}{2}}\right\|^m_{L^2(\mathbb{R}^d)}\right]\right)^{1/m} \leq C\, N^{\frac{\alpha}{2}(d+2\delta+2\gamma)-\frac{1}{2}},$$

and this finishes the proof of (i).

(ii) Let $s \leq t$ and $x \in \mathbb{R}^d$. First, notice that

$$\begin{aligned} |M^N_t(x) - M^N_s(x)| &\leq \left|\frac{1}{N}\sum_{i=1}^N \int_s^t \nabla e^{(t-u)\Delta} V^N(X^{i,N}_u - x)\cdot dW^i_u\right| \\ &\quad + \left|\frac{1}{N}\sum_{i=1}^N \int_0^s \nabla e^{(s-u)\Delta}\left[e^{(t-s)\Delta} V^N(X^{i,N}_u - x) - V^N(X^{i,N}_u - x)\right]\cdot dW^i_u\right| \\ &=: |I^N_{s,t}(x)| + |II^N_{s,t}(x)|. \end{aligned} \tag{A.3}$$

Thus, one has

$$\big\|\|M^N_t - M^N_s\|_\gamma\big\|_{L^m(\Omega)} \leq \big\|\|I^N_{s,t}\|_\gamma\big\|_{L^m(\Omega)} + \big\|\|II^N_{s,t}\|_\gamma\big\|_{L^m(\Omega)}. \tag{A.4}$$

As in the first part of this proof, introducing an auxiliary martingale and applying the BDG inequality from [58] yields

$$\big\|\|I^N_{s,t}\|_\gamma\big\|_{L^m(\Omega)} \leq C\left(\mathbb{E}\left[\left\|\left(\frac{1}{N^2}\sum_{i=1}^N \int_s^t |(I-\Delta)^{\frac{\gamma}{2}} e^{(t-u)\Delta}\nabla V^N(X^{i,N}_u - \cdot)|^2 du\right)^{\frac{1}{2}}\right\|^m_{L^2(\mathbb{R}^d)}\right]\right)^{1/m}, \tag{A.5}$$

and

$$\begin{aligned} &\big\|\|II^N_{s,t}\|_\gamma\big\|_{L^m(\Omega)} \\ &\leq C\left(\mathbb{E}\left[\left\|\left(\frac{1}{N^2}\sum_{i=1}^N \int_0^s \left|(I-\Delta)^{\frac{\gamma}{2}}\nabla e^{(s-u)\Delta}[e^{(t-s)\Delta} V^N(X^{i,N}_u - \cdot) - V^N(X^{i,N}_u - \cdot)\right|^2 ds\right)^{\frac{1}{2}}\right\|^m_{L^2(\mathbb{R}^d)}\right]\right)^{1/m}. \end{aligned}$$

Now as in the proof of Lemma 3.1 of [20], one gets for arbitrary small $\delta > 0$ that

$$\left\| \|I^N_{s,t}\|_\gamma \right\|_{L^m(\Omega)} \leq C N^{\frac{\alpha}{2}(d+2\delta+2\gamma)-\frac{1}{2}} \left(\int_s^t \frac{1}{(t-u)^{1-\delta}} du \right)^{\frac{1}{2}} \lesssim (t-s)^{\frac{\delta}{2}} N^{\frac{\alpha}{2}(d+2\delta+2\gamma)-\frac{1}{2}},$$

while for $II^N_{s,t}$ it comes

$$\left\| \|II^N_{s,t}\|_\gamma \right\|_{L^m(\Omega)} \leq \frac{C}{N^{\frac{1}{2}}} \left(\int_0^s \frac{1}{(s-u)^{1-\delta}} \|(I-\Delta)^{\frac{\gamma+\delta}{2}} (e^{(t-s)\Delta} V^N - V^N)\|^2_{L^2(\mathbb{R}^d)} du \right)^{\frac{1}{2}}. \tag{A.6}$$

It is easy to obtain that, for $f \in H^1(\mathbb{R}^d)$,

$$\|e^{(t-s)\Delta} f - f\|^2_{L^2(\mathbb{R}^d)} \leq C \|\nabla f\|^2_{L^2(\mathbb{R}^d)} (t-s).$$

Hence, choosing $f = (I-\Delta)^{\frac{\gamma+\delta}{2}} V^N$ and plugging the result of the previous inequality in (A.6) yields

$$\begin{aligned} \left\| \|II^N_{s,t}\|_\gamma \right\|_{L^m(\Omega)} &\leq C \frac{(t-s)^{\frac{1}{2}}}{N^{\frac{1}{2}}} \|(I-\Delta)^{\frac{\gamma+\delta}{2}} \nabla V^N\|_{L^2(\mathbb{R}^d)} \\ &\leq C \frac{(t-s)^{\frac{1}{2}}}{N^{\frac{1}{2}}} \|(I-\Delta)^{\frac{\gamma+\delta+1}{2}} V^N\|_{L^2(\mathbb{R}^d)} \\ &\leq C (t-s)^{\frac{1}{2}} N^{\frac{\alpha}{2}(d+2\delta+2\gamma+2)-\frac{1}{2}}. \end{aligned} \tag{A.7}$$

Although the regularity in $(t-s)$ is good in the previous inequality, we paid a factor N^α which will penalise too much the rest of the computations in Propositions A.1 and A.2. Hence we also observe that

$$\begin{aligned} \|(I-\Delta)^{\frac{\gamma+\delta}{2}} (e^{(t-s)\Delta} V^N - V^N)\|_{L^2(\mathbb{R}^d)} &\leq \|(I-\Delta)^{\frac{\gamma+\delta}{2}} e^{(t-s)\Delta} V^N\|_{L^2(\mathbb{R}^d)} + \|(I-\Delta)^{\frac{\gamma+\delta}{2}} V^N\|_{L^2(\mathbb{R}^d)} \\ &\leq 2 \|(I-\Delta)^{\frac{\gamma+\delta}{2}} V^N\|_{L^2(\mathbb{R}^d)} \leq C N^{\frac{\alpha}{2}(d+2\delta+2\gamma)}. \end{aligned}$$

