



**HAL**  
open science

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

Christian Olivera, Alexandre Richard, Milica Tomasevic

► **To cite this version:**

Christian Olivera, Alexandre Richard, Milica Tomasevic. Quantitative particle approximation of nonlinear Fokker-Planck equations with singular kernel. *Annali della Scuola Normale Superiore di Pisa, Classe di Scienze*, 2022, 10.2422/2036-2145.202105\_087 . hal-02914779v4

**HAL Id: hal-02914779**

**<https://inria.hal.science/hal-02914779v4>**

Submitted on 27 Jun 2022

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

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

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

Christian Olivera      Alexandre Richard      Milica Tomašević

## Abstract

In this work, we study the convergence of the empirical measure of moderately interacting particle systems with singular interaction kernels. First, we prove quantitative convergence of the time marginals of the empirical measure of particle positions towards the solution of the limiting nonlinear Fokker-Planck equation. Second, we prove the well-posedness for the McKean-Vlasov SDE involving such singular kernels and the convergence of the empirical measure towards it (propagation of chaos).

Our results only require very weak regularity on the interaction kernel, which permits to treat models for which the mean field particle system is not known to be well-defined. For instance, this includes attractive kernels such as Riesz and Keller-Segel kernels in arbitrary dimension. For some of these important examples, this is the first time that a quantitative approximation of the PDE is obtained by means of a stochastic particle system. 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 regularity of infinite-dimensional stochastic convolution integrals.

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

<b>1</b>	<b>Introduction and main results</b>	<b>2</b>
1.1	Notations . . . . .	3
1.2	Setting . . . . .	5
1.3	First main result: Rate of convergence to the PDE . . . . .	7
1.4	Second main result: Propagation of chaos . . . . .	8
1.5	Organisation of the paper . . . . .	10
<b>2</b>	<b>Rate of convergence</b>	<b>10</b>
2.1	Properties of the regularised empirical measure . . . . .	11
2.2	Proof of Theorem 1.3 . . . . .	12
2.3	Corollaries of Theorem 1.3 . . . . .	15
<b>3</b>	<b>Propagation of chaos</b>	<b>16</b>
3.1	Proof of Proposition 1.8 . . . . .	16
3.2	Proof of Theorem 1.9 . . . . .	17
<b>4</b>	<b>Riesz-like singular kernels</b>	<b>20</b>
4.1	Rate in Sobolev norm: Proof of Corollary 4.4 . . . . .	23
4.2	Convergence in $H_r^\beta$ : Proof of Corollary 4.6 . . . . .	24
4.3	Sketch of the proof of Theorem 4.8 . . . . .	24
<b>5</b>	<b>Examples</b>	<b>25</b>
5.1	A stronger, easier-to-check condition on the kernel . . . . .	25
5.2	Riesz potentials . . . . .	26
5.3	Classical kernels . . . . .	28

<b>Appendix</b>	<b>29</b>
A.1 Properties of the PDE and of the PDE with cut-off . . . . .	29
A.2 Time and space estimates of the stochastic convolution integrals . . . . .	32
A.3 Proof of the boundedness estimate . . . . .	39
 <b>References</b>	 <b>41</b>

## 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

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

by means of moderately interacting particle systems. The main interest here is for kernels  $K$  with a singularity at the origin.

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

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

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 (see [36, 38, 41, 53]) and 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).

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 [35] and Osada [44] on the  $2d$  Navier-Stokes equation (see also [24] more recently); while Sznitman [52] and Bossy and Talay [5] worked 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 and in the modelling of chemotaxis (see [8, 14, 22, 27, 31]).

When the interaction is attractive in addition to being singular, 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) is known to hold only for small values of the critical parameter of the equation (see [22]). For the  $d$ -dimensional parabolic-elliptic Keller-Segel model with  $d \geq 3$  or for 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) in 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:

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

where  $(W^i)_{1 \leq i \leq N}$  are independent 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. This kind of interacting particle system, introduced by Oelschläger [41], relies on a smoothing of the interaction kernel at the scale  $N^{-\alpha}$ ,  $\alpha \in [0, 1]$ . Observe that the interaction kernel  $K * V^N(x)$  is very close to  $K(x)$  when  $|x|$  is sufficiently large compared to  $N^{-\alpha}$ . However, the difficulty is to prove that at the limit  $N \rightarrow \infty$ , the system (1.3) behaves as the mean-field system would. In this work, we show that:

- (a)  $\{\mu_t^N = \frac{1}{N} \sum_{i=1}^N \delta_{X_t^{i,N}}, t \in [0, T]\}$ , the marginals of the empirical measure of (1.3), converge to the solution of the PDE (1.1) for values of the smoothing parameter  $\alpha$  which can be up to  $(\frac{1}{d})^-$  for some of the models we consider, which is the typical distance between the particles in problems coming from statistical physics with repulsive kernels. We get a rate of convergence of order  $N^{-\varrho}$ , with  $\varrho$  which can be up to  $(\frac{1}{d+2})^-$ . Moreover, one can take  $F_A(x) \equiv x$  in (1.3), at the price of a weaker form of convergence.
- (b) The system (1.3) propagates chaos towards the following nonlinear equation (without the cut-off and the mollifier):

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

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

As an application, one can use the rate in (a) and a time discretization of (1.3) to propose a numerical approximation of the PDE (1.1). We leave this line of investigation for a future work.

In the rest of this introduction, we will detail the notations and the assumptions necessary for our study. Then, we will present our main results in Subsections 1.3 and 1.4. The organisation of the paper is then given in Subsection 1.5.

## 1.1 Notations

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

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

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

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

In particular, note that

$$\|u\|_{0,p} = \|u\|_{L^p(\mathbb{R}^d)} \quad \text{and for any } \beta \leq \gamma, \quad \|u\|_{\beta,p} \leq \|u\|_{\gamma,p}. \quad (1.5)$$

The space  $H_p^\beta(\mathbb{R}^d)$  is associated to the fractional operator  $(I - \Delta)^{\frac{\beta}{2}}$  defined as (see e.g. [56, p.180] for more details on this operator):

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

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

$$\langle u, v \rangle_\beta := \int_{\mathbb{R}^d} (1 + |\xi|^2)^\beta \mathcal{F} u(\xi) \overline{\mathcal{F} v(\xi)} d\xi \quad (1.7)$$

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

In addition, for  $p \geq 1$  we will denote by  $p'$  its conjugate, i.e.  $p' \geq 1$  such that  $\frac{1}{p} + \frac{1}{p'} = 1$ .

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

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

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

$$g_{\sigma^2}(x) = \frac{1}{(2\pi\sigma^2)^{\frac{d}{2}}} e^{-\frac{|x|^2}{2\sigma^2}}. \quad (1.8)$$

- For  $\mathcal{X}$  some normed vector space, the space  $\mathcal{C}([0, T]; \mathcal{X})$  of continuous functions from the time interval  $[0, T]$  with values in  $\mathcal{X}$  is classically endowed with the norm

$$\|f\|_{T, \mathcal{X}} = \sup_{s \in [0, T]} \|f_s\|_{\mathcal{X}}.$$

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

$$\mathcal{N}_\delta(f) := \sup_{x \neq y \in \mathbb{R}^d} \frac{|f(x) - f(y)|}{|x - y|^\delta}. \quad (1.9)$$

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

- 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  $x \mapsto u(t, x)$ .
- 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.
- The ball centred in the origin that has radius 1 is denoted by  $\mathcal{B}_1$ .
- Finally, for functions from  $\mathbb{R}^d$  to  $\mathbb{R}$ , the space of  $n$ -times ( $n \in \mathbb{N}$ ) differentiable functions is denoted by  $\mathcal{C}^n(\mathbb{R}^d)$ ; the space of  $n$ -times ( $n \in \mathbb{N}$ ) differentiable functions with bounded derivatives of any order between 0 and  $n$ , denoted by  $\mathcal{C}_b^n(\mathbb{R}^d)$ ; and the space of  $n$ -times ( $n \in \mathbb{N}$ ) differentiable functions with compact support, denoted by  $\mathcal{C}_c^n(\mathbb{R}^d)$ . For  $n = 0$ , we will denote the space of continuous (resp. bounded continuous, and continuous with compact support) by  $\mathcal{C}(\mathbb{R}^d)$  (resp.  $\mathcal{C}_b(\mathbb{R}^d)$  and  $\mathcal{C}_c(\mathbb{R}^d)$ ).
- Throughout the paper, we denote by  $C$  a constant that may change from line to line.

## 1.2 Setting

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

Let  $A > 0$  and  $f_A : \mathbb{R} \rightarrow \mathbb{R}$  be a  $\mathcal{C}_b^2(\mathbb{R})$  function such that

- (i)  $f_A(x) = x$ , for  $x \in [-A, A]$ ,
- (ii)  $f_A(x) = A$ , for  $x > A + 1$  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

$$F_A : (x_1, \dots, x_d)^T \mapsto (f_A(x_1), \dots, f_A(x_d))^T. \quad (1.10)$$

Let  $V : \mathbb{R}^d \rightarrow \mathbb{R}_+$  be a smooth probability density function, and assume further that  $V$  is compactly supported. For any  $x \in \mathbb{R}^d$ , define

$$V^N(x) := N^{d\alpha} V(N^\alpha x), \quad \text{for some } \alpha \in [0, 1]. \quad (1.11)$$

Below,  $\alpha$  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:

$$\begin{cases} dX_t^{i,N} = F_A \left( \frac{1}{N} \sum_{k=1}^N (K * V^N)(X_t^{i,N} - X_t^{k,N}) \right) dt + \sqrt{2} dW_t^i, & t \leq T, \ 1 \leq i \leq N, \\ X_0^{i,N}, \ 1 \leq i \leq N, & \text{are independent of } \{W^i, \ 1 \leq i \leq N\}, \end{cases} \quad (1.12)$$

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  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ .

Let us denote the empirical measure on  $\mathcal{C}([0, T], \mathbb{R}^d)$  of  $N$  particles by

$$\mu^N = \frac{1}{N} \sum_{i=1}^N \delta_{X^{i,N}}, \quad (1.13)$$

and the mollified empirical measure by

$$u_t^N := V^N * \mu_t^N, \quad t \in [0, T].$$

The following properties of the kernel will be assumed:

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

- $(\mathbf{A}_i^K)$   $K \in L^p(\mathcal{B}_1)$ , for some  $\mathbf{p} \in [1, +\infty]$ ;
- $(\mathbf{A}_{ii}^K)$   $K \in L^q(\mathcal{B}_1^c)$ , for some  $\mathbf{q} \in [1, +\infty]$ ;
- $(\mathbf{A}_{iii}^K)$  There exists  $\mathbf{r} \geq \max(\mathbf{p}', \mathbf{q}')$ ,  $\zeta \in (0, 1]$  and  $C > 0$  such that for any  $f \in L^1 \cap L^{\mathbf{r}}(\mathbb{R}^d)$ , one has

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

Throughout the paper, it will be supposed that the parameter  $\mathbf{r}$  is such that  $\mathbf{r} \geq \max(\mathbf{p}', \mathbf{q}')$ , where  $(\mathbf{p}, \mathbf{q})$  are given in  $(\mathbf{A}^K)$ . Hence, even if some statements do not require the regularity assumption  $(\mathbf{A}_{iii}^K)$ , the parameter  $\mathbf{r}$  (when it appears) will always be considered larger than  $\max(\mathbf{p}', \mathbf{q}')$ .

The Assumption  $(\mathbf{A}^K)$  is rather mild, and we provide in Section 5.1 a sufficient condition which is easier to check in concrete examples. In Section 5, we show that Riesz potentials (see (1.2)) up to  $s \leq d - 2$ , whether repulsive or attractive, satisfy  $(\mathbf{A}^K)$ . Note that the case of more singular kernels, e.g. Riesz kernels with  $s \in (d - 2, d - 1)$ , will be treated separately in Section 4.

The restriction with respect to the key parameters is given by the following assumption:

( $\mathbf{A}_\alpha$ ): The parameters  $\alpha$  and  $\mathbf{r}$  (which appear respectively in (1.11) and ( $\mathbf{A}_{iii}^K$ )) satisfy

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

Notice that if the integrability of the kernel is such that  $\mathbf{r}$  in ( $\mathbf{A}_{iii}^K$ ) could be chosen in the interval  $[1, 2]$ , then  $\alpha$  could be arbitrarily close to, but smaller than,  $\frac{1}{d}$  (although this will not yield the best possible rate of convergence, see Remark 1.5). This scaling of order  $N^{-(\frac{1}{d})^-}$  is close to the typical interparticle distance in physical models with repulsive interactions. This is what we get for instance in the case of  $2d$  Navier-Stokes equation and  $2d$  Coulomb potential.

However, in the case of attractive interaction, the typical interparticle distance is not necessarily  $N^{-\frac{1}{d}}$  as the particles may collide or even agglomerate. Indeed, in the case of the parabolic-elliptic Keller-Segel equation, the associated mean-field particles collide with positive probability (see [22]), whilst in the fully parabolic case they seem to agglomerate (numerically), see [55, Fig. 7.1, p. 120].

Finally, let us state the assumptions on the initial conditions of the system:

( $\mathbf{A}$ ): Fix  $\mathbf{r}$  from ( $\mathbf{A}_{iii}^K$ ). For any  $m \geq 1$ ,  $\sup_{1 \leq i \leq N, N \in \mathbb{N}} \mathbb{E}|X_0^{i,N}|^m < \infty$  and

$$\sup_{N \in \mathbb{N}} \mathbb{E} \left[ \|\mu_0^N * V^N\|_{L^r(\mathbb{R}^d)}^m \right] < \infty.$$

A sufficient condition for ( $\mathbf{A}$ ) to hold is that particles are initially i.i.d. with a law which is in  $L^r$  and  $\alpha d < 1$ .

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  $M = 1$ , see Remark A.4.

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

**Definition 1.1.** *Given  $K$  satisfying ( $\mathbf{A}_i^K$ )-( $\mathbf{A}_{ii}^K$ ),  $u_0 \in L^1 \cap L^r(\mathbb{R}^d)$  with  $\mathbf{r} \geq \max(\mathbf{p}', \mathbf{q}')$  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)  $u \in \mathcal{C}([0, T]; L^1 \cap L^r(\mathbb{R}^d))$ ;

(ii)  $u$  satisfies the integral equation

$$u_t = e^{t\Delta} u_0 - \int_0^t \nabla \cdot (e^{(t-s)\Delta} (u_s K * u_s)) ds, \quad 0 \leq t \leq T. \quad (1.14)$$

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 A.1.1 the following result about the (local) well-posedness of the PDE (1.1) is established:

**Proposition 1.2.** *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.*

From here we can explicit a cut-off  $A$  (for the function  $F_A$  defined in (1.10)) to define the particle system. For a local mild solution  $u$  on  $[0, T]$ , we will use the following cut-off:

$$A_T := C_{K,d} \|u\|_{T, L^1 \cap L^r(\mathbb{R}^d)}, \quad (1.15)$$

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 blow up 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

Our first main result is the following claim, whose proof is detailed in Section 2.2:

**Theorem 1.3.** *Assume that the initial conditions  $\{\mu_0^N\}_{N \in \mathbb{N}}$  satisfy **(A)** and that the kernel  $K$  satisfies **(A<sup>K</sup>)**. Moreover, let **(A<sub>α</sub>)** hold true. Let  $T_{max}$  be the maximal existence time for (1.1) and fix  $T \in (0, T_{max})$ . In addition, let the dynamics of the particle system be given by (1.12) with  $A$  greater than  $A_T$  defined in (1.15).*

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

$$\left\| \|u^N - u\|_{T, L^1 \cap L^r(\mathbb{R}^d)} \right\|_{L^m(\Omega)} \leq C \left\| \|u_0^N - u_0\|_{L^1 \cap L^r(\mathbb{R}^d)} \right\|_{L^m(\Omega)} + CN^{-\varrho + \varepsilon},$$

where

$$\varrho = \min \left( \alpha \zeta, \frac{1}{2} \left( 1 - \alpha \left( d + d \left( 1 - \frac{2}{r} \right) \vee 0 \right) \right) \right). \quad (1.16)$$

It is clear from the definition (1.16) of  $\varrho$  that there is a trade-off between choosing  $\alpha$  close to  $\frac{1}{d}$  (as imposed by **(A<sub>α</sub>)**, provided  $r \leq 2$ ) so as to have more physical particles; and choosing  $\alpha$  smaller (such that  $\alpha \zeta = \frac{1}{2} (1 - \alpha d)$ , for a given  $\zeta$  and assuming again that  $r \leq 2$ ) so as to maximize the rate of convergence. The latter case could be of importance in numerical applications. We discuss the possible choices of the parameters in more details on various examples in Section 5.

In Corollary 2.2, we obtain the same rate for the genuine empirical measure of the system (1.3):

$$\left\| \sup_{t \in [0, T]} \|\mu_t^N - u_t\|_0 \right\|_{L^m(\Omega)} \leq C \left\| \|u_0^N - u_0\|_{L^1 \cap L^r(\mathbb{R}^d)} \right\|_{L^m(\Omega)} + CN^{-\varrho + \varepsilon},$$

where  $\|\cdot\|_0$  denotes the Kantorovich-Rubinstein metric (see Subsection 2.3).

In the next result, we remove the cutoff  $F_A$  from the drift of the particle system (i.e. we choose  $F_A = \text{Id}$  in (1.3), or equivalently  $A = \infty$ ), and the convergence is now in probability.

**Corollary 1.4.** *Let the same assumptions as in Theorem 1.3 hold. For each  $N \in \mathbb{N}^*$ , consider the solution  $(X^{i,N})_{1 \leq i \leq N}$  to (1.3) with  $A = \infty$  (i.e.  $F_A = \text{Id}$ ). We still denote by  $\mu_t^N$  the empirical measure of this particle system and  $u_t^N = \mu_t^N * V^N$ . Let  $\varrho$  be as in Theorem 1.3. Then for any  $\varepsilon \in (0, \varrho)$ , any  $\eta > 0$  and any  $m \geq 1$ , there exists  $C > 0$  such that, for any  $N \in \mathbb{N}^*$ ,*

$$\mathbb{P} \left( \|u_t^N - u_t\|_{T, L^1 \cap L^r(\mathbb{R}^d)} \geq \eta \right) \leq \frac{C}{\eta^m} \left( \left\| \|u_0^N - u_0\|_{L^1 \cap L^r(\mathbb{R}^d)} \right\|_{L^m(\Omega)} + CN^{-\varrho + \varepsilon} \right)^m.$$

Note that for fixed  $N \in \mathbb{N}^*$ , the drift term in Equation (1.3) is smooth and bounded, hence the particle system with  $F_A = \text{Id}$  has a unique strong solution. The proof of this result is presented in Subsection 2.3.

