Particle approximation of the 2-d parabolic-elliptic Keller-Segel system in the subcritical regime

Christian Olivera∗ Alexandre Richard† Milica Tomašević‡

April 8, 2020

Abstract

The parabolic-elliptic Keller-Segel partial differential equation is a two-dimensional model for chemotaxis. In this work we introduce a stochastic system of moderately interacting particles which converges, globally in time, to the solution to the Keller-Segel model in 2-d. The advantage of our approach is that we show the convergence in a strong sense for all the subcritical values of the total mass, \( M < 8\pi \).

Key words and phrases: Stochastic differential equations, Keller-Segel partial differential equation, Interacting particle systems, Analytic semigroups.

MSC2010 subject classification: 60K35, 35K55, 60H10.

1 Introduction

In this paper we study a stochastic particle approximation of the two-dimensional parabolic-elliptic Keller-Segel (KS) model that reads

\[
\begin{aligned}
\partial_t \rho(t, x) &= \Delta \rho(t, x) - \nabla (\rho(t, x) \nabla (G * \rho)(t, x)), \quad t > 0, \quad x \in \mathbb{R}^2, \\
\rho(0, x) &= \rho_0(x),
\end{aligned}
\]

(1.1)

where the convolution is with respect to the \( x \)-variable only and \( G \) is the fundamental solution of Poisson’s equation in \( \mathbb{R}^2 \), i.e.

\[
G(x) = -\frac{1}{2\pi} \log |x|^2, \quad x \in \mathbb{R}^2 \setminus \{0\}.
\]

This is a closed formulation of the following problem

\[
\begin{aligned}
\partial_t \rho(t, x) &= \Delta \rho(t, x) - \nabla (\rho(t, x) \nabla c(t, x)), \quad t > 0, \quad x \in \mathbb{R}^2, \\
\Delta c(t, x) + \rho(t, x) &= 0, \quad t > 0, \quad x \in \mathbb{R}^2, \\
\rho(0, x) &= \rho_0(x), \quad c(0, x) = 0,
\end{aligned}
\]

which describes the time evolution of the density \( \rho \) of a cell population whose motion is guided by the gradient of the concentration \( c \) of a chemical stimulus (chemo-attrac-tant). Note that the equation for the chemo-attractant concentration is in the steady state, which justifies to call (1.1) the parabolic-elliptic KS model.

System (1.1) is a special case of the general Keller-Segel model for chemotaxis [14, 15] and it has been widely studied: see for instance Horstmann [12, 13] and Perthame [20] for a thorough review of the literature up to 2000’s. For a recent review, see Biler [2] and the references therein. Noticing that (1.1) admits mass conservation, we denote \( M := \int_{\mathbb{R}^2} \rho_0(x) \, dx = \int_{\mathbb{R}^2} \rho(t, x) \, dx \). Interestingly,
the solutions of (1.1) blow up in finite time when the total mass is large. This is interpreted as a formation of agglomerations of cells in their environment. The critical value for the mass has been established and when $M > 8\pi$, the solutions blow-up in finite time. On the other hand, when $M < 8\pi$, global (in time) existence holds. For these results, see e.g. Blanchet, Dolbeault and Perthame [3], Nagai [17] and Nagai and Ogawa [18]. For more details on the blow-up phenomenon, see Herrero and Velazquez [11].

In this work, we are interested in the stochastic particle approximation of (1.1). This is a problem of noticeable difficulty due to the singularity of the interactions, which has attracted a lot of attention lately. First, Fournier and Jourdain [10] studied the following singularly interacting particle system associated to (1.1):

$$dX_t^{i,N} = \frac{1}{N} \sum_{j=1}^{N} \frac{-(X_t^{i,N} - X_t^{j,N})}{2\pi|X_t^{i,N} - X_t^{j,N}|^2} dt + \sqrt{2} dW_t^i,$$  \hspace{1cm} (1.2)

where $\{W_t^i, i \in \mathbb{N}\}$ is a family of independent standard two-dimensional Brownian motions defined on a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$. Due to the singular interaction kernel, it is not obvious that this particle system is well-defined and that the propagation of chaos holds. Nevertheless, the authors proved its well-posedness when $M < 2\pi{N^{-1}}$. In addition, when the mass $M$ is smaller than $2\pi$, they proved that any weak limit point of the empirical measure of $N$ particles is a.s. the law of the associated non-linear process of McKean-Vlasov type, whose one-dimensional time marginals satisfy (1.1). They also described complex behaviors of the particle trajectories and proved, using generalized Bessel processes, that (1.2) is well-defined until a time in which a 3-particle collision occurs. This time is infinite when $M < 8\pi{N^{-2}}$. The existence of solutions to (1.2) was also studied by Cattiaux and Pédèches [6] using Dirichlet forms, and it was proved that (1.2) is well-posed for $M \leq 8\pi{N^{-2}}$.

Another result regarding the convergence of (1.2) has been established in Bresch, Jabin and Wang [4]. Namely, the authors were interested in the convergence, when $N \to \infty$, of the joint law of $k$ fixed particles at a time $t$ towards $\rho_t^\otimes k$, where $\rho$ solves (1.1). They worked on a periodic domain $\Pi \subset \mathbb{R}^2$. Under the constraint that $M \leq 4\pi$ and the assumption that the particles are well-defined and that $\rho \in L^\infty((0,T); W^{2,\infty}(\Pi))$, they proved using new techniques of relative entropy the above convergence in $L^\infty((0,T); L^1(\Pi^k))$. Moreover, their result is quantitative, in the sense that the rate of convergence is explicit as a function of $N$.

Unlike [4, 6, 10], we present a moderately interacting system of stochastic particles (in the sense of Oelschlager [19] and Méléard and Roelly [16]). Our 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 all the subcritical values of the total mass ($M < 8\pi$). For that purpose, we follow the new approach presented in Flandoli, Leimbach and Olivera [8], based on semigroup theory and developed first with application to the FKPP equation. This technique permits to approximate nonlinear PDEs by smoothed empirical measures in strong functional topologies. It has already found many applications: Flandoli and Leocata [7] for a PDE-ODE system related to aggregation phenomena; Olivera and Simon [21] for non-local conservation laws; and Flandoli, Olivera and Simon [9] for the 2d Navier-Stokes equation. The main difficulty here will be the singular nature of the Keller-Segel equation and finding a suitable functional framework.

Thus, we consider the following particle system:

$$dX_t^{i,N} = F_A \left( \frac{1}{N} \sum_{k=1}^{N} (\nabla G \ast V^N)(X_t^{i,N} - X_t^{k,N}) \right) dt + \sqrt{2} dW_t^i, \hspace{1cm} t \leq T, \hspace{1cm} 1 \leq i \leq N,$$  \hspace{1cm} (1.3)

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 solutions (in the strong sense) for (1.3) is ensured.
Our main result is the convergence in probability, as \( N \to \infty \), of the mollified empirical measure \( \{g^N_t := V^N \ast S^N_t\}_{t \in [0,T]} \), where \( S^N \) is the empirical measure of (1.3), towards the unique mild solution to (1.1) for \( M < 8 \pi \). Suitable conditions on the initial law \( \rho_0 \) are required among which we emphasize that we work with \( \rho_0 \in L^1 \cap H^1(\mathbb{R}^2) \), for some \( \beta > 1 \), where \( H^\beta \) is a fractional Sobolev space. We prove convergence in probability in the following topologies: in the strong topology of \( \mathcal{C} ([0,T]; H^\beta_{loc}(\mathbb{R}^2)) \), for some \( \gamma \in (1, \beta) \) and in the weak topology of \( L^2 ([0,T]; H^\beta(\mathbb{R}^2)) \).

