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

Christian Olivera, Alexandre Richard, Milica Tomasevic

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

Christian Olivera * Alexandre Richard† Milica Tomašević ‡

November 11, 2020

Abstract

We propose a new approach to obtain quantitative convergence of moderately interacting particle systems to solutions of nonlinear Fokker-Planck equations with singular kernels. Our result only requires very weak regularity on the interaction kernel, including the Biot-Savart kernel, the family of Keller-Segel kernels in arbitrary dimension, and more generally singular Riesz kernels. This seems to be the first time that such quantitative convergence results are obtained in Lebesgue and Sobolev norms for the aforementioned kernels. In particular, this convergence holds locally in time for PDEs exhibiting a blow-up in finite time. The proof is based on a semigroup approach combined with stochastic calculus techniques, and we also exploit the regularity of the solutions of the limiting equation.

Furthermore, we obtain well-posedness for the McKean-Vlasov SDEs involving these singular kernels and we prove the trajectorial propagation of chaos for the associated moderately interacting particle systems.

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

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

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

1 Introduction

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

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

where K is a locally integrable kernel that may have a singular behaviour at 0.

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

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

The associated kernel is then $K_s := \nabla V_s$. If $d \geq 2$ and $s = d-2$, this is the Coulomb potential that characterises electrostatic and gravitational forces (depending on the sign).

Our motivation comes, more particularly, from several classical models. First, there is the 2d Navier-Stokes equation, which in vorticity form can be written as in (1.1) with the Biot-Savart kernel

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

Note that $|K(x)| = \frac{1}{\pi}|\nabla V_0(x)|$, where $V_0(x) = -\log|x|$ is the fundamental solution of the Laplace equation in dimension 2 up to a constant.
We are also interested in the parabolic-elliptic Keller-Segel PDE in any dimension $d \geq 1$, for which the kernel is given by

$$K_{KS}(x) = -\chi \frac{x}{|x|^d} \tag{1.3}$$

for some $\chi > 0$. Equation (1.1) then describes the chemotaxis of biological cells subject to diffusive displacement and attracted by the concentration of nutrients in their environment. It entails a

mathematical problem which is particularly tricky due to the fact that the dynamics is attractive and that a blow-up may occur in the PDE, depending on the dimension d and the value of χ (see e.g. Biler [3]).
Equation (1.1) also arises in many different contexts, as for instance swarming, aggregation phenomena, see e.g. [8, 56]. See [28] for an excellent summary.

Note that the drift term may also be of the form $\Phi(x) + K *_x u(t,x)$ where Φ is some bounded force applied to the system. By requiring enough regularity on Φ (typically bounded and Lipschitz continuous), the term $\Phi(x)$ does not bring any particular difficulty and all the results in this paper hold for a system with this additional term. However, the statements and the proofs would become cumbersome with it. Thus, for the sake of notational simplicity, we choose to work with $\Phi \equiv 0$. Besides, Equation (1.1) formally preserves the total mass $M := \int_{\mathbb{R}^d} u_0(x)\, dx$. That is why, we may and will assume from now on that $M = 1$ in (1.1) and we will not change the notation for K.

The problem of deriving a macroscopic equation from a microscopic model of interacting particles can be traced back to the original inspiration of Kac [32], in the context of the Boltzmann equation. Since then, a huge literature has been devoted to interacting particle systems and their convergence to (1.1), mostly in the case of Lipschitz continuous kernels of interaction: When the particles are interacting diffusion processes, this problem is now well-understood, see e.g. Sznitman [54] and Méléard [36] for a general account of the theory.

The case of singular kernels such as the aforementioned ones is more recent and there are fewer works on this topic, despite the great 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, and of Sznitman [53] and Bossy and Talay [5] on Burgers' equation. Cépa and Lépingle [11] 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. Oustide the scope of physics, interesting biological models have arisen, for instance in neuroscience with the work of Delarue et al. [15] (diffusive particles interacting through their hitting times), and several works on the Keller-Segel equation (Fournier and Jourdain [22], Cattiaux and Pédèches [9] and Jabir et al. [30]).

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

$$dX_t^{i,N} = \frac{1}{N}\sum_{j=1}^{N} K(X_t^{i,N} - X_t^{j,N})\, dt + \sqrt{2}\, dW_t^i, \tag{1.4}$$

where $\{W_t^i,\ i \in \mathbb{N}\}$ is a family of independent d-dimensional Brownian motions (with the convention that $K(0) = 0$). Unfortunately, due to the singular interaction kernel, it is not obvious that this particle system is well-defined and that the propagation of chaos holds. The study of this system is usually done on a case-by-case basis.

In the case of similar singular non-diffusive problems (i.e. with deterministic particles), there has been significant progress recently, see e.g. the work of Serfaty [50] who introduced the modulated energy method, Duerinckx [16] and Rosenzweig [47].

Another well-suited approach for tackling the convergence of particles interacting through singular kernels, that we adopt in the present paper, is to consider moderately interacting particles in the sense of Oelschläger [41]. In the moderate interaction setting, one introduces a smoothing of the interaction kernel. This has been successfully applied in the aforementioned works [5], [21], as well as by Méléard [37] and Méléard and Roelly-Coppoletta [38], and more recently with a new semigroup approach developed by Flandoli et al. [18].

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

$$dX_t^{i,N} = F_A\left(\frac{1}{N}\sum_{k=1}^{N}(K * V^N)(X_t^{i,N} - X_t^{k,N})\right) dt + \sqrt{2}\, dW_t^i, \quad t \leq T, \quad 1 \leq i \leq N, \tag{1.5}$$

where V^N is a mollifier, F_A is a smooth cut-off function that ensures that the drift driving each particle remains uniformly bounded in N, and $A > 0$ is the cut-off parameter. As such, the existence of strong solutions for (1.5) is ensured.

Our first objective is to prove the uniform convergence of its mollified empirical measure towards the solution of (1.1) when the number of particles goes to infinity, for a large class of interacting kernels including those mentioned in the beginning.

For that purpose, we follow the new approach presented in Flandoli et al. [18], based on semi-group theory and developed first with application to the FKPP equation. This technique permits to approximate non-linear PDEs by smoothed empirical measures in strong functional topologies. It has already found many applications: see Flandoli and Leocata [17] for a PDE-ODE system related to aggregation phenomena; Simon and Olivera [52] for non-local conservation laws; Flandoli et al. [20] for the 2d Navier-Stokes equation; and Olivera et al. [43] for the $2d$ Keller-Segel systems.

The main difficulty here is the singular nature of the kernel and finding a suitable functional framework in which the convergence takes place. In particular, it is crucial in our approach to get a solution theory for the class of PDEs (1.1). Hence in Section 3, we establish the local existence and uniqueness of mild solutions to (1.1) in the space $\mathcal{C}\left([0,T); L^1 \cap L^\infty\right)$. This is also established for an ansatz PDE with cut-off corresponding to the limit of the particle system (1.5). Then, we obtain the convergence in probability of the mollified empirical measure of the particle system (1.5) towards the solution of (1.1) in spaces of the form $L^2([0,T], H_p^\gamma(\mathbb{R}^d))$, see Theorem 2.5. Here H^γ is a nonhomogeneous Sobolev space (also known as Bessel potential space). Note that even for PDEs having a blow-up in finite time (such as the Keller-Segel model with $d = 2$ and χ large enough), this convergence happens, at least before the blow-up.

The general question of quantifying the convergence of interacting particle systems towards the PDE in non-singular framework has been addressed thoroughly in the literature. See for example [31, 42, 54] or, more recently, Cortez and Fontbona [14] in the case of homogeneous Boltzmann equation for Maxwell molecules. However, in the singular case there are fewer results in the literature: Méléard [37] obtained a non-explicit rate on the density of one particle for the 2d Navier-Stokes equation, Bossy and Talay [5] got a rate for Burger's equation, Fournier and Mischler [23] obtained a rate for the so-called Nanbu particle system approaching the Boltzmann equation, Fournier and Hauray [21] approximated the Landau equation with moderately soft potentials; and for a more systematic study, Jabin and Wang [29] recently obtained a quantitative convergence of the marginal densities with a $N^{-1/2}$ rate for some singular kernels including the Biot-Savart kernel (but not Riesz kernels). Bresch et al. [6] also provide a slower rate of convergence for the tricky attractive case of the Keller-Segel model (in dimension 2). We shall discuss in details and compare some of these works to ours in Section 2.5.

In this work, we obtain a quantitative convergence of the mollified empirical measure in $L^m(\Omega; L^\infty((0,T); \mathcal{X}))$ norm, where either $\mathcal{X} = L^1 \cap L^\infty$ or $\mathcal{X} = H_p^\gamma$ (see Theorem 2.6 and Corollary 2.9). The main tool in obtaining these results is the mild equation for the mollified empirical measure and its limit combined with stochastic calculus techniques. As a consequence in Corollary 2.13, we obtain the same rate for the real empirical measure in a weak topology.

This seems to be the first time that such quantitative convergence results are obtained in Lebesgue and Sobolev norms for kernels such as the Biot-Savart, Riesz or Keller-Segel kernels.

However, the rates we provide are always below $\frac{1}{2}$, which is the rate one would classically hope to obtain for non-singular interactions of mean field type (see e.g. [10]) or even in the Biot-Savart case [29] (although the measured quantities are not the same). The reason for this slower convergence is that our particles do not interact in a mean field way, but rather within a typical range of order $N^{-\alpha}$, for some $\alpha \in (0,1)$. As it happens, we have the constraint $\alpha < \frac{1}{2(d+1)}$ (for any $d \geq 2$) for the convergence to hold, and we do obtain this rate in the end, which is in this sense optimal.

Finally, our last results concern, on one hand, the weak well-posedness of the following McKean-Vlasov SDE:

$$\begin{cases} dX_t = K * u_t(X_t)\, dt + \sqrt{2} dW_t, & t \leq T, \\ \mathcal{L}(X_t) = u_t, \ \mathcal{L}(X_0) = u_0, \end{cases} \tag{1.6}$$

assuming that u_0 has enough regularity ($u_0 \in L^1 \cap L^\infty$). On the other hand, we prove the propagation of chaos of the empirical measure of the particle system (1.5) towards the law of this SDE, at the level of probability measures on the space of trajectories (Proposition 2.16 and Theorem 2.17).

The first result comes from the combination of the fact that the marginal laws of the process are uniquely determined as solutions of (1.1) with the fact that the linearised version of (1.6) admits a unique weak solution. This is one of the classical ways to prove the well-posedness of such distribution dependent SDEs, along with the frequently used fixed point argument and convergence of the empirical measure argument. We choose this approach as we have *a priori* information about the PDE. We remark here that the weak solution is local in time if there is an explosion in finite time in the corresponding PDE, and global if the corresponding PDE is globally well-posed.

To obtain the second result, we first prove that the empirical measure converges to the cut-off version of the SDE (1.6) (where the drift is $F_A(K * u_t)$). This is obtained in the spirit of the work of Méléard and Roelly-Coppoletta [38], that we extend to singular kernels thanks to our result about the convergence of the mollified empirical measure. Then, choosing the parameter A conveniently, we are able to lift the cut-off and conclude the convergence of the empirical measure towards the law of (1.6).

Recently, the well-posedness of McKean-Vlasov SDEs (or distribution dependent SDEs) has gained much attention in the literature (see e.g. [2, 12, 13, 27, 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 (1.6). 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 coefficient (in the space variable). As both of these difficulties appear in the specific distribution dependence in (1.6), our well-posedness result gives a new perspective on the matter. In some particular cases such as the Keller-Segel model, this extends previous results [22] and [46], but it requires more regularity on the initial data.

When it comes to our second result about the propagation of chaos, one important remark is in order. Usually, when dealing with singular non-linear PDEs, the question of propagation of chaos of the related particle system is very demanding as it may happen (due to singular interaction) that it is not even possible to define the particle system. That is why, heavy techniques are used on a case by case basis (see e.g. [22]) or the well-posedness of the particle system is assumed (see e.g. [6]). Passing to the framework of a particle system in moderate interaction, we exhibit a powerful tool to approximate singular PDEs that circumvents the difficulty of well-posedness of the mean-field particle system and we prove the propagation of chaos of the moderately interacting particle system in very general and singular framework (even when there is an explosion in the associated PDE). This is particularly useful for numerical applications that we plan to tackle in a future work.

Plan of the paper. The framework, the assumptions on the kernel K, as well as the main results of this work are presented in Section 2. The existence and uniqueness of (1.1) and its cut-off version are studied in Section 3. In Section 4, we detail the proof that gives the convergence of the mollified empirical measure (Theorem 2.5). The result about the rate of convergence (Theorems 2.6) is proved in Section 5. In Section 6, the existence and uniqueness of the martingale problem associated to the cut-off McKean-Vlasov SDE are proven (Proposition 2.16), as well as the propagation of chaos for the empirical measure of (1.5) (Theorem 2.17). Finally, we present some examples and applications of our results in Section 7. In the Appendix one may find a general time and space estimates for some stochastic convolution integrals and the proof of tightness estimates used in Section 4.

Notations and definitions.

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

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

where $\mathcal{F}u$ denotes the *Fourier transform* of u. We endow this space with the norm

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

In particular, note that

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

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

$$(I-\Delta)^{\frac{\beta}{2}} f := \mathcal{F}^{-1}\left((1+|\cdot|^2)^{\frac{\beta}{2}} \mathcal{F}f\right). \tag{1.7}$$

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

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

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

- Let us now recall the definition of *Sobolev spaces*. Let U be a general, possibly non smooth, open set in $\mathbb{R}^d$. Let $1 \leq p < \infty$. For any positive integer k, the usual Sobolev spaces of k times weakly differentiable elements of $L^p(\mathbb{R}^d)$ whose derivatives are also in $L^p(\mathbb{R}^d)$ is defined by

$$W^{k,p}(U) := \Big\{ f \in L^p(U)\ ;\ \|f\|_{k,p} := \sum_{|\alpha|\leq k} \|D^\alpha f\|_{L^p(U)} < \infty \Big\}.$$

For any $s > 0$ *not* an integer, we define

$$W^{s,p}(U) := \Big\{ f \in W^{[s],p}(U)\ ;\ \|f\|_{s,p} := \sum_{|\alpha|=[s]} \Big(\int_U \int_U \frac{|D^\alpha f(x) - D^\alpha f(y)|^p}{|x-y|^{d+(s-[s])p}} dxdy \Big)^{1/p} < \infty \Big\}.$$

We observe that when $U = \mathbb{R}^d$ and $p = 2$, the Sobolev space $W^{s,2}(\mathbb{R}^d)$ and the Bessel space $H_2^s(\mathbb{R}^d)$ coincide: $W^{s,2}(\mathbb{R}^d) = H_2^s(\mathbb{R}^d)$. Moreover, note that for any open set U, $W^{s,2}(U)$ corresponds to distributions f on U which are restrictions of some $f \in H_2^s(\mathbb{R}^d)$, see [57] for instance.

- 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.