**Remark 1.5.** *Unlike particle systems in mean field interaction, we cannot expect here a  $N^{-1/2}$  rate of convergence, as for regular interactions [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  $\mu$  and its regularisation  $V^N * \mu$  is of order  $N^{-\alpha}$  too. Hence it is reasonable to expect no better than an  $N^{-\alpha}$  rate of convergence.*

*Assuming that the kernel  $K$  is such that  $\zeta$  can be chosen equal or close to 1, the rate in (1.16) is determined by the minimum between  $\alpha$  and  $\frac{1}{2} (1 - \alpha (d + d (1 - \frac{2}{r}) \vee 0))$ , so choosing  $\alpha$  accordingly does lead to a rate of convergence of “optimal” order  $N^{-\alpha}$ .*



**Remark 1.6.** Finally, let us consider the application of Theorem 1.3 to several classes of models (the details are given in Section 5):

- For Coulomb-type kernels, which includes the Biot-Savart kernel in dimension 2, the Riesz kernel with  $s = d - 2$  and the Keller-Segel kernel, the convergence happens for any  $\alpha < \frac{1}{2(d-1)}$  (note that in dimension 2, this accounts for choosing  $\alpha = (\frac{1}{2})^-$ ); the best possible rate of convergence is  $\varrho = \left(\frac{1}{2(d+1)}\right)^-$  for the choice  $\alpha = \left(\frac{1}{2(d+1)}\right)^+$ .
- The 2d Keller-Segel model (see kernel (5.2)). We recall that the PDE has a global solution whenever the critical parameter  $\chi$  satisfies  $\chi < 8\pi$ , and explodes in finite time otherwise (see [40]). In Theorem 1.3, we get a rate for any value of  $\chi$  which is almost  $\frac{1}{2(d+1)}$ . This result holds even if the PDE explodes in finite time ( $\chi > 8\pi$ ). In that case, one works on  $[0, T]$  for any  $T < T_{max}$ .
- The Riesz kernels with  $s > d - 2$  do not satisfy Assumption  $(\mathbf{A}_{ii}^K)$ , see Section 5. However, by imposing more regularity on the initial conditions and smaller values of  $\alpha$  ( $\alpha < \frac{1}{2d}$ ), we present a rate of convergence for singular Riesz kernels with  $s \in (d - 2, d - 1)$ , see Theorem 4.2.

**Related works.** 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, 42, 53] 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: when the particle system is somehow regularised (in a wider sense than (1.3), i.e. with moderate interactions or a more general cut-off in the kernel), Méléard [37] obtained a rate on the density of one particle for the 2d Navier-Stokes equation, and Bossy and Talay [5] got a rate for Burger's equation.

When the kernel is not regularised, let us mention that Fournier and Mischler [23] obtained a rate for the empirical measure of the so-called Nambu 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 Liouville equations associated to the mean field particle system, Jabin and Wang [30] 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 [50], who introduced the modulated energy method (see also [15] and [47]).

The present work is inspired by the new semigroup approach developed by Flandoli et al. [17], which 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 [51] for non-local conservation laws; Flandoli et al. [20] for the 2d Navier-Stokes equation; and Olivera et al. [43] 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.

## 1.4 Second main result: Propagation of chaos

We now tackle the question of propagation of chaos of (1.3) towards (1.4).

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.** Given  $K$  satisfying  $(\mathbf{A}_i^K)$ - $(\mathbf{A}_{ii}^K)$ ,  $u_0 \in L^1 \cap L^r(\mathbb{R}^d)$  with  $r \geq \max(\mathbf{p}', \mathbf{q}')$  and  $T > 0$ , 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  $q \in \mathcal{C}([0, T]; L^1 \cap L^r(\mathbb{R}^d))$ ;
- (iii) For any  $f \in \mathcal{C}_c^2(\mathbb{R}^d)$ , the process  $(M_t)_{t \in [0, T]}$  defined as

$$M_t := f(w_t) - f(w_0) - \int_0^t \left[ \Delta f(w_s) + \nabla f(w_s) \cdot (K * q_s(w_s)) \right] ds$$

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

We will see that for  $f \in L^1 \cap \tilde{L}^r(\mathbb{R}^d)$ ,  $K * f$  is bounded, hence  $(M_t)_{t \geq 0}$  is well-defined. The following claim establishes the well-posedness of the martingale problem and is proven in Section 3.1:

**Proposition 1.8.** *Let  $T < T_{max}$ . 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 linear version of (1.4) admits a unique weak solution. We choose this approach as we have *a priori* information about the PDE.

By Proposition 1.8, McKean-Vlasov Equation (1.4) admits a unique weak solution, so let us consider  $N$  independent copies  $(\tilde{X}^i)_{i \in \{1, \dots, N\}}$  and denote by  $(W^i)_{i \in \{1, \dots, N\}}$  the associated Brownian motions. Consider the particles  $(X^i)_{i \in \{1, \dots, N\}}$  defined as the (strong) solutions of (1.12) driven by the same Brownian motions  $(W^i)_{i \in \{1, \dots, N\}}$  and with initial conditions  $(\tilde{X}_0^i)_{i \in \{1, \dots, N\}}$ . A simple consequence of Theorem 1.3 is that the particles  $(X^i)_{i \in \{1, \dots, N\}}$  (with mollified interaction kernel and cut-off  $F_A$ ) are uniformly close to the non-regularised McKean-Vlasov particles. More precisely,

$$\left\| \max_{i \in \{1, \dots, N\}} \sup_{t \in [0, T]} |X_t^{i, N} - \tilde{X}_t^i| \right\|_{L^m(\Omega)} \leq C N^{-\epsilon + \varepsilon}, \quad \forall N \in \mathbb{N}^*. \quad (1.17)$$

Now, we present our second main result. It will be proven in Section 3.2.

**Theorem 1.9.** *Let the hypotheses of Theorem 1.3 hold. In particular, recall that  $u \in \mathcal{C}([0, T], L^1 \cap L^r(\mathbb{R}^d))$ . Assume further that the family of random variables  $\{X_0^i, i \in \mathbb{N}\}$  is identically distributed and that  $\langle u_0^N, \varphi \rangle \rightarrow \langle u_0, \varphi \rangle$  in probability, for any  $\varphi \in \mathcal{C}_b(\mathbb{R}^d)$ . Then, the empirical measure  $\mu^N$  (defined in (1.13) as a probability measure on  $\mathcal{C}([0, T]; \mathbb{R}^d)$ ) converges in probability towards  $\mathbb{Q}$ , which is the law of the unique weak solution 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 1.3 is very much related to the propagation of chaos and should be considered as the most important ingredient when proving Theorem 1.9.

**Remark 1.10.** *As in the previous section, the case of the singular Riesz kernels is not covered by Proposition 1.8 and Theorem 1.9. However, we are able to give very similar statements of well-posedness of the McKean-Vlasov equations for Riesz kernels with  $s \in (d-2, d-1)$  in Section 4.3, provided that the initial conditions are smoother and  $\alpha < \frac{1}{2d}$ .*

**Related works.** 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, 34, 39, 46] 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 as in the above works one cannot treat singular interactions in the drift term with kernels  $K$  of the order  $\frac{1}{|x|^{s+1}}$  for  $s \in (0, d-1)$ , in any dimension.

Let us focus on the Keller-Segel model, for which the kernel is given in (5.2). 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.

## 1.5 Organisation of the paper

In Section 2, we prove Theorem 1.3 and its corollaries. In Section 3, the existence and uniqueness of the solution of the martingale problem ( $\mathcal{MP}$ ) is proven, as well as the propagation of chaos for the empirical measure of (1.3) (Theorem 1.9). In Section 4 we extend the results of Sections 2 and 3. Finally, we present some examples and applications of our results in Section 5. In the Appendix, the existence and uniqueness of the Fokker-Planck equation (1.1) and its cut-off version are studied in Section A.1. Then, one may find the time and space estimates for some stochastic convolution integrals in Section A.2, and the proof of a result about the boundedness of the mollified empirical measure in  $L^2([0, T]; H_r^\beta(\mathbb{R}^d))$  in Section A.3.

**Acknowledgment.** The authors are very grateful to an anonymous referee for his careful reading of the manuscript, and in particular for suggesting us the result concerning the particle system without cutoff (Corollary 1.4).

C.O. is partially supported by CNPq through the grant 426747/2018-6 and FAPESP by the grants 2020/15691-2 and 2020/04426-6. 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.

## 2 Rate of convergence

In this section, we prove our first main result (Theorem 1.3). More precisely, in Subsection 2.1 we establish the equation satisfied by  $u^N$  and present the boundedness result for it that is proved in the appendix. Then in Subsection 2.2, we prove Theorem 1.3. In Subsection 2.3, we present the corollaries of Theorem 1.3.

First, we give a preliminary lemma that is essential for our calculations and for the analysis of the mild form of (1.1) and the mild form of the cut-off PDE (A.6).

**Lemma 2.1.** *Let  $K$  be satisfying Assumptions  $(\mathbf{A}_i^K)$  and  $(\mathbf{A}_{ii}^K)$  and recall that  $\mathbf{r} \geq \max(\mathbf{p}', \mathbf{q}')$ . There exists  $C_{K,d} > 0$  (which depends on  $K$  and  $d$  only) such that for any  $f \in L^1 \cap L^{\mathbf{r}}(\mathbb{R}^d)$ ,*

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

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

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

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

We recall here some classical properties of the heat kernel that will be used throughout this section. Applying the convolution inequality [7, Th. 4.15] for  $p \geq 1$  and using the equality

$$\left\| \nabla \frac{1}{(4\pi t)^{\frac{d}{2}}} e^{-\frac{|\cdot|^2}{4t}} \right\|_{L^1(\mathbb{R}^d)} = \frac{C}{\sqrt{t}},$$

$$\left\| \nabla e^{t\Delta} \right\|_{L^p \rightarrow L^p} \leq \frac{C}{\sqrt{t}}. \quad (2.1)$$

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 (2.1) extends to

$$\left\| (I - \Delta)^{\frac{\beta}{2}} e^{t\Delta} \right\|_{L^p \rightarrow L^p} \leq C t^{-\frac{\beta}{2}}. \quad (2.2)$$

## 2.1 Properties of the regularised empirical measure

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

$$\begin{aligned} u_t^N(x) &= e^{t\Delta} u_0^N(x) - \int_0^t \nabla \cdot e^{(t-s)\Delta} \langle \mu_s^N, V^N(x - \cdot) F(K * u_s^N(\cdot)) \rangle ds \\ &\quad - \frac{1}{N} \sum_{i=1}^N \int_0^t e^{(t-s)\Delta} \nabla V^N(x - X_s^{i,N}) \cdot dW_s^i, \end{aligned} \quad (2.3)$$

and the following property: For  $q \geq 1$ , one has

$$\sup_{N \in \mathbb{N}^*} \mathbb{E} \left[ \sup_{t \in [0, T]} \|u_t^N\|_{L^r(\mathbb{R}^d)}^q \right] < \infty. \quad (2.4)$$

This estimate will be proven in Proposition A.12 (that we apply here for  $\beta = 0$ ).

Now, let us establish Equation (2.3). 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 (1.12) as

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

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

$$\begin{aligned} u_t^N(x) = & u_0^N(x) - \frac{1}{N} \sum_{i=1}^N \int_0^t \nabla V^N(x - X_s^{i,N}) \cdot F(K * u_s^N(X_s^{i,N})) ds \\ & - \frac{1}{N} \sum_{i=1}^N \int_0^t \nabla V^N(x - X_s^{i,N}) \cdot dW_s^i + \frac{1}{N} \sum_{i=1}^N \int_0^t \Delta V^N(x - X_s^{i,N}) ds. \end{aligned} \quad (2.5)$$

Notice that

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

and

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

The preceding equalities combined with (2.5) lead to

$$\begin{aligned} u_t^N(x) = & u_0^N(x) - \int_0^t \langle \mu_s^N, \nabla V^N(x - \cdot) \cdot F(K * u_s^N(\cdot)) \rangle ds \\ & - \frac{1}{N} \sum_{i=1}^N \int_0^t \nabla V^N(x - X_s^{i,N}) \cdot dW_s^i + \int_0^t \Delta u_s^N(x) ds. \end{aligned} \quad (2.6)$$

The following mild form is immediately derived:

$$\begin{aligned} u_t^N(x) = & e^{t\Delta} u_0^N(x) - \int_0^t e^{(t-s)\Delta} \langle \mu_s^N, \nabla V^N(x - \cdot) \cdot F(K * u_s^N(\cdot)) \rangle ds \\ & - \frac{1}{N} \sum_{i=1}^N \int_0^t e^{(t-s)\Delta} \nabla V^N(x - X_s^{i,N}) \cdot dW_s^i. \end{aligned} \quad (2.7)$$

Finally, developing the scalar product, one has

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

Combining the latter with the fact that  $e^{t\Delta} \nabla \cdot f = \nabla \cdot e^{t\Delta} f$ , we deduce the mild form (2.3) from (2.7).

## 2.2 Proof of Theorem 1.3

Let  $u$  the unique mild solution to (1.1) on  $[0, T]$ . By assumption  $A$  is large enough, so  $u$  also satisfies Equation (A.7).

In view of the mild formulas (2.3) for  $u^N$  and (A.7) for  $u$ , it comes

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

Now we subtract and add the term  $\int_0^t \nabla \cdot e^{(t-s)\Delta} \langle \mu_s^N, V^N(x - \cdot) F(K * u_s^N(x)) \rangle ds$ , which is also equal to  $\int_0^t \nabla \cdot e^{(t-s)\Delta} u_s^N(x) F(K * u_s^N(x)) ds$ . Hence we get

$$\begin{aligned} u_t^N(x) - u_t(x) = & e^{t\Delta} (u_0^N - u_0)(x) + \int_0^t \nabla \cdot e^{(t-s)\Delta} (u_s F(K * u_s) - u_s^N F(K * u_s^N))(x) ds \\ & + E_t(x) - M_t^N(x), \end{aligned}$$

where we have set

$$\begin{aligned} E_t(x) &:= \int_0^t \nabla \cdot e^{(t-s)\Delta} \langle \mu_s^N, V^N(x - \cdot) (F(K * u_s^N(x)) - F(K * u_s^N(\cdot))) \rangle ds, \\ M_t^N(x) &:= \frac{1}{N} \sum_{i=1}^N \int_0^t e^{(t-s)\Delta} \nabla V^N(x - X_s^{i,N}) \cdot dW_s^i. \end{aligned} \quad (2.8)$$

For  $p \in \{1, \mathbf{r}\}$ , in view of the estimate (2.1), one has

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

We will use the bound  $\|e^{t\Delta}(u_0^N - u_0)\|_{L^p(\mathbb{R}^d)} \leq \|u_0^N - u_0\|_{L^p(\mathbb{R}^d)}$ . Observe also that

$$\|(u_s F(K * u_s) - u_s^N F(K * u_s^N))\|_{L^p(\mathbb{R}^d)} \leq \|(u_s (F(K * u_s) - F(K * u_s^N)))\|_{L^p(\mathbb{R}^d)} + \|(F(K * u_s^N)(u_s - u_s^N))\|_{L^p(\mathbb{R}^d)}.$$

Hence, by the Lipschitz property of  $F$  on one term and the boundedness of  $F$  for the other, it follows that

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

Besides, by Proposition A.7 for  $\beta = 0$ , it comes that for some  $C > 0$  which depends on  $\|u\|_{T, L^1 \cap L^{\mathbf{r}}(\mathbb{R}^d)}$ ,  $\|F\|_{L^\infty(\mathbb{R}^d)}$  and  $\|F\|_{\text{Lip}}$ ,

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

Finally we apply Lemma 2.1 and obtain

$$\begin{aligned} \|u_t^N - u_t\|_{L^p(\mathbb{R}^d)} &\leq \|u_0^N - u_0\|_{L^p(\mathbb{R}^d)} + C \int_0^t \frac{1}{\sqrt{t-s}} \|u_s^N - u_s\|_{L^p(\mathbb{R}^d)} ds \\ &\quad + C \int_0^t \frac{1}{\sqrt{t-s}} \|u_s - u_s^N\|_{L^1 \cap L^{\mathbf{r}}(\mathbb{R}^d)} ds \\ &\quad + \|E_t\|_{L^p(\mathbb{R}^d)} + \|M_t^N\|_{L^p(\mathbb{R}^d)}. \end{aligned} \quad (2.9)$$

Therefore, considering (2.9) for both  $p = 1$  and  $p = \mathbf{r}$ , we deduce that

$$\begin{aligned} \|u_t^N - u_t\|_{L^1 \cap L^{\mathbf{r}}(\mathbb{R}^d)} &\leq \|u_0^N - u_0\|_{L^1 \cap L^{\mathbf{r}}(\mathbb{R}^d)} + C \int_0^t \frac{1}{\sqrt{t-s}} \|u_s^N - u_s\|_{L^1 \cap L^{\mathbf{r}}(\mathbb{R}^d)} ds \\ &\quad + \|E_t\|_{L^1 \cap L^{\mathbf{r}}(\mathbb{R}^d)} + \|M_t^N\|_{L^1 \cap L^{\mathbf{r}}(\mathbb{R}^d)}. \end{aligned}$$

Using the Grönwall lemma for convolution integrals (see e.g. [29, Lemma 7.1.1]), we obtain

$$\|u^N - u\|_{t, L^1 \cap L^{\mathbf{r}}(\mathbb{R}^d)} \leq C (\|u_0^N - u_0\|_{L^1 \cap L^{\mathbf{r}}(\mathbb{R}^d)} + \|E\|_{t, L^1 \cap L^{\mathbf{r}}(\mathbb{R}^d)} + \|M^N\|_{t, L^1 \cap L^{\mathbf{r}}(\mathbb{R}^d)}). \quad (2.10)$$

Quantitative convergence of particle systems

• *Now, we estimate the moments of  $\|E\|_{t, L^q(\mathbb{R}^d)}$  for  $1 \leq q \leq r$ .* Using (2.1) for  $p = q$ , we observe that

$$\|E_t\|_{L^q(\mathbb{R}^d)} \leq C \int_0^t \frac{1}{(t-s)^{1/2}} \left( \int_{\mathbb{R}^d} |\langle \mu_s^N, V^N(x-\cdot)(F(K * u_s^N(\cdot)) - F(K * u_s^N(x))) \rangle|^q dx \right)^{\frac{1}{q}} ds.$$