Thus, plugging this bound in (A.6) also gives

$$\left\| \|II^N_{s,t}\|_\gamma \right\|_{L^m(\Omega)} \leq C N^{\frac{\alpha}{2}(d+2\delta+2\gamma)-\frac{1}{2}}. \tag{A.8}$$

Hence, one can interpolate between (A.7) and (A.8) to obtain that for any $\epsilon \in [0,1]$,

$$\left\| \|II^N_{s,t}\|_\gamma \right\|_{L^m(\Omega)} \leq C (t-s)^{\frac{\epsilon}{2}} N^{\frac{\alpha}{2}(d+2\delta+2\gamma)-\frac{1}{2}+\alpha\epsilon}.$$

So the bound (A.5) for $I^N_{s,t}$ and the previous inequality plugged in (A.4) and applied to $\epsilon = \delta$ yield the inequality of (ii). □

Proof of Proposition A.2. This proof will be divided in three according to whether $p \geq 2$, $p \in (1,2)$ or $p = 1$, in that order. This might seem too much since we only need the cases $p = \infty$ and $p = 1$ in Theorem 1.4. However, note that the proof is the same for any $p \in [2, +\infty]$. Besides, we present the proof for $p \in (1,2)$ before the proof for $p = 1$, because our proof for the latter case consists in applying a Hölder inequality with weights so as to use the case $p \in (1,2)$ (we were not able to treat directly the case $p = 1$ because L^1 is not a UMD space).

We will use the decomposition of $M^N_t - M^N_s$ given in Equation (A.3) and apply the BDG inequality following the same approach as in the beginning of the proof of Proposition A.3.

- *First, assume that* $p \in [2, +\infty]$.

Define, for some $\delta > 0$ small enough,

$$\gamma := d \left(\frac{1}{2} - \frac{1}{p} \right) + \frac{\delta}{2},$$

with the convention $\frac{1}{p}=0$ if $p=\infty$. In view of the Sobolev embedding of H^γ into L^p (which holds because $p\geq 2$, see [1, Theorem 1.66]), one has

$$\big\|\|M_t^N\|_{L^p(\mathbb{R}^d)}\big\|_{L^m(\Omega)} \leq \big\|\|M_t^N\|_\gamma\big\|_{L^m(\Omega)}$$

and

$$\big\|\|M_t^N - M_s^N\|_{L^p(\mathbb{R}^d)}\big\|_{L^m(\Omega)} \leq \big\|\|M_t^N - M_s^N\|_\gamma\big\|_{L^m(\Omega)}.$$

Thus Proposition A.3 yields that

$$\big\|\|M_t^N\|_{L^p(\mathbb{R}^d)}\big\|_{L^m(\Omega)} \leq C\, N^{-\frac{1}{2}\left(1-\alpha(d+3\delta+2d(\frac{1}{2}-\frac{1}{p}))\right)}$$

and

$$\big\|\|M_t^N - M_s^N\|_{L^p(\mathbb{R}^d)}\big\|_{L^m(\Omega)} \leq C\,(t-s)^{\frac{\delta}{2}}\, N^{-\frac{1}{2}\left(1-\alpha(d+5\delta+2d(\frac{1}{2}-\frac{1}{p}))\right)}.$$

so we obtained the inequalities (i) and (ii) in the case $p\geq 2$.

- *Assume now that* $p\in(1,2)$.

Then by the same argument that leads to Equation (A.2) in the proof of Proposition A.3 (with the difference that here the UMD space is L^p, with $p>1$), one gets

$$\big\|\|M_t^N\|_{L^p(\mathbb{R}^d)}\big\|_{L^m(\Omega)} \leq \frac{C}{N}\left\|\left\|\left(\sum_{i=1}^N\int_0^t\left|\nabla e^{(t-u)\Delta}V^N(X_u^{i,N}-\cdot)\right|^2 du\right)^{\frac{1}{2}}\right\|_{L^p(\mathbb{R}^d)}\right\|_{L^m(\Omega)} \tag{A.9}$$

and for the decomposition (A.3) of $M_t^N - M_s^N = I_{s,t}^N + II_{s,t}^N$,

$$\big\|\|I_{s,t}^N\|_{L^p(\mathbb{R}^d)}\big\|_{L^m(\Omega)} \leq \frac{C}{N}\left\|\left\|\left(\sum_{i=1}^N\int_s^t|e^{(t-u)\Delta}\nabla V^N(X_u^{i,N}-\cdot)|^2 du\right)^{1/2}\right\|_{L^p(\mathbb{R}^d)}\right\|_{L^m(\Omega)}, \tag{A.10}$$

$$\begin{aligned}&\big\|\|II_{s,t}^N\|_{L^p(\mathbb{R}^d)}\big\|_{L^m(\Omega)}\\ &\leq \frac{C}{N}\left\|\left\|\left(\sum_{i=1}^N\int_0^s\left|\nabla e^{(s-u)\Delta}\left[e^{(t-s)\Delta}V^N(X_u^{i,N}-\cdot)-V^N(X_u^{i,N}-\cdot)\right]\right|^2 du\right)^{1/2}\right\|_{L^p(\mathbb{R}^d)}\right\|_{L^m(\Omega)}.\end{aligned} \tag{A.11}$$