- For any $R > 0$, denote by $\mathcal{B}_R$ the centred ball of $\mathbb{R}^d$ of radius R. For $\beta \in \mathbb{R}$, $p > 1$ and any ball $\mathcal{B}_R \subset \mathbb{R}^d$, the space $H_p^\beta(\mathcal{B}_R)$ is defined in Triebel [57, p.310], and corresponds to distributions f on $\mathcal{B}_R$ which are restrictions of $g \in H_p^\beta(\mathbb{R}^d)$. Then $H_{p,\mathrm{loc}}^\beta(\mathbb{R}^d)$ is the space of distributions f on $\mathbb{R}^d$ such that $f \in H_p^\beta(\mathcal{B}_R)$ for any $R > 0$. One defines similarly $W_{\mathrm{loc}}^{\beta,p}(\mathbb{R}^d)$, which is endowed with the distance induced by the family of seminorms on $W^{\beta,p}(\mathcal{B}_n)$, $n \in \mathbb{N}\setminus\{0\}$:

$$d(f,g) = \sum_{n=1}^\infty 2^{-n} \left(\|f-g\|_{W^{\beta,p}(\mathcal{B}_n)} \wedge 1\right). \tag{1.8}$$

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

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

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

$$g_{\sigma^2}(x) = \frac{1}{(2\pi\sigma^2)^{\frac{d}{2}}} e^{-\frac{|x-y|^2}{2\sigma^2}}. \tag{1.9}$$

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

$$\left\|\nabla e^{t\Delta}\right\|_{L^p\to L^p} \leq \frac{C}{\sqrt{t}}. \tag{1.10}$$

By explicit computations in the Fourier space, we get that

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

hence the inequality (1.10) extends to

$$\left\|(I-\Delta)^{\frac{\beta}{2}} e^{t\Delta}\right\|_{L^p\to L^p} \leq C\, t^{-\frac{\beta}{2}}. \tag{1.11}$$

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

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

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

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

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

$$\mathcal{N}_\delta(f) := \sup_{x\neq y\in\mathbb{R}^d} \frac{|f(x)-f(y)|}{|x-y|^\delta}. \tag{1.12}$$

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

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

2 Main results

The aim of this section is to present and discuss our main results. First we give in Section 2.1 the framework to establish our theorems, then state some general assumptions. Then we state our results about the convergence of particle systems to the PDE (1.1) and give rates, as well as the propagation of chaos and convergence to the associated McKean-Vlasov SDE (1.6).

2.1 Framework and assumptions

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

(i) $f_A(x) = x$, for $x \in [-A, A]$,

(ii) $f_A(x) = A$, for $x > A+1$ and $f_A(x) = -A$, for $x < -(A+1)$,

(iii) $\|f'_A\|_\infty \leq 1$ and $\|f''_A\|_\infty < \infty$.

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

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

Compared to the singular particle system (1.4), we introduce a mollifier that will be used both to regularise the particle system and its empirical measure. Let $V : \mathbb{R}^d \to \mathbb{R}_+$ be a smooth, rapidly decreasing probability density function, and assume further that V is even. For any $x \in \mathbb{R}^d$, define

$$V^N(x) := N^{d\alpha} V(N^\alpha x), \qquad \text{for some } \alpha \in [0,1]. \tag{2.2}$$

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

Let $T > 0$. For each $N \in \mathbb{N}$, we consider the following interacting particle system:

$$\begin{cases} dX_t^{i,N} = F_A\left(\frac{1}{N}\sum_{k=1}^N (K * V^N)(X_t^{i,N} - X_t^{k,N})\right) dt + \sqrt{2}\, dW_t^i, \quad t \leq T,\ 1 \leq i \leq N, \\ X_0^{i,N},\ 1 \leq i \leq N, \quad \text{are independent of } \{W^i,\ 1 \leq i \leq N\}, \end{cases} \tag{2.3}$$

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

Let us denote the empirical measure of N particles by

$$\mu_\cdot^N = \frac{1}{N}\sum_{i=1}^N \delta_{X_\cdot^{i,N}}\ , \tag{2.4}$$

and the mollified empirical measure by

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

The following properties of the kernel will be assumed:

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

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

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

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

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

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

Note that Assumption ($\mathbf{A}_{iii}^K$) is rather mild, and it will be easily verified that all the examples we have in mind satisfy it, as will be detailed in Section 7. Moreover, we provide in Section 7.1 a sufficient condition which is easier to check in the examples.

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

($\mathbf{A}$):

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

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

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

For instance, a sufficient condition for ($\mathbf{A}$) to hold is that particles are initially i.i.d. with a density that is smooth enough (see [17, Lemma 2.9] for a related result). The reader may also find interesting comments on this assumption in Remark 1.2 of [20].

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

($\mathbf{A}_\alpha$): The parameters α, β and $\boldsymbol{r}$ (which appear respectively in (2.2) and ($\mathbf{A}_{iii}^K$)) satisfy

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

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

$$\begin{cases} \partial_t \widetilde{u}(t,x) = \Delta \widetilde{u}(t,x) - \nabla \cdot \left(\widetilde{u}(t,x) F_A(K * \widetilde{u}(t,x))\right), \quad t > 0,\ x \in \mathbb{R}^d \\ \widetilde{u}(0,x) = u_0(x). \end{cases} \tag{2.5}$$

Although this is implicit, $\widetilde{u}$ actually depends on A. Note that if F_A is replaced by the identity function, one recovers (1.1). Solutions to (1.1) and (2.5) will be understood in the following sense:

Definition 2.1. *Given $u_0 \in L^1 \cap L^\infty(\mathbb{R}^d)$, $A > 0$ and $T > 0$, a function u on $[0,T] \times \mathbb{R}^d$ is said to be a mild solution to* (2.5) *on $[0,T]$ if*

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

(ii) u satisfies the integral equation

$$u_t = e^{t\Delta} u_0 - \int_0^t \nabla \cdot e^{(t-s)\Delta}(u_s F_A(K * u_s))\, ds, \quad 0 \leq t \leq T. \tag{2.6}$$

A function u on $[0,\infty) \times \mathbb{R}^d$ is said to be a global mild solution to (2.5) *if it is a mild solution to* (2.5) *on $[0,T]$ for all $T > 0$.*

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

$$u_t = e^{t\Delta} u_0 - \int_0^t \nabla \cdot e^{(t-s)\Delta}(u_s\ K * u_s)\, ds, \quad 0 \leq t \leq T. \tag{2.7}$$

Finally, as the kernel K may be singular, the PDEs we are interested in may only have local in time solutions (i.e. they may explode in finite time). For this reason, we will denote by T_{max} the maximal time of existence of a solution to (1.1) in the sense of Remark 2.2. 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 $T_{max} = \infty$.

2.2 Convergence of the particle system

Before we present the convergence result, we establish the (local) well-posedness of the PDE (1.1). This result is proven in Section 3.2.

Proposition 2.3. *Assume that the kernel $K : \mathbb{R}^d \to \mathbb{R}^d$ satisfies $(\mathbf{A}_i^K)$ and $(\mathbf{A}_{ii}^K)$ and that the initial condition satisfies $u_0 \in L^1 \cap L^\infty(\mathbb{R}^d)$. Then there exists $T > 0$ such that the PDE (1.1) admits a mild solution u in the sense of Remark 2.2 on $[0, T]$. In addition, this mild solution is unique.*

Now, let T_{max} be the maximal existence time of (1.1) and let $T < T_{max}$. For a local mild solution u on $[0, T]$, we will use the cut-off A_T defined by

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

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

Remark 2.4. *In Corollary 3.3, we observe that if T is small enough, then $\|u\|_{T,L^1\cap L^\infty(\mathbb{R}^d)}$ is controlled by $\|u_0\|_{L^1\cap L^\infty(\mathbb{R}^d)}$, so that the cut-off depends explicitly on the initial condition. If there exists a global solution, then in some cases (e.g. the parabolic-elliptic Keller-Segel PDE) , it is possible to define a global cut-off that depends explicitly on the initial condition (see [43] for more details).*

In the absence of information regarding the existence of a global solution, T will always denote a time smaller than the maximal existence time T_{max}.

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

Theorem 2.5. *Assume that the initial conditions $\{u_0^N\}_{N\in\mathbb{N}}$ satisfy $(\mathbf{A})$ and that the kernel K satisfies $(\mathbf{A}^K)$. Moreover, let $(\mathbf{A}_\alpha)$ hold true.*
Let T_{max} be the maximal existence time for (1.1) and fix $T \in (0, T_{max})$. Then let the dynamics of the particle system be given by (2.3) with A greater than A_T.
Then the sequence of mollified empirical measures $\{u_t^N,\ t \in [0, T]\}_{N\in\mathbb{N}}$ converges in probability, as $N \to \infty$, towards the unique mild solution u on $[0, T]$ of the PDE (1.1), in the following sense:

- $\forall \varphi \in L^2\left([0, T]; H_{\boldsymbol{r'}}^{-\beta}(\mathbb{R}^d)\right)$, $\int_0^T \langle u_t^N, \varphi_t\rangle_\beta\, dt \xrightarrow{\mathbb{P}} \int_0^T \langle u_t, \varphi_t\rangle_\beta\, dt$;
- *in the strong topology of* $\mathcal{C}\left([0, T]; W_{\boldsymbol{r},loc}^{\gamma}(\mathbb{R}^d)\right)$, *for* $\gamma < \beta$.

Here $\boldsymbol{r}$ is fixed in $(\mathbf{A}_{iii}^K)$ and $\boldsymbol{r'}$ is such that $\frac{1}{\boldsymbol{r}} + \frac{1}{\boldsymbol{r'}} = 1$.

Sections 4.1, 4.2 an 4.3 are devoted to the proof of Theorem 2.5, generalising the semigroup approach presented in [20].

2.3 Rate of convergence to the PDE

Now, we are interested in completing the above result by quantifying the obtained convergence. Our second main result is the following claim, whose proof is detailed in Section 5:

Theorem 2.6. *Assume that the initial conditions $\{u_0^N\}_{N\in\mathbb{N}}$ satisfy $(\mathbf{A})$ and that the kernel K satisfies $(\mathbf{A}^K)$. Moreover, let $(\mathbf{A}_\alpha)$ hold true. If $d = 1$, assume further that $\beta \in (\frac{1}{2} + \frac{1}{\boldsymbol{r}}, 1)$.*

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

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

where

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

Remark 2.7. *Unlike weakly interacting particle systems ($\alpha = 0$, which corresponds to a mean field interaction), we cannot expect here a $\sqrt{N}$ rate of convergence, which is usually the case for smoothly interacting particle systems [10]. This is due to the short range of interaction of the particles which is of order $N^{-\alpha}$. At the macroscopic level, we also observe that the distance between a finite measure μ and its regularisation $V^N * \mu$ tested against a Lipschitz function ϕ is of order $N^{-\alpha}$ too. Hence it is reasonable to expect no better than an N^α rate of convergence. See [42] for a thorough discussion on this matter.*

Remark 2.8. *Assuming that the kernel K is such that $\boldsymbol{\zeta}$ can be chosen equal or close to 1, the rate is really determined by $\alpha(\beta - \frac{d}{\boldsymbol{r}})$ and $\frac{1}{2}(1 - \alpha(d + [2 \vee d]))$. Then the supremum of $\alpha(\beta - \frac{d}{\boldsymbol{r}})$ under the constraint that $\alpha < (d + 2\beta + 2d(\frac{1}{2} - \frac{1}{\boldsymbol{r}}))^{-1}$ (see $(\mathbf{A}_\alpha)$) and $\beta < 1$ (see $(\mathbf{A}^K_{iii})$) is equal to*

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

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

So when $d \geq 2$, the best possible rate is almost α, which in view of the discussion of Remark 2.7, is optimal (what might not be optimal is the constraint $(\mathbf{A}_\alpha)$ on α).

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

In Section 5.2, using an interpolation inequality between the results of Theorem 2.5 and Theorem 2.6, we obtain the following rate of convergence with respect to Sobolev norms:

Corollary 2.9. *Let the same assumptions as in Theorem 2.6 hold. Let ϱ be as in Theorem 2.6. Then, for any $\varepsilon > 0$ and any $m \geq 1$, there exists a constant $C > 0$ such that for all $N \in \mathbb{N}^*$,*

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

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

Remark 2.10. *It is clear that $\gamma < \beta$. It will also be important (in particular for the propagation of chaos in the next section) to ensure that $\gamma > \frac{d}{\boldsymbol{r}-\delta}$ so as to have an embedding in a space of Hölder continuous functions. This is indeed the case if δ is chosen small enough, see condition* (5.24).

The following corollary is a direct consequence of Theorem 2.6, Corollary 2.9 and of Borel-Cantelli's lemma (see e.g. Lemma 2.1 in [33]):

Corollary 2.11. *Let the same assumptions as in Theorem 2.6 hold. Let ϱ be as in Theorem 2.6, and γ, δ be as in Corollary 2.9. Let $m \geq 1$ and assume further that*

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

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

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

Remark 2.12. *By a classical embedding recalled in Section 3.1, the results of Corollaries 2.9 and 2.11 imply the same rates in η-Hölder norm, with $\eta = \gamma - \frac{d}{\mathbf{r}-\delta}$, provided that this quantity is positive (see condition* (5.24)*).*

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

$$\|\mu - \nu\|_0 = \sup \left\{ \int_{\mathbb{R}^d} \phi \, d(\mu - \nu)\,;\ \phi \text{ Lipschitz with } \|\phi\|_{L^\infty(\mathbb{R}^d)} \le 1 \text{ and } \|\phi\|_{\text{Lip}} \le 1 \right\}. \tag{2.9}$$

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

Corollary 2.13. *Let ϱ be as in Theorem 2.6 and $m \ge 1$. Let the same assumptions as in Corollary 2.11 hold. Then for any $\varepsilon \in (0, \varrho)$, there exists $C > 0$ such that, for any $N \in \mathbb{N}^*$,*

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

As in Corollary 2.11, this gives an almost sure rate of convergence for $\sup_{t\in[0,T]} \|\mu_t^N - u_t\|_0$.

The proof is given in Section 5.3.

Remark 2.14. *Observe that the previous result implies the convergence in law of μ_t^N to u_t for a fixed $t \in (0, T_{max})$, which is equivalent to saying that the law of $(X_t^1, \dots, X_t^N)$ is u_t-chaotic (in the sense of [54, Def. 2.1]).*

2.4 Propagation of chaos

Finally, our third objective is to establish the well-posedness of the McKean-Vlasov equation (1.6) on any time interval $[0, T]$ with $T < T_{max}$, and to obtain the propagation of chaos property of the particle system (2.3). To achieve this, we will prove the convergence of the empirical measure μ^N towards the law of the weak solution to (1.6) with a cut-off and we will then lift the cut-off by choosing A large enough. First, we need to ensure (1.6) admits a unique weak solution. For that purpose, we will solve the associated nonlinear martingale problem. By classical arguments, one can then pass from a solution to this martingale problem to the existence of a weak solution to (1.6)

We consider the following nonlinear martingale problem related to (1.6):

Definition 2.15. *Consider the canonical space $\mathcal{C}([0,T];\mathbb{R}^d)$ equipped with its canonical filtration. Let $\mathbb{Q}$ be a probability measure on the canonical space and denote by $\mathbb{Q}_t$ its one-dimensional time marginals. We say that $\mathbb{Q}$ solves the nonlinear martingale problem $(\mathcal{MP})$ if:*

(i) $\mathbb{Q}_0 = u_0$*;*

(ii) For any $t \in (0, T]$, $\mathbb{Q}_t$ has a density q_t w.r.t. Lebesgue measure on $\mathbb{R}^d$. In addition, it satisfies $q \in \mathcal{C}([0,T]; L^1 \cap L^\infty(\mathbb{R}^d))$;

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

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

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

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

Proposition 2.16. *Let $T < T_{max}$. Assume that u_0 is a probability density function belonging to $L^\infty(\mathbb{R}^d)$ and that the kernel K satisfies $(\mathbf{A}^K)$. Then, the martingale problem $(\mathcal{MP})$ admits a unique solution in the sense of Definition 2.15.*

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

Theorem 2.17. *Let the hypotheses of Theorem 2.5 hold and assume further that the family of random variables $\{X_0^i,\ i \in \mathbb{N}\}$ is identically distributed. Then, the empirical measure $\mu_\cdot^N$ (defined in (2.4) on $\mathcal{C}([0,T];\mathbb{R}^d)$) converges in law towards the unique weak solution of (1.6).*

2.5 Comparison with previous works

The approximation of nonlinear Fokker-Planck equations of the type (1.1) has been widely covered in the literature, but such convergence of a mollified empirical measure in strong topologies, for singular kernels, are quite new. In this sense, we extend and improve previous results of [20]. The extension to several singular models, in particular to singular Riesz kernels and to attractive kernels (e.g. Keller-Segel) is completely new.