By using the positivity of  $V^N$ , we get

$$\|E_t\|_{L^q(\mathbb{R}^d)} \leq C \int_0^t \frac{1}{(t-s)^{\frac{1}{2}}} \left( \int_{\mathbb{R}^d} \langle \mu_s^N, V^N(x-\cdot) |F(K * u_s^N(\cdot)) - F(K * u_s^N(x))| \rangle^q dx \right)^{\frac{1}{q}} ds.$$

Using the Lipschitz regularity of  $F$  and the  $\zeta$ -Hölder continuity of  $K * u^N$  from Assumption  $(\mathbf{A}_{iii}^K)$ , we get

$$\|E_t\|_{L^q(\mathbb{R}^d)} \leq C \int_0^t \frac{\|u_s^N\|_{L^1 \cap L^r(\mathbb{R}^d)}}{(t-s)^{\frac{1}{2}}} \left( \int_{\mathbb{R}^d} \langle \mu_s^N, V^N(x-\cdot) |\cdot - x|^\zeta \rangle^q dx \right)^{\frac{1}{q}} ds.$$

Since  $V$  is compactly supported (without loss of generality, assume that the support of  $V$  is included in the unit ball), we have that  $V^N(x-y) |y-x|^\zeta \leq N^{-\alpha\zeta} V^N(x-y)$ . Thus

$$\begin{aligned} \|E_t\|_{L^q(\mathbb{R}^d)} &\leq \frac{C}{N^{\alpha\zeta}} \int_0^t \frac{1}{(t-s)^{\frac{1}{2}}} \|u_s^N\|_{L^1 \cap L^r(\mathbb{R}^d)} \|u_s^N\|_{L^q(\mathbb{R}^d)} ds \\ &\leq \frac{C}{N^{\alpha\zeta}} \int_0^t \frac{1}{(t-s)^{\frac{1}{2}}} \|u_s^N\|_{L^1 \cap L^r(\mathbb{R}^d)}^2 ds, \end{aligned}$$

where the last inequality holds by interpolation. Apply Hölder's inequality with  $p = \frac{3}{2}$  to obtain

$$\|E_t\|_{L^q(\mathbb{R}^d)} \leq \frac{C}{N^{\alpha\zeta}} \left( \int_0^t (t-s)^{-\frac{3}{4}} ds \right)^{\frac{2}{3}} \left( \int_0^t \|u_s^N\|_{L^1 \cap L^r(\mathbb{R}^d)}^6 ds \right)^{\frac{1}{3}}.$$

Finally, we have from Jensen's inequality that for  $m \geq 3$ ,

$$\| \|E\|_{t, L^q(\mathbb{R}^d)} \|_{L^m(\Omega)} \leq \frac{C}{N^{\alpha\zeta}} \left( \int_0^t \mathbb{E} \left[ \|u_s^N\|_{L^1 \cap L^r(\mathbb{R}^d)}^{2m} \right] ds \right)^{\frac{1}{m}}$$

and the bound (2.4) permits to conclude that

$$\| \|E\|_{t, L^q(\mathbb{R}^d)} \|_{L^m(\Omega)} \leq \frac{C}{N^{\alpha\zeta}}. \quad (2.11)$$

This inequality immediately extends to  $1 \leq m < 3$ .

• *We turn to the moments of  $\|M_t^N\|_{L^1(\mathbb{R}^d)}$ .* Note that  $(\|M_s^N\|_{L^1(\mathbb{R}^d)})_{s \geq 0}$  is not a martingale, but a stochastic convolution integral. We explain how to deal with this term in Appendix A.2. By Proposition A.8, we have: for any  $\varepsilon > 0$  arbitrary small, there exists  $C > 0$  such that

$$\left\| \sup_{s \in [0, t]} \|M_s^N\|_{L^1(\mathbb{R}^d)} \right\|_{L^m(\Omega)} \leq C N^{-\frac{1}{2}(1-\alpha d) + \varepsilon}, \quad \forall N \in \mathbb{N}^*. \quad (2.12)$$

• *It remains to estimate the moments of  $\|M_t^N\|_{L^r(\mathbb{R}^d)}$ .* By Proposition A.8 for  $p = r$ , we have: for any  $\varepsilon > 0$ , there exists  $C > 0$  such that

$$\left\| \sup_{s \in [0, t]} \|M_s^N\|_{L^r(\mathbb{R}^d)} \right\|_{L^m(\Omega)} \leq C N^{-\frac{1}{2}(1-\alpha(d+\kappa_r)) + \varepsilon}, \quad (2.13)$$

where  $\varkappa_r = d(1 - \frac{2}{r}) \vee 0$ . By the assumption  $(\mathbf{A}_\alpha)$ , the above exponent in  $N$  is indeed negative.

• *Conclusion.* Plugging Inequalities (2.11)-(2.13) in (2.10), we conclude that for any  $\varepsilon > 0$  small enough, there exists  $C > 0$  such that for any  $N \in \mathbb{N}^*$ ,

$$\left\| \|u^N - u\|_{t, L^1 \cap L^r(\mathbb{R}^d)} \right\|_{L^m(\Omega)} \leq C \left( \left\| \|u_0^N - u_0\|_{L^1 \cap L^r(\mathbb{R}^d)} \right\|_{L^m(\Omega)} + N^{-\alpha\zeta} + N^{-\frac{1}{2}(1-\alpha(d+\varkappa_r))+\varepsilon} \right).$$

### 2.3 Corollaries of Theorem 1.3

In view of the previous result, we 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  $\mu$  and  $\nu$  on  $\mathbb{R}^d$ ,

$$\|\mu - \nu\|_0 = \sup \left\{ \int_{\mathbb{R}^d} \phi d(\mu - \nu); \phi \text{ Lipschitz with } \|\phi\|_{L^\infty(\mathbb{R}^d)} \leq 1 \text{ and } \|\phi\|_{\text{Lip}} \leq 1 \right\}. \quad (2.14)$$

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

**Corollary 2.2.** *Let the same assumptions as in Theorem 1.3 hold. Let  $\varrho$  be as in Theorem 1.3. Then for any  $\varepsilon \in (0, \varrho)$ , there exists  $C > 0$  such that, for any  $N \in \mathbb{N}^*$ ,*

$$\left\| \sup_{t \in [0, T]} \|\mu_t^N - u_t\|_0 \right\|_{L^m(\Omega)} \leq C \left( \left\| \|u_0^N - u_0\|_{L^1 \cap L^r(\mathbb{R}^d)} \right\|_{L^m(\Omega)} + N^{-\varrho+\varepsilon} \right).$$

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

$$|\langle u_t^N, \phi \rangle - \langle \mu_t^N, \phi \rangle| \leq \frac{C \|\phi\|_{\text{Lip}}}{N^\alpha} \quad a.s. \quad (2.15)$$

Indeed,

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

Recalling the definition (2.14) 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 (2.15) to the first term on the right-hand side of the above inequality, and Theorem 1.3 to the second term, we obtain the inequality of Corollary 2.2.  $\square$

We now carry out the proof of Corollary 1.4, for the particle system without cutoff.

*Proof of Corollary 1.4.* Let us introduce some notations to distinguish the particle systems with and without cutoff. Recall that  $A_T$  is given in (1.15), and fix  $A > A_T$  to be precisely chosen later. The particles which solve (1.3) with cutoff  $A$  will be denoted here by  $(X^{i, N, (A)})_{1 \leq i \leq N}$  and those which solve (1.3) with  $A = \infty$  will be denoted by  $(X^{i, N, (\infty)})_{1 \leq i \leq N}$ . Accordingly, we denote by



$\mu^{N,(A)}$  the empirical measure of  $(X_t^{i,N,(A)})_{1 \leq i \leq N}$ ,  $\mu^{N,(\infty)}$  the empirical measure of  $(X_t^{i,N,(\infty)})_{1 \leq i \leq N}$ , as well as  $u^{N,(A)} = \mu^{N,(A)} * V^N$  and  $u^{N,(\infty)} = \mu^{N,(\infty)} * V^N$ . We assume these particles have the same initial conditions and are driven by the same family of independent  $\mathbb{R}^d$ -valued Brownian motions  $(W^i)_{i \in \mathbb{N}^*}$ . For any  $N \in \mathbb{N}^*$ , define

$$\Omega_N = \left\{ \sup_{\substack{t \in [0, T] \\ i \in \{1, \dots, N\}}} \frac{1}{N} \left| \sum_{k=1}^N (K * V^N)(X_t^{i,N,(A)} - X_t^{k,N,(A)}) \right| \leq A \right\}$$

and observe that on  $\Omega_N$ , we have  $X_t^{i,N,(A)} = X_t^{i,N,(\infty)}$  for all  $t \in [0, T]$  and all  $i \in \{1, \dots, N\}$ . Since  $\frac{1}{N} \sum_{k=1}^N (K * V^N)(X_t^{i,N,(A)} - X_t^{k,N,(A)}) = K * u_t^{N,(A)}(X_t^{i,N,(A)})$ , we also get that on  $\Omega_N$ ,  $u_t^{N,(A)} = u_t^{N,(\infty)}$  for all  $t \in [0, T]$  and all  $i \in \{1, \dots, N\}$ . Hence

$$\begin{aligned} \mathbb{P} \left( \|u_t^{N,(\infty)} - u_t\|_{T, L^1 \cap L^r(\mathbb{R}^d)} \geq \eta \right) &= \mathbb{P} \left( \Omega_N^c \cap \{ \|u_t^{N,(\infty)} - u_t\|_{T, L^1 \cap L^r(\mathbb{R}^d)} \geq \eta \} \right) \\ &\quad + \mathbb{P} \left( \Omega_N \cap \{ \|u_t^{N,(\infty)} - u_t\|_{T, L^1 \cap L^r(\mathbb{R}^d)} \geq \eta \} \right) \\ &\leq \mathbb{P}(\Omega_N^c) + \mathbb{P} \left( \|u_t^{N,(A)} - u_t\|_{T, L^1 \cap L^r(\mathbb{R}^d)} \geq \eta \right). \end{aligned}$$

Now by Lemma 2.1 we have that  $|K * u_t^{N,(A)}(X_t^{i,N,(A)})| \leq C_{K,d} \|u_t^{N,(A)}\|_{L^1 \cap L^r(\mathbb{R}^d)}$ . Thus we get that for  $A = C_{K,d}(\eta + \|u\|_{T, L^1 \cap L^r(\mathbb{R}^d)})$ ,

$$\begin{aligned} \mathbb{P}(\Omega_N^c) &\leq \mathbb{P} \left( \|u^{N,(A)}\|_{T, L^1 \cap L^r(\mathbb{R}^d)} > \frac{A}{C_{K,d}} \right) \\ &\leq \mathbb{P} \left( \|u\|_{T, L^1 \cap L^r(\mathbb{R}^d)} + \|u^{N,(A)} - u\|_{T, L^1 \cap L^r(\mathbb{R}^d)} > \frac{A}{C_{K,d}} \right) \\ &\leq \mathbb{P} \left( \|u^{N,(A)} - u\|_{T, L^1 \cap L^r(\mathbb{R}^d)} > \eta \right). \end{aligned}$$

Hence

$$\mathbb{P} \left( \|u_t^{N,(\infty)} - u_t\|_{T, L^1 \cap L^r(\mathbb{R}^d)} \geq \eta \right) \leq 2\mathbb{P} \left( \|u^{N,(A)} - u\|_{T, L^1 \cap L^r(\mathbb{R}^d)} > \eta \right).$$

Now using Markov's inequality and Theorem 1.3, we obtain the desired result.  $\square$

### 3 Propagation of chaos

In this section we study the well-posedness of the nonlinear SDE (1.4) and then the propagation of chaos of the particle system (1.12).

#### 3.1 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. Assuming there is a solution to the martingale problem, we study the mild equation of its time-marginals. We will see that this equation admits a unique solution in a suitable functional space (**Step 1**). This will enable us to study a linear version of the martingale problem ( $\mathcal{MP}$ ) (**Step 2**). Analysing this linear martingale problem, we will get the uniqueness and existence for ( $\mathcal{MP}$ ) (**Steps 3 and 4**).

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

$$\sup_{t \leq T} \|K * q_t\|_{L^\infty(\mathbb{R}^d)} \leq C_{K,d} \sup_{t \leq T} \|q_t\|_{L^1 \cap L^r(\mathbb{R}^d)}. \quad (3.1)$$

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. [54, 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 (1.14) and we know it admits a unique solution in the sense of Definition 1.1 up to time  $T$ . Meaning, as  $q \in \mathcal{C}([0, T]; L^1 \cap L^r(\mathbb{R}^d))$ , the one-dimensional time marginals of  $\mathbb{Q}$  are uniquely determined.

**Step 2.** Define the corresponding linear martingale problem  $(\mathcal{MP}_{\text{lin}})$  in the same way as  $(\mathcal{MP})$ , except that in the definition of the process  $(M_t)_{t \leq T}$  from  $(\mathcal{MP})$ -(iii), fix  $q$  to be the unique mild solution  $u$  to (1.1). 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, weak uniqueness holds. Let us show that the probability measure  $\mathbb{P} := \mathcal{L}(Y)$  solves  $(\mathcal{MP}_{\text{lin}})$ . Assume for a moment that the family  $(\mathbb{P}_t(dx))_{t \geq 0} = (p_t(x)dx)_{t \geq 0}$  (absolute continuity follows from bounded drift and Girsanov transformation) belongs to  $\mathcal{C}([0, T]; L^1 \cap L^r(\mathbb{R}^d))$ . We will prove this fact in **Step 4**. Then, as  $\mathbb{P}$  is  $\mathcal{L}(Y)$ , all the requirements of  $(\mathcal{MP}_{\text{lin}})$  are satisfied. In addition, this solution is unique.

**Step 3.** The previous step immediately yields the uniqueness of solutions to  $(\mathcal{MP})$  (as any solution to  $(\mathcal{MP})$  is also a solution to  $(\mathcal{MP}_{\text{lin}})$  by the uniqueness of the one-dimensional time marginals). We now turn to the question of existence.

A candidate for a solution to the problem  $(\mathcal{MP})$  is the probability measure  $\mathbb{P}$  defined above. To prove the latter solves  $(\mathcal{MP})$ , we need to ensure that the family of marginal laws  $(\mathbb{P}_t)_{0 \leq t \leq T}$  is exactly the family  $(u_t)_{0 \leq t \leq T}$  we used to define the drift in  $(\mathcal{MP}_{\text{lin}})$ .

To do so, for  $0 < t \leq T$ , one derives the mild equation for  $\mathbb{P}_t(dx) = p_t(x)dx$ . Following the same arguments as in [54, Section 4], as the drift is bounded, we have that for a.e.  $x \in \mathbb{R}^d$ ,

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

Assume again that  $p \in \mathcal{C}([0, T]; L^1 \cap L^r(\mathbb{R}^d))$ . The previous equation is a linearized version of Eq. (1.14) and, by the same arguments as in Proposition 1.2, it admits a unique solution in  $\mathcal{C}([0, T]; L^1 \cap L^r(\mathbb{R}^d))$ . Since both  $u$  and  $p$  solve this equation, they must coincide and we have the desired result :  $(p_t)_{t \in [0, T]} = (u_t)_{t \in [0, T]}$ .

**Step 4.** It only remains to prove that  $p \in \mathcal{C}([0, T]; L^1 \cap L^r(\mathbb{R}^d))$ . Obviously, as we work with a family of probability density functions, we only need to prove that  $p \in \mathcal{C}([0, T]; L^r(\mathbb{R}^d))$ . Performing the same calculations as in the proof of Proposition A.6, we get that

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

In view of Lemma 2.1, one has

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

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

### 3.2 Proof of Theorem 1.9

To prove Theorem 1.9, we will show that  $\mu^N$  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:

$$M_t := f(w_t) - f(w_0) - \int_0^t \left[ \Delta f(w_s) + \nabla f(w_s) \cdot F_A(K * q_s(w_s)) \right] ds.$$

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  $\mu^N$  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  $\Pi^\infty$  of  $\Pi^N$  is  $\delta_{\mathbb{Q}}$ . The latter is done by showing that under  $\Pi^\infty$  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  $\Pi^N$  and use the convergence of  $\Pi^N$  to  $\Pi^\infty$  to get the desired result. This is where  $\mu^N$  and the particle system appear.

However, here the situation is slightly modified. Namely, at the level of  $\Pi^N$ , we need to keep track not just of  $\mu^N$ , 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  $(\mu^N, u^N)$ . 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

$$\mathcal{Y} = \mathcal{C}([0, T]; L^1 \cap L^r(\mathbb{R}^d)). \quad (3.2)$$

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  $(\mu^N, u^N)$  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 1.3).

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  $\Pi^\infty$  is concentrated around  $\mathbb{Q}$ .

**Lemma 3.1.**  *$\tilde{\Pi}^\infty$ -almost surely,  $\boldsymbol{\mu}_t$  is absolutely continuous w.r.t. the Lebesgue measure and its density is  $\boldsymbol{\mu}_t(dx) = \mathbf{u}_t(x)dx$ .*

*Proof.* This is a consequence of Inequality (2.15). 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,

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

where the last inequality comes from (2.15). 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  $t \in [0, T]$ ,

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

□

The following proposition will be the last ingredient needed for the proof of Theorem 1.9.