(i) We consider first M_t^N and look for an upper bound on the right-hand side of (A.9). Since $p<2$, we add the weights $(1+|x|)^{-p}\times(1+|x|)^p$ in the integral over $\mathbb{R}^d$ to perform a Hölder inequality and we get

$$\begin{aligned}\big\|\|M_t^N\|_{L^p(\mathbb{R}^d)}\big\|_{L^m(\Omega)} \leq{}& \frac{C}{N}\left(\int_{\mathbb{R}^d}\frac{1}{(1+|x|)^{\frac{2p}{2-p}}}dx\right)^{\frac{2-p}{2p}}\\ &\times\left\|\left(\int_{\mathbb{R}^d}(1+|x|)^2\sum_{i=1}^N\int_0^t\left|\nabla e^{(t-s)\Delta}V^N(X_s^{i,N}-x)\right|^2 ds\,dx\right)^{\frac{1}{2}}\right\|_{L^m(\Omega)},\end{aligned}$$

where the first integral in the right-hand side of the previous inequality is finite. By the simple inequality $(1+|a+b|) \leq (1+|a|)(1+|b|)$ and Fubini's theorem, we then have

$$
\begin{aligned}
&\mathbb{E}\|M_t^N\|_{L^p(\mathbb{R}^d)}^m \\
&\leq \frac{C}{N^m}\mathbb{E}\left(\sum_{i=1}^N \int_0^t \int_{\mathbb{R}^d} \left(1+|X_s^{i,N}|\right)^2 \left(1+|X_s^{i,N}-x|\right)^2 \left|\nabla e^{(t-s)\Delta} V^N(X_s^{i,N}-x)\right|^2 dx\, ds\right)^{\frac{m}{2}} \\
&\leq \frac{C}{N^m}\mathbb{E}\left(\sum_{i=1}^N \int_0^t \left(1+|X_s^{i,N}|\right)^2 \int_{\mathbb{R}^d} \left(1+|y|\right)^2 \left|\nabla e^{(t-s)\Delta} V^N(y)\right|^2 dy\, ds\right)^{\frac{m}{2}} \\
&\leq \frac{C}{N^m}\left(\int_0^t \int_{\mathbb{R}^d} (1+|y|)^2 \left|\nabla e^{(t-s)\Delta} V^N(y)\right|^2 dy\, ds\right)^{\frac{m}{2}} \mathbb{E}\left(\sum_{i=1}^N \left(1+\sup_{s\in[0,t]}|X_s^{i,N}|\right)^2\right)^{\frac{m}{2}}, \quad \text{(A.12)}
\end{aligned}
$$

performing a simple change of variables in the second inequality. Since $X^{i,N}$ is a diffusion with bounded coefficients, a classical argument gives that for any $p>0$, there exists a constant $C>0$ which depends only on p and T such that $\mathbb{E}\sup_{s\in[0,T]}|X_s^{i,N}|^p \leq C$. Then, it is not difficult to verify that

$$
\mathbb{E}\left(\sum_{i=1}^N \left(1+\sup_{s\in[0,t]}|X_s^{i,N}|\right)^2\right)^{\frac{m}{2}} \leq C\, N^{\frac{m}{2}}. \quad \text{(A.13)}
$$

Now, the Cauchy-Schwarz inequality, Fubini's theorem and simple changes of variables lead to

$$
\begin{aligned}
\int_0^t \int_{\mathbb{R}^d} (1+|y|)^2 \left|\nabla e^{(t-s)\Delta} V^N(y)\right|^2 dy\, ds \leq N^{d\alpha+2\alpha} &\int_0^t \int_{\mathbb{R}^d} |\nabla V(x)|^2 \\
&\times \int_{\mathbb{R}^d} \left(1+|y+\frac{x}{N^\alpha}|\right)^2 g_{2(t-s)}(y)\, dy\, dx\, ds,
\end{aligned}
$$

where we recall that the notation g for the heat kernel was introduced in (1.6). By the simple inequality $(1+|a+b|) \leq (1+|a|)(1+|b|)$ and Fubini's theorem, we then have

$$
\begin{aligned}
\int_0^t \int_{\mathbb{R}^d} (1+|y|)^2 \left|\nabla e^{(t-s)\Delta} V^N(y)\right|^2 dy\, ds \leq C N^{d\alpha+2\alpha} &\int_{\mathbb{R}^d} |\nabla V(x)|^2 \left(1+|\frac{x}{N^\alpha}|\right)^2 dx \\
&\times \int_0^t \int_{\mathbb{R}^d} (1+|y|)^2 g_{2(t-s)}(y)\, dy\, ds,
\end{aligned}
$$

and therefore

$$
\left(\int_0^t \int_{\mathbb{R}^d} (1+|y|)^2 \left|\nabla e^{(t-s)\Delta} V^N(y)\right|^2 dy\, ds\right)^{\frac{m}{2}} \leq C N^{\frac{m(d\alpha+2\alpha)}{2}}. \quad \text{(A.14)}
$$

Combining (A.12)-(A.14), we obtain the desired property (i):

$$
\big\|\|M_t^N\|_{L^p(\mathbb{R}^d)}\big\|_{L^m(\Omega)} \leq C\, N^{-\frac{1}{2}(1-\alpha(d+2))}.
$$

(ii) Consider now $I_{s,t}^N$ and the inequality (A.10). The same computations as in (i) yield

$$
\begin{aligned}
\big\|\|I_{s,t}^N\|_{L^p(\mathbb{R}^d)}\big\|_{L^m(\Omega)} \leq \frac{C}{N} N^{\frac{1}{2}} N^{\frac{d\alpha+2\alpha}{2}} &\left(\int_{\mathbb{R}^d} |\nabla V(x)|^2 \left(1+|\frac{x}{N^\alpha}|\right)^2 dx\right. \\
&\left.\times \int_s^t \int_{\mathbb{R}^d} (1+|y|)^2 g_{2(t-u)}(y)\, dy\, du\right)^{\frac{1}{2}}.
\end{aligned}
$$