Compared to the results of [4, 6, 10], the main difference is that we start from a smoothed version of (1.2) and that we obtain the convergence for the whole range of subcritical parameter \( M \). In addition, we will compare in Subsection 2.2 the modes of convergence of the empirical measures in those works and ours.

Finally, let us briefly describe our approach and point out the main difficulties arising in this work. In the definition of the particle system (1.3) and its convergence, it is very convenient to have a bounded drift term, which is ensured by the smooth cut-off function \( F_A(x) \approx \text{sign}(x) \times (|x| \land A) \). However this implies that the particle system will not converge to the true Keller-Segel PDE (1.1) but rather to a PDE with a modified reaction term involving \( F_A(\nabla G \ast \rho) \) (see precisely Equation (2.2)). We recall here from [3] that assuming mild conditions on the initial data yields weak solutions of (1.1) in \( L^\infty_{loc}, (\varepsilon, \infty); L^p(\mathbb{R}^2) \) for any \( p \in (1, \infty) \), or from [17] that assuming an initial condition \( \rho_0 \in L^1 \cap H^1(\mathbb{R}^2) \) implies that there exists a unique local (in time) mild solution to (1.1) that belongs to \( C_A([0,T]; L^1 \cap H^1(\mathbb{R}^2)) \). In both cases, it is not clear that \( \nabla G \ast \rho \) remains bounded on \([0,T] \times \mathbb{R}^2\), for an arbitrary \( T > 0 \). Hence, we proved that if the initial condition \( \rho_0 \in L^1 \cap H^1(\mathbb{R}^2) \), the solution to the original PDE (1.1) satisfies \( \|\nabla G \ast \rho\|_{L^\infty([0,T] \times \mathbb{R}^2)} < \infty \).

Thanks to this new estimate, one can choose \( A \) larger than \( \|\nabla G \ast \rho\|_{L^\infty([0,T] \times \mathbb{R}^2)} \), and it follows that the solution to (1.1) is a solution to the PDE with cut-off.

As for the convergence of \( g^N \) (the mollified empirical measure of the particle system), we emphasize the closeness of our computations with those that are used in studying the vorticity formulation of the 2d Navier-Stokes equation in [9] (and indeed, this equation has an interaction kernel very close to KS), which rely on semigroup techniques. A new ingredient here compared to the previous literature on probabilistic KS seems to be a functional inequality of Calderón-Zygmund type for \( \nabla^2 G \). The convergence of \( g^N \) is obtained by tightness in \( \mathcal{C} ([0,T]; H^\beta_{loc}(\mathbb{R}^2)) \cap L^2_{\text{loc}} ([0,T]; H^\beta(\mathbb{R}^2)) \), where \( L^2_{\text{loc}} \) denotes the weak topology of \( L^2 \). Any limit point is then identified as a mild solution of the Keller-Segel PDE with cut-off \( F_A \). Therefore, we provide a uniqueness result for the Keller-Segel PDE with cut-off \( F_A \). Hence, in view of the discussion of the previous paragraph, it follows that for \( A \) large enough, any limit point of \( g^N \) is a mild solution to the original PDE (1.1).

**Related works and perspectives.** Probabilistic interpretation of the Keller-Segel system in its parabolic-parabolic version yields a non-linear McKean-Vlasov stochastic process proposed by Talay and Tomašević [24] and studied in the 2-d case in Tomašević [25]. The fact that the equation for the chemo-attractor concentration is not in steady state introduces a memory component in the non-linear term and the process interacts with all its past time marginal densities in a singular way. The well-posedness of the non-linear process (and the KS system) is proved under an explicit constraint on \( M \) and for \( \rho_0 \in L^1(\mathbb{R}^2) \) and \( c_0 \in H^1(\mathbb{R}^2) \).

Using the methods developed in this work in the parabolic-elliptic framework, we are currently investigating the moderately interacting particle system related to the non-linear process in [25].

Finally, we point out that Stevens [23] studied the convergence of a moderately interacting stochastic particle system towards a generalized version of the parabolic-parabolic Keller-Segel equation in \( \mathbb{R}^d \). Her particle system is slightly different than the one proposed in [24], as she considers 2 sub-populations of particles, one for the cell population and one for the chemo-attractor. Assuming, among other conditions, that the solution to the parabolic-parabolic KS system is such that \( \rho, c \in C^1_{\text{loc}}([0,T] \times \mathbb{R}^d) \cap C^0([0,T]; L^2(\mathbb{R}^d)) \), Stevens proves the convergence in probability of the regularized empirical measure of the particle system towards the solution of the Keller-Segel model in the strong topology of \( C^0([0,T]; L^2(\mathbb{R}^d)) \cap L^2([0,T]; H^1(\mathbb{R}^d)) \).
Plan of the paper. In Section 2, we present our framework in more details and state our main result, Theorem 2.3. Then we compare it to previously known results on the particle approximation of the 2d parabolic-elliptic Keller-Segel PDE. The rest of Section 2 is dedicated to stating important intermediate results and exhibiting the organisation of the proof of Theorem 2.3. In Section 3, we treat the well-posedness of (1.1) and its cut-off version: that is, we prove Theorem 2.7 about the existence of a mild solution to (1.1) in $C_b([0, T], L^1 \cap H^1(\mathbb{R}^2))$, its Corollary 3.4 that gives an explicit bound on $\|\nabla G \ast \rho\|_{L^\infty([0, T] \times \mathbb{R}^2)}$ and finally the proof of Theorem 2.9 that establishes uniqueness for the cut-off PDE. In Section 4, we develop the computations that yield the mild formulation of $g^N$ and its tightness, thus establishing Proposition 2.4. Finally, we gather in an Appendix some technical computations related to the boundedness of $g^N$ in a space that is compactly embedded in $C \left([0, T]; H^1_{\text{loc}}(\mathbb{R}^2)\right) \cap L^2_w \left([0, T]; H^2(\mathbb{R}^2)\right)$, as well as a couple of useful inequalities related to $G$, including the Calderón-Zygmund inequality for $\nabla^2 G$.

Notations and definitions. For any $\alpha \in \mathbb{R}$ and $p, \ldots$, we denote by $H^\alpha(\mathbb{R}^d)$ the Bessel potential space

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

where $\mathcal{F} u$ denotes the Fourier transform of $u$. These spaces are endowed with their norm

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

Note that

$$\|u\|_{\alpha, 2} = \|u\|_{L^2(\mathbb{R}^d)}$$

and, for any $\alpha \leq \beta$,

$$\|u\|_{\alpha, 2} \leq \|u\|_{\beta, 2}.$$ 

For positive $\alpha$ and any ball $B(0, R) \subset \mathbb{R}^d$, the space $H^\alpha(B(0, R))$ is defined in Triebel [26, p.310], and corresponds roughly to distributions $f$ on $B(0, R)$ which are restrictions of $g \in H^\alpha(\mathbb{R}^d)$. Then $H^\alpha_{\text{loc}}(\mathbb{R}^d)$ is the space of distributions $f$ on $\mathbb{R}^d$ such that $f \in H^\alpha(B(0, R))$ for any $R > 0$.