**Proposition 3.2.** *Let  $p \in \mathbb{N}$ ,  $f \in \mathcal{C}_b^2(\mathbb{R}^d)$ ,  $\Phi \in \mathcal{C}_b(\mathbb{R}^{dp})$  and  $0 < s_1 < \dots < s_p \leq s < t \leq T$ . Define  $\Gamma$  as the following function on  $\mathcal{H}$  :*

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

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

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

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

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

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

$$\begin{aligned} I_n & \leq \|\Phi\|_\infty \|\nabla f\|_\infty \langle \boldsymbol{\mu}^n, \int_s^t |F(K * \mathbf{u}_\sigma^n(\cdot)) - F(K * \mathbf{u}_\sigma(\cdot))| d\sigma \rangle \\ & \leq C \int_s^t \langle \boldsymbol{\mu}_\sigma^n, |K * (\mathbf{u}_\sigma^n - \mathbf{u}_\sigma)| \rangle d\sigma \leq C \int_s^t \|K * (\mathbf{u}_\sigma^n - \mathbf{u}_\sigma)\|_\infty d\sigma. \end{aligned} \quad (3.3)$$

In view of Lemma 2.1, one has

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

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  $G : \mathcal{C}([0, T]; \mathbb{R}^d) \rightarrow \mathbb{R}$  defined by

$$G(x) = \Phi(x_{s_1}, \dots, x_{s_p}) \left[ f(x_t) - f(x_s) - \int_s^t \Delta f(x_\sigma) d\sigma - \int_s^t F_A(K * \mathbf{u}_\sigma(x_\sigma)) \cdot \nabla f(x_\sigma) d\sigma \right].$$

Let  $(x^n)_{n \geq 1}$  a sequence converging in  $\mathcal{C}([0, T]; \mathbb{R}^d)$  to  $x$ . To prove  $G(x^n) \rightarrow G(x)$  as  $n \rightarrow \infty$ , having in mind the properties of  $f$  and  $\Phi$ , we should only concentrate on the term  $\int_s^t F_A(K * \mathbf{u}_\sigma(x_\sigma^n)) \cdot \nabla f(x_\sigma^n) d\sigma$ . Here we use the continuity property  $(\mathbf{A}_{iii}^K)$  to deduce that  $K * \mathbf{u}_\sigma(x_\sigma^n)$  converges to  $K * \mathbf{u}_\sigma(x_\sigma)$  and by dominated convergence,

$$\int_s^t F_A(K * \mathbf{u}_\sigma(x_\sigma^n)) \cdot \nabla f(x_\sigma^n) d\sigma \rightarrow \int_s^t F_A(K * \mathbf{u}_\sigma(x_\sigma)) \cdot \nabla f(x_\sigma) d\sigma, \quad \text{as } n \rightarrow \infty.$$

**Conclusion.** Combine Step 1 and Step 2 to finish the proof.  $\square$

We have all the elements in hand to finish the proof of Theorem 1.9. By Lemma 3.1 and Proposition 3.2, we get that  $\boldsymbol{\mu} \in \text{supp}(\tilde{\Pi}^\infty)$  solves the nonlinear martingale problem  $(\mathcal{MP}_A)$ . Choose  $A > A_T := C_{k,D} \|q\|_{T, L^1 \cap L^r(\mathbb{R}^d)}$  and lift the cut-off (see (3.1)). 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  $\tilde{\Pi}^N$  which is  $\delta_{\mathbb{Q}}$ .

## 4 Riesz-like singular kernels

In this section, we deal with kernels that are too singular to be covered by Assumption  $(\mathbf{A}^K)$ , typically deriving from Riesz potentials (1.2) with  $s \in (d-2, d-1)$  (see Section 5 for more details).

The main results are the analogous of Theorem 1.3 (see Theorem 4.2) and Theorem 1.9 (see Theorem 4.8).

The following alternative to Assumption  $(\mathbf{A}^K)$  will now be considered:

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

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

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

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

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

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

$(\tilde{\mathbf{A}}_\alpha)$ : The parameters  $\alpha$ ,  $\beta$  and  $\tilde{r}$  (which appear respectively in (1.11) and  $(\tilde{\mathbf{A}}_{iii}^K)$ ) satisfy

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

A kernel which would satisfy Assumption  $(\mathbf{A}^K)$  also satisfies  $(\tilde{\mathbf{A}}^K)$  (by Hölder's inequality in  $(\tilde{\mathbf{A}}_i^K)$  and a Sobolev embedding in  $(\tilde{\mathbf{A}}_{iii}^K)$ ), hence this new assumption is more general. However  $(\tilde{\mathbf{A}}^K)$ - $(\tilde{\mathbf{A}}_\alpha)$  impose more restrictions on the choice of the parameters  $\alpha$ ,  $\beta$  and  $\tilde{r}$ . For instance, we are no longer able to choose  $\alpha = (\frac{1}{d})^-$ , but instead we have  $\alpha = (\frac{1}{2d})^-$  at best.

Finally, let us now state the assumptions on the initial conditions of the system:

$(\tilde{\mathbf{A}})$ : Let  $\tilde{r} \in (d, +\infty)$  and  $\beta \in (\frac{d}{\tilde{r}}, 1)$ . For any  $m \geq 1$ ,  $\sup_{1 \leq i \leq N, N \in \mathbb{N}} \mathbb{E}|X_0^{i,N}|^m < \infty$  and

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

A sufficient condition for  $(\tilde{\mathbf{A}})$  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 a similar assumption in [20, Remark 1.2].

We adopt a definition of mild solutions to (1.1) in  $\mathcal{C}([0, T]; L^1 \cap L^\infty(\mathbb{R}^d))$  which is analogous to Definition 1.1. We get the following result, which is an immediate adaptation of Proposition 1.2. The proof is omitted.

**Proposition 4.1.** *Assume that the kernel  $K : \mathbb{R}^d \rightarrow \mathbb{R}^d$  satisfies  $(\tilde{\mathbf{A}}_i^K)$  and  $(\tilde{\mathbf{A}}_{ii}^K)$  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  $\mathcal{C}([0, T]; L^1 \cap L^\infty(\mathbb{R}^d))$ . In addition, this mild solution is unique.*

The above proposition and the theorem below rely in particular on the Lemma 2.1 applied with  $r = \infty$ .

**Theorem 4.2.** *Assume that the initial conditions  $\{\mu_0^N\}_{N \in \mathbb{N}}$  satisfy  $(\tilde{\mathbf{A}})$  and that the kernel  $K$  satisfies  $(\tilde{\mathbf{A}}^K)$ . Moreover, let  $(\tilde{\mathbf{A}}_\alpha)$  hold true. Let  $T_{\max}$  be the maximal existence time for (1.1) in the space  $\mathcal{C}([0, T], L^1 \cap L^\infty(\mathbb{R}^d))$  and fix  $T \in (0, T_{\max})$ . In addition, let the dynamics of the particle system be given by (1.12) with  $A$  greater than  $A_T$  (defined in (1.15) with the  $L^1 \cap L^\infty(\mathbb{R}^d)$  norm instead of the  $L^1 \cap L^r(\mathbb{R}^d)$  norm).*

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

$$\| \|u^N - u\|_{T, L^1 \cap L^\infty(\mathbb{R}^d)} \|_{L^m(\Omega)} \leq C \left( \| \|u_0^N - u_0\|_{L^1 \cap L^\infty(\mathbb{R}^d)} \|_{L^m(\Omega)} + N^{-\varrho + \varepsilon} \right),$$

where

$$\tilde{\varrho} = \min \left( \alpha \zeta, \frac{1}{2} - \alpha d \right).$$

**Elements of proof of Theorem 4.2.** We do not detail the proof of this theorem, as it follows the same lines as the proof of Theorem 1.3. We only mention how the key estimates evolve.

First, the following embeddings justify that we assume  $\beta > \frac{d}{\tilde{r}}$  in  $(\tilde{\mathbf{A}}_{iii}^K)$ . Since  $\beta - \frac{d}{\tilde{r}} > 0$ ,  $H_{\tilde{r}}^\beta(\mathbb{R}^d)$  is continuously embedded into  $\mathcal{C}^{\beta - \frac{d}{\tilde{r}}}(\mathbb{R}^d)$  (see [56, p.203]). In particular  $H_{\tilde{r}}^\beta$  is continuously embedded into  $L^{\tilde{r}} \cap L^\infty$ . That is, there exists  $C > 0$  such that

$$\|f\|_{L^{\tilde{r}} \cap L^\infty(\mathbb{R}^d)} \leq C \|f\|_{\beta, \tilde{r}} \quad \text{and} \quad \|f\|_{\mathcal{C}^{\beta - \frac{d}{\tilde{r}}}(\mathbb{R}^d)} \leq C \|f\|_{\beta, \tilde{r}}, \quad \forall f \in H_{\tilde{r}}^\beta(\mathbb{R}^d). \quad (4.1)$$

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

$$\|f\|_{L^1 \cap L^\infty(\mathbb{R}^d)} \leq C_{d, \beta, \tilde{r}} \|f\|_{L^1 \cap H_{\tilde{r}}^\beta(\mathbb{R}^d)}, \quad \forall f \in L^1 \cap H_{\tilde{r}}^\beta(\mathbb{R}^d). \quad (4.2)$$

The Assumption  $(\tilde{\mathbf{A}}_\alpha)$ , which strengthens  $(\mathbf{A}_\alpha)$ , is also partially understood in view of these embeddings. This assumption yields the boundedness result for  $u^N$  given in Proposition A.12 (it generalises to  $\beta > 0$  the result given in (2.4)) and reads, for any  $q \geq 1$ ,

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

It is then possible to control  $\mathbb{E} \|K * u_s^N\|_{L^\infty(\mathbb{R}^d)}^m$  by  $\mathbb{E} \|u_s^N\|_{L^1 \cap L^\infty(\mathbb{R}^d)}^m$  thanks to Lemma 2.1 (applied with  $\mathbf{r} = \infty$ ), then using the embedding (4.2) and finally the previous bound ensures that

$$\sup_{N \in \mathbb{N}^*} \sup_{s \in [0, T]} \mathbb{E} \|K * u_s^N\|_{L^\infty(\mathbb{R}^d)}^m < \infty.$$

A similar control is obtained on  $\sup_{N \in \mathbb{N}^*} \sup_{s \in [0, T]} \|K * u_s\|_{L^\infty(\mathbb{R}^d)}$  using Proposition A.7 instead of Proposition A.12.

Therefore, we obtain a bound analogous to (2.10):

$$\begin{aligned} & \| \|u^N - u\|_{t, L^1 \cap L^\infty(\mathbb{R}^d)} \|_{L^m(\Omega)} \leq \| \|u_0^N - u_0\|_{L^1 \cap L^\infty(\mathbb{R}^d)} \|_{L^m(\Omega)} \\ & + C \int_0^t \frac{1}{(t-s)^{\frac{1}{2}}} \| \|u^N - u\|_{s, L^1 \cap L^\infty(\mathbb{R}^d)} \|_{L^m(\Omega)} ds + \| \|E\|_{t, L^1 \cap L^\infty(\mathbb{R}^d)} \|_{L^m(\Omega)} \\ & + \| \|M^N\|_{t, L^1 \cap L^\infty(\mathbb{R}^d)} \|_{L^m(\Omega)}, \end{aligned}$$

and to (2.11) for the new error term:

$$\| \|E\|_{t, L^1 \cap L^\infty(\mathbb{R}^d)} \|_{L^m(\Omega)} \leq \frac{C}{N \alpha \zeta}.$$

It remains to bound the stochastic integral term in  $L^1 \cap L^\infty(\mathbb{R}^d)$  norm, which is given immediately by Proposition A.8: for any  $\varepsilon > 0$ , there exists  $C > 0$  such that

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

Using the three previous inequalities, Grönwall's lemma gives the desired result.  $\square$

**Remark 4.3.** Assuming that the kernel  $K$  is such that  $\zeta$  can be chosen equal or close to 1, the rate is really determined by the minimum between  $\alpha$  and  $\frac{1}{2} - \alpha d$ , under the constraint that  $\alpha < (d + 2\beta + 2d(\frac{1}{2} - \frac{1}{\tilde{r}}))^{-1}$  (see  $(\tilde{\mathbf{A}}_\alpha)$ ) and  $\beta < 1$  (see  $(\tilde{\mathbf{A}}_{iii}^K)$ ). We will see in Section 5.2 that the best rate for a Riesz kernel of parameter  $s \in (d - 2, d - 1)$  is

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

which is obtained for  $\alpha = \frac{1}{2(d+1)}$  and initial conditions satisfying  $(\tilde{\mathbf{A}})$  with any  $\beta \in (\frac{d}{\tilde{r}}, 1)$  and  $\tilde{r} > \frac{d}{\beta - 2 + d - s}$ .

In Section 4.1, using an interpolation inequality between the results of Proposition A.12 and Theorem 1.3, we obtain the following rate of convergence with respect to Sobolev norms:

**Corollary 4.4.** Let the assumptions of Theorem 4.2 hold and assume further that  $u_0 \in H_{\tilde{r}}^\beta(\mathbb{R}^d)$ . Then, for any  $\varepsilon > 0$  and any  $m \geq 1$ , there exists a constant  $C > 0$  such that for all  $N \in \mathbb{N}^*$ ,

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

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

**Remark 4.5.** It is clear that  $\gamma < \beta$ . It will also be important (in particular for the propagation of chaos in the Section 4.3) to ensure that  $\gamma > \frac{d}{\tilde{r} - \delta}$  so as to have an embedding in a space of Hölder continuous functions. This is indeed the case if  $\delta$  is chosen small enough, see condition (4.6).

In addition, we get by the classical embedding recalled in (4.1) that the result of Corollary 4.4 imply the same rates in  $\eta$ -Hölder norm, with  $\eta = \gamma - \frac{d}{\tilde{r} - \delta}$ , provided that this quantity is positive (see condition (4.6)).

Observing that a rate of convergence was obtained in  $H_{\tilde{r} - \delta}^\gamma$ , one could wonder if the convergence also happens in  $H_{\tilde{r}}^\beta$ . Corollary 4.6 answers positively, thus extending the main convergence results in [20, 43] to general kernels.

**Corollary 4.6.** Let the same assumptions as in Corollary 4.4 hold, with  $\gamma, \delta$  as in Corollary 4.4. Let  $m \geq 1$  and assume further that

$$\left\| \|u_0^N - u_0\|_{L^1 \cap L^\infty(\mathbb{R}^d)} \right\|_{L^m(\Omega)} \longrightarrow 0. \quad (4.3)$$

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

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

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

The proof is given in Section 4.2. Theorem 4.2 and Corollary 4.6 cover the convergence results for the mollified empirical measures obtained in a case by case basis in [20, 43]. 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 those presented in [16, 17, 18].

We also notice without giving the full statement, that as in Corollary 1.4, a rate of convergence can be obtained for the particle system without cutoff.

Finally, it remains to give the results about the propagation of chaos in the context of the hypothesis  $(\tilde{\mathbf{A}}^K)$ .

Firstly, one must redefine the martingale problem  $(\mathcal{MP})$ . Let us denote by  $(\widetilde{\mathcal{MP}})$  a martingale problem analogous to  $(\mathcal{MP})$  except that in point (ii) we replace  $L^1 \cap L^r(\mathbb{R}^d)$  by  $L^1 \cap L^\infty(\mathbb{R}^d)$ .

Then, adapt the arguments of Proof of Proposition 1.8 do this functional setting by using the well-posedness of the mild solution given by Proposition 4.1 and Lemma 2.1 (with  $r = \infty$  as mentioned above). This yields the following result:

**Proposition 4.7.** *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  $(\tilde{\mathbf{A}}^K)$ . Then, the martingale problem  $(\widetilde{\mathcal{MP}})$  admits a unique solution.*

Now, we are in the position to state the propagation of chaos result.

**Theorem 4.8.** *Let the hypotheses of Theorem 4.2 hold. Assume further that the family of random variables  $\{X_0^i, i \in \mathbb{N}\}$  is identically distributed, that  $u_0 \in H_{\tilde{r}}^\beta(\mathbb{R}^d)$  and that  $\langle u_0^N, \varphi \rangle \rightarrow \langle u_0, \varphi \rangle$  in probability, for any  $\varphi \in C_b(\mathbb{R}^d)$ . Then, the empirical measure  $\mu^N$  converges in probability towards  $\mathbb{Q}$ , which is the law of the unique weak solution of (1.4).*

The scheme of proof is the same as for Theorem 1.9, but there are some additional technical difficulties that need to be adressed. Hence we will give the sketch of the proof in Section 4.3.

#### 4.1 Rate in Sobolev norm: Proof of Corollary 4.4

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

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

$$\gamma = \beta \frac{\tilde{r}(\tilde{r} - \delta - 1)}{(\tilde{r} - \delta)(\tilde{r} - 1)}. \quad (4.4)$$

The interpolation theorem for Bessel potential spaces, see [56, p.185], gives that for any  $f \in H_1^0 \cap H_{\tilde{r}}^\beta(\mathbb{R}^d) (\equiv L^1 \cap H_{\tilde{r}}^\beta(\mathbb{R}^d))$ ,

$$\|f\|_{\gamma, \tilde{r} - \delta} \leq \|f\|_{0,1}^\theta \|f\|_{\beta, \tilde{r}}^{1-\theta}, \quad (4.5)$$

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

Hence it follows from (4.5) that for any  $m \geq 1$ ,

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

and we deduce from Hölder's inequality that

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



In view of the previous inequality and using Theorem 4.2, Proposition A.7 and Proposition A.12, we deduce the rate of convergence in  $L^m(\Omega; L^\infty([0, T], H_{\tilde{r}-\delta}^\gamma(\mathbb{R}^d)))$ .

Finally, note that it is always true that  $\gamma < \beta$ . Besides, it will be important to ensure that  $\gamma > \frac{d}{\tilde{r}-\delta}$  to have an embedding in the space of Hölder continuous functions (see (4.1)). For this, it suffices to choose  $\delta$  which satisfies:

$$\frac{\tilde{r} - \delta - 1}{\tilde{r} - 1} > \frac{d/\tilde{r}}{\beta}. \quad (4.6)$$

## 4.2 Convergence in $H_{\tilde{r}}^\beta$ : Proof of Corollary 4.6

Let us introduce the space

$$\mathcal{X} = L^2([0, T]; H_{\tilde{r}}^\beta(\mathbb{R}^d))$$

with the strong topology, and denote by  $\mathcal{X}_w$  the same space endowed with the weak topology. Notice first that  $L^2([0, T]; H_{\tilde{r}}^\beta(\mathbb{R}^d))$  is a reflexive Banach space for  $1 < \tilde{r} < \infty$  (see [56, p.198-199]). Hence by the Banach-Alaoglu theorem, it is compactly embedded in  $\mathcal{X}_w$ .