The integrability properties of V and classical estimates on g yield

$$\big\| \| I^N_{s,t} \|_{L^p(\mathbb{R}^d)} \big\|_{L^m(\Omega)} \leq C N^{-\frac{1}{2}(1-\alpha(d+2))} \sqrt{t-s}. \tag{A.15}$$

In view of $(\mathbf{A}_\alpha)$, one has : if $d \geq 2$, then $\alpha(d+2) < \frac{d+2}{2d} \leq 1$; if $d = 1$, then we assumed further that $\beta \in (\frac{1}{2} + \frac{1}{\boldsymbol{r}}, 1)$, thus $\alpha(d+2) < 1$. Hence, the power of N in the previous expression is negative.

Now, similarly for $II^N_{s,t}$ we deduce from (A.11) that

$$\big\| \| II^N_{s,t} \|_{L^p(\mathbb{R}^d)} \big\|_{L^m(\Omega)} \leq \frac{C}{N} N^{\frac{1}{2}} \left(\int_0^s \int_{\mathbb{R}^d} (1+|y|)^2 \left| \nabla e^{(s-u)\Delta} \left[e^{(t-s)\Delta} V^N(y) - V^N(y) \right] \right|^2 dy\, du \right)^{\frac{1}{2}}. \tag{A.16}$$

We will first estimate $\left| \nabla e^{(s-u)\Delta} \left[e^{(t-s)\Delta} V^N(y) - V^N(y) \right] \right|^2 = \left| \nabla g_{2(s-u)} * \left[g_{2(t-s)} * V^N(y) - V^N(y) \right] \right|^2$ by introducing a small $\varepsilon > 0$ and separating it into two terms,

$$\begin{aligned} \left| \nabla g_{2(s-u)} * \left[g_{2(t-s)} * V^N(y) - V^N(y) \right] \right|^2 &= \left| \nabla g_{2(s-u)} * \left[g_{2(t-s)} * V^N(y) - V^N(y) \right] \right|^{2-\varepsilon} \\ &\quad \times \left| g_{2(s-u)} * \left[g_{2(t-s)} * \nabla V^N(y) - \nabla V^N(y) \right] \right|^{\varepsilon}. \end{aligned} \tag{A.17}$$

For the first term above, use the triangular inequality and the simple inequality $|\nabla g_{2(s-u)}| \leq \frac{C}{\sqrt{s-u}} g_{s-u}$. Then, applying Hölder's inequality w.r.t. to a probability measure (several times), it comes

$$\begin{aligned} &\left| \nabla g_{2(s-u)} * \left[g_{2(t-s)} * V^N(y) - V^N(y) \right] \right|^{2-\varepsilon} \leq \frac{C}{(s-u)^{1-\frac{\varepsilon}{2}}} \left(g_{s-u} * \left| g_{2(t-s)} * V^N(y) - V^N(y) \right| \right)^{2-\varepsilon} \\ &\leq \frac{C}{(s-u)^{1-\frac{\varepsilon}{2}}} \int_{\mathbb{R}^d} g_{s-u}(z) \int_{\mathbb{R}^d} g_{2(t-s)}(x) \left| V^N(y-z-x) - V^N(y-z) \right|^{2-\varepsilon} dx\ dz \\ &= \frac{C}{(s-u)^{1-\frac{\varepsilon}{2}}} \int_{\mathbb{R}^d} g_{s-u}(z) \int_{\mathbb{R}^d} g_{2(t-s)}(x) \left| \int_0^1 \nabla V^N(y-z-rx) \cdot x\ dr \right|^{2-\varepsilon} dx\ dz \\ &\leq \frac{C}{(s-u)^{1-\frac{\varepsilon}{2}}} \int_{\mathbb{R}^d} g_{s-u}(z) \int_{\mathbb{R}^d} g_{2(t-s)}(x) |x|^{2-\varepsilon} \int_0^1 \left| \nabla V^N(y-z-rx) \right|^{2-\varepsilon} dr\ dx\ dz. \end{aligned} \tag{A.18}$$

For the second term in (A.17), the triangular inequality and the Lipschitz regularity of ∇V^N lead to

$$\begin{aligned} \left| g_{2(s-u)} * \left[g_{2(t-s)} * \nabla V^N(y) - \nabla V^N(y) \right] \right|^{\varepsilon} &\leq N^{(d\alpha+2\alpha)\varepsilon} \left| \int_{\mathbb{R}^d} g_{2(s-u)}(z) \int_{\mathbb{R}^d} g_{2(t-s)}(x) |x|\, dx\, dz \right|^{\varepsilon} \\ &\leq C N^{(d\alpha+2\alpha)\varepsilon} (t-s)^{\frac{\varepsilon}{2}}. \end{aligned} \tag{A.19}$$

After plugging (A.18) and (A.19) in (A.17), one gets

$$\begin{aligned} \left| \nabla g_{2(s-u)} * \left[g_{2(t-s)} * V^N(y) - V^N(y) \right] \right|^2 &\leq \frac{C}{(s-u)^{1-\frac{\varepsilon}{2}}} N^{(d\alpha+2\alpha)\varepsilon} (t-s)^{\frac{\varepsilon}{2}} \\ &\int_{\mathbb{R}^d} g_{s-u}(z) \int_{\mathbb{R}^d} g_{2(t-s)}(x) |x|^{2-\varepsilon} \int_0^1 \left| \nabla V^N(y-z-rx) \right|^{2-\varepsilon} dr\, dx\, dz. \end{aligned} \tag{A.20}$$