Concerning the rate, we consider the same kind of PDE (1.1) as Jabin and Wang [29] and Bresch et al. [6], with a different set of assumptions on the interaction kernel. In [29], the authors were interested in the convergence, when $N \to \infty$, of the joint law of k fixed particles at a time t towards $u_t^{\otimes k}$, where u solves (1.1). They worked on a periodic domain $\Pi \subset \mathbb{R}^d$. Under the assumptions that the particles are well-defined, that $u \in L^\infty((0,T);W^{2,\infty}(\Pi))$, that $K \in W^{-1,\infty}$ and $\nabla \cdot K \in W^{-1,\infty}$ (which includes the Biot-Savart kernel, but does not cover the Coulomb and Keller-Segel kernel), they proved using new techniques of relative entropy the above convergence in $L^\infty((0,T);L^1(\Pi^k))$. Moreover, the rate in [29] is $N^{-1/2}$ which is considered as optimal. In another work, Bresch et al. [6] combined the relative entropy with the modulated energy of Serfaty [50], to deal with attractive, gradient-flow kernels including $2d$-Keller-Segel model.

Let us mention several improvements compared to these papers. More general attractive or repulsive kernels can be dealt with here (e.g. Coulomb kernels, and even more singular Riesz kernels with potential (1.2)). However, note that in [6] they can treat repulsive Riesz kernels with $s \in [d-1, d)$. Furthermore, our approximation enables us to obtain a convergence rate in Sobolev topology. Milder assumptions on the PDE solution are imposed here, and in particular it is not required to work on a bounded domain (although we work in $\mathbb{R}^d$, all the results can be easily adapted for the periodical domain case).
More specifically, let us discuss the case of the $2d$ Keller-Segel model (see kernel (1.3)). We recall that the PDE has a global solution whenever $\chi < 8\pi$, and explodes in finite time otherwise (see [40]). The rate obtained in [6] depends on $8\pi - \chi$ provided that the latter is positive, while we get a rate for any value of χ even if the PDE explodes in finite time (in that case one works on $[0,T]$ for any $T < T_{\max}$). Moreover, our rate is independent of χ and we do not assume that particles are initially i.i.d.
However, recall from Remark 2.7 that we cannot get a better rate than $N^{-\alpha}$, and the condition $(\mathbf{A}_\alpha)$ that we impose to get convergence leads to $\alpha < \frac{1}{2}$. Moreover, our approach does not cover the case of vanishing diffusion which for some kernels could be done by [29].

Recently, some quantitative results have also been obtained for the Landau equation by Fournier and Hauray [21]. Although this work has some specific difficulties that we do not encounter here (mostly related to a nonlinear diffusion term which may vanish), it is interesting to notice that the kernel used in the drift is of order $|\nabla b(v)| \approx |v|^\gamma$, with $\gamma \in (-2,0)$ and $d = 3$. The authors obtain propagation of chaos for a mollified system when $\gamma \in (-2,0)$ and a rate for $\gamma \in (-1,0)$. We may work with the same range of parameter in Theorem 2.17 for the propagation of chaos ($s \equiv -\gamma \in (0,2)$ if s is the parameter of a Riesz kernel, see Section 7.3) and for the convergence rate of Theorem 2.6 we are still able to consider the whole range of parameter $s \in (0,2)$.

Finally, we obtained new well-posedness results for a large class of McKean-Vlasov SDEs given by (1.6). Let us emphasize on the Keller-Segel model for which the kernel is given in (1.3) (see also Section 7.4). This model has attracted a lot of attention lately, for instance Fournier and Jourdain [22] proved the well-posedness of the associated McKean-Vlasov SDE for a value of the sensitivity parameter $\chi < 2\pi$ in dimension 2. They also proved tightness and consistency result for the associated particle system (we 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. This behaviour depends on χ and the dimension. In particular in dimension 2, we get the global well-posedness of (1.6) whenever $\chi < 8\pi$, and local well-posedness of (1.6) when $\chi \geq 8\pi$. In both cases, we obtain the propagation of chaos for the moderately interacting particle system.

Another singular drift given in the literature is the one in [46], where (roughly speaking) the interaction kernel is only integrable in the $L^q([0,T];L^p(\mathbb{R}^d))$ space where $\frac{d}{p}+\frac{2}{q}<1$. A typical example of such interaction given in [46] is of order $\frac{1}{|x|^r}$ where $r \in [0,1)$. A kernel with this kind of irregularity will satisfy $(\mathbf{A}^K)$. Actually, we can treat the more singular cases of kernels K of the order $\frac{1}{|x|^{s+1}}$ for $s \in (0, d-1)$, in any dimension. However, in our framework one needs to be more flexible with the initial condition and our drift is not time dependent (though our techniques could support time dependence in the drift up to the order of magnitude that could allow us to use the Grönwall lemma when needed, we preferred not to introduce such a technicality in our computations).

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

We start this section with some classical embeddings that will be used throughout the article. Then in Subsection 3.2, we derive some general inequalities and prove Proposition 2.3. In Subsection 3.3, we prove that the PDE with cut-off (2.5) can have at most one mild solution.

3.1 Some classical embeddings

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

$$\|f\|_{L^{\boldsymbol{r}}\cap L^\infty(\mathbb{R}^d)} \leq C\|f\|_{\beta,\boldsymbol{r}} \quad \text{and} \quad \|f\|_{\mathcal{C}^{\beta-\frac{d}{\boldsymbol{r}}}(\mathbb{R}^d)} \leq C'\|f\|_{\beta,\boldsymbol{r}}, \quad \forall f \in H^\beta_{\boldsymbol{r}}(\mathbb{R}^d). \tag{3.1}$$

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

$$\|f\|_{L^1\cap L^\infty(\mathbb{R}^d)} \leq C_{d,\beta,\boldsymbol{r}}\|f\|_{L^1\cap H^\beta_{\boldsymbol{r}}(\mathbb{R}^d)}, \quad \forall f \in L^1 \cap H^\beta_{\boldsymbol{r}}(\mathbb{R}^d). \tag{3.2}$$

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

3.2 Properties of mild solutions of the PDE

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

Lemma 3.1. *Let K be satisfying the Assumptions $(\mathbf{A}^K_i)$ and $(\mathbf{A}^K_{ii})$. There exists $C_{K,d} > 0$ (which depends on K and d only) such that for any $f \in L^1 \cap L^\infty(\mathbb{R}^d)$,*

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

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

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

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

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

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

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

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

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

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

$$\|B(u,v)(t)\|_{L^1\cap L^\infty(\mathbb{R}^d)} \leq C\sqrt{t}\ \|u\|_{\mathcal{X}}\|v\|_{\mathcal{X}}, \tag{3.3}$$

which suffices to prove the continuity of B.

First, using the property (1.10) of the Gaussian kernel with $p = \infty$, observe that for any $t \in [0,T]$,

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

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

$$\begin{aligned} \|B(u,v)(t)\|_{L^\infty(\mathbb{R}^d)} &\leq \int_0^t \frac{C}{\sqrt{t-s}} \|K * v_s\|_{L^\infty(\mathbb{R}^d)}\ \|u_s\|_{L^\infty(\mathbb{R}^d)}\ ds \\ &\leq C\sqrt{t}\ \|u\|_{\mathcal{X}}\|v\|_{\mathcal{X}}. \end{aligned} \tag{3.4}$$

On other hand we have, using similarly the property (1.10) of the Gaussian kernel with $p = 1$ and Lemma 3.1, that

$$\begin{aligned} \|B(u,v)(t)\|_{L^1(\mathbb{R}^d)} &\leq \int_0^t \frac{C}{\sqrt{t-s}} \|u_s\ K * v_s\|_{L^1(\mathbb{R}^d)}\ ds \\ &\leq \int_0^t \frac{C}{\sqrt{t-s}} \|u_s\|_{L^1(\mathbb{R}^d)} \|K * v_s\|_{L^\infty(\mathbb{R}^d)}\ ds \\ &\leq C\sqrt{t}\ \|u\|_{\mathcal{X}}\|v\|_{\mathcal{X}}. \end{aligned} \tag{3.5}$$

Therefore, combining Equations (3.4), (3.5) one obtains (3.3). □

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

Proof of Proposition 2.3. 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 3.2, such a local mild solution is obtained by a standard contraction argument (Banach fixed-point Theorem).

We will prove the uniqueness in the (slightly more complicated) case of the cut-off PDE (2.5), for any value of the cut-off A (see Proposition 3.5). 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 3.1 implies that for $i = 1, 2$,

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

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

Corollary 3.3. *Assume that* $(\mathbf{A}_i^K)$ *and* $(\mathbf{A}_{ii}^K)$ *hold. Let* C *be the constant that appears in* (3.3)*. Then for* $T > 0$ *such that*

$$4C\sqrt{T}\ \|u_0\|_{L^1\cap L^\infty(\mathbb{R}^d)} < 1, \tag{3.6}$$

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

$$\|u\|_{T,L^1\cap L^\infty(\mathbb{R}^d)} \leq \frac{1-\sqrt{1-4C\sqrt{T}\ \|u_0\|_{L^1\cap L^\infty(\mathbb{R}^d)}}}{2C\sqrt{T}}. \tag{3.7}$$

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

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

Proof. We rely on the bound (3.3), in order to get that for $t \leq T$,

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

Then by a standard argument (see e.g. Lemma 2.3 in [40]), choosing $T > 0$ which satisfies (3.6) ensures that (3.7) holds true. □

Remark 3.4. *Since* u^N *is a probability density function, we expect that its limit, whenever it exists, stays nonnegative and has mass* 1*.*
Indeed, assume that u *is a local mild solution on* $[0,T]$ *to* (1.1)*. Then in view of Definition 2.1-i) and by the inequality* $\|e^{(t-s)\Delta}\|_{L^p\to L^p} \leq C$*,*

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

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

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

Hence by integration-by-parts

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

and it follows that

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

Moreover, when the initial data is such that $u_0 \geq 0$ *and* $u_0 \not\equiv 0$*, then by an argument similar to [40, Prop. 2.7] (although using* $(\mathbf{A}_{iii}^K)$ *instead of the Poisson kernel),* u *is such that* $u_t \geq 0$*, for* $(t,x) \in (0,T)\times\mathbb{R}^d$ *(we can also deduce this fact later from the convergence of* u^N *to* u*). Hence the mass is preserved.*

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

In this section, we consider the cut-off PDE (2.5) and its mild solution from Definition 2.1. Here, F_A is given in (2.1), but we denote it simply by F for the sake of readability.

Note that due to the boundedness of the reaction term in (2.5), any mild solution will always be global. This global solution will be rigorously obtained as the limit of the particle system (2.3). Thus we only consider global mild solutions when it comes to the PDE (2.5).

Proposition 3.5. *Assume that K satisfies Assumptions $(\mathbf{A}_i^K)$ and $(\mathbf{A}_{ii}^K)$. Let $u_0 \in L^1 \cap L^\infty(\mathbb{R}^d)$. Then for any $A > 0$ and F defined in* (2.1)*, there is at most one mild solution to the cut-off PDE* (2.5)*.*

Proof. Assume there are two mild solutions u^1 and u^2 to (2.5). Then,

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

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

$$\begin{aligned} \|u_t^1 - u_t^2\|_{L^1(\mathbb{R}^d)} + \|u_t^1 - u_t^2\|_{L^\infty(\mathbb{R}^d)} &\le C \int_0^t \frac{1}{\sqrt{t-s}} \left(\|u_s^1 - u_s^2\|_{L^1(\mathbb{R}^d)} + \|u_s^2 \; K * (u_s^1 - u_s^2)\|_{L^1(\mathbb{R}^d)} \right) \, ds \\ &\quad + C \int_0^t \frac{1}{\sqrt{t-s}} \left(\|u_s^1 - u_s^2\|_{L^\infty(\mathbb{R}^d)} + \|u_s^2 \; K * (u_s^1 - u_s^2)\|_{L^\infty(\mathbb{R}^d)} \right) \, ds \\ &\le C\sqrt{t} \, \|u^1 - u^2\|_{t, L^1 \cap L^\infty(\mathbb{R}^d)} \\ &\quad + C \int_0^t \frac{\|u_s^2\|_{L^\infty(\mathbb{R}^d)} + \|u_s^2\|_{L^1(\mathbb{R}^d)}}{\sqrt{t-s}} \, \|K * \left(u_s^1 - u_s^2\right)\|_{L^\infty(\mathbb{R}^d)} \, ds. \end{aligned} \tag{3.8}$$

In view of Lemma 3.1, one has $\|K * (u_s^1 - u_s^2)\|_{L^\infty(\mathbb{R}^d)} \le C \|u_s^1 - u_s^2\|_{L^1 \cap L^\infty(\mathbb{R}^d)}$, thus plugging this upper bound in (3.8) gives

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

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

4 Convergence of the mollified empirical measure

We now proceed to the proof of Theorem 2.5, which is split between Sections 4.1, 4.2 and 4.3. In Section 4.1, we present the two functional spaces and the compact embedding between them that will be used in our tightness argument. In Section 4.2, we detail the steps of the proof and state the most important intermediate results that are used to prove the main theorem, and refer to Section 4.3 for the proof of these results.

In this section, it will always be supposed that $T < T_{max}$.

4.1 Preliminary remark: A compact embedding

In Theorem 2.5, we aim to establish the convergence of the particle system in the following space:

$$\mathcal{Y} := L_w^2 \left([0, T]; \, H_{\boldsymbol{r}}^\beta(\mathbb{R}^d)\right) \cap \mathcal{C}\left([0, T]; \, W_{\boldsymbol{r}, \mathrm{loc}}^\gamma(\mathbb{R}^d)\right), \quad \frac{d}{\boldsymbol{r}} < \gamma < \beta, \tag{4.1}$$

where L_w^2 denotes the $L^2\left([0, T]; \, H_{\boldsymbol{r}}^\beta(\mathbb{R}^d)\right)$ space endowed with the weak topology. It will be convenient to work within a compactly embedded subspace of $\mathcal{Y}$ in order to establish tightness. Namely, we will use the space

$$\mathcal{Y}_0 := L^{q_0}\left([0, T]; \; H_{\boldsymbol{r}}^\beta(\mathbb{R}^d)\right) \cap W^{\eta, q_1}\left([0, T]; \; H_2^{-2}(\mathbb{R}^d)\right), \tag{4.2}$$

where $\eta\vartheta \ge \frac{1-\vartheta}{q_0} + \frac{\vartheta}{q_1}$ and $\vartheta < \min \frac{\beta - \gamma}{\beta + 2 + \frac{d}{2} - \frac{d}{r}}$.