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

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

Obviously, $\nabla e^{t \Delta} f = e^{t \Delta} \nabla f$. Applying the convolution inequality [5, Th. 4.15] for $p = 2$ and using the equality $\left\| \nabla \frac{1}{4 \pi t} e^{-\frac{|x|^2}{4t}} \right\|_{L^1(\mathbb{R}^2)} = \frac{C}{\sqrt{t}}$, it comes that

$$\left\| \nabla e^{t \Delta} \right\|_{L^2 \to L^2} \leq \frac{C}{\sqrt{t-s}} \quad (1.4)$$

The space $C(I; L^1 \cap H^1(\mathbb{R}^2))$ of continuous functions from the time interval $I$ with values in $L^1 \cap H^1(\mathbb{R}^2)$ is endowed with the norm

$$\|f\|_{L^1 \cap H^1} = \sup_{s \in I} \left( \|f_s\|_{L^1(\mathbb{R}^2)} + \|f_s\|_{H^1(\mathbb{R}^2)} \right).$$

For any $t > 0$, we will also need the norm

$$\|f\|_{L^1 \cap H^1} = \sup_{s \in [0, t]} \left( \|f_s\|_{L^1(\mathbb{R}^2)} + \|f_s\|_{H^1(\mathbb{R}^2)} \right).$$

Finally, if $u$ is a function or stochastic process defined on $[0, T] \times \mathbb{R}^2$, we will most of the time use the notation $u_t$ to denote the mapping $x \mapsto u(t, x)$.

2 Main result and proof

The aim of this section is to present our main result and the organisation of its proof, whose technical details are presented in separate sections.
2.1 Statement of the theorem

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 $C^2_b(\mathbb{R})$ function such that:

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

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

$$F_A : \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} f_A(x_1) \\ f_A(x_2) \end{pmatrix}. \quad (2.1)$$

The modified Keller-Segel PDE with cut-off now reads, in closed form:

$$\begin{cases} 
\partial_t \tilde{\rho}(t, x) = \Delta \tilde{\rho}(t, x) - \nabla \cdot (\tilde{\rho}(t, x) F_A(\nabla G * \tilde{\rho}(t, x))), & t > 0, x \in \mathbb{R}^2 \\
\tilde{\rho}(0, x) = \rho_0(x).
\end{cases} \quad (2.2)$$

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

**Definition 2.1.** Given $u_0 \in L^1 \cap H^1(\mathbb{R}^2)$ and $A > 0$, a function $u$ on $[0, T) \times \mathbb{R}^2$ is said to be a mild solution to (2.2) on $[0, T)$ if

(i) $u \in C_b([0, T); L^1 \cap H^1(\mathbb{R}^2));$

(ii) $u$ satisfies the integral equation

$$u_t = e^{tA} \rho_0 - \int_0^t \nabla \cdot e^{(t-s)A} (u_s F_A(\nabla G * u_s)) \, ds, \quad 0 < t < T. \quad (2.3)$$

A function $u$ on $[0, \infty) \times \mathbb{R}^2$ is said to be a global mild solution to (2.2) if it is a mild solution to (2.2) on $[0, T)$ for all $0 < T < \infty$.

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

$$u_t = e^{tA} \rho_0 - \int_0^t \nabla \cdot e^{(t-s)A} (u_s \nabla G * u_s) \, ds, \quad 0 < t < T. \quad (2.4)$$

Compared to the singular particle system (1.2), we introduce a mollifier that will be used both to regularise the particle system and its empirical measure. Let $V : \mathbb{R}^3 \to \mathbb{R}_+$ be a smooth probability density function. For any $x \in \mathbb{R}^2$, define

$$V^N(x) := N^{2\alpha} V(N^{\alpha} x), \quad \text{for some } \alpha \in [0, 1]. \quad (2.5)$$

To cancel out the self-interaction term of a particle, we further assume that $V$ is even (hence $\nabla G * V^N$ is odd, so that $\nabla G * V^N(0) = 0$ and the self-interaction does vanish, see below).

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

$$\begin{cases} 
\frac{dX^{i,N}_t}{dt} = F_A \left( \frac{1}{N} \sum_{k=1}^N (\nabla G * V^N)(X^{i,N}_t - X^{k,N}_t) \right) + \sqrt{2} \, dW^i_t, & t \leq T, \quad 1 \leq i \leq N, \\
X^{i,N}_0, & 1 \leq i \leq N, \quad \text{i.i.d. and independent of } \{W^i_t\},
\end{cases} \quad (2.6)$$

where $\{W^i_t, \ i \in \mathbb{N} \}$ is a family of independent standard two-dimensional 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

$$S^N_t = \frac{1}{N} \sum_{i=1}^N \delta_{X^i,N}_t, \quad 0 \leq t \leq T. \quad (2.7)$$
and the mollified empirical measure by
\[ g^N := V^N * S^N. \]
The following hypotheses on the initial conditions of the system will be assumed:

(C0):

(C0) There exists \( \beta > 1 \) such that for all \( p \geq 2 \), \( \sup_{N \in \mathbb{N}} \mathbb{E} \left[ \| g^N_0 \|_{\beta,2}^p \right] < \infty \).

(C0i) Let \( \rho_0 \in L^1 \cap H^\beta (\mathbb{R}^2) \) such that \( \rho_0 \geq 0 \). Then \( \langle g^N_0, \varphi \rangle \rightarrow \langle \rho_0, \varphi \rangle \) in probability, for any \( \varphi \in C_b(\mathbb{R}^2) \).

(C0ii) The initial total mass \( M = \| \rho_0 \|_{L^1(\mathbb{R}^2)} \) satisfies \( M < 8\pi \).

(C0iii) For the parameters \( \alpha \) and \( \beta \) (which appear respectively in (2.5) and (C0i)), assume that
\[ 0 < \alpha < \frac{1}{2 + 2\beta}. \]

The main result of this paper is the following theorem. It involves a value of the cut-off \( A_0 \) which depends only on \( M \) and is given precisely in Equation (2.10), and the notion of mild solution of the PDE (1.1) which is given in Definition 2.1.

**Theorem 2.3.** Assume that the initial conditions \( \{ S^N_0 \}_{N \in \mathbb{N}} \) satisfy (C0) and that the dynamics of the particle system is given by (2.6) with \( A \) greater than \( A_0 \).
Then for any \( \gamma \in (1, \beta) \), the sequence of mollified empirical measures \( \{ g^N_t, t \in [0,T] \}_{N \in \mathbb{N}} \) converges in probability, as \( N \rightarrow \infty \), towards the unique mild solution \( \rho \) of the parabolic-elliptic Keller-Segel PDE (1.1), in the following senses:

- \( \forall \varphi \in L^2([0,T]; H^\beta(\mathbb{R}^2)) \), \( \int_0^T \langle g^N_t, \varphi \rangle_{H^\beta} \ dt \xrightarrow{p} \int_0^T \langle \rho_t, \varphi \rangle_{H^\beta} \ dt; \)
- in the strong topology of \( C([0,T]; H^\gamma_{\text{loc}}(\mathbb{R}^2)) \).

We now proceed to the proof of Theorem 2.3. We will state the most important intermediate results that are used to prove the main theorem, and refer to subsequent sections for the proof of these results.

### 2.2 Comparison with previous results

Fournier and Jourdain [10] proved a tightness – consistency result, not only on the level of time marginal laws, but on the level of the laws on the space of trajectories. In addition, they analyse fine properties of the particle trajectories and obtain existence results for the particle system. However, to prove convergence, they remain in the very subcritical case \( M < 2\pi \). On the other hand, Bresch et al. [4] are able to improve the constraint on \( M \) by passing from \( 2\pi \) to \( 4\pi \) and have a quantitative result, but the convergence is only on the level of the time marginal laws (although it is in a stronger sense). In addition, it relies on the assumptions that the particles are well defined for all \( M < 4\pi \) and all \( N \geq 1 \) and that the solution of (1.1) belongs to an appropriate functional space.