Now one can plug (A.20) in (A.16), and from Fubini's theorem and a change of variables, it comes

$$\begin{aligned} &\big\| \| II^N_{s,t} \|_{L^p(\mathbb{R}^d)} \big\|_{L^m(\Omega)} \leq C\, N^{-\frac{1}{2}(1-\alpha\varepsilon(d+2))} (t-s)^{\frac{\varepsilon}{4}} N^{\frac{1}{2}(\alpha(2-\varepsilon)+d\alpha(1-\varepsilon))} \\ &\times \left(\int_0^s \frac{1}{(s-u)^{1-\frac{\varepsilon}{2}}} \int_{\mathbb{R}^d} g_{s-u}(z) \int_{\mathbb{R}^d} g_{2(t-s)}(x) |x|^{2-\varepsilon} \int_0^1 \int_{\mathbb{R}^d} \left(1 + \left| \frac{y}{N^\alpha} + z + rx \right| \right)^2 |\nabla V(y)|^{2-\varepsilon} dy\, dr\, dx\, dz\, du \right)^{\frac{1}{2}}. \end{aligned}$$

Then it follows from the simple identity $(1+|a+b|) \leq (1+|a|)(1+|b|)$ that

$$\begin{aligned} \big\| \| II^N_{s,t} \|_{L^p(\mathbb{R}^d)} \big\|_{L^m(\Omega)} &\leq C\, N^{-\frac{1}{2}(1-\alpha(d+2+\varepsilon))} (t-s)^{\frac{\varepsilon}{4}} \left(\int_{\mathbb{R}^d} \left(1 + \left| \frac{y}{N^\alpha} \right| \right)^2 |\nabla V(y)|^2 dy \right)^{\frac{1}{2}} \\ &\times \left(\int_0^s \frac{1}{(s-u)^{1-\frac{\varepsilon}{2}}} \int_{\mathbb{R}^d} g_{s-u}(z) (1+|z|)^2 dz\, du \right)^{\frac{1}{2}} \left(\int_{\mathbb{R}^d} (1+|x|)^2 |x|^{2-\varepsilon} g_{2(t-s)}(x)\, dx \right)^{\frac{1}{2}}. \end{aligned}$$

The latter implies that

$$\big\| \|II_{s,t}^N\|_{L^p(\mathbb{R}^d)} \big\|_{L^m(\Omega)} \leq C\, N^{-\frac{1}{2}(1-\alpha(d+2+\varepsilon))} \sqrt{t-s}. \tag{A.21}$$

Now in view of (A.3), $\big\| \|M_t^N - M_s^N\|_{L^p(\mathbb{R}^d)} \big\|_{L^m(\Omega)} \leq \big\| \|I_{s,t}^N\|_{L^p(\mathbb{R}^d)} \big\|_{L^m(\Omega)} + \big\| \|II_{s,t}^N\|_{L^p(\mathbb{R}^d)} \big\|_{L^m(\Omega)}$, hence the desired result in the case $p \in (1,2)$ follows from (A.15) and (A.21).

- *Assume finally that* $p=1$. $L^1(\mathbb{R}^d)$ is not a UMD Banach space, so in the case $p=1$, we proceed as follows.

(i) Fix $r \in (1,2)$. Applying Hölder's inequality, one has

$$\|M_t^N\|_{L^1(\mathbb{R}^d)} \leq C\|(1+|\cdot|)M_t^N\|_{L^r(\mathbb{R}^d)}.$$

Now, we can repeat the computations from the previous part with a slight modification as follows. First, apply the BDG inequality to the new process $\overline{M}^N = (1+|\cdot|)\widetilde{M}^N$. It comes

$$\big\| \|(1+|\cdot|)M_t^N\|_{L^r(\mathbb{R}^d)} \big\|_{L^m(\Omega)} \leq \frac{C}{N} \left\| \left\| \left((1+|\cdot|)^2 \sum_{i=1}^N \int_0^t \left| \nabla e^{(t-u)\Delta} V^N(X_u^{i,N} - \cdot) \right|^2 du \right)^{\frac{1}{2}} \right\|_{L^r(\mathbb{R}^d)} \right\|_{L^m(\Omega)}.$$

Now, we add weights and perform a Hölder inequality as before. One has

$$\begin{aligned}\big\| \|M_t^N\|_{L^1(\mathbb{R}^d)} \big\|_{L^m(\Omega)} \leq \frac{C}{N} &\left(\int_{\mathbb{R}^d} \frac{1}{(1+|x|)^{\frac{2r}{2-r}}} dx \right)^{\frac{2-r}{2r}} \\ &\times \left\| \left(\int_{\mathbb{R}^d} (1+|x|)^4 \sum_{i=1}^N \int_0^t \left| \nabla e^{(t-s)\Delta} V^N(X_s^{i,N} - x) \right|^2 ds\, dx \right)^{\frac{1}{2}} \right\|_{L^m(\Omega)}.\end{aligned}$$

From this point, nothing changes in the computation except that before we had $(1+|x|)^2$ and now we have $(1+|x|)^4$. Following (A.12), (A.13) and (A.14) and taking into account this modification, one gets that for any $m \geq 1$, there exists $C>0$ such that

$$\big\| \|M_t^N\|_{L^1(\mathbb{R}^d)} \big\|_{L^m(\Omega)} \leq C\, N^{-\frac{1}{2}(1-\alpha(d+2))}.$$