Notice first that $L^{q_0}\big([0,T];\ H^{\beta}_{\boldsymbol{r}}(\mathbb{R}^d)\big)$ is a reflexive Banach space for $1<\boldsymbol{r}<\infty$ (see [57, p.198-199]), hence by the Banach-Alaoglu theorem, it is compactly embedded in itself endowed with the weak topology. If q_0 is chosen larger than 2, this ensures that $L^{q_0}\big([0,T];\ H^{\beta}_{\boldsymbol{r}}(\mathbb{R}^d)\big)$ is compactly embedded into $L^2_w\left([0,T];\, H^{\beta}_{\boldsymbol{r}}(\mathbb{R}^d)\right)$.

Then, let us explain why $\mathcal{Y}_0$ is compactly embedded into $\mathcal{C}\left([0,T];\, W^{\gamma}_{\boldsymbol{r},\mathrm{loc}}(\mathbb{R}^d)\right)$. We aim to apply [51, Corollary 9] to the spaces $X=W^{\beta,\boldsymbol{r}}(\mathcal{B}_R)$, $B=W^{\gamma,\boldsymbol{r}}(\mathcal{B}_R)$ and $Y=H^{-2}_2(\mathcal{B}_R)$. Note that all these spaces are Besov spaces (see [57, p.310]) and as such, we have the following embeddings:

- by [48, Theorem 2, p.82], $X\subset B$ and the embedding is compact;
- B is continuously embedded into $L^{\boldsymbol{r}}(\mathcal{B}_R)$, which is itself continuously embedded into $L^2(\mathcal{B}_R)$ since we assumed in $(\mathbf{A}^K_{iii})$ that $\boldsymbol{r}>2$. Then $L^2(\mathcal{B}_R)$ is continuously embedded into Y.

Hence X is compactly embedded into Y. Besides, Lemma 12 of [51] states that for all $\vartheta<\min\left\{\frac{\beta-\gamma}{\beta+2},\frac{\beta-\gamma}{\beta+2+\frac{d}{2}-\frac{d}{\boldsymbol{r}}}\right\}=\frac{\beta-\gamma}{\beta+2+\frac{d}{2}-\frac{d}{\boldsymbol{r}}}$ (since $\boldsymbol{r}>2$),

$$\|f\|_B\le C\|f\|_X^{1-\vartheta}\|f\|_Y^{\vartheta}.$$

Hence all the hypotheses of [51, Corollary 9] are met, and it comes that for all $R>0$,

$$L^{q_0}\big([0,T];\ W^{\beta,\boldsymbol{r}}(\mathcal{B}_R)\big)\cap W^{\eta,q_1}\big([0,T];\ H^{-2}_2(\mathcal{B}_R)\big)$$

is compactly embedded into

$$\mathcal{C}\left([0,T];\, W^{\gamma,\boldsymbol{r}}(\mathcal{B}_R)\right),$$

provided that $\eta\vartheta\ge\frac{1-\vartheta}{q_0}+\frac{\vartheta}{q_1}$ and $\vartheta<\frac{\beta-\gamma}{\beta+2+\frac{d}{2}-\frac{d}{\boldsymbol{r}}}$. Besides, recalling the continuous embedding $H^{\beta}_{\boldsymbol{r}}(\mathcal{B}_R)\subset W^{\beta,\boldsymbol{r}}(\mathcal{B}_R)$ (which holds since $\boldsymbol{r}>2$, see Section 3.1), we obtain that for all $R>0$,

$$L^{q_0}\big([0,T];\ H^{\beta}_{\boldsymbol{r}}(\mathcal{B}_R)\big)\cap W^{\eta,q_1}\big([0,T];\ H^{-2}_2(\mathcal{B}_R)\big)$$

is compactly embedded into

$$\mathcal{C}\left([0,T];\, W^{\gamma,\boldsymbol{r}}(\mathcal{B}_R)\right),$$

with the same constraints on η,ϑ,q_0 and q_1 as above.

Finally, if $(f_k)_{k\in\mathbb{N}}$ is a bounded family in $\mathcal{Y}_0$, then for each $n\in\mathbb{N}\setminus\{0\}$, consider $f^{(n)}_k$ the restriction (in the sense of distributions) of f_k to $L^{q_0}\big([0,T];\ H^{\beta}_{\boldsymbol{r}}(\mathcal{B}_n)\big)\cap W^{\eta,q_1}\big([0,T];\ H^{-2}_2(\mathcal{B}_n)\big)$. Up to taking a subsequence, $(f^{(n)}_k)_{k\in\mathbb{N}}$ converges in $\mathcal{C}\left([0,T];\, W^{\gamma,\boldsymbol{r}}(\mathcal{B}_n)\right)$. Hence by a diagonal argument, we can find a subsequence φ such that for each n, $(f^{n}_{\varphi(k)})_{k\in\mathbb{N}}$ converges in $\mathcal{C}\left([0,T];\, W^{\gamma,\boldsymbol{r}}(\mathcal{B}_n)\right)$. Now recalling the distance on $W^{\gamma,\boldsymbol{r}}_{\mathrm{loc}}(\mathbb{R}^d)$ which is given by (1.8), it comes that $(f_k)_{k\in\mathbb{N}}$ converges in $\mathcal{C}\big([0,T];\, W^{\gamma,\boldsymbol{r}}_{\mathrm{loc}}(\mathbb{R}^d)\big)$.

Hence $\mathcal{Y}_0$ is compactly embedded into $\mathcal{Y}$.

4.2 Proof of Theorem 2.5: Strategy of the proof

Step 1. First, it will be established in Subsection 4.3.1 that for any test function φ, u^N satisfies the following equation

$$\begin{aligned}\langle u^N_t,\varphi\rangle=&\langle u^N_0,\varphi\rangle+\int_0^t\langle S^N_s,\nabla(V^N*\varphi)\cdot F_A\big(K*u^N_s\big)\rangle\,ds\\&+\frac{1}{N}\sum_{i=1}^N\int_0^t\nabla(V^N*\varphi)(X^{i,N}_s)\cdot dW^i_s+\int_0^t\langle u^N_s,\Delta\varphi\rangle\,ds.\end{aligned}\tag{4.3}$$

In Subsection 4.3.2, we will then prove that $\{u^N\}$ is tight in the space $\mathcal{Y}$. It suffices for that to prove the boundedness of the sequence in $\mathcal{Y}_0$, since this space is compactly embedded in $\mathcal{Y}$ (see previous section).

By Prokhorov's theorem, the tightness of $\{u^N\}$ implies that it is relatively compact in a sense that we precise now (because Prokhorov's theorem applies only in Polish spaces, and L^2_w is not metrizable). Indeed, we will make a slight abuse of language in the following when we say that u^N converges in law (resp. in probability, or almost surely) in $\mathcal{Y}$: it will be understood that for any $\varphi \in L^2\left([0,T]; H^{-\beta}_{\boldsymbol{r}'}(\mathbb{R}^d)\right)$, $\langle u^N, \varphi\rangle$ converges in law (resp. in probability or a.s.), and of course that u^N converges in law (resp. in probability or a.s.) in $\mathcal{C}\left([0,T]; W^{\gamma}_{\boldsymbol{r},\mathrm{loc}}(\mathbb{R}^d)\right)$.

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

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

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

Remark 4.2. *For each N and $t \in [0,T]$, the definition of u^N_t yields that $u^N_t \in L^\infty(\mathbb{R}^d)$ a.s., and since u^N_t is a probability density function, it is also in $L^1(\mathbb{R}^d)$ a.s. Hence by interpolation, $u^N_t \in \bigcap_{p=1}^\infty L^p(\mathbb{R}^d)$ a.s.*
Now by Fatou's lemma, one gets that a.s., for almost all t, $\widetilde{u}_t \in L^1(\mathbb{R}^d)$. Moreover, by Sobolev embedding, $\widetilde{u}_t \in H^\beta_{\boldsymbol{r}}(\mathbb{R}^d)$ with $\beta > d/\boldsymbol{r}$ implies that $\widetilde{u}_t \in L^\infty(\mathbb{R}^d)$ (see embedding (3.1)). Hence by interpolation, $\widetilde{u}_t \in \bigcap_{p=1}^\infty L^p(\mathbb{R}^d)$ a.s.

In Subsection 4.3.3, using (4.3) and the convergence result of Proposition 4.1, we will prove that the following equality holds for all $t > 0$ and all $\varphi \in \mathcal{C}^\infty_c(\mathbb{R}^d)$,

$$\langle \widetilde{u}_t, \varphi\rangle = \langle u_0, \varphi\rangle + \int_0^t \int_{\mathbb{R}^d} \widetilde{u}_s \, \nabla\varphi(x) \cdot F_A(K * \widetilde{u}_s)(x) \, dx \, ds \, + \int_0^t \langle \widetilde{u}_s, \Delta\varphi\rangle \, ds. \tag{4.4}$$

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

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

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

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

The next proposition will be useful in identifying $\widetilde{u}_s$ as a mild solution to (2.5). Its proof is given in Subsection 4.3.4.

Proposition 4.3. *Let $T > 0$ chosen arbitrarily and let $\widetilde{u}$ be as in Proposition 4.1. Then, $\widetilde{u} \in \mathcal{C}([0,T]; L^1 \cap L^\infty(\mathbb{R}^d))$. As a consequence, there exists a global mild solution to the cut-off PDE (2.5).*

It follows from this result and Proposition 3.5 that $\widetilde{u}$ is the unique global mild solution of (2.5).

Step 2. In Step 1 we have obtained that on the probability space $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{P}})$, u^N converges almost surely in $\mathcal{Y}$ to $\widetilde{u}$, which satisfies the mild formulation of (2.5).

Observe now that as $T < T_{max}$ and if the cut-off value A is chosen larger than A_T (which was defined in (2.8)), it comes that a mild solution $\widetilde{u}$ to (2.5) is also a mild solution to (1.1). Hence by the uniqueness in (1.1) (Proposition 2.3 again), $\widetilde{u}$ (hereafter denoted by u) is the unique mild solution to (1.1) on $[0,T]$, as claimed in Theorem 2.5.

Let us now come back to the original probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We have obtained that every subsequence of $\{u^N\}$ has a further subsequence that converges in law to u, the unique mild solution of (1.1), in $\mathcal{Y}$. Hence u^N converges in law to u, and since u is non-random, the convergence also happens in probability for the topology of $\mathcal{Y}$, which concludes the proof of Theorem 2.5.

4.3 Proof of Theorem 2.5: Proofs of intermediate results

Recall that $V : \mathbb{R}^d \to \mathbb{R}_+$ is an even smooth probability density function and that V^N is defined by Equation (2.2), that $\{X^{i,N}\}$ is the particle system defined by (2.3) with cutoff F_A given in (2.1). In this section, we use again the notation F instead of F_A, for the sake of readability.

4.3.1 Equation satisfied by the regularised empirical measure: Proof of Equality (4.3)

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 (2.3) as

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

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

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

Notice that

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

and

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

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

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

Now for $\varphi \in \mathcal{D}(\mathbb{R}^d)$, Equation (4.6) implies the desired Equality (4.3).

For further use in Section 4.3.2, we also get the following mild form

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

Finally, developing the scalar product, one has

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

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

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

4.3.2 Tightness of u^N: Proof of Proposition 4.1

Let us now prove the tightness of $\{u^N\}$ in the space $\mathcal{Y}$ defined in (4.1). In view of Section 4.1, this will be achieved by proving boundedness in the space $\mathcal{Y}_0$ that is compactly embedded in $\mathcal{Y}$. Recall that $\mathcal{Y}_0$ was defined in (4.2).
The norm of $W^{\eta,q_1}\big([0,T]\ ;\ H_2^{-2}(\mathbb{R}^d)\big)$, for $\eta\in(0,1)$ and $q_1>1$, is equivalent to the following norm (see [57, p.323]):

$$\|f\|^{q_1}_{W^{\eta,q_1}([0,T];H_2^{-2}(\mathbb{R}^d))} \sim \|f\|^{q_1}_{L^{q_1}([0,T];H_2^{-2}(\mathbb{R}^d))} + \int_0^T\int_0^T \frac{\|f_t - f_s\|^{q_1}_{-2,2}}{|t-s|^{1+q_1\eta}} ds dt.$$

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

Proposition 4.4. *Let the assumptions of Theorem 2.5 hold. Let $q\geq 1$. Then there exists a constant $C>0$ such that, for all $t\in[0,T]$ and $N\in\mathbb{N}$, it holds:*

$$\mathbb{E}\left[\|u_t^N\|^q_{\beta,\boldsymbol{r}}\right] \leq C.$$

Proposition 4.5. *Let the assumptions of Theorem 2.5 hold. Let $\eta\in(0,\frac{1}{2})$ and $q\geq 1$. There exists a constant $C>0$ such that, for any $N\in\mathbb{N}$, it holds:*

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

The proofs of these two results are similar to the proofs of Propositions 2.1 and 2.2 in [20] (the kernel plays no role here). We present them in Appendix A.2 and note that this is where the restriction ($\mathbf{A}_\alpha$) on α appears.

The Chebyshev inequality then ensures that

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

Thus by Proposition 4.4 and Proposition 4.5, we obtain

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

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

4.3.3 Weak convergence to the PDE solution: Proof of Equality (4.4)

First recall from Assumption (**A**) on the initial condition that

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

In view of Proposition 4.1, we have that $u^N \to \widetilde{u}$ a.s. in the space $\mathcal{Y}$ which was defined in Equation (4.1). It is clear that this result implies that we can pass to the limit in (4.3):

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

and

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

To conclude that $\widetilde{u}$ satisfies Equation (4.4), it remains to prove:

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

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

Proof. First, let $\epsilon > 0$ and let $R > 0$ be large enough to ensure, thanks to Assumption ($\mathbf{A}_{ii}^K$), that

$$\int_{\mathbb{R}^d} \mathbb{1}_{\mathcal{B}_R^c}(y) \, |K(y)|^{\boldsymbol{q}} \, dy \le \epsilon^{\boldsymbol{q}}, \tag{4.9}$$

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

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

It follows, using the Cauchy-Schwarz inequality in the second inequality and the bound (4.9) in the third, that

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

Observe that by the Sobolev embedding Theorem, $L^2([0,T], L^1 \cap H_{\boldsymbol{r}}^\beta(\mathbb{R}^d))$ is continuously embedded into $L^2([0,T], L^1 \cap L^\infty(\mathbb{R}^d))$ (see (3.1)), which is itself continuously embedded into $L^2([0,T], L^q(\mathbb{R}^d))$, for any $q \in [1, +\infty)$. Hence, $s \mapsto \|u_s^N - \widetilde{u}_s\|_{L^{\boldsymbol{q}'}(\mathbb{R}^d)}$ is a.s. bounded in L^2 (see also Remark 4.2) and it follows from the previous inequality that almost surely,

$$\lim_{N\to\infty} \int_0^t \int_{\mathbb{R}^d} \widetilde{u}_s(x) \; \nabla\varphi(x) \cdot F(K * u_s^N)(x) \, dx \, ds = \int_0^t \int_{\mathbb{R}^d} \widetilde{u}_s(x) \; \nabla\varphi(x) \cdot F(K * \widetilde{u}_s)(x) \, dx \, ds. \tag{4.10}$$

Next, observe that

$$\begin{aligned}&\left|\left\langle \mu_s^N, \nabla(V^N * \varphi) \cdot F(K * u_s^N)\right\rangle - \left\langle u_s^N, \nabla(V^N * \varphi) \cdot F(K * u_s^N)\right\rangle\right| \\ &\quad \le \sup_{x\in\mathbb{R}^d} \left|\nabla(V^N * \varphi)(x) \cdot F(K * u_s^N(x)) - \big(\nabla(V^N * \varphi) \cdot F(K * u_s^N)\big) * V^N(x)\right|. \end{aligned} \tag{4.11}$$