When comparing our result with the one in [10], we can only speak about the comparison of the result of convergence on the level of one-dimensional time marginals of the empirical measure. In that sense, the notion of convergence we obtain is stronger and it fills-in the gap \((2\pi, 8\pi)\) for the values of the critical parameter. However, we do analyse a smoothed version of the system and the empirical measure, which explains the stronger notion of convergence.

When comparing our result with the one in [4], notice first that their result is implies the convergence in law of the time marginals of the empirical measure of \( N \) particles, uniformly in time. In addition, it is quantitative. We do not have a quantitative estimate for the convergence and we do work with the mollified empirical measure. However, they work on periodic domains in \( \mathbb{R}^2 \), we rather work on the whole domain. To prove the convergence they suppose
\[ \rho \in L^\infty((0, T); W^{2, \infty}(\Sigma)) \]. Our procedure shows that it is enough to find the solution of (1.1) that belongs to \( \rho \in \mathcal{C}_b([0, T); L^1 \cap W^{1, 2}(\mathbb{R}^2)) \). Once again, we fill in the gap for the mass constraint that is this time [4\pi, 8\pi].

2.3 First step of the proof: convergence of \( \{g^N\} \)

First, it will be established in Section 4.2 that \( \{g^N\} \) is tight in the space

\[ \mathcal{Y} := L^2_w \left([0, T]; H^\beta(\mathbb{R}^2)\right) \cap \mathcal{C} \left([0, T]; H^\infty_{\text{loc}}(\mathbb{R}^2)\right), \tag{2.7} \]

where \( L^2_w(\mathbb{R}^2) \) denotes the \( L^2(\mathbb{R}^2) \) space endowed with the weak topology. It suffices for that to prove the boundedness of the sequence in a space that is compactly embedded in \( \mathcal{Y} \). By Prokhorov’s theorem, the tightness of \( \{g^N\} \) 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 \( g^N \) converges in law (resp. in probability, or almost surely) in \( \mathcal{Y} \): it will be understood that for any \( \varphi \in L^2 \left([0, T]; H^\beta(\mathbb{R}^2)\right) \), \( (g^N, \varphi) \) converges in law (resp. in probability or a.s.), and of course that \( g^N \) converges in law (resp. in probability or a.s.) in \( \mathcal{C} \left([0, T]; H^\infty_{\text{loc}}(\mathbb{R}^2)\right) \).

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

**Proposition 2.4.** There exists a probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \) rich enough to support \( \{g^N\}_{N \in \mathbb{N}} \) and there exists a \( \mathcal{Y} \)-valued random variable \( \xi \) defined on \( (\Omega, \mathcal{F}, \mathbb{P}) \) such that

\[ g^N \xrightarrow{\mathcal{Y}} \xi \quad \text{a.s.} \]

**Remark 2.5.** For each \( N \) and \( t \in [0, T] \), the definition of \( g^N_t \) yields that \( g^N_t \in L^\infty(\mathbb{R}^2) \), and since \( g^N_t \) is a probability density function, it is also in \( L^1(\mathbb{R}^2) \). Hence by interpolation, \( g^N_t \in \bigcap_{p=1}^{\infty} L^p(\mathbb{R}^2) \).

Now by Fatou’s lemma, one gets that \( \xi_t \in L^1(\mathbb{R}^2) \). Moreover, by Sobolev embedding in dimension 2, \( \xi_t \in H^\beta \) with \( \beta > 1 \) implies that \( \xi_t \in L^\infty(\mathbb{R}^2) \) (see e.g. [1, Thm 1.66]). Hence by interpolation, \( \xi_t \in \bigcap_{p=1}^{\infty} L^p(\mathbb{R}^2) \).

In Subsection 4.1, we will then prove that for any test function \( \varphi \), \( g^N \) satisfies the following equation

\[ \langle g^N_t, \varphi \rangle = \langle g^N_0, \varphi \rangle + \int_0^t \left( S_N^x, \nabla (V^N * \varphi) \cdot F_A(\nabla G * g^N_s) \right) \, ds \]

\[ + \frac{1}{N} \sum_{i=1}^N \int_0^t \nabla (V^N * \varphi)(X^{i,N}_s) \cdot dW^i_s + \int_0^t \langle g^N_s, \Delta \varphi \rangle \, ds. \tag{2.8} \]

In Subsection 4.3, using (2.8) and the convergence result of Proposition 2.4, we will prove that the following equality holds for all \( t > 0 \) and all \( \varphi \in C_c^\infty(\mathbb{R}^2) \),

\[ \langle \xi(t, \cdot), \varphi \rangle = \langle \rho_0, \varphi \rangle + \int_0^t \int_{\mathbb{R}^2} \xi(s, x) \nabla \varphi(x) \cdot F_A(\nabla G * \xi(s, \cdot))(x) \, dx \, ds + \int_0^t \langle \xi(s, \cdot), \Delta \varphi \rangle \, ds. \tag{2.9} \]

Observing that \( \xi \in \mathcal{Y} \), one deduces that \( \int_0^t \nabla \langle \xi(s, \cdot), F_A(\nabla G * \xi(s, \cdot)) \rangle \, ds \in L^1_{\text{loc}}(\mathbb{R}^2) \), hence the following mild formulation in distribution holds: for any \( \varphi \in C_c^\infty(\mathbb{R}^2) \),

\[ \langle \xi(t, \cdot), \varphi \rangle = \langle e^{t \Delta} \rho_0, \varphi \rangle - \int_0^t \nabla e^{(t-s)\Delta} \langle \xi(s, \cdot), F_A(\nabla G * \xi(s, \cdot)) \rangle \, ds, \varphi \rangle. \]
Notice from the above equation, that $\xi$ is non-random and that, for any $t \in (0, T)$, it satisfies almost surely in $\mathbb{R}^2$ the following equation:

$$
\xi(t, \cdot) = e^{t\Delta} \rho_0 - \int_0^t \nabla e^{(t-s)\Delta} (\xi(s, \cdot) F_A(\nabla G * \xi(s, \cdot))) \, ds.
$$

The next proposition will be useful in identifying $\xi$ as a mild solution to (1.1). Its proof is given in Subsection 4.4.

**Proposition 2.6.** Let $\xi$ be as in Proposition 2.4. Then, $\xi \in C_b((0, T); L^1 \cap H^1(\mathbb{R}^2))$.

In the next subsection, we will state the existence and uniqueness of a function $\xi \in C([0, T], L^1 \cap H^1)$ that satisfies the previous equation. Thus $\xi$ will be called a mild solution of the cut-off PDE (2.2).

### 2.4 Second step of the proof: identification of the limit by uniqueness in the cut-off PDE

In Section 3, we will study the mild solutions of the PDEs (1.1) and (2.2) with $L^1 \cap H^1(\mathbb{R}^2)$ space regularity. Although our results are close to the work of Nagai [17], it seems that they do not appear as such in the previous literature. Hence in Subsection 3.1, we will prove the following theorem for the Keller-Segel PDE (1.1).

**Theorem 2.7.** Let $\rho_0 \in L^1 \cap H^1(\mathbb{R}^2)$ be a non-negative initial data. Then, there exists a unique non-negative mild solution to (1.1) locally in time.

Assuming further that $M < 8\pi$, the non-negative mild solution of (1.1) exists globally in time.