(ii) Fix again some $r \in (1,2)$. Consider now $I_{s,t}^N$ from the decomposition in (A.3) of $M_t^N - M_s^N = I_{s,t}^N + II_{s,t}^N$. As above,

$$\|I_{s,t}^N\|_{L^1(\mathbb{R}^d)} \leq C\|(1+|\cdot|)I_{s,t}^N\|_{L^r(\mathbb{R}^d)}.$$

Now,

$$\big\| \|(1+|\cdot|)I_{s,t}^N\|_{L^r(\mathbb{R}^d)} \big\|_{L^m(\Omega)} \leq \frac{C}{N} \left\| \left\| \left((1+|\cdot|)^2 \sum_{i=1}^N \int_s^t |e^{(t-u)\Delta} \nabla V^N(X_u^{i,N} - \cdot)|^2 du \right)^{1/2} \right\|_{L^r(\mathbb{R}^d)} \right\|_{L^m(\Omega)}.$$

The same computations as in the case $p \in (1,2)$ part (i), with the modification mentioned above ($(1+|x|)^4$ instead of $(1+|x|)^2$ in the beginning) yield

$$\begin{aligned}\big\| \|I_{s,t}^N\|_{L^1(\mathbb{R}^d)} \big\|_{L^m(\Omega)} \leq \frac{C}{N} N^{\frac{1}{2}} N^{\frac{d\alpha+2\alpha}{2}} &\left(\int_{\mathbb{R}^d} |\nabla V(x)|^2 \left(1 + |\frac{x}{N^\alpha}|\right)^4 dx \right. \\ &\left. \times \int_s^t \int_{\mathbb{R}^d} (1+|y|)^4\, g_{2(t-u)}(y)\, dy\, du \right)^{\frac{1}{2}}.\end{aligned}$$

Besides, the integrability properties of V and classical estimates on g yield

$$\big\| \|I_{s,t}^N\|_{L^1(\mathbb{R}^d)} \big\|_{L^m(\Omega)} \leq C N^{-\frac{1}{2}(1-\alpha(d+2))} \sqrt{t-s}. \tag{A.22}$$

It remains to treat the term $II^N_{s,t}$. Again the same computations as above lead us to

$$\|II^N_{s,t}\|_{L^1(\mathbb{R}^d)} \le C\|(1+|\cdot|)II^N_{s,t}\|_{L^r(\mathbb{R}^d)},$$

and we have

$$\begin{aligned}&\big\|\|(1+|\cdot|)II^N_{s,t}\|_{L^r(\mathbb{R}^d)}\big\|_{L^m(\Omega)}\\ &\le \frac{C}{N}\left\|\left\|\left((1+|\cdot|)^2\sum_{i=1}^N\int_0^s\left|\nabla e^{(s-u)\Delta}\left[e^{(t-s)\Delta}V^N(X_u^{i,N}-\cdot)-V^N(X_u^{i,N}-\cdot)\right]\right|^2du\right)^{1/2}\right\|_{L^r(\mathbb{R}^d)}\right\|_{L^m(\Omega)}.\end{aligned}$$

Now, start from the line (A.16) where $(1+|y|)^2$ is replaced by $(1+|y|)^4$ and repeat the computations line by line. Eventually, it comes

$$\big\|\|II^N_{s,t}\|_{L^1(\mathbb{R}^d)}\big\|_{L^m(\Omega)} \le C\,N^{-\frac12(1-\alpha(d+2+\varepsilon))}\sqrt{t-s}. \tag{A.23}$$

Now in view of (A.3), $\big\|\|M^N_t-M^N_s\|_{L^1(\mathbb{R}^d)}\big\|_{L^m(\Omega)} \le \big\|\|I^N_{s,t}\|_{L^1(\mathbb{R}^d)}\big\|_{L^m(\Omega)}+\big\|\|II^N_{s,t}\|_{L^1(\mathbb{R}^d)}\big\|_{L^m(\Omega)}$, hence the desired result in the case $p=1$ follows from (A.22) and (A.23). □

A.2 Proof of the boundedness estimate

Proof of Proposition 3.1. **Step 1**. Recall that the operators $(I-\Delta)^\beta,\ \beta\in\mathbb{R}$ were defined in the Notations section, see Equation (1.5), with a clear link with the Sobolev norm $\|\cdot\|_{\beta,\boldsymbol{r}}$.

Let F stand for the function F_A defined in (1.7). From (3.1) after applying $(\mathrm{I}-\Delta)^{\frac{\beta}{2}}$ and by the triangular inequality we have

$$\left\|(\mathrm{I}-\Delta)^{\frac{\beta}{2}}u_t^N\right\|_{L^q(\Omega;L^r(\mathbb{R}^d))} \le \left\|(\mathrm{I}-\Delta)^{\frac{\beta}{2}}e^{t\Delta}u_0^N\right\|_{L^q(\Omega;L^r(\mathbb{R}^d))} \tag{A.24}$$

$$+\int_0^t\left\|(\mathrm{I}-\Delta)^{\frac{\beta}{2}}\nabla\cdot e^{(t-s)\Delta}\left(V^N*\left(F(K*u_s^N)\mu_s^N\right)\right)\right\|_{L^q(\Omega;L^r(\mathbb{R}^d))}ds \tag{A.25}$$

$$+\left\|\frac{1}{N}\sum_{i=1}^N\int_0^t(\mathrm{I}-\Delta)^{\frac{\beta}{2}}\nabla e^{(t-s)\Delta}\left(V^N\left(X_s^{i,N}-\cdot\right)\right)\cdot dW_s^i\right\|_{L^q(\Omega;L^r(\mathbb{R}^d))}. \tag{A.26}$$