Now the Chebyshev inequality ensures that

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

Thus by Proposition A.12, we obtain

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

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. [48, 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  $F \in \mathcal{C}_b^0(\mathbb{R})$  and  $\varphi \in L^2([0, T]; H_{\tilde{r}}^{-\beta}(\mathbb{R}^d))$ , we get that

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

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

On the other hand, from Theorem 4.2 with assumption (4.3), we know that  $u^N$  converges to  $u$  in  $L^m(\Omega; L^\infty([0, T], L^1 \cap L^\infty(\mathbb{R}^d)))$ . Hence  $u^N$  converges almost surely to  $u$  in  $L^\infty([0, T], L^1 \cap L^\infty(\mathbb{R}^d))$ , up to a subsequence. Without loss of generality, we still denote by  $u^N$  this subsequence. By testing against any function  $\varphi \in \mathcal{C}^\infty([0, T] \times \mathbb{R}^d)$ , we deduce that  $\mathbb{E}F(\langle u^N, \varphi \rangle) \rightarrow F(\langle u, \varphi \rangle)$  for any  $F \in \mathcal{C}_b^0(\mathbb{R})$ . Thus by uniqueness of the limit, we get that  $\mathbb{E}F(\langle u^\infty, \varphi \rangle) = F(\langle u, \varphi \rangle)$ . Since  $\mathcal{C}^\infty([0, T] \times \mathbb{R}^d)$  is dense in  $L^2([0, T]; H_{\tilde{r}}^{-\beta}(\mathbb{R}^d))$ , 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 4.6.

## 4.3 Sketch of the proof of Theorem 4.8

We will show that  $\mu^N$  converges to the unique solution  $\mathbb{Q}$  of the martingale problem  $(\widetilde{\mathcal{M}\mathcal{P}})$ . The proof closely follows the proof of Theorem 1.9 (see Section 3.2) in a new functional framework. Namely, we redefine the space  $\mathcal{Y}$  (see (3.2)) in the following way:

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

Recall from Section 4.2 that  $\mathcal{X}_w$  is the space  $L^2([0, T]; H_{\tilde{\mathbf{r}}}^\beta(\mathbb{R}^d))$  endowed with the weak topology. This choice of a new functional framework will be explained at the end of this section.

Adopting the notation of Section 3.2, the sequence  $(\tilde{\Pi}^N, N \geq 1)$  is tight. Indeed, the convergence of Theorem 4.2 gives the tightness in  $\mathcal{C}([0, T]; L^1(\mathbb{R}^d))$  and the discussion in Section 4.2 gives the tightness in  $\mathcal{X}_w$ . Besides, despite the fact that  $\mathcal{X}_w$  is not metrisable, we explained in Section 4.2 why it is still possible to extract a converging subsequence out of  $(\tilde{\Pi}^N, N \geq 1)$ .

Now repeat the arguments that follow line by line to obtain that an analogue of Lemma 3.1 holds. It remains to obtain the analogue of Proposition 3.2. Step 1 remains unchanged. In Step 2, when dealing with the term  $I_n$  one must do the following: In view of Lemma 2.1 (with  $\mathbf{r} = \infty$  now), one has

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

Now recall  $\tilde{\mathbf{r}}$  and  $\beta$  are fixed in  $(\tilde{\mathbf{A}}_{iii}^K)$ , and let  $\gamma$  and  $\delta$  satisfy (4.4) and (4.6), so that  $\frac{d}{\tilde{\mathbf{r}}-\delta} < \gamma < \beta$ . Then, use the Sobolev embedding  $L^1 \cap H_{\tilde{\mathbf{r}}-\varepsilon}^\gamma(\mathbb{R}^d) \subset L^1 \cap L^\infty(\mathbb{R}^d)$  to get

$$\|K * (\mathbf{u}_s^n - \mathbf{u}_s)\|_{L^\infty(\mathbb{R}^d)} \leq C_{K,d} \|\mathbf{u}_s^n - \mathbf{u}_s\|_{L^1 \cap H_{\tilde{\mathbf{r}}-\varepsilon}^\gamma(\mathbb{R}^d)}.$$

Plug the latter in (3.3) to obtain

$$I_n \leq C_{K,d} \int_s^t \|\mathbf{u}_\sigma^n - \mathbf{u}_\sigma\|_{L^1 \cap H_{\tilde{\mathbf{r}}-\varepsilon}^\gamma(\mathbb{R}^d)} d\sigma.$$

By the interpolation inequality (4.5),

$$\|\mathbf{u}_s^n - \mathbf{u}_s\|_{L^1 \cap H_{\tilde{\mathbf{r}}-\varepsilon}^\gamma(\mathbb{R}^d)} \leq \|\mathbf{u}_s^n - \mathbf{u}_s\|_{L^1(\mathbb{R}^d)} + \|\mathbf{u}_s^n - \mathbf{u}_s\|_{L^1(\mathbb{R}^d)}^\theta \|\mathbf{u}_s^n - \mathbf{u}_s\|_{H_{\tilde{\mathbf{r}}}^\beta(\mathbb{R}^d)}^{1-\theta},$$

for  $\theta$  as in Section 4.1. Now since  $\mathbf{u}^n$  converges in  $\mathcal{Y}$ , and converges in particular weakly in  $L^2([0, T], H_{\tilde{\mathbf{r}}}^\beta(\mathbb{R}^d))$ , the uniform boundedness principle tells us that it is bounded in this space. Gathering this fact with the convergence in  $\mathcal{C}([0, T]; L^1(\mathbb{R}^d))$  (by assumption), the previous inequality yields the convergence of  $\mathbf{u}^n$  in  $L^2([0, T], L^1 \cap H_{\tilde{\mathbf{r}}-\varepsilon}^\gamma(\mathbb{R}^d))$ . Hence  $I_n$  converges to 0.

For  $II_n$  the arguments remains the same using the hypothesis  $(\tilde{\mathbf{A}}_{iii}^K)$  instead of  $(\mathbf{A}_{iii}^K)$  and having in mind that as  $\mathbf{u}^n$  converges in  $\mathcal{Y}$ ,  $\mathbf{u}_t^n$  is uniformly bounded w.r.t.  $n$  in  $L^1 \cap L^\infty(\mathbb{R}^d)$ . Then, the proof is finished.

Here is a good place to explain why we could not define  $\mathcal{Y}$  to be just  $\mathcal{C}([0, T]; L^1 \cap L^\infty(\mathbb{R}^d))$ . Indeed, the computations for  $I_n$  would have been straightforward in that case (as in the original proof). However, this space is incompatible with the assumption  $(\tilde{\mathbf{A}}_{iii}^K)$ . Hence, when dealing with  $II_n$ , nothing would guarantee that  $K * \mathbf{u}_t^n$  is continuous if  $\mathbf{u}$  was in  $\mathcal{C}([0, T]; L^1 \cap L^\infty(\mathbb{R}^d))$ . To ensure the latter, we must be in the position to apply  $(\tilde{\mathbf{A}}_{iii}^K)$ . That is  $\mathbf{u}_t^n$  must be in  $L^1 \cap H_{\tilde{\mathbf{r}}}^\beta(\mathbb{R}^d)$ . The reason we took the space  $\mathcal{X}_w$  and not directly a space of type  $L^2([0, T]; H_{\tilde{\mathbf{r}}}^\beta(\mathbb{R}^d))$  is that, by the proof of Corollary 4.6, we only have the convergence of  $u^N$  in  $\mathcal{X}_w$  and not in the strong topology of  $L^2([0, T]; H_{\tilde{\mathbf{r}}}^\beta(\mathbb{R}^d))$ . Hence, we would not have the tightness of  $(\tilde{\Pi}^N, N \geq 1)$  in such space.

## 5 Examples

In this section we first focus on the set of assumptions  $(\mathbf{A}^K)$  and we present a sufficient condition for them to be satisfied that will be used in practice. Then, we turn to specific singular kernels of Riesz type and discuss how our results can be applied to them.

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

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. We give with the following lemma an easier-to-check condition that in practice replaces  $(\mathbf{A}_{iii}^K)$ .

**Lemma 5.1.** *Assume that  $K$  satisfies  $(\mathbf{A}_i^K)$  and  $(\mathbf{A}_{ii}^K)$ . Assume further that*

$(\widetilde{\mathbf{A}}_{iii}^K)$  *There exists  $\mathbf{r} \geq \max(\mathbf{p}', \mathbf{q}')$  and  $z \in [\mathbf{p} \vee \mathbf{q}, +\infty] \cap (d, +\infty]$  such that the matrix-valued kernel  $\nabla K$  defines a convolution operator which is bounded component-wise from  $L^1 \cap L^{\mathbf{r}}(\mathbb{R}^d)$  to  $L^z(\mathbb{R}^d)$ .*

*Then,  $K$  satisfies  $(\mathbf{A}_{iii}^K)$  with the same parameter  $\mathbf{r}$  and with  $\zeta = 1 - \frac{d}{z}$ .*

*Proof.* First, we will make use of  $(\mathbf{A}_i^K)$  and  $(\mathbf{A}_{ii}^K)$ . Young's convolution inequality states that for  $\mathbf{p}'_z := (1 + \frac{1}{z} - \frac{1}{\mathbf{p}})^{-1}$  and  $\mathbf{q}'_z := (1 + \frac{1}{z} - \frac{1}{\mathbf{q}})^{-1}$ , for any  $f \in L^{\mathbf{p}'_z} \cap L^{\mathbf{q}'_z}(\mathbb{R}^d)$  we have

$$\begin{aligned} \|K * f\|_{L^z(\mathbb{R}^d)} &\leq \|(\mathbb{1}_{\mathcal{B}_1} K) * f\|_{L^z(\mathbb{R}^d)} + \|(\mathbb{1}_{\mathcal{B}_1^c} K) * f\|_{L^z(\mathbb{R}^d)} \\ &\leq \|\mathbb{1}_{\mathcal{B}_1} K\|_{L^{\mathbf{p}}(\mathbb{R}^d)} \|f\|_{L^{\mathbf{p}'_z}(\mathbb{R}^d)} + \|\mathbb{1}_{\mathcal{B}_1^c} K\|_{L^{\mathbf{q}}(\mathbb{R}^d)} \|f\|_{L^{\mathbf{q}'_z}(\mathbb{R}^d)} \\ &\leq C_K \|f\|_{L^{\mathbf{p}'_z} \cap L^{\mathbf{q}'_z}(\mathbb{R}^d)}. \end{aligned}$$

In particular the previous inequality holds true if  $f \in L^1 \cap L^{\mathbf{r}}(\mathbb{R}^d)$ , because  $\mathbf{p}'_z \leq \mathbf{p}' \leq \mathbf{r}$  (and similarly  $\mathbf{q}'_z \leq \mathbf{r}$ ) and then by interpolation,  $f$  is in  $L^{\mathbf{p}'_z} \cap L^{\mathbf{q}'_z}(\mathbb{R}^d)$ . Now in view of the previous fact and using the property  $(\widetilde{\mathbf{A}}_{iii}^K)$  of  $\nabla K$ , one deduces that if  $f \in L^1 \cap L^{\mathbf{r}}(\mathbb{R}^d)$ , then  $K * f \in H_z^1(\mathbb{R}^d)$ . Hence it follows from Morrey's inequality [7, Th. 9.12] that there exists  $C_{z,d} > 0$  such that for any  $f \in L^1 \cap L^{\mathbf{r}}(\mathbb{R}^d)$ ,

$$\begin{aligned} \mathcal{N}_\eta(K * f) &\leq C \|K * f\|_{H_z^1(\mathbb{R}^d)} \\ &\leq C_{z,d,K} \|f\|_{L^1 \cap L^{\mathbf{r}}(\mathbb{R}^d)}, \end{aligned}$$

where  $\eta = 1 - \frac{d}{z}$ . Hence the desired result.  $\square$

A simple example that satisfies  $(\widetilde{\mathbf{A}}_{iii}^K)$  is a kernel  $K$  such that  $\nabla K$  is integrable. Then  $\nabla K$  defines a convolution operator and by a convolution inequality, this operator is bounded in any  $L^z(\mathbb{R}^d)$ ,  $z \in [1, +\infty]$ . Hence, by interpolation inequality, if  $\mathbf{r} \geq z$ , one has

$$\|\nabla K * f\|_{L^z(\mathbb{R}^d)} \leq C \|f\|_{L^z(\mathbb{R}^d)} \leq C \|f\|_{L^1 \cap L^{\mathbf{r}}(\mathbb{R}^d)}.$$

As a consequence,  $\nabla K$  satisfies  $(\widetilde{\mathbf{A}}_{iii}^K)$  for any  $z \in [\mathbf{p} \vee \mathbf{q}, +\infty] \cap (d, +\infty]$  and any  $\mathbf{r} \geq \max(\mathbf{p}', \mathbf{q}', z)$ . Hence it satisfies  $(\mathbf{A}_{iii}^K)$  with  $\zeta = 1 - \frac{d}{z}$  and  $\mathbf{r} = \max(\mathbf{p}', \mathbf{q}', z)$ .

Nevertheless, for kernels  $K$  such that  $\nabla K$  is not integrable, it may be possible to define the convolution operator of kernel  $\nabla K$  as the Principal Value integral acting on the space of smooth, rapidly decaying functions (i.e. the Schwartz space), thus defining a tempered distribution. In the next section we will see some interesting examples of such kernels.

## 5.2 Riesz potentials

In their general form, the Riesz potentials were defined in (1.2). We denote the associated kernel by  $K_s := \pm \nabla V_s$ .  $K_s$  satisfies Assumption  $(\mathbf{A}^K)$ , provided that  $d \geq 2$  and  $s \in [0, d-1)$ : Indeed, in the order of the hypotheses in  $(\mathbf{A}^K)$ , we have that

- $K_s \in L^1(\mathcal{B}_1)$  if and only if  $0 \leq s < d-1$ . Moreover,  $K_s \in L^p(\mathcal{B}_1)$  for any  $p < \frac{d}{s+1}$ . Then one can choose  $\mathbf{p} = \left(\frac{d}{s+1}\right)^-$ .
- $K_s \in L^q(\mathcal{B}_1^c)$  for any  $q > \frac{d}{s+1}$ , so one can choose  $\mathbf{q} = \left(\frac{d}{s+1}\right)^+$ .

- – If  $d \geq 3$  and  $s < d - 2$ , then  $\nabla K_s$  is not bounded in any  $L^p$  but it is bounded from  $L^{\tilde{z}}$  to  $L^z$  whenever  $\tilde{z} \in (1, \frac{d}{d-(s+2)})$  and  $\frac{1}{z} = \frac{1}{\tilde{z}} + \frac{s+2}{d} - 1$  (see [49, Theorem 25.2]). By letting  $\tilde{z}$  be close enough to  $\frac{d}{d(s+2)}$ , one can always choose  $z \geq \mathbf{p} \vee \mathbf{q}$  as large as desired, in particular one can find  $z \geq \mathbf{p} \vee \mathbf{q}$ ,  $z > d$  and  $\mathbf{r} \geq \max(\tilde{z}, \mathbf{p}', \mathbf{q}')$  to ensure property  $(\tilde{\mathbf{A}}_{iii}^K)$  is verified. Then  $(\mathbf{A}_{iii}^K)$  holds for  $\zeta = 1 - \frac{d}{z}$  (see Section 5.1). Hence all our results can be applied to this kernel.
- If  $s = d - 2$ , then  $\nabla K_s$  is a typical kernel satisfying the conditions of [28, Chapter 4.4], and therefore it defines a bounded operator in any  $L^z(\mathbb{R}^d)$ ,  $z \in (1, \infty)$ . Hence all our results apply to this particular kernel (see the previous discussion in Section 5.1). The choice of parameters will then be  $\mathbf{p} = \left(\frac{d}{d-1}\right)^-$ ,  $\mathbf{q} = \left(\frac{d}{d-1}\right)^+$ ,  $\mathbf{r} = z$  and  $\zeta = 1 - \frac{d}{z}$ , for some  $z$  to be chosen in  $(d, +\infty)$ .
- If  $s \in [d - 2, d - 1)$ , then one can verify (see e.g. [15, Lemma 2.5]) that  $\nabla K_s$  defines a convolution operator from  $L^1 \cap L^\infty \cap \mathcal{C}^\sigma(\mathbb{R}^d)$  to  $L^\infty(\mathbb{R}^d)$ , with  $\sigma \in (2 - d + s, 1)$ :

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

for some  $\beta$  and  $\tilde{\mathbf{r}}$  such that  $\sigma = \beta - \frac{d}{\tilde{\mathbf{r}}}$ , using the embedding (4.1) in the second inequality above, and (4.2) in the third. Hence it follows that  $\mathcal{N}_1(K * f) \leq \|\nabla K_s * f\|_{L^\infty(\mathbb{R}^d)} \lesssim \|f\|_{L^1 \cap H_{\tilde{\mathbf{r}}}^\beta(\mathbb{R}^d)}$ , i.e. Assumption  $(\tilde{\mathbf{A}}^K)$  is satisfied for  $\zeta = 1$ , for any  $\tilde{\mathbf{r}} > d$  and any  $\beta \in (\frac{d}{\tilde{\mathbf{r}}}, 1)$  such that  $\beta - \frac{d}{\tilde{\mathbf{r}}} \in (2 - d + s, 1)$ .

Hence, Theorem 1.3 and its Corollaries 2.2 and 1.4 are applicable for  $s \leq d - 2$  and Theorem 4.2 and its corollaries are applicable for  $s \in (d - 2, d - 1)$ . In particular,

- For  $s = d - 2$  we have the following:
  - Convergence holds for the widest possible range of parameter  $\alpha$  by choosing  $\mathbf{r}$  as small as possible. This means here choosing  $z = \mathbf{r} = d^+$ . Under the constraint  $(\mathbf{A}_\alpha)$ , one can therefore obtain the convergence for any  $\alpha < \frac{1}{2(d-1)}$ ;
  - On the other hand, if one wishes to maximize the rate of convergence, then one must choose  $z$  very large and  $\mathbf{r} = +\infty$ , then the rate will be in  $L^1 \cap L^\infty$  norm and is  $\varrho = \left(\frac{1}{2(d+1)}\right)^-$  for  $\alpha = \left(\frac{1}{2(d+1)}\right)^+$  and  $z$  very large. More precisely, the rate is given by

$$\varrho = \min \left( \left(1 - \frac{d}{z}\right)\alpha, \frac{1}{2} - \alpha d \right),$$

which gives  $\varrho = \frac{1-d/z}{2(d+1-\frac{d}{z})}$  for  $\alpha = \frac{1}{2(d+1-\frac{d}{z})}$ , whose supremum is  $\frac{1}{2(d+1)}$  for  $z \rightarrow +\infty$ .

- For  $s \in (d - 2, d - 1)$  we have the following:
  - Convergence holds for the widest possible range of parameter  $\alpha$  by maximizing the constraint  $(d + 2\beta + 2d(\frac{1}{2} - \frac{1}{\tilde{\mathbf{r}}}))^{-1} = (2(d + \sigma))^{-1}$  from the Assumption  $(\tilde{\mathbf{A}}_\alpha)$ . Hence, letting  $\sigma = (2 - d + s)^+$ , one can therefore obtain the convergence for any  $\alpha < \frac{1}{2(s+2)}$ ;
  - On the other hand, if one wishes to maximize the rate of convergence, it is clear that one must choose  $\alpha$  such that  $\alpha\zeta = \frac{1}{2} - \alpha d$  (see the definition of  $\tilde{\varrho}$ ). Since  $\zeta = 1$ , this gives  $\tilde{\varrho} \leq \frac{1}{2(d+1)}$  and this maximum is attained for  $\alpha = \frac{1}{2(d+1)}$ .