**Remark 2.8.** In Corollary 3.4, we will obtain the following useful bound to compare the solutions of (1.1) and a solution of (2.2) for a given $A > 0$: There exists a universal constant $C > 0$ such that the unique mild solution $\rho$ of (1.1) satisfies

$$
\forall t > 0, \quad \|\nabla G * \rho_t\|_{L^\infty(\mathbb{R}^2)} \leq C(M^2 \|\rho_0\|_{L^1 \cap H^1(\mathbb{R}^2)}) \lor (\|\rho_0\|_{L^1 \cap H^1(\mathbb{R}^2)}) =: A_0.
$$

As for the Keller-Segel PDE with cut-off, we will obtain in Subsection 3.2 the following uniqueness result.

**Theorem 2.9.** Let $\rho_0 \in L^1 \cap H^1(\mathbb{R}^2)$. Then for any $A > 0$ and $F$ defined in (2.1), there is at most one mild solution to the cut-off PDE (2.2).

In Section 2.3, we have obtained that on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, $g^N$ converges almost surely in $\mathcal{Y}$ to $\xi$, which satisfies the mild formulation (2.4) of the Keller-Segel equation (Proposition 2.4 and the discussion below it). Thanks to Propostion 2.6, $\xi$ is the unique mild solution to the PDE (2.2).

Observe now that when $M < 8\pi$ and $A \geq A_0$, it follows from Equation (2.10) that a mild solution $\rho$ to (1.1) is also a mild solution to (2.2). Hence by uniqueness, $\xi = \rho$ and now, $\xi$ is a mild solution to (1.1), as claimed in Theorem 2.3.

Let us now come back to the original probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We have obtained that every subsequence of $\{g^N\}$ has a further subsequence that converges in law to $\rho$, the unique mild solution of (1.1), in $\mathcal{Y}$. Hence $g^N$ converges in law to $\rho$, and since $\rho$ is non-random, the convergence also happens in probability for the topology of $\mathcal{Y}$, which concludes the proof of Theorem 2.3.

### 3 Existence and uniqueness for KS parabolic-elliptic PDE and modified KS parabolic-elliptic PDE

In this section we first prove Theorem 2.7 combining the results obtained in Nagai [17]. We choose to work in the functional space $C_b([0, T]; L^1 \cap H^1(\mathbb{R}^2))$ for $\rho$. The latter will imply that $\nabla c$ belongs...
to the space $C_b([0, T]; L^\infty(\mathbb{R}^2))$ as seen below. This choice enables us to adapt all the techniques (described in the introduction) in order to obtain the uniform convergence of the mollified empirical measure towards the solution of the KS model.

Then, we focus on the cut-off equation (2.2) for some $A > 0$ and prove the uniqueness of mild solutions (Theorem (2.9)).

3.1 Mild solutions of the Keller-Segel PDE

We recall that the definition of a mild solution to (1.1) is given in Remark 2.2.

**Lemma 3.1.** Let $u$ be a mild solution to (1.1) on $[0, T)$. Then

$$\sup_{t \in [0, T)} \|\nabla G \ast u_t\|_{L^\infty(\mathbb{R}^2)} < \infty.$$  

**Proof.** For any $t > 0$ and any $p \in [1, \infty]$, we apply Lemma 2.5 of [17] with $q = 3$ to get

$$\|\nabla G \ast u_t\|_{L^\infty(\mathbb{R}^2)} \leq C \|u_t\|_{L^p(\mathbb{R}^2)}^{\frac{1}{2}} \|u_t\|_{L^3(\mathbb{R}^2)}^{\frac{1}{2}}. \quad (3.1)$$

By applying Inequality (20) of [5, p.280] (with $m = \frac{3}{2}$ in the notation of [5]) , one obtains

$$\|u_t\|_{L^3(\mathbb{R}^2)} \leq \frac{3}{2} \|u_t\|_{L^1(\mathbb{R}^2)} \|\nabla u_t\|_{L^2(\mathbb{R}^2)}. \quad (3.2)$$

Thus

$$\|\nabla G \ast u_t\|_{L^\infty(\mathbb{R}^2)} \leq C \|u_t\|_{L^1(\mathbb{R}^2)}^{\frac{1}{2}} \|\nabla u_t\|_{L^2(\mathbb{R}^2)}^{\frac{1}{2}} \leq C \|u\|_T \|L^1 \cap H^1\|.$$  

Having in mind the fact that $u$ is a mild solution to (1.1) and as such it belongs to $C_b([0, T); L^1 \cap H^1(\mathbb{R}^2))$, the proof is finished.

**Remark 3.2.** Let $u$ a mild solution on $(0, T)$ to (1.1). Repeat the arguments of [17, Prop. 2.4] with the following modification. Everytime one needs to control $\|u(\nabla G \ast u)\|_{L^1(\mathbb{R}^2)}$, it is possible to use the previous lemma and the fact that $u$ satisfies Definition 2.1-i). Then, one obtains that

$$\int_{\mathbb{R}^2} u_t(x) \, dx = \int_{\mathbb{R}^2} u_0(x) \, dx.$$  

Moreover, when the initial data is such that $u_0 \geq 0$ and $u_0 \not\equiv 0$, then by repeating the arguments of [17, Prop. 2.7], $u$ is such that $u(t, x) > 0$, for $(t, x) \in (0, T) \times \mathbb{R}^2$.

**Remark 3.3.** In [17], mild solutions are considered in the space $C_{1-\frac{A}{p}, T}(L^p(\mathbb{R}^2))$ of functions such that $\sup_{t \in [0, T]} t^{1-\frac{A}{p}} \|u_t\|_{L^p(\mathbb{R}^2)} < \infty$, with $p = \frac{3}{2}$. Observe that if $u \in C_b([0, T), L^1 \cap H^1(\mathbb{R}^2))$, then by (3.2), $\sup_{t \in [0, T]} \|u_t\|_{L^3(\mathbb{R}^2)}$ is finite, and by an interpolation inequality, so is $\sup_{t \in [0, T]} \|u_t\|_{L^4(\mathbb{R}^2)}$. Hence $u \in C_{1-\frac{A}{p}, T}(L^p(\mathbb{R}^2))$ for $p = \frac{3}{2}$, therefore a mild solution in the sense of Nagai [17, Def. 2.1],

It thus follows from the uniqueness result of Nagai ([17, Prop. 2.1]) that there can be at most one mild solution in the sense of Definition 2.1 (see also the Proof of Proposition 3.6).

Now we are in a position to prove the existence result given in Theorem 2.7.