Step 2. Noticing that by a convolution inequality

$$\|(\mathrm{I}-\Delta)^{\frac{\beta}{2}}e^{t\Delta}u_0^N\|_{L^r(\mathbb{R}^d)} \le \|e^{t\Delta}\|_{L^p\to L^p}\|(\mathrm{I}-\Delta)^{\frac{\beta}{2}}u_0^N\|_{L^r(\mathbb{R}^d)},$$

one gets that the first term (A.24) can be estimated by

$$\left\|(\mathrm{I}-\Delta)^{\frac{\beta}{2}}e^{t\Delta}u_0^N\right\|_{L^q(\Omega;L^r(\mathbb{R}^d))} \le \left\|(\mathrm{I}-\Delta)^{\frac{\beta}{2}}u_0^N\right\|_{L^q(\Omega;L^r(\mathbb{R}^d))} \le C_\beta,$$

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

Step 3. Let us come to the second term (A.25):

$$\begin{aligned}&\int_0^t\|(\mathrm{I}-\Delta)^{\frac{\beta}{2}}\nabla\cdot e^{(t-s)\Delta}\left(V^N*\left(F(K*u_s^N)\mu_s^N\right)\right)\|_{L^q(\Omega;L^r(\mathbb{R}^d))}ds\\ &\le C\int_0^t\|(\mathrm{I}-\Delta)^{\frac{\beta}{2}}\nabla\cdot e^{(t-s)\Delta}\|_{L^r\to L^r}\|\left(V^N*\left(F(K*u_s^N)\mu_s^N\right)\right)\|_{L^q(\Omega;L^r(\mathbb{R}^d))}ds.\end{aligned}$$

In view of Inequality (1.14), we have that

$$\|(\mathrm{I}-\Delta)^{\frac{\beta}{2}}\nabla\cdot e^{(t-s)\Delta}\|_{L^r\to L^r} \le C\frac{1}{(t-s)^{\frac{(1+\beta)}{2}}}.$$

Thus,

$$\begin{aligned}
&\int_0^t \left\| (\mathrm{I}-\Delta)^{\frac{\beta}{2}} \nabla \cdot e^{(t-s)\Delta} \left(V^N * \left(F(K*u_s^N)\mu_s^N\right)\right) \right\|_{L^q(\Omega;L^r(\mathbb{R}^d))} ds \\
&\leq C \int_0^t \frac{1}{(t-s)^{(1+\beta)/2}} \|u_t^N\|_{L^q(\Omega;L^r(\mathbb{R}^d))} ds \\
&\leq C \int_0^t \frac{1}{(t-s)^{(1+\beta)/2}} \| (\mathrm{I}-\Delta)^{\frac{\beta}{2}} u_t^N \|_{L^q(\Omega;L^r(\mathbb{R}^d))} ds.
\end{aligned}$$

This bounds the second term.

Step 4. For the third term (A.26), recalling the notation introduced in (3.5), we need to control $\left\| \|M_t^N\|_{\beta,\boldsymbol{r}} \right\|_{L^q(\Omega)}$. The embedding for Bessel potential spaces of [57, p.203] gives that $H^{\beta+d(\frac{1}{2}-\frac{1}{r})}(\mathbb{R}^d)$ is continuously embedded into $H^\beta_{\boldsymbol{r}}(\mathbb{R}^d)$, thus we obtain

$$\left\| \|M_t^N\|_{\beta,\boldsymbol{r}} \right\|_{L^q(\Omega)} \leq C \left\| \|M_t^N\|_{\beta+d(\frac{1}{2}-\frac{1}{r}),2} \right\|_{L^q(\Omega)}.$$

Now Proposition A.3-(i) permits to bound the previous upper bound, hence we get $\left\| \|M_t^N\|_{\beta,\boldsymbol{r}} \right\|_{L^q(\Omega)} \leq C\, N^{-\frac{1}{2}\left(1-\alpha(d+2\delta+2\beta+2d(\frac{1}{2}-\frac{1}{r}))\right)}$, where δ is arbitrarily small. In view of Assumption $(\mathbf{A}_\alpha)$, it thus follows that

$$\left\| \|M_t^N\|_{\beta,\boldsymbol{r}} \right\|_{L^q(\Omega)} \leq C.$$

From the three bounds obtained in Steps 2 to 4 and the Grönwall lemma, there exists a deterministic constant $C > 0$ (which depends only on β, T, A and $\boldsymbol{r}$) such that

$$\left\| (\mathrm{I}-\Delta)^{\frac{\beta}{2}} u_t^N \right\|_{L^q(\Omega;L^r(\mathbb{R}^d))} \leq C, \quad \forall N \in \mathbb{N}^*,$$

which proves the desired result. □

References

[1] H. Bahouri, J.-Y. Chemin, and R. Danchin. *Fourier Analysis and Nonlinear Partial Differential Equations*, volume 343. Grundlehren der Mathematischen Wisserchaften, 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. II.* 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. Mean-field limit and quantitative estimates with singular attractive kernels. *Preprint arXiv:2011.08022*, 2020.

[7] H. Brezis. *Functional analysis, Sobolev Spaces and Partial Differential Equations.* Springer, 2010.

[8] P. Cattiaux and L. Pédèches. The 2-D stochastic Keller-Segel particle model: existence and uniqueness. *ALEA*, 13:447–463, 2016.

[9] 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.

[10] E. Cépa and D. Lépingle. Diffusing particles with electrostatic repulsion. *Probab. Theory Related Fields*, 107(4):429–449, 1997.

[11] 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.

[12] P.-E. Chaudru de Raynal and N. Frikha. Well-posedness for some non-linear diffusion processes and related PDE on the Wasserstein space. *Preprint arXiv:1811.06904*, 2018.

[13] R. Cortez and J. Fontbona. Quantitative uniform propagation of chaos for Maxwell molecules. *Comm. Math. Phys.*, 357(3):913–941, 2018.