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

$$\begin{aligned}
&\Big|\nabla(V^N * \varphi)(x) \cdot F(K * u_s^N)(x) - \big(\nabla(V^N * \varphi) \cdot F(K * u_s^N)\big) * V^N(x)\Big| \\
&\leq \int_{\mathbb{R}^2} V(y)\ \left|\nabla(V^N * \varphi)(x)\right|\ \left|F(K * u_s^N)(x) - F(K * u_s^N)\big(x - \tfrac{y}{N^\alpha}\big)\right| dy \\
&\quad + \int_{\mathbb{R}^d} V(y)\ \left|\nabla(V^N * \varphi)(x) - \nabla(V^N * \varphi)\big(x - \tfrac{y}{N^\alpha}\big)\right| \left|F(K * u_s^N)\big(x - \tfrac{y}{N^\alpha}\big)\right| dy \\
&\leq C \int_{\mathbb{R}^d} V(y)\ \left|\nabla(V^N * \varphi)(x)\right|\ \left|K * u_s^N(x) - K * u_s^N\big(x - \tfrac{y}{N^\alpha}\big)\right| dy \\
&\quad + \frac{C}{N^\alpha} \int_{\mathbb{R}^2} V(y)|y| dy,
\end{aligned}$$

where the second inequality comes using the Lipschitz continuity and boundedness of F. Now in view of $(\mathbf{A}_{iii}^K)$

$$\begin{aligned}
\left|K * u_s^N(x) - K * u_s^N\big(x - \tfrac{y}{N^\alpha}\big)\right| &\leq \left(\tfrac{|y|}{N^\alpha}\right)^{\boldsymbol{\zeta}} \mathcal{N}_{\boldsymbol{\zeta}}\left(K * u_s^N\right) \\
&\leq C_{\boldsymbol{p},d} \|u_s^N\|_{L^1 \cap H_{\boldsymbol{r}}^\beta(\mathbb{R}^d)} \frac{|y|^{\boldsymbol{\zeta}}}{N^{\boldsymbol{\zeta}\alpha}}.
\end{aligned}$$

Therefore,

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

Thus we have obtained

$$\left|\nabla(V^N * \varphi)(x) \cdot F(K * u_s^N)(x) - \big(\nabla(V^N * \varphi) \cdot F(K * u_s^N)\big) * V^N(x)\right| \leq C \left(\frac{1}{N^\alpha} + \frac{\|u_s^N\|_{L^1 \cap H_{\boldsymbol{r}}^\beta(\mathbb{R}^d)}}{N^{\boldsymbol{\zeta}\alpha}} \right),$$

so that the left-hand side of (4.11) converges to 0.

Recall that $\{u^N\}_{N\in\mathbb{N}}$ converges almost surely in $L^2([0,T], H_{\boldsymbol{r}}^\beta(\mathbb{R}^d))$ for the weak topology, hence it is bounded in this space (by the uniform boundedness principle). Thus, $\sup_N \int_0^T \|u_s^N\|_{H_{\boldsymbol{r}}^\beta(\mathbb{R}^d)}^2\ ds < \infty$, and since $\|u_s^N\|_{L^1(\mathbb{R}^d)} = 1$, it follows from (3.2) that $\sup_N \int_0^T \|u_s^N\|_{L^1 \cap H_{\boldsymbol{r}}^\beta(\mathbb{R}^d)}^2\ ds < \infty$. Hence $\sup_N \int_0^T \|u_s^N\|_{L^1 \cap H_{\boldsymbol{r}}^\beta(\mathbb{R}^d)}\ ds < \infty$ a.s. and therefore

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

where in the last equality we used:

1. $u^N \xrightarrow{a.s.} \widetilde{u}$ strongly in $L^2([0,T]; \mathcal{C}(D))$ for D the compact support of φ (recall that u^N converges a.s. in $\mathcal{C}([0,T], W_{\boldsymbol{r},\mathrm{loc}}^\gamma(\mathbb{R}^d))$, hence by Sobolev embedding and dominated convergence, the convergence in $L^2([0,T]; \mathcal{C}(D))$ holds);
2. the convergence established in (4.10).

□

4.3.4 Time and space regularity of $\widetilde{u}$: Proof of Proposition 4.3

As $\widetilde{u} \in \mathcal{Y}$, we know that for almost all $t \in [0,T]$, $\widetilde{u}_t \in L^1 \cap H^\beta_{\boldsymbol{r}}(\mathbb{R}^d)$. Observe that for any $t \in [0,T]$ and $p \geq 1$, we have $\mathbb{E}\|u^N_t\|_{L^p(\mathbb{R}^d)} \leq C_{T,p}$: indeed, for $p=1$, this is because u^N_t is a probability density function; for $p=\infty$, recall that when $\beta > \frac{d}{\boldsymbol{r}}$, $H^\beta_{\boldsymbol{r}}(\mathbb{R}^d)$ is continuously embedded into L^∞ (see embedding (3.1)), hence it follows from Proposition 4.4 that $\mathbb{E}\|u^N_t\|_{L^1\cap L^\infty(\mathbb{R}^d)} \leq C_{T,p}$. Then Fatou's lemma implies that

$$\sup_{t\leq T} \|\widetilde{u}_t\|_{L^1\cap L^\infty(\mathbb{R}^d)} < \infty. \tag{4.12}$$

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

$$\lim_{\substack{s\to t\\ s\in(0,T)}} \|\widetilde{u}_t - \widetilde{u}_s\|_{L^1\cap L^\infty(\mathbb{R}^d)} = 0. \tag{4.13}$$

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

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

To check (4.13), we need to ensure that

$$\lim_{s\to t} \int_s^t \|\nabla\cdot e^{(t-r)\Delta}(\widetilde{u}_r F(K*\widetilde{u}_r))\|_{L^1\cap L^\infty(\mathbb{R}^d)}\, dr = 0.$$

This will follow from the continuity of the integral if the integral is well-defined. We have that

$$\int_s^t \|\nabla\cdot e^{(t-r)\Delta}(\widetilde{u}_r F(K*\widetilde{u}_r))\|_{L^1\cap L^\infty(\mathbb{R}^d)}\, dr \leq \int_s^t \frac{C_A}{\sqrt{t-r}} \|\widetilde{u}_r\|_{L^1\cap L^\infty(\mathbb{R}^d)}\, dr.$$

In view of (4.12), the integral $\int_s^t \frac{C_A}{\sqrt{t-r}}\|\widetilde{u}_r\|_{L^1\cap L^\infty(\mathbb{R}^d)}\, dr$ is well-defined. Hence the proof is complete.

5 Rate of convergence

5.1 Rate in $L^1 \cap L^\infty$ norm: Proof of Theorem 2.6

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

In view of the mild formulas (4.8) for u^N and (2.6) for u, it comes

$$\begin{aligned}
u^N_t(x) - u_t(x) &= e^{t\Delta}(u^N_0 - u_0)(x) - \int_0^t \nabla\cdot e^{(t-s)\Delta}\left(\langle \mu^N_s, V^N(\cdot - x)F\big(K*u^N_s(\cdot)\big)\rangle - u_sF(K*u_s)(x)\right) ds \\
&\quad + \frac{1}{N}\sum_{i=1}^N \int_0^t e^{(t-s)\Delta}\nabla V^N(X^{i,N}_s - x)\cdot dW^i_s \\
&= e^{t\Delta}(u^N_0 - u_0)(x) + \int_0^t \nabla\cdot e^{(t-s)\Delta}\left((u_sF(K*u_s) - u^N_sF(K*u^N_s))*V^N\right)(x)\, ds \\
&\quad + E^{(1)}_t(x) + E^{(2)}_t(x) + M^N_t(x),
\end{aligned}$$

where we have set

$$\begin{aligned}
E^{(1)}_t(x) &:= \int_0^t \nabla\cdot e^{(t-s)\Delta}\left(u_sF(K*u_s) - (u_sF(K*u_s))*V^N\right)(x)\, ds, \\
E^{(2)}_t(x) &:= \int_0^t \nabla\cdot e^{(t-s)\Delta}\langle u^N_s - \mu^N_s, V^N(\cdot - x)F\big(K*u^N_s(\cdot)\big)\rangle\, ds, \\
M^N_t(x) &:= \frac{1}{N}\sum_{i=1}^N \int_0^t e^{(t-s)\Delta}\nabla V^N(X^{i,N}_s - x)\cdot dW^i_s.
\end{aligned} \tag{5.1}$$

For any $p \in [1, +\infty]$, in view of the estimate (1.10), one has

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

and it follows that

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

Recall from Proposition 4.4 that we have, for any $m \ge 1$,

$$\sup_{N\in\mathbb{N}} \sup_{s\in[0,T]} \mathbb{E}[\|u_s^N\|_{\beta,\boldsymbol{r}}^m] < \infty. \tag{5.2}$$

Hence by Fatou's lemma, one has

$$\sup_{t\in[0,T]} \|u_t\|_{\beta,\boldsymbol{r}} < \infty, \tag{5.3}$$

which by Sobolev embedding (see Section 3.1) implies that $\sup_{t\in[0,T]} \|u_t\|_{L^\infty(\mathbb{R}^d)} < \infty$. Besides, u_t is a density for each t, so $\|u\|_{T,L^1\cap L^\infty(\mathbb{R}^d)} < \infty$, and by interpolation it follows that

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

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

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

Finally we apply Lemma 3.1 and obtain

$$\begin{aligned}\|u_t^N - u_t\|_{L^p(\mathbb{R}^d)} \le \|e^{t\Delta}(u_0^N - u_0)\|_{L^p(\mathbb{R}^d)} + C\int_0^t \frac{1}{\sqrt{t-s}} \|u_s^N - u_s\|_{L^p(\mathbb{R}^d)} ds \\ + C\int_0^t \frac{1}{\sqrt{t-s}} \|u_s - u_s^N\|_{L^1\cap L^\infty(\mathbb{R}^d)} ds \\ + \|E_t^{(1)}\|_{L^p(\mathbb{R}^d)} + \|E_t^{(2)}\|_{L^p(\mathbb{R}^d)} + \|M_t^N\|_{L^p(\mathbb{R}^d)}.\end{aligned} \tag{5.4}$$

Therefore, considering (5.4) for both $p = 1$ and $p = \infty$, we deduce that

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

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

$$\begin{aligned}
&\left\| \|u^N - u\|_{t,L^1\cap L^\infty(\mathbb{R}^d)} \right\|_{L^m(\Omega)} \\
&\quad \leq \left\| \sup_{s\in[0,t]} \|e^{s\Delta}(u_0^N - u_0)\|_{L^1\cap L^\infty(\mathbb{R}^d)} \right\|_{L^m(\Omega)} \\
&\quad + C\int_0^t \frac{1}{(t-s)^{\frac{1}{2}}} \left\| \|u^N - u\|_{s,L^1\cap L^\infty(\mathbb{R}^d)} \right\|_{L^m(\Omega)} ds \\
&\quad + \left\| \|E^{(1)}\|_{t,L^1\cap L^\infty(\mathbb{R}^d)} \right\|_{L^m(\Omega)} + \left\| \|E^{(2)}\|_{t,L^1\cap L^\infty(\mathbb{R}^d)} \right\|_{L^m(\Omega)} + \left\| \|M^N\|_{t,L^1\cap L^\infty(\mathbb{R}^d)} \right\|_{L^m(\Omega)} .
\end{aligned} \tag{5.5}$$

Step 2: L^1-estimates.

• *First, let us estimate* $\|E_t^{(1)}\|_{L^1(\mathbb{R}^d)}$. The property (1.10) on the derivative of the heat kernel gives

$$\sup_{s\in[0,t]} \|E_s^{(1)}\|_{L^1(\mathbb{R}^d)} \leq \int_0^t \frac{C}{(t-s)^{\frac{1}{2}}} \|e^{\frac{t-s}{2}\Delta}\left(u_sF(K*u_s) - (u_sF(K*u_s))*V^N\right)\|_{L^1(\mathbb{R}^d)} ds. \tag{5.6}$$

Recall that the d-dimensional Gaussian probability density function g_{t-s} (defined in (1.9)) is the kernel associated to the operator $e^{\frac{t-s}{2}\Delta}$ and observe that

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

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

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

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

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

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

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

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

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

Hence plugging this bound into (5.6) yields

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

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

$$\sup_{s\in[0,t]} \|E_s^{(1)}\|_{L^1(\mathbb{R}^d)} \leq C\, N^{-(1-\varepsilon)\alpha}. \tag{5.7}$$

• *We now search for a bound on* $\| \sup_{s\in[0,t]} \|E_s^{(2)}\|_{L^1(\mathbb{R}^d)}\|_{L^m(\Omega)}$. First, observe that due to the convolution inequality (1.10),

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

Hence by Fubini's theorem, one gets

$$\begin{aligned}&\sup_{s\in[0,t]} \|E_s^{(2)}\|_{L^1(\mathbb{R}^d)} \\ &\leq C\int_0^t \frac{1}{(t-s)^{\frac{1}{2}}} \Bigg(\Bigg(\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} V(y) \left| g_{t-s} * V^N(z-x) \left(F(K*u_s^N(z)) - F(K*u_s^N(z+\tfrac{y}{N^\alpha}))\right)\right| dy\, \mu_s^N(dz)\, dx \Bigg) \\ &+ \Bigg(\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} V(y) \left| F(K*u_s^N(z+\tfrac{y}{N^\alpha})) \left(g_{t-s} * V^N(z-x) - g_{t-s} * V^N(z-x+\tfrac{y}{N^\alpha})\right)\right| dy\, \mu_s^N(dz)\, dx \Bigg) \Bigg) ds \\ &=: E^{(2,1)} + E^{(2,2)}.\end{aligned} \tag{5.8}$$

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

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

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

where in the first inequality, we used the fact that V is rapidly decreasing and therefore the integral with respect to y is finite. Then by the standard convolution inequality, $\|g_{t-s} * V^N\|_{L^1(\mathbb{R}^d)} \leq 1$. Hence it follows from (5.2) that

$$\begin{aligned}\|E^{(2,1)}\|_{L^m(\Omega)} &\leq \frac{C}{N^{\alpha\boldsymbol{\zeta}}} \sup_{s\in[0,T]} \Big\| \|u_s^N\|_{L^1\cap H_{\boldsymbol{r}}^\beta(\mathbb{R}^d)} \Big\|_{L^m(\Omega)} \int_0^t (t-s)^{-\frac{1}{2}} ds \\ &\leq \frac{C}{N^{\alpha\boldsymbol{\zeta}}}.\end{aligned} \tag{5.9}$$

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

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

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

$$E^{(2,2)} \leq \frac{C}{N^\alpha} \int_0^t \frac{1}{(t-s)^{\frac{1}{2}}} \|\nabla(g_{t-s} * V^N)\|_{L^1(\mathbb{R}^d)} ds. \tag{5.10}$$

The estimation of $\|\nabla(g_{t-s} * V^N)\|_{L^1(\mathbb{R}^d)}$ is twofold, depending on which term of the convolution we apply the gradient. First, by the convolution inequality (1.10),

$$\|\nabla(g_{t-s} * V^N)\|_{L^1(\mathbb{R}^d)} = \|\nabla g_{t-s} * V^N\|_{L^1(\mathbb{R}^d)} \leq \frac{C}{\sqrt{t-s}}. \tag{5.11}$$

Second, still by applying a convolution inequality, it comes

$$\begin{aligned} \|\nabla(g_{t-s} * V^N)\|_{L^1(\mathbb{R}^d)} = \|g_{t-s} * (\nabla V^N)\|_{L^1(\mathbb{R}^d)} &\leq \|g_{t-s}\|_{L^1(\mathbb{R}^d)} \|\nabla V^N\|_{L^1(\mathbb{R}^d)} \\ &\leq C N^\alpha. \end{aligned} \tag{5.12}$$

Hence, combining (5.11) and (5.12), we deduce that for any $\varepsilon \in [0,1]$,

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

Plugging the previous bound in (5.10) yields

$$E^{(2,2)} \leq C N^{-\alpha(1-\varepsilon)}. \tag{5.13}$$

In view of deterministic bounds obtained in Inequalities (5.9) and (5.13), we deduce from (5.8) that

$$\left\| \sup_{s\in[0,t]} \|E_s^{(2)}\|_{L^1(\mathbb{R}^d)} \right\|_{L^m(\Omega)} \leq C \left(N^{-\alpha\boldsymbol{\zeta}} + N^{-\alpha(1-\varepsilon)} \right). \tag{5.14}$$

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

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

$$\left\| \sup_{s\in[0,t]} \|M_s^N\|_{L^1(\mathbb{R}^d)} \right\|_{L^m(\Omega)} \leq C\, N^{-\frac{1}{2}(1-\alpha(d+2))+\varepsilon}, \quad \forall N \in \mathbb{N}^*. \tag{5.15}$$

Step 3: L^∞-estimates.