**Proof of Theorem 2.7.** In view of the above remarks it only remains to discuss the existence.

The existence of a solution $\rho$ to (2.4) in the sense of [17], i.e. such that $\rho \in C_{1-\frac{A}{p}, T}(L^p(\mathbb{R}^2))$ for $p = \frac{3}{2}$ (see Remark 3.3) on some $[0, T_0]$, is given by [17, Prop. 2.6]. We need to prove that $\rho \in C([0, T_0], L^1 \cap H^1(\mathbb{R}^2))$. We rely on the explicit bounds given by Inequality (2.24) of [17], in order to get

$$\|\rho\|_{L^1 \cap H^1} \leq \|\rho_0\|_{H^1 \cap L^1} + C(t + \sqrt{t}) \|\rho\|_{L^1 \cap H^1}^2.$$
By the standard arguments of [17, Lemma 2.3] choosing $T_0$ such that
\[ 4C(T_0 + \sqrt{T_0})\|\rho_0\|_{H^1 \cap L^1(R^2)} < 1, \]
we have that
\[ \|\rho\|_{T_0, H^1 \cap L^1} \leq \frac{1 - \sqrt{1 - 4C(T_0 + \sqrt{T_0})\|\rho_0\|_{H^1 \cap L^1(R^2)}}}{2C(T_0 + \sqrt{T_0})}. \]
Choosing $T_0$ such that, for example, $4C(T_0 + \sqrt{T_0})\|\rho_0\|_{H^1 \cap L^1(R^2)} = \frac{1}{4}$, one has
\[ \|\rho\|_{T_0, H^1 \cap L^1} < 4\|\rho_0\|_{H^1 \cap L^1(R^2)}. \]  \hfill (3.3)

Now, [17, Thm. 5.2] implies the global existence in time of the solution to (2.4) that satisfies, for any $1 \leq p \leq \infty$,
\[ \|\rho_t\|_{L^p(R^2)} \leq \frac{C_p}{t^{1-\frac{1}{p}}}, \quad t > 0. \]  \hfill (3.4)

Now, noticing that $\|\rho_{t+T_0}\|_{L^2(R^2)} \leq \frac{C}{\sqrt{T_0}}$ for $t \geq 0$, it remains to control $\|\nabla \rho_{t+T_0}\|_{L^2(R^2)}$ for $t > 0$.

Simple manipulations of (2.4) lead us to
\[ \rho_{t+T_0} = e^{t\Delta} \rho_{T_0} - \int_0^t \nabla \cdot \left( e^{(t-s)\Delta} (\rho_{s+T_0}(\nabla G \ast \rho_{s+T_0})) \right) ds. \]

As $\nabla (\nabla G \ast \rho) = -\rho$, we have, in view of (1.4), that
\[ \|\nabla \rho_{t+T_0}\|_{L^2(R^2)} \leq \|\nabla \rho_{T_0}\|_{L^2(R^2)} + \int_0^t \frac{C}{\sqrt{t-s}} \|\nabla \rho_{s+T_0}\|_{L^2(R^2)} \|\nabla G \ast \rho_{s+T_0}\|_{L^\infty} ds \\
+ \int_0^t \frac{C}{\sqrt{t-s}} \|\rho_{s+T_0}\|_{L^2(R^2)} ds. \]

In view of (3.1) and (3.4), we have that $\|\nabla G \ast \rho_{s+T_0}\|_{L^\infty} \leq C \frac{M}{\sqrt{T_0}}$ and
\[ \|\rho_{s+T_0}\|_{L^2(R^2)} \leq \frac{C}{T_0} \|\rho_{s+T_0}\|_{L^2(R^2)} \leq \frac{C}{T_0 \sqrt{s}}. \]

Plugging this in the above inequality we obtain
\[ \|\nabla \rho_{t+T_0}\|_{L^2(R^2)} \leq 4\|\rho_0\|_{H^1 \cap L^1(R^2)} + \frac{C}{T_0} \beta \left( \frac{1}{2}, \frac{1}{2} \right) + C \frac{M}{\sqrt{T_0}} \int_0^t \frac{\|\nabla \rho_{s+T_0}\|_{L^2(R^2)}}{\sqrt{t-s}} ds, \]
where $\beta$ denotes the usual beta function.

Singular Gronwall’s lemma allows us to conclude that $\rho \in C_b([0,T); L^1 \cap H^1(R^2))$, for any $T > 0$. \hfill \square

**Corollary 3.4.** Let $T > 0$. Then, for any $t \in [0,T]$, one has
\[ \|\nabla G \ast \rho_t\|_{L^\infty(R^2)} \leq C(M^+ \beta \|\rho_0\|_{H^1 \cap L^1(R^2)}) \vee (\|\rho_0\|_{L^1 \cap H^1(R^2)}). \]

**Proof.** Fix $T_0$ as in the proof of Theorem 2.7 and let $t < T_0$. In view of Lemma 3.1 and (3.3), one has
\[ \|\nabla G \ast \rho_t\|_{L^\infty(R^2)} \leq 4\|\rho_0\|_{H^1 \cap L^1(R^2)}. \]

Now, let $t \in [T_0,T]$. Combine (3.1) and (3.4). It comes
\[ \|\nabla G \ast \rho_t\|_{L^\infty(R^2)} \leq \frac{CM^{1/4}}{T_0^{1/2}}. \]

Given the choice of $T_0$, one obtains the desired estimate. \hfill \square
3.2 Mild solutions of the modified Keller-Segel PDE

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

**Remark 3.5.** Let $\rho_0 \in L^1 \cap H^1(\mathbb{R}^2)$ such that $\rho_0 \geq 0$ and $\rho_0 \not\equiv 0$. Then, same arguments as in the Remark 3.2 enable us to conclude that a solution to (2.3) is non-negative and that it admits the mass conservation.

**Proposition 3.6.** Let $\rho_0 \in L^1 \cap H^1(\mathbb{R}^2)$. Then there is at most one mild solution to (2.2).

**Remark 3.7.** In view of Theorem 2.7 and Corollary 3.4, if one chooses $A > A_0$, it follows that the mild solution $\rho$ to (1.1) is a mild solution to (2.2).

We are now ready to prove Theorem 2.9 about the uniqueness of mild solutions to (2.2).

**Proof of Theorem 2.9.** Assume there are two mild solutions $\rho^1$ and $\rho^2$ to (2.2). Then,

$$\rho^1 - \rho^2 = -\int_0^t \nabla \cdot e^{(t-s)\Delta} \left\{ \rho^1_s F(\nabla (G \ast \rho^1_s)) - \rho^2_s F(\nabla (G \ast \rho^2_s)) \right\} \, ds$$

$$= -\int_0^t \nabla \cdot e^{(t-s)\Delta} \left\{ (\rho^1_s - \rho^2_s) F(\nabla (G \ast \rho^1_s)) + \rho^2_s F(\nabla (G \ast \rho^2_s)) - F(\nabla (G \ast \rho^2_s)) \right\} \, ds.$$

Hence

$$\|\rho^1 - \rho^2\|_{L^1(\mathbb{R}^2)} + \|\rho^1 - \rho^2\|_{L^2(\mathbb{R}^2)} \leq C_A \int_0^t \frac{1}{\sqrt{t-s}} \left( \|\rho^1_s - \rho^2_s\|_{L^1(\mathbb{R}^2)} + \|\rho^2_s \nabla G \ast (\rho^1_s - \rho^2_s)\|_{L^1(\mathbb{R}^2)} \right) \, ds$$

$$+ C_A \int_0^t \frac{1}{\sqrt{t-s}} \|\rho^1_s - \rho^2_s\|_{L^2(\mathbb{R}^2)} \, ds + \int_0^t \frac{C}{\sqrt{t-s}} \|\nabla G \ast (\rho^1_s - \rho^2_s)\|_{L^2(\mathbb{R}^2)} \, ds$$

$$\leq C\sqrt{t} \|\rho^1 - \rho^2\|_{\tau, L^1 \cap H^1} + C \int_0^t \frac{\|\rho^2_s\|_{L^1(\mathbb{R}^2)} + \|\rho^1_s\|_{L^2(\mathbb{R}^2)}}{\sqrt{t-s}} \|\nabla G \ast (\rho^1_s - \rho^2_s)\|_{L^\infty(\mathbb{R}^2)} \, ds. \quad (3.5)$$