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 these values, 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.

### 5.3 Classical kernels

We present here the following important kernels that enter in the above framework for  $s = d - 2$ : Coulomb, Keller-Segel and Biot-Savart kernels. In particular, for these kernels we get the convergence for  $\alpha < \frac{1}{2(d-1)}$  and the best possible rate is  $\varrho = (\frac{1}{2(d+1)})^-$  for  $\alpha = (\frac{1}{2(d+1)})^+$ . At the end of this section we present a fourth example of an attractive-repulsive kernel.

**Coulomb kernel.** The Coulomb interaction kernel is given by

$$K_C := -\nabla V_{d-2},$$

where  $V_{d-2}$  is defined in (1.2). It is a generalisation in any dimension of the classical Coulomb force and an example of repulsive kernel, in the sense that

$$x \cdot K_C(x) \geq 0, \quad \text{on the domain of definition of } K_C. \quad (5.1)$$

It models for instance the interaction between particles with identical electrical charges.

**Parabolic-elliptic Keller-Segel models.** An important and tricky example covered by this paper is the parabolic-elliptic Keller-Segel PDE, which takes the form (1.1) with the kernel defined, for some  $\chi > 0$ , by

$$K_{KS}(x) = -\chi \frac{x}{|x|^d}. \quad (5.2)$$

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

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

This leads to important issues that we discuss in more details in [43]. 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 possible to choose a value of the cut-off  $A_T$  independently of  $T$  (see [43]).

Theorem 1.3, Corollaries 2.2 and 1.4 give rates of convergence of the particle system to the Keller-Segel PDE, even for solutions that exhibit a blow-up (with the parameters discussed above for  $s = d - 2$ ).

As a consequence of Theorem 1.8, we also deduce the local-in-time weak well-posedness of the McKean-Vlasov SDE (1.4) for all values of the concentration parameter  $\chi$ , and global-in-time weak well-posedness for  $\chi < 8\pi$ , which is a novelty (see [43] for a thorough discussion and comparison with previous results). We obtain the propagation of chaos result for our particle system to this SDE.

**Biot-Savart kernel and the 2d Navier-Stokes equation.** By considering the vorticity field  $\xi$  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 of (5.1). In this case, the Biot-Savart kernel is merely repulsive since  $x \cdot K(x) = 0$ .

It is well-known that with such kernel, Eq. (1.1) has a unique global solution, and that  $\|K * \xi_t\|_{L^\infty(\mathbb{R}^2)}$  can be bounded by  $C(1 + \|\xi_0\|_{L^\infty(\mathbb{R}^2)})$  (see [20] and references therein). This permits to choose the cut-off value  $A_T$  from (1.15) independently of  $T$ .

The kernel  $K_{BS}$  is covered by our assumption ( $\mathbf{A}^K$ ), for the same reason as the Coulomb kernel with  $d = 2$  ( $K_0(x) = \frac{1}{|x|}$  and we can choose  $\alpha < \frac{1}{2}$ ). In that case, we improve within Theorem 1.3

and Corollary 4.6 the Theorem 1.3 of Flandoli et al. [20], by allowing  $\alpha \in [\frac{1}{4}, \frac{1}{2})$  and by providing a rate of convergence.

All the other results of this paper apply. In particular, if the initial condition is regular enough, we obtain a rate  $\varrho$  in  $L^1 \cap L^\infty$  norm (the rate is maximized with  $\mathbf{r} = +\infty$  when  $s = d - 2$ ), and this rate is almost  $\frac{1}{6}$ .

**Attractive-repulsive kernels.** There is at least one other very interesting class of kernels that enters our framework. The attractive-repulsive kernels are attractive in a region of space, i.e. they satisfy (5.3) on a subdomain  $D$  of the domain of definition of  $K$ , and repulsive (i.e. satisfying (5.1)) on the complement of  $D$ .

The most famous example of such attractive-repulsive kernels might be the Lennard-Jones potential in molecular dynamics: this isotropic potential reads

$$V(x) = V_0 (|x|^{-12} - |x|^{-6}),$$

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

$$V(x) = V_a |x|^{-a} - V_b |x|^{-b},$$

with  $a, b > 0$  and  $V_a, V_b > 0$ . One can now refer to the discussion on Riesz kernels in Section 5.2 to determine the values of  $a$  and  $b$  that ensure the applicability of our results.

## Appendix

### A.1 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 A.1.1, we derive some general inequalities and prove Proposition 1.2. In Subsection A.1.2, we prove that the PDE with cut-off (A.6) can have at most one mild solution.

#### A.1.1 Properties of mild solutions of the PDE

We recall that whenever  $(\mathbf{A}_i^K)$  and  $(\mathbf{A}_{ii}^K)$  are assumed to hold, we suppose  $\mathbf{r} \geq \max(\mathbf{p}', \mathbf{q}')$ . For each  $T > 0$ , let us consider the space

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

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

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

**Lemma A.2.** *The bilinear mapping  $B$  is continuous from  $\mathcal{X} \times \mathcal{X}$  to  $\mathcal{X}$ .*

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

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

which suffices to prove the continuity of  $B$ .

First, using the property (2.1) of the Gaussian kernel with  $p = r$ , observe that for any  $t \in [0, T]$ ,

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

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

$$\begin{aligned} \|B(u, v)(t)\|_{L^r(\mathbb{R}^d)} &\leq \int_0^t \frac{C}{\sqrt{t-s}} \|K * v_s\|_{L^\infty(\mathbb{R}^d)} \|u_s\|_{L^r(\mathbb{R}^d)} ds \\ &\leq C\sqrt{t} \|u\|_{\mathcal{X}} \|v\|_{\mathcal{X}}. \end{aligned} \quad (\text{A.2})$$

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

$$\begin{aligned} \|B(u, v)(t)\|_{L^1(\mathbb{R}^d)} &\leq \int_0^t \frac{C}{\sqrt{t-s}} \|u_s K * v_s\|_{L^1(\mathbb{R}^d)} ds \\ &\leq \int_0^t \frac{C}{\sqrt{t-s}} \|u_s\|_{L^1(\mathbb{R}^d)} \|K * v_s\|_{L^\infty(\mathbb{R}^d)} ds \\ &\leq C\sqrt{t} \|u\|_{\mathcal{X}} \|v\|_{\mathcal{X}}. \end{aligned} \quad (\text{A.3})$$

Therefore, combining Equations (A.2), (A.3) one obtains (A.1).  $\square$

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 A.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 (A.6), for any value of the cut-off  $A$  (see Proposition A.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  $i = 1, 2$ ,

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

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

**Corollary A.3.** *Assume that  $(\mathbf{A}_i^K)$  and  $(\mathbf{A}_{ii}^K)$  hold. Let  $C$  be the constant that appears in (A.1). Then for  $T > 0$  such that*

$$4C\sqrt{T} \|u_0\|_{L^1 \cap L^r(\mathbb{R}^d)} < 1, \quad (\text{A.4})$$

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

$$\|u\|_{T, L^1 \cap L^r(\mathbb{R}^d)} \leq \frac{1 - \sqrt{1 - 4C\sqrt{T} \|u_0\|_{L^1 \cap L^r(\mathbb{R}^d)}}}{2C\sqrt{T}}. \quad (\text{A.5})$$

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

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

*Proof.* We rely on the bound (A.1), in order to get that for  $t \leq T$ ,

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

Then by a standard argument (see e.g. Lemma 2.3 in [40]), choosing  $T > 0$  which satisfies (A.4) ensures that (A.5) holds true.  $\square$

**Remark A.4.** Here we recall that if  $u$  is a local mild solution on  $[0, T]$  to (1.1), then  $u$  preserves mass and sign: if  $u_0 \geq 0$ , then  $u_t \geq 0$  for any  $t \in [0, T]$  and  $\int_{\mathbb{R}^d} u_t(x) dx = \int_{\mathbb{R}^d} u_0(x) dx$ . The sign preservation is obtained by an argument similar to [40, Proposition 2.7] (although using  $(\mathbf{A}_{iii}^K)$  instead of the Poisson kernel). Then, the mass preservation is obtained by a standard integration-by-parts argument (see [40, Proposition 2.4] for a similar result in a different functional framework).

### A.1.2 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:

$$\begin{cases} \partial_t \tilde{u}(t, x) = \Delta \tilde{u}(t, x) - \nabla \cdot (\tilde{u}(t, x) F_A(K * \tilde{u}(t, x))), & t > 0, x \in \mathbb{R}^d \\ \tilde{u}(0, x) = u_0(x). \end{cases} \quad (\text{A.6})$$

Note that if  $F_A$  is replaced by the identity function, one recovers (1.1).

**Remark A.5.** Similarly to Definition 1.1, a mild solution to (A.6) satisfies Definition 1.1(i) and solves

$$u_t = e^{t\Delta} u_0 - \int_0^t \nabla \cdot e^{(t-s)\Delta} (u_s F_A(K * u_s)) ds, \quad 0 \leq t \leq T. \quad (\text{A.7})$$

In this section, we consider the cut-off PDE (A.6) and its mild solution defined above. Here,  $F_A$  is given in (1.10), but we denote it simply by  $F$  for the sake of readability.

Note that due to the boundedness of the reaction term in (A.6), any mild solution will always be global. This global solution will be rigorously obtained as the limit of the particle system (1.12).

**Proposition A.6.** Assume that  $K$  satisfies Assumptions  $(\mathbf{A}_i^K)$  and  $(\mathbf{A}_{ii}^K)$  and recall that  $\mathbf{r} \geq \max(\mathbf{p}', \mathbf{q}')$ . Let  $u_0 \in L^1 \cap L^{\mathbf{r}}(\mathbb{R}^d)$ . Then for any  $A > 0$  and  $F$  defined in (1.10), there is at most one mild solution to the cut-off PDE (A.6).

*Proof.* Assume there are two mild solutions  $u^1$  and  $u^2$  to (A.6). Then,

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

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

$$\begin{aligned} \|u_t^1 - u_t^2\|_{L^1(\mathbb{R}^d)} + \|u_t^1 - u_t^2\|_{L^{\mathbf{r}}(\mathbb{R}^d)} &\leq C \int_0^t \frac{1}{\sqrt{t-s}} (\|u_s^1 - u_s^2\|_{L^1(\mathbb{R}^d)} + \|u_s^2 K * (u_s^1 - u_s^2)\|_{L^1(\mathbb{R}^d)}) ds \\ &\quad + C \int_0^t \frac{1}{\sqrt{t-s}} (\|u_s^1 - u_s^2\|_{L^{\mathbf{r}}(\mathbb{R}^d)} + \|u_s^2 K * (u_s^1 - u_s^2)\|_{L^{\mathbf{r}}(\mathbb{R}^d)}) ds \\ &\leq C\sqrt{t} \|u^1 - u^2\|_{t, L^1 \cap L^{\mathbf{r}}(\mathbb{R}^d)} \\ &\quad + C \int_0^t \frac{\|u_s^2\|_{L^{\mathbf{r}}(\mathbb{R}^d)} + \|u_s^2\|_{L^1(\mathbb{R}^d)}}{\sqrt{t-s}} \|K * (u_s^1 - u_s^2)\|_{L^\infty(\mathbb{R}^d)} ds. \end{aligned} \quad (\text{A.8})$$



In view of Lemma 2.1, one has  $\|K * (u_s^1 - u_s^2)\|_{L^\infty(\mathbb{R}^d)} \leq C\|u_s^1 - u_s^2\|_{L^1 \cap L^r(\mathbb{R}^d)}$ , thus plugging this upper bound in (A.8) gives

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

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

Finally, we get the following stability estimate in  $H_r^\beta$  for the solution of the PDE.

**Proposition A.7.** *Assume that  $K$  satisfies Assumptions  $(\mathbf{A}_i^K)$  and  $(\mathbf{A}_{ii}^K)$ . Let  $u_0 \in H_r^\beta(\mathbb{R}^d)$  and  $0 \leq \beta < 1$ . Then for any  $A > 0$  and  $F$  defined in (1.10), the unique solution of the cut-off PDE (A.6) verifies*

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

*Proof.* Let  $u$  be the unique solution of the cut-off PDE (A.6). Then using (1.6) and the fractional estimate (2.2), we have

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

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

## A.2 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 (2.8). Such estimates will be used in the proof of Theorem 1.3 for  $p = 1$  and  $p = \infty$  only, but the result is established for any  $p$  without any additional effort.

To achieve this, we use a generalization of the Burkholder-Davis-Gundy (BDG) inequality 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  $\| \|M_t^N\|_{L^1(\mathbb{R}^d)} \|_{L^m(\Omega)}$  for a fixed  $t > 0$ , instead of a bound on  $\|\sup_{s \in [0, t]} \|M_s^N\|_{L^1(\mathbb{R}^d)}\|_{L^m(\Omega)}$ . In order to keep the supremum in time inside the expectation, we will also use the lemma of Garsia, Rodemich and Rumsey [26]. 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.

In this section, the kernel  $K$  is only assumed to be locally integrable. We keep this very general assumption (which is in particular implied by  $(\mathbf{A}_i^K)$ - $(\mathbf{A}_{ii}^K)$  or  $(\tilde{\mathbf{A}}_i^K)$ - $(\tilde{\mathbf{A}}_{ii}^K)$ ), since we believe it might be useful in other contexts.

**Proposition A.8.** *Assume that the initial conditions satisfy  $(\mathbf{A})$  or  $(\tilde{\mathbf{A}})$  and that  $K \in L_{loc}^1(\mathbb{R}^d)$ . Let  $m \geq 1$ ,  $p \in [1, +\infty]$  and  $\varepsilon > 0$ . Then there exists  $C > 0$  such that for any  $t \in [0, T]$  and  $N \in \mathbb{N}^*$ ,*

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

where  $\varkappa_p = \max\left(0, d\left(1 - \frac{2}{p}\right)\right)$ .

The proof of Proposition A.8 relies on the following proposition, which we prove at the end of this section:

**Proposition A.9.** *For any  $p \in [1, \infty]$ , any  $m \geq 1$  and any  $\delta \in (0, 1]$ , there exists  $C > 0$  such that*

$$\left\| \|M_t^N - M_s^N\|_{L^p(\mathbb{R}^d)} \right\|_{L^m(\Omega)} \leq C (t-s)^{\frac{\delta}{2}} N^{-\frac{1}{2}(1-\alpha(d+6\delta+\varkappa_p))}, \quad \forall s \leq t \in [0, T], \quad \forall N \in \mathbb{N}^*,$$

where  $\varkappa_p$  was defined in Proposition A.8.

When  $p \geq 2$ , Proposition A.9 relies itself on the following result. Recall here the notation  $\|\cdot\|_\gamma$  for the Hilbert norm  $\|\cdot\|_{\gamma,2}$ , which was defined in (1.7).

**Proposition A.10.** *Let  $\gamma \in \mathbb{R}$  and  $m \geq 1$ . For any  $\delta \in (0, 1]$ , there exists  $C > 0$  such that*

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

The final ingredient in the proof of Proposition A.8 is a consequence of Garsia-Rodemich-Rumsey's Lemma [26], given in the following lemma (for  $\mathbb{R}$ -valued processes, it already appears in [45, Corollary 4.4], and the extension to Banach spaces is consistent with Garsia-Rodemich-Rumsey's Lemma with no additional difficulty, see e.g. [25, Theorem A.1]).

**Lemma A.11.** *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 a constant  $C_0 > 0$  and a sequence  $(\delta_n)_{n \geq 1}$  of positive real numbers such that*

$$\left( \mathbb{E} \left[ \|Y_s^n - Y_t^n\|_E^m \right] \right)^{\frac{1}{m}} \leq C_0 |s-t|^\eta \delta_n, \quad \forall s, t \in [0, T], \quad \forall n \geq 1.$$

Then for any  $\bar{m}_0 \in (0, m]$ , there exists a constant  $C$ , depending only on  $C_0, \bar{m}, m_0, \eta$  and  $T$ , such that  $\forall n \geq 1$ ,

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

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

*Proof of Proposition A.8.* We aim to apply Lemma A.11 to  $M^N$  in the Banach space  $L^p(\mathbb{R}^d)$ , for some  $p \in [1, +\infty]$ . Let  $\varepsilon > 0$  and  $m_0 > 0$ . With the notations of Proposition A.9, let us choose  $\delta = \frac{\varepsilon}{3\alpha}$ ,  $\eta = \frac{\delta}{2}$  and  $\delta_N = N^{-\rho}$  with  $\rho = -\frac{1}{2}(1-\alpha(d+\varkappa_p)) + 3\alpha\delta = -\frac{1}{2}(1-\alpha(d+\varkappa_p)) + \varepsilon$ . Hence, choosing  $m \geq 1 \vee m_0$  large enough so that  $m\eta > 1$ , the inequality in Proposition A.9 shows that  $M^N$  satisfies the conditions of Lemma A.11, and it follows that, for some constant  $C > 0$ ,

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

which is the desired result.  $\square$

It remains to prove Propositions A.10 and A.9.

*Proof of Proposition A.10.* Let  $s \leq t$  and  $x \in \mathbb{R}^d$ . First, notice that

$$\begin{aligned} |M_t^N(x) - M_s^N(x)| &\leq \left| \frac{1}{N} \sum_{i=1}^N \int_s^t \nabla e^{(t-u)\Delta} V^N(X_u^{i,N} - x) \cdot dW_u^i \right| \\ &\quad + \left| \frac{1}{N} \sum_{i=1}^N \int_0^s \nabla e^{(s-u)\Delta} \left[ e^{(t-s)\Delta} V^N(X_u^{i,N} - x) - V^N(X_u^{i,N} - x) \right] \cdot dW_u^i \right| \\ &=: |I_{s,t}^N(x)| + |II_{s,t}^N(x)|. \end{aligned} \tag{A.9}$$

Thus, one has

$$\| \|M_t^N - M_s^N\|_\gamma \|_{L^m(\Omega)} \leq \| \|I_{s,t}^N\|_\gamma \|_{L^m(\Omega)} + \| \|II_{s,t}^N\|_\gamma \|_{L^m(\Omega)}. \quad (\text{A.10})$$

Let us formulate an important remark here. As the above semigroup acts as a convolution in time within the stochastic integral,  $(M_t^N)_{t \geq 0}$  is not a martingale, and neither are  $I^N$  and  $II^N$ .