- *We estimate the quantity* $\|E_t^{(1)}\|_\infty$. Applying (1.10), one has

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

where $\mathcal{N}_{\boldsymbol{\zeta}}$ (resp. $\mathcal{N}_\delta$) is the Hölder seminorm of parameter $\boldsymbol{\zeta}$ (resp. δ) defined in (1.12). In view of the embedding (3.1) and the inequality (5.3), the Hölder regularity of u_s is $\delta=\beta-\frac{d}{\boldsymbol{r}}$ and

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

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

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

Hence the boundedness of u in $L^\infty([0,T];L^1\cap H^\beta_{\boldsymbol{r}}(\mathbb{R}^d))$ (see again (5.3)) yields

$$\sup_{s\in[0,t]}\|E_s^{(1)}\|_{L^\infty(\mathbb{R}^d)}\le C\left(N^{-\alpha\boldsymbol{\zeta}}+N^{-\alpha(\beta-\frac{d}{\boldsymbol{r}})}\right). \tag{5.16}$$

- *Now, we turn to* $\|E_t^{(2)}\|_\infty$. In view of (1.10), one has

$$\|E_s^{(2)}\|_{L^\infty(\mathbb{R}^d)}\le C\int_0^t\frac{1}{(t-s)^{\frac12}}\Big\|\int_{\mathbb{R}^d}(g_{t-s}*V^N)(z-\cdot)F\big(K*u_s^N(z)\big)(u_s^N-\mu_s^N)(dz)\Big\|_{L^\infty(\mathbb{R}^d)}ds$$
$$\le C\int_0^t\frac{1}{(t-s)^{\frac12}}\Big\|\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}V(y)(g_{t-s}*V^N)(z-\frac{y}{N^\alpha}-\cdot)$$
$$\times\Big(F\big(K*u_s^N(z-\frac{y}{N^\alpha})\big)-F\big(K*u_s^N(z)\big)\Big)\,dy\;\mu_s^N(dz)\Big\|_{L^\infty(\mathbb{R}^d)}ds$$
$$+C\int_0^t\frac{1}{(t-s)^{\frac12}}\Big\|\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}V(y)\Big((g_{t-s}*V^N)(z-\frac{y}{N^\alpha}-\cdot)-(g_{t-s}*V^N)(z-\cdot)\Big)$$
$$\times F\big((K*u_s^N)(z)\big)dy\mu_s^N(dz)\Big\|_{L^\infty(\mathbb{R}^d)}ds$$
$$=:E_t^{(2,1,\infty)}+E_t^{(2,2,\infty)}. \tag{5.17}$$

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

$$E_t^{(2,1,\infty)}$$
$$\le\frac{C}{N^{\alpha\boldsymbol{\zeta}}}\int_0^t\frac{1}{(t-s)^{\frac12}}\|u_s^N\|_{L^1\cap H^\beta_{\boldsymbol{r}}(\mathbb{R}^d)}\Big\|\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}V(y)|y|^{\alpha\boldsymbol{\zeta}}(g_{t-s}*V^N)(z-\frac{y}{N^\alpha}-\cdot)dy\mu_s^N(dz)\Big\|_{L^\infty(\mathbb{R}^d)}ds$$
$$\le\frac{C}{N^{\alpha\boldsymbol{\zeta}}}\int_0^t\frac{1}{(t-s)^{\frac12}}\|u_s^N\|_{L^1\cap H^\beta_{\boldsymbol{r}}(\mathbb{R}^d)}\int_{\mathbb{R}^d}V(y)|y|^{\alpha\boldsymbol{\zeta}}\Big\|\int_{\mathbb{R}^d}(g_{t-s}*V^N)(z-\frac{y}{N^\alpha}-\cdot)\mu_s^N(dz)\Big\|_{L^\infty(\mathbb{R}^d)}dy\,ds$$
$$\le\frac{C}{N^{\alpha\boldsymbol{\zeta}}}\int_0^t\frac{1}{(t-s)^{\frac12}}\|u_s^N\|_{L^1\cap H^\beta_{\boldsymbol{r}}(\mathbb{R}^d)}\int_{\mathbb{R}^d}V(y)|y|^{\alpha\boldsymbol{\zeta}}\,\big\|g_{t-s}*u_s^N\big\|_{L^\infty(\mathbb{R}^d)}\,dy\,ds$$
$$\le\frac{C}{N^{\alpha\boldsymbol{\zeta}}}\int_0^t\frac{1}{(t-s)^{\frac12}}\|u_s^N\|^2_{L^1\cap H^\beta_{\boldsymbol{r}}(\mathbb{R}^d)}\,ds. \tag{5.18}$$

Observe that

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

where $F(K * u_s^N)\mu_s^N$ denotes the weighted empirical measure. Now, we shall prove that $(F(K * u_s^N)\mu_s^N) * (g_{t-s} * V^N)$ is bounded in $H_{\boldsymbol{r}}^\beta(\mathbb{R}^d)$. Recall the following representation for the $H_{\boldsymbol{r}}^\beta(\mathbb{R}^d)$ norm (see (1.7)):

$$\begin{aligned} \|(F(K * u_s^N)\mu_s^N) * (g_{t-s} * V^N)\|_{\beta,\boldsymbol{r}} &= \left\| (I-\Delta)^{\frac{\beta}{2}} \left((F(K * u_s^N)\mu_s^N) * (g_{t-s} * V^N) \right) \right\|_{L^{\boldsymbol{r}}(\mathbb{R}^d)} \\ &= \left\| (I-\Delta)^{\frac{\beta}{2}} g_{t-s} * \left((F(K * u_s^N)\mu_s^N) * V^N \right) \right\|_{L^{\boldsymbol{r}}(\mathbb{R}^d)} \end{aligned}$$

where the second equality holds because $(I-\Delta)^{\frac{\beta}{2}}$ acts as a convolution. Then from the inequality $\|(I-\Delta)^{\frac{\beta}{2}} g_{t-s}\|_{L^1(\mathbb{R}^d)} \leq C(t-s)^{-\frac{\beta}{2}}$ (see Equation (1.11)) and a convolution inequality, it comes

$$\begin{aligned} \|(F(K * u_s^N)\mu_s^N) * (g_{t-s} * V^N)\|_{\beta,\boldsymbol{r}} &\leq \frac{C}{(t-s)^{\frac{\beta}{2}}} \|(F(K * u_s^N)\mu_s^N) * V^N)\|_{L^{\boldsymbol{r}}(\mathbb{R}^d)} \\ &\leq \frac{C}{(t-s)^{\frac{\beta}{2}}} \|u_s^N\|_{L^{\boldsymbol{r}}(\mathbb{R}^d)}. \end{aligned}$$

Thus by the Sobolev embedding (3.1), we obtain that for $\delta = (\beta - \frac{d}{\boldsymbol{r}})$,

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

which combined with the previous inequality yields

$$E_t^{(2,2,\infty)} \leq \frac{C}{N^{\alpha(\beta - \frac{d}{\boldsymbol{r}})}} \int_0^t \frac{1}{(t-s)^{\frac{1+\beta}{2}}} \|u_s^N\|_{L^{\boldsymbol{r}}(\mathbb{R}^d)}\, ds. \tag{5.19}$$

From (5.17), (5.18) and (5.19), it comes that

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

Hence, in view of the fact that $\|u_s^N\|_{L^1(\mathbb{R}^d)} = 1$, of the uniform bound on $\left\| \|u_s^N\|^2_{\beta,\boldsymbol{r}} \right\|_{L^m(\Omega)}$ (Eq. (5.2)) and the assumption $\beta < 1$ (see $(\mathbf{A}_{iii}^K)$), we conclude that

$$\left\| \sup_{s\in[0,t]} \|E_s^{(2)}\|_{L^\infty(\mathbb{R}^d)} \right\|_{L^m(\Omega)} \leq C \left(N^{-\alpha(\beta - \frac{d}{\boldsymbol{r}})} + N^{-\alpha\boldsymbol{\zeta}} \right). \tag{5.20}$$

• *It remains to estimate M_t^N.* We proceed with the same care as when we got the bound in (5.15). The details may be found in Appendix A.1 and here we only apply Proposition A.1 for $p = \infty$ and conclude that for any $\varepsilon > 0$, there exists $C > 0$ such that

$$\left\| \sup_{s\in[0,t]} \|M_s^N\|_{L^\infty(\mathbb{R}^d)} \right\|_{L^m(\Omega)} \leq C\, N^{-\frac{1}{2}(1-\alpha(d+[2\vee d]))+\varepsilon}. \tag{5.21}$$

Step 4 : Conclusion.

From the Inequalities (5.5), (5.7), (5.14), (5.15), (5.16), (5.20) and (5.21), and using the Grönwall lemma, we conclude that for any $\varepsilon > 0$ small enough, there exists $C > 0$ such that for any $N \in \mathbb{N}^*$,

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

5.2 Rate in Sobolev norm: Proof of Corollary 2.9

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

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

$$\gamma = \beta\frac{\boldsymbol{r}(\boldsymbol{r}-\delta-1)}{(\boldsymbol{r}-\delta)(r-1)}. \tag{5.22}$$

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

$$\|f\|_{\gamma,\boldsymbol{r}-\delta} \leq \|f\|_{0,1}^\theta \, \|f\|_{\beta,\boldsymbol{r}}^{1-\theta}, \tag{5.23}$$

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

Hence it follows from (5.23) that for any $m \geq 1$,

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

and we deduce from Hölder's inequality that

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

In view of the previous inequality and using Theorem 2.6 and Proposition 4.4, we deduce the rate of convergence in $L^m\left(\Omega; L^\infty([0,T], H_{\boldsymbol{r}-\delta}^\gamma(\mathbb{R}^d))\right)$.

Finally, note that it is always true that $\gamma < \beta$. Besides, it will be important to ensure that $\gamma > \frac{d}{\boldsymbol{r}-\delta}$ to have an embedding in the space of Hölder continuous functions (see (3.1)). For this, it suffices to choose δ which satisfies:

$$\frac{\boldsymbol{r}-\delta-1}{\boldsymbol{r}-1} > \frac{d/\boldsymbol{r}}{\beta}. \tag{5.24}$$

5.3 Propagation of chaos for the marginals: Proof of Corollary 2.13

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

$$\left|\langle u_t^N, \phi\rangle - \langle \mu_t^N, \phi\rangle\right| \leq \frac{C\|\phi\|_{\text{Lip}}}{N^\alpha} \quad a.s. \tag{5.25}$$

Indeed,

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

Recalling the definition (2.9) 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 (5.25) to the first term on the right-hand side of the above inequality, and Theorem 2.6 to the second term, we obtain the inequality of Corollary 2.13.

6 Propagation of chaos

In this section we study the well-posededness of the nonlinear SDE (1.6) and then the propagation of chaos of the particle system (2.3). More precisely, we prove Proposition 2.16 about the well-posedness of the martingale problem $(\mathcal{MP})$ related to (1.6). Then, we prove the convergence in law, when $N\to\infty$, of the empirical measure μ^N of the particle system (2.3) towards the unique solution of the martingale problem $(\mathcal{MP})$.

6.1 Proof of Proposition 2.16

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

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

Let $\mathbb{Q}$ be a solution to $(\mathcal{MP})$. Notice first that as the family of marginal laws $(q_t)_{t\leq T}$ belongs to $q\in\mathcal{C}([0,T];L^1\cap L^\infty(\mathbb{R}^d))$, one has according to Lemma 3.1 that

$$\sup_{t\leq T}\|K*q_t\|_{L^\infty(\mathbb{R}^d)}\leq C_{K,d}\sup_{t\leq T}\|q_t\|_{L^1\cap L^\infty(\mathbb{R}^d)}. \tag{6.1}$$

To obtain the equation satisfied by $(q_t)_{t\leq T}$, one derives the mild equation for the marginal distributions of the corresponding nonlinear process. This is done in the usual way as the drift component is bounded (see e.g. [55, Section 4]). One has

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

This equation is exactly (1.1) and we know it admits a unique mild solution in the sense of Remark 2.2 up to T. Meaning, as $q\in\mathcal{C}([0,T];L^1\cap L^\infty(\mathbb{R}^d))$, the one-dimensional time marginals of $\mathbb{Q}$ are uniquely determined.

Define the corresponding linear martingale problem by fixing q to be the unique mild solution u to (1.1) in the definition of the process $(M_t)_{t\leq T}$ from $(\mathcal{MP})$.

By Girsanov transformation, the equation

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

admits a weak solution. In addition, strong uniqueness holds (see [38, Remark 1.6]). Thus, the associated linear martingale problem admits a unique solution. This immediately implies the uniqueness of solutions to $(\mathcal{MP})$.

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

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

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

Assume for a moment that $p \in \mathcal{C}([0,T]; L^1 \cap L^\infty(\mathbb{R}^d))$. The previous equation is a linearized version of Eq. (2.7) and with the same arguments as in Proposition 3.5, its cut off version admits a unique solution in $\mathcal{C}([0,T]; L^1 \cap L^\infty(\mathbb{R}^d))$. Then, by the same arguments as in Proposition 2.3, the above equation admits a unique solution in $\mathcal{C}([0,T]; L^1 \cap L^\infty(\mathbb{R}^d))$. Since both u and p solve this equation, they must coincide and we have the desired result : $(p_t)_{t\in[0,T]} = (u_t)_{t\in[0,T]}$.

It only remains to prove that $p \in \mathcal{C}([0,T]; L^1 \cap L^\infty(\mathbb{R}^d))$. Obviously, as we work with a family of probability density functions, we only need to prove that $p \in \mathcal{C}([0,T]; L^\infty(\mathbb{R}^d))$. Performing the same calculations as in the proof of Proposition 3.5, we get that

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

In view of Lemma 3.1, one has

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

One gets that

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

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

6.2 Proof of Theorem 2.17

To prove Theorem 2.17, we will show that μ^N converges to the unique solution $\mathbb{Q}$ of the martingale problem $(\mathcal{MP})$. To do so, we will first prove the convergence towards an auxiliary martingale problem which is identical to $(\mathcal{MP})$ except that in the point *iii*) the process $(M_t)_{t\leq T}$ is the following:

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

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

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

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

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

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

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

We will denote by $(\boldsymbol{\mu}, \mathbf{u})$ the canonical projections in $\mathcal{H}$.