As in (3.2),

$$\|\rho^1_s - \rho^2_s\|^2_{L^2(\mathbb{R}^2)} \leq \frac{3}{2}\|\rho^1_s - \rho^2_s\|^2_{L^2(\mathbb{R}^2)} \leq \frac{1}{2}\|\rho^1_s - \rho^2_s\|^2_{L^2(\mathbb{R}^2)} \leq \frac{1}{2}\|\nabla G \ast (\rho^1_s - \rho^2_s)\|^2_{L^2(\mathbb{R}^2)} \leq C\|\rho^1 - \rho^2\|_{\tau, L^1 \cap H^1}. \quad (3.6)$$

Plugging this upper bound in (3.5) gives

$$\|\rho^1 - \rho^2\|_{L^1(\mathbb{R}^2)} + \|\rho^1 - \rho^2\|_{L^2(\mathbb{R}^2)} \leq \|\rho^1 - \rho^2\|_{\tau, L^1 \cap H^1} \sqrt{t} \left( 1 + \|\rho^2\|_{\tau, L^1 \cap H^1} \right). \quad (3.7)$$

Consider now

$$\nabla (\rho^1_s - \rho^2_s) = -\int_0^t \nabla \cdot e^{(t-s)\Delta} \left\{ \nabla \cdot \{ (\rho^1_s - \rho^2_s) F(\nabla (G \ast \rho^1_s)) + \rho^2_s F(\nabla (G \ast \rho^2_s)) \} \right\} \, ds$$

$$= -\int_0^t \nabla \cdot e^{(t-s)\Delta} (\nabla (\rho^1_s - \rho^2_s) \cdot F(\nabla (G \ast \rho^1_s))) \, ds - \int_0^t \nabla \cdot e^{(t-s)\Delta} ((\rho^1_s - \rho^2_s) \nabla \cdot F(\nabla (G \ast \rho^1_s))) \, ds$$

$$- \int_0^t \nabla \cdot e^{(t-s)\Delta} (\rho^2_s (\nabla \cdot F(\nabla (G \ast \rho^1_s)) - \nabla \cdot F(\nabla (G \ast \rho^2_s)))) \, ds.$$
Thus, multiple applications of Cauchy-Schwarz inequality lead to,

\[ \| \nabla (\rho_1^2 - \rho_2^2) \|_{L^2(\mathbb{R}^2)} \leq CA \int_0^t \| \nabla e^{(t-s)A} \|_{L^2 \rightarrow L^2}(\| \nabla (\rho_1^2 - \rho_2^2) \|_{L^2(\mathbb{R}^2)} + \| (\rho_1^2 - \rho_2^2) \nabla : F(\nabla(G \ast \rho_1^2)) \|_{L^2(\mathbb{R}^2)} + \| \nabla \rho_2^2 \nabla G \ast (\rho_1^2 - \rho_2^2) \|_{L^2(\mathbb{R}^2)} + \| \rho_2^2 (\nabla \cdot F(\nabla(G \ast \rho_2^2)) - \nabla \cdot F(\nabla(G \ast \rho_1^2))) \|_{L^2(\mathbb{R}^2)}) \|_{L^2(\mathbb{R}^2)} \| \nabla : F(\nabla(G \ast \rho_1^2)) \|_{L^2(\mathbb{R}^2)} \| + \| \nabla \rho_2^2 \nabla G \ast (\rho_1^2 - \rho_2^2) \|_{L^2(\mathbb{R}^2)} + \| \rho_2^2 (\nabla \cdot F(\nabla(G \ast \rho_2^2)) - \nabla \cdot F(\nabla(G \ast \rho_1^2))) \|_{L^2(\mathbb{R}^2)} \|_{L^2(\mathbb{R}^2)} \] 

\[ \leq CA \int_0^t \frac{C}{(t-s)^{1/2}} (\| \rho_1^2 - \rho_2^2 \|_{\tau, L^1 \cap H^1} + \| \rho_1^2 - \rho_2^2 \|_{L^2(\mathbb{R}^2)} \| \nabla : F(\nabla(G \ast \rho_1^2)) \|_{L^2(\mathbb{R}^2)} + \| \nabla \rho_2^2 \|_{L^2(\mathbb{R}^2)} \| \nabla \cdot F(\nabla(G \ast \rho_2^2)) - \nabla \cdot F(\nabla(G \ast \rho_1^2))) \|_{L^2(\mathbb{R}^2)} \|_{L^2(\mathbb{R}^2)} \] 

\[ = CA \int_0^t \frac{C}{(t-s)^{1/2}} (\| \rho_1^2 - \rho_2^2 \|_{\tau, L^1 \cap H^1} + I_1(s) + I_2(s) + I_3(s)) \| ds. \]  

(3.8)

In view of (3.6), one has

\[ I_2(s) \leq C\| \rho_2^2 \|_{\tau, L^1 \cap H^1} \| \rho_1^2 - \rho_2^2 \|_{\tau, L^1 \cap H^1}. \]  

(3.9)

To treat the other terms we first notice that, for \( i = 1, 2 \),

\[ \nabla : F(\nabla(G \ast \rho_i)) = f_A'((\partial_1 G \ast \rho_i') \partial_1^2 (G \ast \rho_i) + f_A'((\partial_2 G \ast \rho_i') \partial_2^2 (G \ast \rho_i)). \]

Now, we need the following general result. For a function \( u \in C([0, T]; L^1 \cap H^1(\mathbb{R}^2)) \), [5, Cor. 9.11] (more precisely the first inequality in the proof, with \( m = N = 2 \)) implies that

\[ \| u_t \|_{L^1(\mathbb{R}^2)} \leq 2\| u \|_{L^2(\mathbb{R}^2)} \| \nabla u \|_{L^2(\mathbb{R}^2)} \leq C\| u \|_{\tau, L^1 \cap H^1}. \]  

(3.10)

Apply (3.10) for \( u = \rho_1^2 - \rho_2^2 \), (A.20) for \( p = 4 \) and again (3.10) for \( u = \rho_1^2 \). It comes

\[ I_1(s) \leq C\| \rho_1^2 - \rho_2^2 \|_{\tau, L^1 \cap H^1} \| f_A'((\partial_1 G \ast \rho_1') \partial_1^2 (G \ast \rho_1) + f_A'((\partial_2 G \ast \rho_1') \partial_2^2 (G \ast \rho_1)) \|_{L^1(\mathbb{R}^2)} \]

\[ \leq C\| \rho_1^2 - \rho_2^2 \|_{\tau, L^1 \cap H^1} \| \nabla \nabla(G \ast \rho_1') \|_{L^1(\mathbb{R}^2)} \leq C\| \rho_1^2 - \rho_2^2 \|_{\tau, L^1 \cap H^1} \| \rho_1^2 \|_{L^4(\mathbb{R}^2)} \]

\[ \leq C\| \rho_1^2 - \rho_2^2 \|_{\tau, L^1 \cap H^1} \| \rho_1^2 \|_{\tau, L^1 \cap H^1}. \]  

(3.11)

It remains to treat \( I_3(s) \). Similarly as above,

\[ \| \nabla : F(\nabla(G \ast \rho_1')) - \nabla : F(\nabla(G \ast \rho_2')) \|_{L^4(\mathbb{R}^2)} \]

\[ \leq \| f'''_A \|_{L^\infty(\mathbb{R})} (\| \nabla^3 (G \ast (\rho_1' - \rho_2')) \|_{L^1(\mathbb{R}^2)} + \| \nabla^2 (G \ast (\rho_1' - \rho_2')) \|_{L^1(\mathbb{R}^2)}) \]

\[ \leq C\| \nabla (G \ast (\rho_1' - \rho_2')) \|_{L^1(\mathbb{R}^2)} \leq C\| \rho_1^2 - \rho_2^2 \|_{\tau, L^1 \cap H^1}. \]  