Thus, we define the processes  $\widetilde{I}^N$  and  $\widetilde{II}^N$  in the following way: For any fixed  $x \in \mathbb{R}^d$  and fixed  $0 \leq s < t$ , set:

$$\begin{aligned} \forall r \in [s, t], \quad \widetilde{I}_r^N &:= \frac{1}{N} \sum_{i=1}^N \int_s^r \nabla e^{(t-u)\Delta} V^N(X_u^{i,N} - x) \cdot dW_u^i, \\ \forall r \in [0, s], \quad \widetilde{II}_r^N &:= \frac{1}{N} \sum_{i=1}^N \int_0^r \nabla e^{(s-u)\Delta} \left[ e^{(t-s)\Delta} V^N(X_u^{i,N} - x) - V^N(X_u^{i,N} - x) \right] \cdot dW_u^i. \end{aligned}$$

Now,  $\{\widetilde{I}_r^N\}_{r \in [s,t]}$  and  $\{\widetilde{II}_r^N\}_{r \in [0,s]}$  are martingales that take values in an infinite-dimensional space, and we have  $\widetilde{I}_t^N = I_{s,t}^N$  and  $\widetilde{II}_s^N = II_{s,t}^N$ . Recall that the operators  $(I - \Delta)^{\frac{\gamma}{2}}$ ,  $\gamma \in \mathbb{R}$  were defined in the Notations section, see Equation (1.6), with the relation  $\|(I - \Delta)^{\frac{\gamma}{2}} f\|_{L^2(\mathbb{R}^d)} = \|f\|_\gamma$ .

We aim at evaluating the  $L^2(\mathbb{R}^d)$  norm of  $(I - \Delta)^{\frac{\gamma}{2}}(M_t^N - M_s^N) = (I - \Delta)^{\frac{\gamma}{2}} \widetilde{I}_t^N + (I - \Delta)^{\frac{\gamma}{2}} \widetilde{II}_s^N$ . Hence, applying the BDG inequality in UMD spaces [58, Cor. 3.11] to  $(I - \Delta)^{\frac{\gamma}{2}} \widetilde{I}^N$  and  $(I - \Delta)^{\frac{\gamma}{2}} \widetilde{II}^N$  in  $L^2(\mathbb{R}^d)$  (which is UMD) yields

$$\begin{aligned} \| \|I_{s,t}^N\|_\gamma \|_{L^m(\Omega)} &= \| \|\widetilde{I}_t^N\|_\gamma \|_{L^m(\Omega)} \\ &\leq C \left( \mathbb{E} \left[ \left\| \left( \frac{1}{N^2} \sum_{i=1}^N \int_s^t |(I - \Delta)^{\frac{\gamma}{2}} e^{(t-u)\Delta} \nabla V^N(X_u^{i,N} - \cdot)|^2 du \right)^{\frac{1}{2}} \right\|_{L^2(\mathbb{R}^d)}^m \right] \right)^{1/m}, \end{aligned} \quad (\text{A.11})$$

and

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

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

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

while for  $II_{s,t}^N$  it comes

$$\| \|II_{s,t}^N\|_\gamma \|_{L^m(\Omega)} \leq \frac{C}{N^{\frac{1}{2}}} \left( \int_0^s \frac{1}{(s-u)^{1-\delta}} \|(I - \Delta)^{\frac{\gamma+\delta}{2}} (e^{(t-s)\Delta} V^N - V^N)\|_{L^2(\mathbb{R}^d)}^2 du \right)^{\frac{1}{2}}. \quad (\text{A.12})$$

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

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

Hence, choosing  $f = (I - \Delta)^{\frac{\gamma+\delta}{2}} V^N$  and plugging the result of the previous inequality in (A.12) yields

$$\| \|II_{s,t}^N\|_\gamma \|_{L^m(\Omega)} \leq C \frac{(t-s)^{\frac{1}{2}}}{N^{\frac{1}{2}}} \|(I - \Delta)^{\frac{\gamma+\delta}{2}} \nabla V^N\|_{L^2(\mathbb{R}^d)}.$$

Using now the equivalence of norms described in [57, Eq. (3) p.59] (note that the space  $F_{p,2}^s$  in [57] is the Bessel space used here), we get

$$\| \|II_{s,t}^N\|_\gamma \|_{L^m(\Omega)} \leq C \frac{(t-s)^{\frac{1}{2}}}{N^{\frac{1}{2}}} \|(I-\Delta)^{\frac{\gamma+\delta+1}{2}} V^N\|_{L^2(\mathbb{R}^d)},$$

and using the estimate on  $\|V^N\|_{\gamma+\delta+1}$  from [17, Lemma 15], it finally comes that

$$\| \|II_{s,t}^N\|_\gamma \|_{L^m(\Omega)} \leq C(t-s)^{\frac{1}{2}} N^{\frac{\alpha}{2}(d+2\delta+2\gamma+2)-\frac{1}{2}}. \quad (\text{A.13})$$

Although the regularity in  $(t-s)$  is good in the previous inequality, we paid a factor  $N^\alpha$  which will penalise too much the rest of the computations in Propositions A.8 and A.9. Hence we also observe that

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

Thus, plugging this bound in (A.12) also gives

$$\| \|II_{s,t}^N\|_\gamma \|_{L^m(\Omega)} \leq CN^{\frac{\alpha}{2}(d+2\delta+2\gamma)-\frac{1}{2}}. \quad (\text{A.14})$$

Hence, one can interpolate between (A.13) and (A.14) to obtain that for any  $\epsilon \in [0, 1]$ ,

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

So the bound (A.11) for  $I_{s,t}^N$  and the previous inequality plugged in (A.10) and applied to  $\epsilon = \delta$  yield the desired inequality.  $\square$

*Proof of Proposition A.9.* This proof will be divided in two according to whether  $p \geq 2$  or  $p = 1$ , in that order. The result for  $p \in (1, 2)$  can be obtained by interpolation between the two previous cases.

- *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  $H^\gamma$  into  $L^p$  (which holds because  $p \geq 2$ , see [1, Theorem 1.66]), one has

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

Thus Proposition A.10 yields the result in the case  $p \geq 2$ :

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

- *Assume now that  $p = 1$ .*

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

Recall the decomposition in (A.9) of  $M_t^N - M_s^N = I_{s,t}^N + II_{s,t}^N$ . Apply Cauchy-Schwarz's inequality with weights  $(1+|x|)^{-\frac{d+1}{2}} \times (1+|x|)^{\frac{d+1}{2}}$  to get:

$$\begin{aligned} \|M_t^N - M_s^N\|_{L^1(\mathbb{R}^d)} &\leq \|I_{s,t}^N\|_{L^1(\mathbb{R}^d)} + \|II_{s,t}^N\|_{L^1(\mathbb{R}^d)} \\ &\leq C \left( \|(1+|\cdot|)^{\frac{d+1}{2}} I_{s,t}^N\|_{L^2(\mathbb{R}^d)} + \|(1+|\cdot|)^{\frac{d+1}{2}} II_{s,t}^N\|_{L^2(\mathbb{R}^d)} \right). \end{aligned} \quad (\text{A.15})$$

We will treat separately the two terms involving  $I^N$  and  $II^N$ .

Consider first the term  $I^N$  in Equation (A.15).

Using again the BDG inequality (as explained in the proof of Proposition A.10),

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

By the simple inequality  $(1 + |a + b|) \leq (1 + |a|)(1 + |b|)$  and Fubini's theorem, we then have

$$\begin{aligned} &\left\| \left\| (1 + |\cdot|)^{\frac{d+1}{2}} I_{s,t}^N \right\|_{L^2(\mathbb{R}^d)} \right\|_{L^m(\Omega)} \\ &\leq \frac{C}{N} \left\| \left\| \left( \sum_{i=1}^N \int_s^t \int_{\mathbb{R}^d} (1 + |X_s^{i,N}|)^{d+1} (1 + |X_s^{i,N} - x|)^{d+1} |\nabla e^{(t-u)\Delta} V^N(X_u^{i,N} - x)|^2 dx du \right)^{1/2} \right\|_{L^m(\Omega)} \right\| \\ &\leq \frac{C}{N} \left\| \left\| \left( \sum_{i=1}^N \int_s^t (1 + |X_s^{i,N}|)^{d+1} \int_{\mathbb{R}^d} (1 + |y|)^{d+1} |\nabla e^{(t-u)\Delta} V^N(y)|^2 dy du \right)^{1/2} \right\|_{L^m(\Omega)} \right\| \\ &\leq \frac{C}{N} \left( \int_s^t \int_{\mathbb{R}^d} (1 + |y|)^{d+1} |\nabla e^{(t-u)\Delta} V^N(y)|^2 dy du \right)^{\frac{1}{2}} \left\| \left( \sum_{i=1}^N \left( 1 + \sup_{u \in [s,t]} |X_s^{i,N}| \right)^{d+1} \right)^{\frac{1}{2}} \right\|_{L^m(\Omega)}, \end{aligned} \tag{A.16}$$

performing a simple change of variables in the second inequality. Since  $X^{i,N}$  is a diffusion with bounded coefficients (uniformly in  $N$ ), a classical argument gives that for any  $p > 0$ , there exists a constant  $C > 0$  which depends only on  $p$  and  $T$  such that  $\mathbb{E} \sup_{s \in [0,T]} |X_s^{i,N}|^p \leq C$ . It follows that

$$\left\| \left( \sum_{i=1}^N \left( 1 + \sup_{u \in [s,t]} |X_s^{i,N}| \right)^{d+1} \right)^{\frac{1}{2}} \right\|_{L^m(\Omega)} \leq C N^{\frac{1}{2}}. \tag{A.17}$$

Equations (A.16) and (A.17) now give

$$\begin{aligned} \left\| \left\| (1 + |\cdot|)^{\frac{d+1}{2}} I_{s,t}^N \right\|_{L^2(\mathbb{R}^d)} \right\|_{L^m(\Omega)} &\leq \frac{C}{\sqrt{N}} \left( \int_s^t \int_{\mathbb{R}^d} (1 + |x|)^{d+1} |\nabla e^{(t-u)\Delta} V^N(x)|^2 dx du \right)^{1/2} \\ &\leq \frac{C}{\sqrt{N}} \left( \int_s^t \int_{\mathbb{R}^d} |\nabla e^{(t-u)\Delta} V^N(x)|^2 dx du \right)^{1/2} \end{aligned} \tag{A.18}$$

$$+ \frac{C}{\sqrt{N}} \left( \int_s^t \int_{\mathbb{R}^d} |x|^{d+1} |\nabla e^{(t-u)\Delta} V^N(x)|^2 dx du \right)^{1/2}. \tag{A.19}$$

Let  $\delta > 0$ . Proceeding as in the first part of this proof with  $p = 2$  and the Sobolev embedding of  $H^0$  in  $L^2$  ( $H^0 = L^2$ ), we obtain for the term (A.18) of the previous sum:

$$\frac{C}{\sqrt{N}} \left( \int_s^t \int_{\mathbb{R}^d} |\nabla e^{(t-u)\Delta} V^N(x)|^2 dx du \right)^{1/2} \leq C (t-s)^{\frac{\delta}{2}} N^{-\frac{1}{2}(1-\alpha(d+5\delta))}.$$

Consider now the term (A.19): We will bound it successively in two different ways, by applying first the gradient to  $V^N$  and then leaving it on the heat kernel, and finally interpolating between these two results.

First, the Cauchy-Schwarz inequality yields

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

where we recall that the notation  $g$  for the heat kernel was introduced in (1.8). Then by Fubini's theorem and a simple change of variables,

$$\begin{aligned} \int_{\mathbb{R}^d} |x|^{d+1} \left| \nabla e^{(t-u)\Delta} V^N(x) \right|^2 dx &\leq N^{d\alpha+2\alpha} \int_{\mathbb{R}^d} |\nabla V(y)|^2 \int_{\mathbb{R}^d} |x + \frac{y}{N^\alpha}|^{d+1} g_{2(t-u)}(x) dx dy \\ &\leq CN^{d\alpha+2\alpha} \left( \int_{\mathbb{R}^d} |\nabla V(y)|^2 \left| \frac{y}{N^\alpha} \right|^{d+1} dy + \int_{\mathbb{R}^d} |x|^{d+1} g_{2(t-u)}(x) dx \right) \\ &\leq CN^{d\alpha+2\alpha} \left( N^{-\alpha(d+1)} + (t-u)^{\frac{d+1}{2}} \right) \\ &\leq CN^{d\alpha+2\alpha}. \end{aligned} \tag{A.20}$$

Second, another way to treat the above quantity is to leave the derivative on the Gaussian density. Notice that  $\nabla g_{2(t-u)}(x) = C \frac{x}{t-u} g_{2(t-u)}(x)$ . Hence, using the Cauchy-Schwarz inequality,

$$\begin{aligned} \left| \nabla e^{(t-u)\Delta} V^N(x) \right|^2 &= \frac{C}{(t-u)^2} \left| \int_{\mathbb{R}^d} (x-y) g_{2(t-u)}(x-y) V^N(y) dy \right|^2 \\ &\leq \frac{C}{(t-u)^2} \int_{\mathbb{R}^d} |x-y|^2 V^N(y)^2 g_{2(t-u)}(x-y) dy. \end{aligned}$$

After applying Fubini's theorem and the inequality  $|x|^{d+1} \lesssim |y|^{d+1} + |x-y|^{d+1}$ , this gives

$$\begin{aligned} \int_{\mathbb{R}^d} |x|^{d+1} \left| \nabla e^{(t-u)\Delta} V^N(x) \right|^2 dx &\leq \frac{C}{(t-u)^2} \int_{\mathbb{R}^d} V^N(y)^2 \int_{\mathbb{R}^d} |x|^{d+1} |x-y|^2 g_{2(t-u)}(x-y) dx dy \\ &\leq \frac{C}{(t-u)^2} \left( \int_{\mathbb{R}^d} |y|^{d+1} (V^N(y))^2 dy \int_{\mathbb{R}^d} |x|^2 g_{2(t-u)}(x) dx \right. \\ &\quad \left. + \int_{\mathbb{R}^d} (V^N(y))^2 dy \int_{\mathbb{R}^d} |x|^{d+3} g_{2(t-u)}(x) dx \right) \\ &\leq \frac{C}{(t-u)^2} \left( N^{-\alpha} (t-u) + N^{d\alpha} (t-u)^{\frac{d+3}{2}} \right) \\ &\leq CN^{d\alpha} (t-u)^{-1}. \end{aligned} \tag{A.21}$$

Interpolating between the bounds (A.20) and (A.21), respectively with powers  $\frac{\delta}{2}$  and  $1 - \frac{\delta}{2}$ , we get

$$\begin{aligned} \int_s^t \int_{\mathbb{R}^d} |x|^{d+1} \left| \nabla e^{(t-u)\Delta} V^N(x) \right|^2 dx du &\leq CN^{\alpha(d+\delta)} \int_s^t (t-u)^{-(1-\frac{\delta}{2})} du \\ &\leq CN^{\alpha(d+\delta)} (t-s)^{\frac{\delta}{2}}. \end{aligned}$$

Hence, in view of this bound and the bound on (A.18), we have finally obtained that for any  $\delta > 0$ , there exists  $C > 0$  such that

$$\left\| \left\| (1 + |\cdot|)^{\frac{d+1}{2}} I_{s,t}^N \right\|_{L^2(\mathbb{R}^d)} \right\|_{L^m(\Omega)} \leq C (t-s)^{\frac{\delta}{2}} N^{-\frac{1}{2}(1-\alpha(d+5\delta))}. \tag{A.22}$$

Consider now the term  $II^N$  in Equation (A.15).

Using again the BDG inequality (as explained in the proof of Proposition A.10),

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

Performing the same computations (A.16)-(A.17) as for the term  $I^N$ , one gets

$$\left\| \left\| (1 + |\cdot|)^{\frac{d+1}{2}} II_{s,t}^N \right\|_{L^2(\mathbb{R}^d)} \right\|_{L^m(\Omega)} \leq \frac{C}{\sqrt{N}} \left( \int_0^s \int_{\mathbb{R}^d} (1 + |x|)^{d+1} \left| \nabla e^{(s-u)\Delta} \left[ e^{(t-s)\Delta} V^N(x) - V^N(x) \right] \right|^2 dx du \right)^{\frac{1}{2}}. \quad (\text{A.23})$$

Here again, we will estimate the previous integral in two different ways and then interpolate.