Now, for $N \geq 1$ we denote by $\tilde{\Pi}^N$ the law of the random variables (μ^N, u^N) that take values in $\mathcal{H}$. The sequence $(\tilde{\Pi}^N, N \geq 1)$ is tight if and only if $(\tilde{\Pi}^N \circ \boldsymbol{\mu}, N \geq 1)$ and $(\tilde{\Pi}^N \circ \mathbf{u}, N \geq 1)$ are tight. The tightness of $(\tilde{\Pi}^N \circ \boldsymbol{\mu}, N \geq 1)$ is classical, as the drift of the particle system is bounded. As for $(\tilde{\Pi}^N \circ \mathbf{u}, N \geq 1)$, we have already proven the convergence of $(u^N, N \geq 1)$ in $\tilde{\mathcal{Y}}$ (see Theorem 2.5 and Theorem 2.6).

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

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

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

Proof. This is a consequence of Inequality (5.25). Take a test function $\varphi \in \mathcal{C}_c^\infty([0,T] \times \mathbb{R}^d)$ and define a functional $\phi(t,x) = \varphi(t, x_t)$, for $x \in \mathcal{C}([0,T];\mathbb{R}^d)$. Then,

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

where the last inequality comes from (5.25). Thus, we obtain that $\tilde{\Pi}^\infty$-a.s. the following measures on $\mathbb{R}^d \times [0,T]$ are equal:

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

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

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

□

The following proposition will be the last ingredient needed for the proof of Theorem 2.17.

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

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

Then

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

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

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

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

We decompose

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

Notice that

$$\begin{aligned} I_n &\leq \|\Phi\|_\infty \|\nabla f\|_\infty \langle \boldsymbol{\mu}^n, \int_s^t |F(K * \mathbf{u}_r^n(\cdot_r)) - F(K * \mathbf{u}_r(\cdot_r))| dr\rangle \\ &\leq C \int_s^t \langle \boldsymbol{\mu}_r^n, |K * (\mathbf{u}_r^n - \mathbf{u}_r)|\rangle \, dr \leq C \int_s^t \|K * (\mathbf{u}_r^n - \mathbf{u}_r)\|_\infty dr. \end{aligned} \tag{6.2}$$

In view of Lemma 3.1, one has

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

Now recall $\boldsymbol{r}$ and β are fixed in $(\mathbf{A}_{iii}^K)$, and let γ and δ satisfy (5.22) and (5.24), so that $\frac{d}{\boldsymbol{r}-\delta} < \gamma < \beta$. Then, use the Sobolev embedding $L^1 \cap H^\gamma_{\boldsymbol{r}-\varepsilon}(\mathbb{R}^d) \subset L^1 \cap L^\infty(\mathbb{R}^d)$ to get

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

Plug the latter in (6.2) to obtain

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

By the interpolation inequality (5.23),

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

for θ as in Section 5.2. Now since $\mathbf{u}^n$ converges in $\mathcal{Y}$, and converges in particular weakly in $L^2\left([0,T], H^\beta_{\boldsymbol{r}}(\mathbb{R}^d)\right)$, the uniform boundedness principle tells us that it is bounded in this space. Gathering this fact with the convergence in $L^\infty([0,T]; L^1(\mathbb{R}^d))$ (by assumption), the previous inequality yields the convergence of $\mathbf{u}^n$ in $L^2\left([0,T], L^1 \cap H^\gamma_{\boldsymbol{r}-\varepsilon}(\mathbb{R}^d)\right)$. Hence I_n converges to 0.

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

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

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

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

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

We have all the elements in hand to finish the proof of Theorem 2.17. By Lemma 6.1 and Proposition 6.2, we get that $\boldsymbol{\mu} \in \text{supp}(\Pi^\infty)$ solves the nonlinear martingale problem $(\mathcal{MP}_A)$. Choose $A > A_T := C_{k,D} \|q\|_{T, L^1 \cap L^\infty(\mathbb{R}^d)}$ and lift the cut-off (see (6.1)). Then, $\boldsymbol{\mu}$ solves the nonlinear martingale problem $(\mathcal{MP})$. As we have the uniqueness for $(\mathcal{MP})$, we get that there is only one limit value of the sequence Π^N which is $\delta_{\mathbb{Q}}$.

7 Examples

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

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

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

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

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

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

In particular the previous inequality holds true if $f \in L^1 \cap H_{\boldsymbol{r}}^{\beta}(\mathbb{R}^d)$, because by Sobolev embedding, f is in $L^1 \cap L^\infty$ (see embedding (3.2)), and then the result holds by interpolation. Now in view of the previous fact and using the property $(\widetilde{\mathbf{A}}_{iii}^K)$ of ∇K, one deduces that if $f \in L^1 \cap H_{\boldsymbol{r}}^{\beta}(\mathbb{R}^d)$, then $K * f \in H_{\boldsymbol{p}}^1(\mathbb{R}^d)$. Hence it follows from Morrey's inequality [7, Th. 9.12] that there exists $C_{\boldsymbol{p},d} > 0$ such that for any $f \in L^1 \cap H_{\boldsymbol{r}}^{\beta}(\mathbb{R}^d)$,

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

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

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

The first two points of Assumption $(\mathbf{A}^K)$ are simple technical conditions and may not require specific comments, except that it would be interesting to lift the first integrability condition in order to be able to consider more singular kernels. The third assumption is much more interesting. Let us start with the simple example of a kernel K such that ∇K is integrable. Then ∇K defines a convolution operator and by a convolution inequality, this operator is bounded in any $L^p(\mathbb{R}^d)$, $p \in [1, +\infty]$. As a consequence of the embedding (3.2), ∇K satisfies $(\widetilde{\mathbf{A}}_{iii}^K)$ for any $\boldsymbol{p} \in [1, \infty]$ and any β and $\boldsymbol{r}$ such that $\beta - \frac{d}{\boldsymbol{r}} > 0$. Hence it satisfies $(\mathbf{A}_{iii}^K)$ with $\boldsymbol{\zeta} = 1$ (see Section 7.1).

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

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

The associated kernel

$$K_C := \nabla V_C$$

is a generalisation in any dimension of the classical Coulomb force, and ∇K_C is not integrable. Nevertheless, it is possible to define the convolution operator of kernel ∇K_C as the Principal Value integral acting on the space of smooth, rapidly decaying functions (i.e. the Schwartz space), thus defining a tempered distribution.

The previous example is a special case of operator defined as a singular integral, which under certain assumptions on the kernel (see the three conditions of [26, Chapter 4.4]) extends to a bounded operator in $L^p(\mathbb{R}^d)$, for any $p \in (1, +\infty)$. In particular, ∇K_C verifies these conditions and therefore K_C satisfies $(\mathbf{A}^K)$ with $\boldsymbol{\zeta} = 1 - \frac{d}{\boldsymbol{p}}$ (see Section 7.1), for any $\boldsymbol{p} \in (1, \infty)$ and any β and $\boldsymbol{r}$ such that $\beta - \frac{d}{\boldsymbol{r}} > 0$.

7.3 Riesz and Coulomb potentials

The Coulomb potential belongs to a more general class of interaction potentials, called Riesz potentials, which were defined in (1.2). If $d \geq 3$ and $s = d - 2$, this is the Coulomb potential presented in the previous subsection. We denote the associated kernel by $K_s := \nabla V_s$. K_s satisfies Assumption $(\mathbf{A}^K)$, provided that $d \geq 2$ and $s \in [0, d-1)$: Indeed,

- These kernels are locally integrable if and only if $0 \leq s < d - 1$.
- K_s is integrable outside the unit ball for any $q > \frac{d}{s+1}$.
-
  - If $d \geq 3$ and $s < d - 2$, then ∇K_s is not bounded in any L^p but it is however bounded from L^p to L^q whenever $p \in (1, \frac{d}{d-(s+2)})$ and $\frac{1}{q} = \frac{1}{p} + \frac{s+2}{d} - 1$ (see [49, Theorem 25.2]). This is still enough, thanks to the embedding (3.2), to ensure property $(\widetilde{\mathbf{A}}^K_{iii})$. In particular, one can choose any $\boldsymbol{p} \in (1, \frac{d}{d-(s+2)})$ and any β and $\boldsymbol{r}$ such that $\beta - \frac{d}{\boldsymbol{r}} > 0$ for these kernels. In particular $(\mathbf{A}^K_{iii})$ holds for $\boldsymbol{\zeta} = 1 - \frac{d}{\boldsymbol{p}}$ (see Section 7.1). Hence all our results can be applied to this kernel.
  - If $s = d - 2$, then ∇K_s is a typical kernel satisfying the conditions of [26, Chapter 4.4], and therefore it defines a bounded operator in any $L^p(\mathbb{R}^d)$, $p \in (1, \infty)$. Hence all our results apply to this particular kernel. Again, one can choose $\boldsymbol{\zeta} = 1 - \frac{d}{\boldsymbol{p}}$ for any $\boldsymbol{p} \in (1, \infty)$, and any β and $\boldsymbol{r}$ such that $\beta - \frac{d}{\boldsymbol{r}} > 0$.
    In particular, Theorem 2.6 and its corollaries 2.9, 2.11 and 2.13 are applicable, and choosing $\boldsymbol{q}$ and $\boldsymbol{p}$ very large, one can obtain with an appropriate choice of α and β (see Remark 2.8) a rate which is as closed as desired to $\frac{1}{2(d+1)}$.
  - If $s \in [d-2, d-1)$, then one can verify (see e.g. [16, Lemma 2.5]) that ∇K_s defines a convolution operator from $L^1 \cap L^\infty \cap \mathcal{C}^\sigma(\mathbb{R}^d)$ to $L^\infty(\mathbb{R}^d)$, with $\sigma \in (2 - d + s, 1)$:
    $$\begin{aligned} \|\nabla K_s * f\|_{L^\infty(\mathbb{R}^d)} &\lesssim \|f\|_{L^1(\mathbb{R}^d)} + \|f\|_{L^\infty(\mathbb{R}^d)} + \mathcal{N}_\sigma(f) \\ &\lesssim \|f\|_{L^1 \cap L^\infty(\mathbb{R}^d)} + \|f\|_{\beta, \boldsymbol{r}}, \end{aligned}$$
    for some β and $\boldsymbol{r}$ such that $\sigma = \beta - \frac{d}{\boldsymbol{r}}$, thanks to the embedding (3.1). Hence, one must choose $\boldsymbol{p} = \infty$, and $\beta < 1$ and $\boldsymbol{r}$ such that $\beta - \frac{d}{\boldsymbol{r}} > 2 - d + s$.
    Theorem 2.6 and its corollaries 2.9, 2.11 and 2.13 are applicable, choosing $\boldsymbol{q} = \boldsymbol{p} = \infty$ is allowed, so with an appropriate choice of α and β (see Remark 2.8), one obtains a rate which is as closed as desired to $\frac{1}{2(d+1)}$.

Besides obtaining rates of convergence, Proposition 2.16 proves the well-posedness of the McKean-Vlasov SDE (1.6) for all Riesz kernels with $s \in (0, d-1)$, which is new for the values of the parameter s, and most notably for the largest values $s \geq d - 2$. The trajectorial propagation of chaos (Theorem 2.17) is also new for this whole class of particle systems.

7.4 Parabolic-elliptic Keller-Segel models

An important and tricky example covered by this paper is the Keller-Segel PDE, which takes the form (1.1) with the kernel defined in (1.3) ($K_{KS}(x) = -\chi \frac{x}{|x|^d}$, for some $\chi > 0$). This is a model of chemotaxis, and we refer to the companion paper [43] for more information on this equation.

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

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

This leads to important issues that we discuss in more details in [43], but let us just mention as an example that in dimension 2, the PDE admits a global solution if and only if $\chi < 8\pi$ (see e.g. Biler [3] for a recent review). Note that in that case ($d = 2$ and $\chi < 8\pi$) it is again possible to choose a value of the cut-off A_T independently of T (see [43]).

Theorem 2.6, Corollaries 2.9, 2.11 and 2.13 give rates of convergence of the particle system to the Keller-Segel PDE, whenever a local or global solution exists, which is, as far as we know, the first time that such quantitative results appear for this class of PDEs (see [6] for related results and Section 2.5 to discuss how this compares to the present work).

As a consequence of Theorem 2.16, we also deduce the local-in-time weak well-posedness of the McKean-Vlasov (1.6) for all values of the concentration parameter χ, and global-in-time weak well-posedness for $\chi < 8\pi$, which is a new result (see [43] for a thorough discussion and comparison with previous results).

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

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

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

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

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

It is well-known that with such kernel, Eq. (1.1) has a unique global solution, and that $\|K * \xi_t\|_{L^\infty(\mathbb{R}^2)}$ can be bounded by $C\left(1 + \|\xi_0\|_{L^\infty(\mathbb{R}^2)}\right)$ (see [20] and references therein). This permits to choose the cut-off value A_T from (2.8) independently on T.

The kernel K_{BS} is covered by our assumption ($\mathbf{A}^K$), for the same reason as the Coulomb kernel with $d = 2$ ($K_0(x) = \frac{1}{|x|}$). In that case, we recover the Theorem 1.3 of Flandoli et al. [20] within our Theorem 2.5.

All the other results of this paper apply, and in particular, if the initial condition is smooth enough (i.e. $\beta \approx 1$ for Theorem 2.6), we obtain a rate ρ in $L^1 \cap L^\infty$ norm which is almost $\frac{1}{6}$.

7.6 Attractive-repulsive kernels

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

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

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

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

A similar, but less singular potential is proposed by Flandoli et al. [19] to model the adhesion of cells in biology. It can be expressed in general as

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

with $a, b > 0$ and $V_a, V_b > 0$. One can now refer to the discussion on Riesz kernels in Section 7.3 to determine the values of a and b that ensure the applicability of our results.

Appendix

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

In this section we study the moments of the supremum in time of $\|M_t^N\|_{L^p(\mathbb{R}^d)}$, where the stochastic convolution integral M_t^N was defined in (5.1). Such estimates will be used in the proof of Theorem 2.6 for $p = 1$ and $p = \infty$ only, but the result is established for any p without any additional effort.

Proposition A.1. *Let the assumption of Theorem 2.6 hold. Let $m_0 \geq 1$, $p \in [1, +\infty]$ and $\varepsilon > 0$. Then there exists $C > 0$ such that for any $t \in [0, T]$ and $N \in \mathbb{N}^*$,*

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

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

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

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

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

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

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

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

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

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

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

The final ingredient in the proof of Proposition A.1 is a consequence of Garsia-Rodemich-Rumsey's Lemma [25], given in the following lemma (for $\mathbb{R}$-valued processes, it already appears in [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. [24, Theorem A.1]).

Lemma A.4. *Let E be a Banach space and $(Y^n)_{n\geq1}$ be a sequence of E-valued continuous processes on $[0, T]$. Let $m \geq 1$ and $\eta > 0$ such that $m\eta > 1$ and assume that there exists constants $\rho > 0$, $C > 0$, and a sequence $(\delta_n)_{n\geq1}$ of positive real numbers such that*

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

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

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

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

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

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

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

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

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

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

which is the desired result. □

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

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

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

Now, $\widetilde{M}^N$ is a martingale that takes values in an inifinite-dimensional space, and $\widetilde{M}_t^N = M_t^N$. Recall that the operators $(I-\Delta)^{\frac{\gamma}{2}}$, $\gamma \in \mathbb{R}$ were defined in the Notations section, see Equation (1.7), with the relation $\|(I-\Delta)^{\frac{\gamma}{2}} f\|_{L^2(\mathbb{R}^d)} = \|f\|_\gamma$.