[14] 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.

[15] M. Duerinckx. Mean-field limits for some Riesz interaction gradient flows. *SIAM J. Math. Anal.*, 48(3):2269–2300, 2016.

[16] F. Flandoli and M. Leocata. A particle system approach to aggregation phenomena. *J. Appl. Probab.*, 56(1):282–306, 2019.

[17] 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.

[18] F. Flandoli, M. Leocata, and C. Ricci. The Vlasov-Navier-Stokes equations as a mean field limit. *Discrete Contin. Dyn. Syst. Ser. B*, 24(8):3741–3753, 2019.

[19] F. Flandoli, M. Leocata, and C. Ricci. On the macroscopic limit of Brownian particles with local interaction. *Preprint arXiv:2006.01781*, 2020.

[20] F. Flandoli, C. Olivera, and M. Simon. Uniform approximation of 2 dimensional Navier-Stokes equation by stochastic interacting particle systems. *SIAM J. Math. Anal.*, 52(6):5339–5362, 2020.

[21] N. Fournier and M. Hauray. Propagation of chaos for the Landau equation with moderately soft potentials. *Ann. Probab.*, 44(6):3581–3660, 2016.

[22] 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.

[23] 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.

[24] P. K. Friz and N. B. Victoir. *Multidimensional stochastic processes as rough paths*, volume 120 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010.

[25] A. M. Garsia, E. Rodemich, and H. Rumsey, Jr. A real variable lemma and the continuity of paths of some Gaussian processes. *Indiana Univ. Math. J.*, 20:565–578, 1970/71.

[26] L. Grafakos. *Classical Fourier analysis*, volume 249 of *Graduate Texts in Mathematics*. Springer, New York, third edition, 2014.

[27] J. Hoeksema, M. Maurelli, T. Holding, and O. Tse. Large deviations for singularly interacting diffusions. *Preprint arXiv:2002.01295*, 2020.

[28] X. Huang and F.-Y. Wang. McKean-Vlasov SDEs with drifts discontinuous under Wasserstein distance. *Preprint* arXiv:2002.06877, 2020.

[29] 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.

[30] 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.

[31] B. Jourdain. Convergence of moderately interacting particle systems to a diffusion-convection equation. *Stochastic Process. Appl.*, 73(2):247–270, 1998.

[32] 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.

[33] 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.

[34] P. E. Kloeden and A. Neuenkirch. The pathwise convergence of approximation schemes for stochastic differential equations. *LMS J. Comput. Math.*, 10:235–253, 2007.

[35] D. Lacker. On a strong form of propagation of chaos for McKean-Vlasov equations. *Electron. Commun. Probab.*, 23:Paper No. 45, 2018.

[36] C. Marchioro and M. Pulvirenti. Hydrodynamics in two dimensions and vortex theory. *Comm. Math. Phys.*, 84(4):483–503, 1982.

[37] 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.

[38] S. Méléard. A trajectorial proof of the vortex method for the two-dimensional Navier-Stokes equation. *Ann. Appl. Probab.*, 10(4):1197–1211, 2000.

[39] 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.

[40] Y. S. Mishura and A. Y. Veretennikov. Existence and uniqueness theorems for solutions of McKean–Vlasov stochastic equations. *Preprint* arXiv:1603.02212, 2016.

[41] 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.

[42] K. Oelschläger. A law of large numbers for moderately interacting diffusion processes. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 69(2):279–322, 1985.

[43] K. Oelschläger. A fluctuation theorem for moderately interacting diffusion processes. *Probab. Theory Related Fields*, 74(4):591–616, 1987.

[44] C. Olivera, A. Richard, and M. Tomašević. Particle approximation of the 2-d parabolic-elliptic Keller-Segel system in the subcritical regime. *Preprint* arXiv:2004.03177, 2020.

[45] H. Osada. Propagation of chaos for the two-dimensional Navier-Stokes equation. *Proc. Japan Acad. Ser. A Math. Sci.*, 62(1):8–11, 1986.

[46] A. Richard, X. Tan, and F. Yang. Discrete-time Simulation of Stochastic Volterra Equations. *Preprint* arXiv:2004.00340, 2020.

[47] M. Röckner and X. Zhang. Well-posedness of distribution dependent SDEs with singular drifts. *Preprint arXiv:1809.02216*, 2018.

[48] M. Rosenzweig. Mean-field convergence of systems of particles with Coulomb interactions in higher dimensions without regularity. *Preprint arXiv:2010.10009*, 2020.

[49] W. Rudin. *Functional analysis.* International Series in Pure and Applied Mathematics. McGraw-Hill, Inc., New York, second edition, 1991.

[50] 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.

[51] S. Serfaty. Mean field limit for Coulomb-type flows. *Duke Math. J.*, 169(15):2887–2935, 2020.

[52] M. Simon and C. Olivera. Non-local conservation law from stochastic particle systems. *J. Dynam. Differential Equations*, 30(4):1661–1682, 2018.

[53] A.-S. Sznitman. A propagation of chaos result for Burgers' equation. *Probab. Theory Related Fields*, 71(4):581–613, 1986.

[54] 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.

[55] 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.

[56] M. Tomašević. Propagation of chaos for stochastic particle systems with singular mean-field interaction of $L^p - L^q$ type. *Preprint hal-03086253v1*, 2020.

[57] H. Triebel. *Interpolation theory, function spaces, differential operators*, volume 18 of *North-Holland Mathematical Library.* North-Holland Publishing Co., Amsterdam-New York, 1978.

[58] J. M. A. M. van Neerven, M. C. Veraar, and L. Weis. Stochastic integration in UMD Banach spaces. *Ann. Probab.*, 35(4):1438–1478, 2007.