Notice first that

$$\left| \nabla e^{(s-u)\Delta} \left[ e^{(t-s)\Delta} V^N(x) - V^N(x) \right] \right|^2 \leq \left| \nabla e^{(t-u)\Delta} V^N(x) \right|^2 + \left| \nabla e^{(s-u)\Delta} V^N(x) \right|^2.$$

Hence,

$$\begin{aligned} & \int_{\mathbb{R}^d} (1 + |x|)^{d+1} \left| \nabla e^{(s-u)\Delta} \left[ e^{(t-s)\Delta} V^N(x) - V^N(x) \right] \right|^2 dx \\ & \leq \int_{\mathbb{R}^d} (1 + |x|)^{d+1} \left[ \left| \nabla e^{(t-u)\Delta} V^N(x) \right|^2 + \left| \nabla e^{(s-u)\Delta} V^N(x) \right|^2 \right] dx \\ & =: A + B. \end{aligned}$$

The two terms  $A$  and  $B$  can be treated in the same manner as in (A.21). This leads us to the following inequalities:

$$A \leq C \frac{N^{d\alpha}}{(t-u)} \quad \text{and} \quad B \leq C \frac{N^{d\alpha}}{(s-u)}.$$

Hence, as  $s < t$ , on the one hand we have

$$\int_{\mathbb{R}^d} (1 + |x|)^{d+1} \left| \nabla e^{(s-u)\Delta} \left[ e^{(t-s)\Delta} V^N(x) - V^N(x) \right] \right|^2 dx \leq C \frac{N^{d\alpha}}{(s-u)}. \quad (\text{A.24})$$

On the other hand, applying the gradient on  $V^N$ , one gets

$$\begin{aligned} \left| \nabla e^{(s-u)\Delta} \left[ e^{(t-s)\Delta} V^N(x) - V^N(x) \right] \right| & = \left| g_{2(s-u)} * [g_{2(t-s)} * \nabla V^N(x) - \nabla V^N(x)] \right| \\ & \leq \int_{\mathbb{R}^d} g_{2(s-u)}(x-y) \int_{\mathbb{R}^d} g_{2(t-s)}(z) \left| \nabla V^N(y-z) - \nabla V^N(y) \right| dz dy \\ & \leq \|\nabla^2 V^N\|_{\infty} \int_{\mathbb{R}^d} g_{2(s-u)}(x-y) \int_{\mathbb{R}^d} g_{2(t-s)}(z) |z| dz dy. \end{aligned}$$

Since  $\|\nabla^2 V^N\|_{\infty} \lesssim N^{d\alpha+2\alpha}$  (using the regularity of  $V$ ), it comes

$$\left| \nabla e^{(s-u)\Delta} \left[ e^{(t-s)\Delta} V^N(x) - V^N(x) \right] \right| \leq C N^{d\alpha+2\alpha} \sqrt{t-s}.$$

Hence

$$\begin{aligned}
 & \int_{\mathbb{R}^d} (1 + |x|)^{d+1} \left| \nabla e^{(s-u)\Delta} \left[ e^{(t-s)\Delta} V^N(x) - V^N(x) \right] \right|^2 dx \\
 & \leq C N^{d\alpha+2\alpha} \sqrt{t-s} \int_{\mathbb{R}^d} (1 + |x|)^{d+1} \left| \nabla e^{(s-u)\Delta} \left[ e^{(t-s)\Delta} V^N(x) - V^N(x) \right] \right| dx \\
 & \leq C N^{d\alpha+2\alpha} \sqrt{t-s} \left\{ \int_{\mathbb{R}^d} (1 + |x|)^{d+1} \left| e^{(t-u)\Delta} \nabla V^N(x) \right| dx \right. \\
 & \quad \left. + \int_{\mathbb{R}^d} (1 + |x|)^{d+1} \left| e^{(s-u)\Delta} \nabla V^N(x) \right| dx \right\}
 \end{aligned}$$

and by a change of variables,

$$\begin{aligned}
 & \int_{\mathbb{R}^d} (1 + |x|)^{d+1} \left| \nabla e^{(s-u)\Delta} \left[ e^{(t-s)\Delta} V^N(x) - V^N(x) \right] \right|^2 dx \\
 & \leq C N^{d\alpha+3\alpha} \sqrt{t-s} \left\{ \int_{\mathbb{R}^d} \left(1 + \left|\frac{x}{N^\alpha}\right|\right)^{d+1} \left| e^{(t-u)\Delta} \nabla V(x) \right| dx + \int_{\mathbb{R}^d} \left(1 + \left|\frac{x}{N^\alpha}\right|\right)^{d+1} \left| e^{(s-u)\Delta} \nabla V(x) \right| dx \right\} \\
 & \leq C N^{d\alpha+3\alpha} \sqrt{t-s} \left\{ \int_{\mathbb{R}^d} (1 + |x|)^{d+1} \left| e^{(t-u)\Delta} \nabla V(x) \right| dx + \int_{\mathbb{R}^d} (1 + |x|)^{d+1} \left| e^{(s-u)\Delta} \nabla V(x) \right| dx \right\}.
 \end{aligned}$$

Observe that for any  $\sigma > 0$ , using the inequality  $(1 + |x|) \leq (1 + |y|)(1 + |x - y|)$ , one gets

$$\int_{\mathbb{R}^d} (1 + |x|)^{d+1} \left| e^{\sigma\Delta} \nabla V(x) \right| dx \leq \int_{\mathbb{R}^d} (1 + |y|)^{d+1} \left| \nabla V(y) \right| \int_{\mathbb{R}^d} (1 + |x - y|)^{d+1} g_{2\sigma}(x - y) dx dy,$$

which is bounded by a constant independent of  $\sigma$ . Thus

$$\int_{\mathbb{R}^d} (1 + |x|)^{d+1} \left| \nabla e^{(s-u)\Delta} \left[ e^{(t-s)\Delta} V^N(x) - V^N(x) \right] \right|^2 dx \leq C N^{d\alpha+3\alpha} \sqrt{t-s}. \quad (\text{A.25})$$

Now for  $\delta > 0$ , one can interpolate between (A.24) and (A.25) to get:

$$\int_{\mathbb{R}^d} (1 + |x|)^{d+1} \left| \nabla e^{(s-u)\Delta} \left[ e^{(t-s)\Delta} V^N(x) - V^N(x) \right] \right|^2 dx \leq C N^{\alpha(d+6\delta)} \frac{(t-s)^\delta}{(s-u)^{1-2\delta}}.$$

Plugging this bound into (A.23) now provides

$$\left\| \left\| (1 + |\cdot|)^{\frac{d+1}{2}} I I_{s,t}^N \right\|_{L^2(\mathbb{R}^d)} \right\|_{L^m(\Omega)} \leq C N^{-\frac{1}{2}(1-\alpha(d+6\delta))} (t-s)^{\frac{\delta}{2}}. \quad (\text{A.26})$$

Putting together Equations (A.22) and (A.26), we deduce from (A.15) that

$$\|M_t^N - M_s^N\|_{L^1(\mathbb{R}^d)} \leq C N^{-\frac{1}{2}(1-\alpha(d+6\delta))} (t-s)^{\frac{\delta}{2}}.$$

Finally for  $p \in [1, 2]$ , we obtain the same estimate by interpolating between the  $L^1$  norm and the  $L^2$  norm.  $\square$

### A.3 Proof of the boundedness estimate

**Proposition A.12.** *Let the assumptions of Proposition A.8 hold. Let  $q \geq 1$  and  $\beta \in [0, 1)$ . We assume that  $\sup_{N \in \mathbb{N}^*} \mathbb{E} \|u_0^N\|_{\beta, r}^q < \infty$  and that*

$$0 < \alpha < \frac{1}{d + 2\beta + 2d\left(\frac{1}{2} - \frac{1}{r}\right) \vee 0}. \quad (\text{A.27})$$

Then

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



*Proof. Step 1.* Recall that the operators  $(I - \Delta)^\beta$ ,  $\beta \in \mathbb{R}$  were defined in the Notations section, see Equation (1.6), with a clear link with the Sobolev norm  $\|\cdot\|_{\beta,r}$ .

Let  $F$  stand for the function  $F_A$  defined in (1.10). From (2.3) after applying  $(I - \Delta)^{\frac{\beta}{2}}$  and by the triangular inequality we have

$$\left\| (I - \Delta)^{\frac{\beta}{2}} u_t^N \right\|_{L^r(\mathbb{R}^d)} \leq \left\| (I - \Delta)^{\frac{\beta}{2}} e^{t\Delta} u_0^N \right\|_{L^r(\mathbb{R}^d)} \quad (\text{A.28})$$

$$+ \int_0^t \left\| (I - \Delta)^{\frac{\beta}{2}} \nabla \cdot e^{(t-s)\Delta} (V^N * (F(K * u_s^N) \mu_s^N)) \right\|_{L^r(\mathbb{R}^d)} ds \quad (\text{A.29})$$

$$+ \left\| \frac{1}{N} \sum_{i=1}^N \int_0^t (I - \Delta)^{\frac{\beta}{2}} \nabla e^{(t-s)\Delta} (V^N (X_s^{i,N} - \cdot)) \cdot dW_s^i \right\|_{L^r(\mathbb{R}^d)}. \quad (\text{A.30})$$

**Step 2.** By a convolution inequality, we have

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

**Step 3.** Let us come to the second term (A.29):

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

Using the equivalence of norms described in [57, Eq. (3) p.59] (note that the space  $F_{p,2}^s$  in [57] is the Bessel space used here), we get

$$\left\| (I - \Delta)^{\frac{\beta}{2}} \nabla \cdot e^{(t-s)\Delta} \right\|_{L^r \rightarrow L^r} \leq C \left\| (I - \Delta)^{\frac{1+\beta}{2}} e^{(t-s)\Delta} \right\|_{L^r \rightarrow L^r}.$$

Hence in view of Inequality (2.2), we have that

$$\left\| (I - \Delta)^{\frac{\beta}{2}} \nabla \cdot e^{(t-s)\Delta} \right\|_{L^r \rightarrow L^r} \leq C \frac{1}{(t-s)^{\frac{(1+\beta)}{2}}}.$$

Thus,

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

using (1.5) in the last inequality.

**Step 4.** Recalling the notation introduced in (2.8), the third term (A.30) is nothing but  $\|M_t^N\|_{\beta,r}$ . Plugging the result of Steps 2 and 3 in (A.28)-(A.29)-(A.30), we can now use the Grönwall lemma for convolution integrals (see e.g. [29, Lemma 7.1.1]). Then we obtain that there exists a constant  $C > 0$  such that

$$\sup_{t \in [0,T]} \|u_t^N\|_{\beta,r} \leq C \left( \|u_0^N\|_{\beta,r} + \sup_{t \in [0,T]} \|M_t^N\|_{\beta,r} \right). \quad (\text{A.31})$$

**Step 5.** The embedding for Bessel potential spaces of [56, p.203] gives that  $H^{\beta+d(\frac{1}{2}-\frac{1}{r})}(\mathbb{R}^d)$  is continuously embedded into  $H_r^\beta(\mathbb{R}^d)$ , thus we obtain

$$\|M_t^N\|_{\beta,r} \leq C \|M_t^N\|_{\beta+d(\frac{1}{2}-\frac{1}{r}),2}.$$

Now combining Proposition A.10 with Lemma A.11 as in the proof of Proposition A.8, we get that for any  $\varepsilon > 0$ , there exists  $C > 0$  such that

$$\left\| \sup_{t \in [0, T]} \|M_t^N\|_{\beta, \mathbf{r}} \right\|_{L^q(\Omega)} \leq C N^{-\frac{1}{2}(1 - \alpha(d + 2\beta + 2d(\frac{1}{2} - \frac{1}{\mathbf{r}}))) + \varepsilon}.$$

In view of the constraint in (A.27), it follows that

$$\left\| \sup_{t \in [0, T]} \|M_t^N\|_{\beta, \mathbf{r}} \right\|_{L^q(\Omega)} \leq C.$$

Taking the  $L^q(\Omega)$  norm in (A.31) and applying the previous inequality, we obtain the desired result.  $\square$

## References

- [1] H. Bahouri, J.-Y. Chemin and R. Danchin, “Fourier Analysis and Nonlinear Partial Differential Equations”, Grundlehren der Mathematischen Wissenschaften, volume 343, 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** (2018), 1–35.
- [3] P. Biler, *Mathematical challenges in the theory of chemotaxis*, Ann. Math. Sil. **32** (2018), 43–63.
- [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** (1997), 157–192.
- [6] D. Bresch, P.-E. Jabin and Z. Wang, *Mean-field limit and quantitative estimates with singular attractive kernels*, Preprint [arXiv:2011.08022](https://arxiv.org/abs/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** (2016), 447–463.
- [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** (2008), 19–40.
- [10] E. Cépa and D. Lépingle, *Diffusing particles with electrostatic repulsion*, Probab. Theory Related Fields **107** (1997), 429–449.
- [11] P.-E. Chaudru de Raynal, *Strong well posedness of McKean-Vlasov stochastic differential equations with Hölder drift*, Stochastic Process. Appl. **130** (2020), 79–107.
- [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](https://arxiv.org/abs/1811.06904) (2018).
- [13] R. Cortez and J. Fontbona, *Quantitative uniform propagation of chaos for Maxwell molecules*, Comm. Math. Phys. **357** (2018), 913–941.

- [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** (2015), 2451–2492.
- [15] M. Duerinckx, *Mean-field limits for some Riesz interaction gradient flows*, SIAM J. Math. Anal. **48** (2016), 2269–2300.
- [16] F. Flandoli and M. Leocata, *A particle system approach to aggregation phenomena*, J. Appl. Probab. **56** (2019), 282–306.
- [17] F. Flandoli, M. Leimbach and C. Olivera, *Uniform convergence of proliferating particles to the FKPP equation*, J. Math. Anal. Appl. **473** (2019), 27–52.
- [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** (2019), 3741–3753.
- [19] F. Flandoli, M. Leocata and C. Ricci, *On the macroscopic limit of Brownian particles with local interaction*, Stoch. Dyn. **20** (2020), 2040007.
- [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** (2020), 5339–5362.
- [21] N. Fournier and M. Hauray, *Propagation of chaos for the Landau equation with moderately soft potentials*, Ann. Probab. **44** (2016), 3581–3660.
- [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** (2017), 2807–2861.
- [23] N. Fournier and S. Mischler, *Rate of convergence of the Nanbu particle system for hard potentials and Maxwell molecules*, Ann. Probab. **44** (2016), 589–627.
- [24] N. Fournier, M. Hauray and S. Mischler, *Propagation of chaos for the 2D viscous vortex model*, J. Eur. Math. Soc. (JEMS) **16** (2014), 1423–1466.
- [25] 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.
- [26] 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** (1970/71), 565–578.
- [27] D. Godinho and C. Quiñinao, *Propagation of chaos for a subcritical Keller-Segel model*, Ann. Inst. Henri Poincaré Probab. Stat. **51** (2015), 965–992.
- [28] L. Grafakos, “Classical Fourier Analysis”, volume 249 of Graduate Texts in Mathematics, Springer, New York, third edition, 2014.
- [29] D. Henry, “Geometric Theory of Semilinear Parabolic Equations”, volume 840 of Lecture Notes in Mathematics, Springer-Verlag, Berlin-New York, 1981.
- [30] P.-E. Jabin and Z. Wang, *Quantitative estimates of propagation of chaos for stochastic systems with  $W^{-1,\infty}$  kernels*, Invent. Math. **214** (2018), 523–591.
- [31] 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).

- [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** (1998), 727–766.
- [33] M. Kac, *Foundations of kinetic theory*, In: “Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability”, 1954–1955, vol. III, University of California Press, Berkeley and Los Angeles, 1956, 171–197.
- [34] D. Lacker, *On a strong form of propagation of chaos for McKean-Vlasov equations*, Electron. Commun. Probab. **23** (2018).
- [35] C. Marchioro and M. Pulvirenti, *Hydrodynamics in two dimensions and vortex theory*, Comm. Math. Phys. **84** (1982), 483–503.
- [36] 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., Springer, Berlin, 1996, 42–95.
- [37] S. Méléard, *A trajectorial proof of the vortex method for the two-dimensional Navier-Stokes equation*, Ann. Appl. Probab. **10** (2000), 1197–1211.
- [38] S. Méléard and S. Roelly-Coppoletta, *A propagation of chaos result for a system of particles with moderate interaction*, Stochastic Process. Appl. **26** (1987), 317–332.
- [39] Y. Mishura and A. Veretennikov, *Existence and uniqueness theorems for solutions of McKean-Vlasov stochastic equations*, Theor. Probability and Math. Statist. **103** (2020).
- [40] 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** (2011), 29–68.
- [41] K. Oelschläger, *A law of large numbers for moderately interacting diffusion processes*, Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete **69** (1985), 279–322.
- [42] K. Oelschläger, *A fluctuation theorem for moderately interacting diffusion processes*, Probab. Theory Related Fields **74** (1987), 591–616.
- [43] 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](https://arxiv.org/abs/2004.03177) (2020).
- [44] H. Osada, *Propagation of chaos for the two-dimensional Navier-Stokes equation*, Proc. Japan Acad. Ser. A Math. Sci. **62** (1986), 8–11.
- [45] A. Richard, X. Tan and F. Yang, *Discrete-time simulation of stochastic Volterra equations*, Stochastic Process. Appl. **141** (2021), 109–138.
- [46] M. Röckner and X. Zhang, *Well-posedness of distribution dependent SDEs with singular drifts*, Bernoulli **27** (2021), 1131 – 1158.
- [47] M. Rosenzweig, *Mean-field convergence of systems of particles with Coulomb interactions in higher dimensions without regularity*, Preprint [arXiv:2010.10009](https://arxiv.org/abs/2010.10009) (2020).
- [48] W. Rudin, “Functional Analysis”, International Series in Pure and Applied Mathematics, McGraw-Hill, Inc., New York, second edition, 1991.
- [49] S. G. Samko, A. A. Kilbas and O. I. Marichev, “Fractional Integrals and Derivatives”, Gordon and Breach Science Publishers, Yverdon, 1993. Edited and with a foreword by S. M. Nikolskii, Translated from the 1987 Russian original, Revised by the authors.
- [50] S. Serfaty, *Mean field limit for Coulomb-type flows*, Duke Math. J. **169** (2020), 2887–2935.

- [51] M. Simon and C. Olivera, *Non-local conservation law from stochastic particle systems*, J. Dynam. Differential Equations **30** (2018), 1661–1682.
- [52] A.-S. Sznitman, *A propagation of chaos result for Burgers' equation*, Probab. Theory Related Fields **71** (1986), 581–613.
- [53] 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., Springer, Berlin, 1991, 165–251.
- [54] D. Talay and M. Tomašević, *A new McKean–Vlasov stochastic interpretation of the parabolic–parabolic Keller–Segel model: The one-dimensional case*, Bernoulli **26** (2020), 1323–1353.
- [55] M. Tomašević, *On a probabilistic interpretation of the parabolic–parabolic Keller–Segel model*, Ph.D. Thesis, 2018.
- [56] H. Triebel, “Interpolation Theory, Function Spaces, Differential Operators”, volume 18 of North-Holland Mathematical Library, North-Holland Publishing Co., Amsterdam-New York, 1978.
- [57] H. Triebel, “Theory of Function Spaces”, volume 78 of Monographs in Mathematics, Birkhäuser Verlag, Basel, 1983.
- [58] J. M. A. M. van Neerven, M. C. Veraar and L. Weis, *Stochastic integration in UMD Banach spaces*, Ann. Probab. **35** (2007), 1438–1478.

CHRISTIAN OLIVERA

Departamento de Matemática, Universidade Estadual de Campinas, Brazil  
*E-mail address:* [colivera@ime.unicamp.br](mailto:colivera@ime.unicamp.br)

ALEXANDRE RICHARD

Université Paris-Saclay, CentraleSupélec, MICS and CNRS FR-3487, 91190 Gif-sur-Yvette, France  
*E-mail address:* [alexandre.richard@centralesupelec.fr](mailto:alexandre.richard@centralesupelec.fr)

MILICA TOMAŠEVIĆ

CNRS, CMAP, Ecole Polytechnique, I.P. Paris, 91128 Palaiseau, France  
*E-mail address:* [milica.tomasevic@polytechnique.edu](mailto:milica.tomasevic@polytechnique.edu)