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

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

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

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

and this finishes the proof of (i).

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

$$
\begin{aligned}
|M_t^N(x) - M_s^N(x)| &\le \left| \frac{1}{N} \sum_{i=1}^N \int_s^t \nabla e^{(t-u)\Delta} V^N(X_u^{i,N} - x) \cdot dW_u^i \right| \\
&+ \left| \frac{1}{N} \sum_{i=1}^N \int_0^s \nabla e^{(s-u)\Delta} \left[e^{(t-s)\Delta} V^N(X_u^{i,N} - x) - V^N(X_u^{i,N} - x) \right] \cdot dW_u^i \right| \\
&=: |I_{s,t}^N(x)| + |II_{s,t}^N(x)|.
\end{aligned} \tag{A.3}
$$

Thus, one has

$$
\left\| \| M_t^N - M_s^N \|_\gamma \right\|_{L^m(\Omega)} \le \left\| \| I_{s,t}^N \|_\gamma \right\|_{L^m(\Omega)} + \left\| \| II_{s,t}^N \|_\gamma \right\|_{L^m(\Omega)}. \tag{A.4}
$$

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

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

and

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Proof of Proposition A.2. This proof will be divided in three according to whether $p\geq 2$, $p\in(1,2)$ or $p=1$, in that order. This might seem too much since we only need the cases $p=\infty$ and $p=1$ in Theorem 2.6. However, note that the proof is the same for any $p\in[2,+\infty]$. Besides, we present the proof for $p\in(1,2)$ before the proof for $p=1$, because our proof for the latter case consists in applying a Hölder inequality with weights so as to use the case $p\in(1,2)$ (we were not able to treat directly the case $p=1$ because L^1 is not a UMD space).

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

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

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

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

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

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

and

$$\big\|\|M^N_t - M^N_s\|_{L^p(\mathbb{R}^d)}\big\|_{L^m(\Omega)} \leq \big\|\|M^N_t - M^N_s\|_\gamma\big\|_{L^m(\Omega)}.$$

Thus Proposition A.3 yields that

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

and

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

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

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

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

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

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

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

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

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

$$\big\|\|M_t^N\|_{L^p(\mathbb{R}^d)}\big\|_{L^m(\Omega)} \leq \frac{C}{N}\left(\int_{\mathbb{R}^d}\frac{1}{(1+|x|)^{\frac{2p}{2-p}}}dx\right)^{\frac{2-p}{2p}}$$
$$\times\left\|\left(\int_{\mathbb{R}^d}(1+|x|)^2\sum_{i=1}^N\int_0^t\left|\nabla e^{(t-s)\Delta}V^N(X_s^{i,N}-x)\right|^2ds\,dx\right)^{\frac12}\right\|_{L^m(\Omega)},$$

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

$$\mathbb{E}\|M_t^N\|_{L^p(\mathbb{R}^d)}^m$$
$$\leq \frac{C}{N^m}\mathbb{E}\left(\sum_{i=1}^N\int_0^t\int_{\mathbb{R}^d}\left(1+|X_s^{i,N}|\right)^2\left(1+|X_s^{i,N}-x|\right)^2\left|\nabla e^{(t-s)\Delta}V^N(X_s^{i,N}-x)\right|^2dx\,ds\right)^{\frac m2}$$
$$\leq \frac{C}{N^m}\mathbb{E}\left(\sum_{i=1}^N\int_0^t\left(1+|X_s^{i,N}|\right)^2\int_{\mathbb{R}^d}(1+|y|)^2\left|\nabla e^{(t-s)\Delta}V^N(y)\right|^2dy\,ds\right)^{\frac m2}$$
$$\leq \frac{C}{N^m}\left(\int_0^t\int_{\mathbb{R}^d}(1+|y|)^2\left|\nabla e^{(t-s)\Delta}V^N(y)\right|^2dy\,ds\right)^{\frac m2}\mathbb{E}\left(\sum_{i=1}^N\left(1+\sup_{s\in[0,t]}|X_s^{i,N}|\right)^2\right)^{\frac m2}, \tag{A.12}$$

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

$$\mathbb{E}\left(\sum_{i=1}^N\left(1+\sup_{s\in[0,t]}|X_s^{i,N}|\right)^2\right)^{\frac m2}\leq C\,N^{\frac m2}. \tag{A.13}$$

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

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

where we recall that the notation g for the heat kernel was introduced in (1.9). By the simple inequality $(1+|a+b|) \le (1+|a|)(1+|b|)$ and Fubini's theorem, we then have

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

and therefore

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

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

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

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

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

The integrability properties of V and classical estimates on g yield

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

The latter implies that

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

Now in view of (A.3), $\left\|\|M^N_t - M^N_s\|_{L^p(\mathbb{R}^d)}\right\|_{L^m(\Omega)} \leq \left\|\|I^N_{s,t}\|_{L^p(\mathbb{R}^d)}\right\|_{L^m(\Omega)} + \left\|\|II^N_{s,t}\|_{L^p(\mathbb{R}^d)}\right\|_{L^m(\Omega)}$, hence the desired result in the case $p \in (1,2)$ follows from (A.15) and (A.21).

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

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

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

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

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

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

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

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

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

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

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

Now,

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

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

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

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

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

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

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

and we have

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

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

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

Now in view of (A.3), $\big\| \|M_t^N - M_s^N\|_{L^1(\mathbb{R}^d)} \big\|_{L^m(\Omega)} \leq \big\| \|I_{s,t}^N\|_{L^1(\mathbb{R}^d)} \big\|_{L^m(\Omega)} + \big\| \|II_{s,t}^N\|_{L^1(\mathbb{R}^d)} \big\|_{L^m(\Omega)}$, hence the desired result in the case $p=1$ follows from (A.22) and (A.23). □

A.2 Proofs of the tightness estimates

Proof of Proposition 4.4. **Step 1**. Recall that the operators $(I-\Delta)^\beta,\ \beta \in \mathbb{R}$ were defined in the Notations section, see Equation (1.7), with a clear link with the Sobolev norm $\|\cdot\|_{\beta,\boldsymbol{r}}$.

Let F stand for the function F_A defined in (2.1). From (4.8) after applying $(\mathrm{I}-\Delta)^{\frac{\beta}{2}}$ and by the triangular inequality we have

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

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

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

Step 2. Noticing that by a convolution inequality

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

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

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

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

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

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

In view of Inequality (1.11), we have that

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

Thus,

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

This bounds the second term.

Step 4. For the third term (A.26), recalling the notation introduced in (5.1), we need to control $\left\|\|M_t^N\|_{\beta,\boldsymbol{r}}\right\|_{L^q(\Omega)}$. The embedding for Bessel potential spaces of [57, p.203] gives that $H^{\beta+d(\frac{1}{2}-\frac{1}{r})}(\mathbb{R}^d)$ is continuously embedded into $H_{\boldsymbol{r}}^\beta(\mathbb{R}^d)$, thus we obtain

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

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

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

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

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

which proves the desired result. □

Let us now prove the second estimate on u^N given in Proposition 4.5.

Proof of Proposition 4.5. In this proof we use the fact that $L^2(\mathbb{R}^d) \subset H^{-2}(\mathbb{R}^d)$ with continuous embedding, and that the linear operator Δ is bounded from $L^2(\mathbb{R}^d)$ to $H^{-2}(\mathbb{R}^d)$. We recall that by interpolation, Proposition 4.4 and $\|u_t^N\|_{L^1(\mathbb{R}^d)} = 1$ we have

$$\mathbb{E}\|u_t^N\|_{0,2}^q \le \mathbb{E}\left[\|u_t^N\|_{0,p}^{\theta q}\,\|u_t^N\|_{L^1(\mathbb{R}^d)}^{(1-\theta)q}\right] \le \mathbb{E}\|u_t^N\|_{0,p}^{\theta q} \le C_T.$$

In view of (4.6), we first observe that

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

Thus we obtain (below the symbol '▪' denotes the variable of integration with respect to the measure μ_r^N)

$$\mathbb{E}\Big[\,\|u_t^N - u_s^N\|_{-2,2}^q\Big] \le (t-s)^{q-1}\int_s^t \mathbb{E}\Big[\|\langle \mu_r^N, \nabla V^N(\text{▪} - \cdot)\cdot F(K*u_r^N)\rangle\|_{-2,2}^q\Big]dr \tag{A.27}$$

$$+ (t-s)^{q-1}\,\frac{1}{2}\int_s^t \mathbb{E}\Big[\,\|\Delta u_r^N\|_{-2,2}^q\Big]\,dr \tag{A.28}$$

$$+\mathbb{E}\left[\left\|\frac{1}{N}\sum_{i=1}^N\int_s^t \nabla V^N\left(X_r^{i,N} - \cdot\right)\cdot dW_r^i\right\|_{-2,2}^q\right]. \tag{A.29}$$

To estimate the first term (A.27), we observe first that

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

Moreover, for the second term (A.28) we have

$$\mathbb{E}\Big[\,\|\Delta u_r^N\|_{-2,2}^q\Big] \le C\mathbb{E}\Big[\,\|u_r^N\|_{L^2(\mathbb{R}^d)}^q\Big] \le C.$$

Now, we bound the last term (A.29):

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

Then we have

$$\frac{1}{N^2}\int_{\mathbb{R}^d}\sum_{i=1}^{N}\int_s^t \left|(\mathrm{I}-\Delta)^{-1}\nabla V^N\left(X_r^{i,N}-x\right)\right|^2 drdx$$
$$= (t-s)\frac{1}{N}\left\|V^N\right\|_{-1}^2 \le (t-s)\frac{1}{N}\left\|V^N\right\|_0^2 \le CN^{d\alpha-1}(t-s) \le C(t-s),$$

since Assumption $(\mathbf{A}_\alpha)$ implies that $d\alpha - 1 < 0$. In order to conclude the lemma, we need to divide (A.27)–(A.29) by $|t-s|^{1+q\eta}$. From the previous estimates, we always get a term of the form $|t-s|^\rho$ with $\rho < 1$ (using the assumption $\eta < \frac{1}{2}$). □

References

[1] H. Bahouri, J.-Y. Chemin, and R. Danchin. *Fourier Analysis and Nonlinear Partial Differential Equations*, volume 343. Grundlehren der Mathematischen Wisserchaften, Springer Verlag, 2011.

[2] M. Bauer, T. Meyer-Brandis, and F. Proske. Strong solutions of mean-field stochastic differential equations with irregular drift. *Electron. J. Probab.*, 23:Paper No. 132, 35, 2018.

[3] P. Biler. Mathematical challenges in the theory of chemotaxis. *Ann. Math. Sil.*, 32(1):43–63, 2018.

[4] V. I. Bogachev. *Measure theory. Vol. II.* Springer-Verlag, Berlin, 2007.

[5] M. Bossy and D. Talay. A stochastic particle method for the McKean-Vlasov and the Burgers equation. *Math. Comp.*, 66(217):157–192, 1997.

[6] D. Bresch, P.-E. Jabin, and Z. Wang. On mean-field limits and quantitative estimates with a large class of singular kernels: Application to the Patlak–Keller–Segel model. *C. R. Math. Acad. Sci. Paris*, 357(9):708–720, 2019.

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

[8] J. A. Carrillo, M. Fornasier, G. Toscani, and F. Vecil. Particle, kinetic, and hydrodynamic models of swarming. In *Mathematical modeling of collective behavior in socio-economic and life sciences*, Model. Simul. Sci. Eng. Technol., pages 297–336. Birkhäuser Boston, Boston, MA, 2010.

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

[10] P. Cattiaux, A. Guillin, and F. Malrieu. Probabilistic approach for granular media equations in the non-uniformly convex case. *Probab. Theory Related Fields*, 140(1-2):19–40, 2008.

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

[12] P.-E. Chaudru de Raynal. Strong well posedness of McKean-Vlasov stochastic differential equations with Hölder drift. *Stochastic Process. Appl.*, 130(1):79–107, 2020.

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

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

[15] F. Delarue, J. Inglis, S. Rubenthaler, and E. Tanré. Particle systems with a singular mean-field self-excitation. Application to neuronal networks. *Stochastic Process. Appl.*, 125(6):2451–2492, 2015.

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

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

[18] F. Flandoli, M. Leimbach, and C. Olivera. Uniform convergence of proliferating particles to the FKPP equation. *J. Math. Anal. Appl.*, 473(1):27–52, 2019.

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

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

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

[22] N. Fournier and B. Jourdain. Stochastic particle approximation of the Keller–Segel equation and two-dimensional generalization of Bessel processes. *Ann. Appl. Probab.*, 27(5):2807–2861, 2017.

[23] N. Fournier and S. Mischler. Rate of convergence of the Nanbu particle system for hard potentials and Maxwell molecules. *Ann. Probab.*, 44(1):589–627, 2016.

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

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

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

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

[28] P.-E. Jabin and Z. Wang. Mean field limit for stochastic particle systems. In *Active particles. Vol. 1. Advances in theory, models, and applications*, Model. Simul. Sci. Eng. Technol., pages 379–402. Birkhäuser/Springer, Cham, 2017.

[29] P.-E. Jabin and Z. Wang. Quantitative estimates of propagation of chaos for stochastic systems with $W^{-1,\infty}$ kernels. *Invent. Math.*, 214(1):523–591, 2018.

[30] J.-F. Jabir, D. Talay, and M. Tomašević. Mean-field limit of a particle approximation of the one-dimensional parabolic-parabolic Keller-Segel model without smoothing. *Electron. Commun. Probab.*, 23, 2018.

[31] B. Jourdain and S. Méléard. Propagation of chaos and fluctuations for a moderate model with smooth initial data. *Ann. Inst. H. Poincaré Probab. Statist.*, 34(6):727–766, 1998.

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

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

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

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

[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.*, pages 42–95. Springer, Berlin, 1996.

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

[38] S. Méléard and S. Roelly-Coppoletta. A propagation of chaos result for a system of particles with moderate interaction. *Stochastic Process. Appl.*, 26(2):317–332, 1987.

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

[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(1–2):29–68, 2011.

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

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

[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*, 2020.

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

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

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

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

[48] T. Runst and W. Sickel. *Sobolev spaces of fractional order, Nemytskij operators, and nonlinear partial differential equations*, volume 3 of *De Gruyter Series in Nonlinear Analysis and Applications*. Walter de Gruyter & Co., Berlin, 1996.

[49] S. G. Samko, A. A. Kilbas, and O. I. Marichev. *Fractional integrals and derivatives*. Gordon and Breach Science Publishers, Yverdon, 1993. Theory and applications, Edited and with a foreword by S. M. Nikolskii, Translated from the 1987 Russian original, Revised by the authors.

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

[51] J. Simon. Compact sets in the space $L^p(0,T;B)$. *Ann. Mat. Pura Appl. (4)*, 146:65–96, 1987.

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

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

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

[55] D. Talay and M. Tomašević. A new McKean–Vlasov stochastic interpretation of the parabolic–parabolic Keller–Segel model: The one-dimensional case. *Bernoulli*, 26(2):1323–1353, 2020.

[56] C. M. Topaz, A. L. Bertozzi, and M. A. Lewis. A nonlocal continuum model for biological aggregation. *Bull. Math. Biol.*, 68(7):1601–1623, 2006.

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

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