(3.12)

Thus, in view of (3.10) for \( u = \rho_2^2 \), one has

\[ I_3(s) \leq C\| \rho_2^2 \|_{\tau, L^1 \cap H^1} \| \rho_1^2 - \rho_2^2 \|_{\tau, L^1 \cap H^1}. \]  

(3.13)

Thus in view of (3.11), (3.9) and (3.13), we obtain

\[ \| \nabla (\rho_1^2 - \rho_2^2) \|_{L^2(\mathbb{R}^2)} \leq C\sqrt{t} \| \rho_1^2 - \rho_2^2 \|_{\tau, L^1 \cap H^1} (1 + \| \rho_1^2 \|_{\tau, L^1 \cap H^1} + \| \rho_2^2 \|_{\tau, L^1 \cap H^1}). \]  

(3.14)

Therefore (3.7) and (3.14) yield

\[ \| \rho_1^2 - \rho_2^2 \|_{\tau, L^1 \cap H^1} \leq C\sqrt{t} \| \rho_1^2 - \rho_2^2 \|_{\tau, L^1 \cap H^1} (1 + \| \rho_1^2 \|_{\tau, L^1 \cap H^1} + \| \rho_2^2 \|_{\tau, L^1 \cap H^1}). \]

Hence for \( \tau \) small enough, we deduce that \( \| \rho_1^2 - \rho_2^2 \|_{\tau, L^1 \cap H^1} = 0 \). Therefore the uniqueness holds for mild solutions on \([0, \tau]\). Then by restarting the equation and using the same arguments as above combined with similar arguments at the end of the proof of Theorem 2.7, one gets global uniqueness.

4 Convergence of the regularised empirical measure

Recall that \( V : \mathbb{R}^2 \rightarrow \mathbb{R}_+ \) is an even smooth probability density function and that \( V^N \) is defined by Equation (2.5), that \( \{ X^i \} \) is the particle system defined by (2.6) with cutoff \( F_A \) given in (2.1). In this section, we use again the notation \( F \) instead of \( F_A \), for the sake of readability.
4.1 Equation satisfied by the regularised empirical measure: Proof of Equality (2.8)

Consider the mollified empirical measure

\[ g_{i}^{N} := V^{N} \ast S_{i}^{N} : x \in \mathbb{R}^{2} \mapsto \int_{\mathbb{R}^{2}} V^{N}(x-y) dS_{i}^{N}(y) = \frac{1}{N} \sum_{k=1}^{N} V^{N}(x-X_{i,k}^{N}). \]

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

\[ dX_{i,N}^{t} = F((\nabla G \ast g_{i}^{N})(X_{i,N}^{t})) \, dt + \sqrt{2} \, dW_{i}^{t}, \quad t \in [0,T], \quad 1 \leq i \leq N. \] \hspace{1cm} (4.1)

Fix \( x \in \mathbb{R}^{2} \) and \( 1 \leq i \leq N \). Apply Itô’s formula to the function \( V^{N}(x-\cdot) \) and the particle \( X_{i,N}^{t} \).

Then, sum for all \( 1 \leq i \leq N \) and divide by \( N \). It comes

\[ g_{i}^{N}(x) = g_{0}^{N}(x) - \frac{1}{N} \sum_{i=1}^{N} \int_{0}^{t} \nabla V^{N}(x-X_{i}^{N}) \cdot F((\nabla G \ast g_{i}^{N})(X_{i,N}^{t})) \, ds \]

\[ - \frac{1}{N} \sum_{i=1}^{N} \int_{0}^{t} \nabla V^{N}(x-X_{s}^{N}) \cdot dW_{i}^{t} + \frac{1}{N} \sum_{i=1}^{N} \int_{0}^{t} \Delta V^{N}(x-X_{i,N}^{t}) \, ds. \] \hspace{1cm} (4.2)

Notice that

\[ \frac{1}{N} \sum_{i=1}^{N} \int_{0}^{t} \nabla V^{N}(x-X_{i}^{N}) \cdot F((\nabla G \ast g_{i}^{N})(X_{i,N}^{t})) \, ds = \int_{0}^{t} \langle S_{i}^{N}, \nabla V^{N}(x-\cdot) \rangle F((\nabla G \ast g_{i}^{N})(\cdot)) \, ds \]

and

\[ \frac{1}{N} \sum_{i=1}^{N} \int_{0}^{t} \Delta V^{N}(x-X_{i}^{N}) \, ds = \int_{0}^{t} \Delta g_{i}^{N}(x) \, ds. \]

The preceding equalities combined with (4.2) and the fact that \( \nabla V^{N}(-x) = -\nabla V^{N}(x) \) (because \( V^{N} \) is even) lead to

\[ g_{i}^{N}(x) = g_{0}^{N}(x) + \int_{0}^{t} \langle S_{i}^{N}, \nabla V^{N}(\cdot-x) \rangle F((\nabla G \ast g_{i}^{N})(\cdot)) \, ds \]

\[ + \frac{1}{N} \sum_{i=1}^{N} \int_{0}^{t} \nabla V^{N}(X_{i,N}^{t}-x) \cdot dW_{i}^{t} + \int_{0}^{t} \Delta g_{i}^{N}(x) \, ds. \] \hspace{1cm} (4.3)

and for \( \varphi \in D(\mathbb{R}^{2}) \), then (4.3) implies (2.8).

For further use in Section 4.2, we also get the following mild form

\[ g_{i}^{N}(x) = e^{t} g_{0}^{N}(x) + \int_{0}^{t} e^{(t-s)} \langle S_{i}^{N}, \nabla V^{N}(\cdot-x) \rangle F((\nabla G \ast g_{i}^{N})(\cdot)) \, ds \]

\[ + \frac{1}{N} \sum_{i=1}^{N} \int_{0}^{t} e^{(t-s)} \Delta \nabla V^{N}(X_{i,N}^{t}-x) \cdot dW_{i}^{t}. \] \hspace{1cm} (4.4)

Finally, developing the scalar product, one has

\[ \langle S_{i}^{N}, \nabla V^{N}(\cdot-x) F((\nabla G \ast g_{i}^{N})(\cdot)) \rangle = \nabla_{x} \cdot \langle S_{i}^{N}, V^{N}(\cdot-x) F((\nabla G \ast g_{i}^{N})(\cdot)) \rangle. \]

Combining the latter with the fact that \( e^{t} \Delta f = \nabla e^{t} f \), (4.4) reads

\[ g_{i}^{N}(x) = e^{t} g_{0}^{N}(x) + \int_{0}^{t} e^{(t-s)} \langle S_{i}^{N}, V^{N}(\cdot-x) F((\nabla G \ast g_{i}^{N})(\cdot)) \rangle \, ds \]

\[ + \frac{1}{N} \sum_{i=1}^{N} \int_{0}^{t} e^{(t-s)} \Delta \nabla V^{N}(X_{i,N}^{t}-x) \cdot dW_{i}^{t}. \] \hspace{1cm} (4.5)
4.2 Tightness of $g^N$: Proof of Proposition 2.4

First, we point to [26, p.169, p.310] for the definition of the Sobolev spaces $W^{n,q}$. Here, we will need more particularly the space $W^{n,q}([0,T]; H^{-2}(\mathbb{R}^2))$, for $\eta > 0$ and $p > 1$, with norm given in [26, p.323]:

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

Let us now prove the tightness of $\{g^N\}$ in the space $\mathcal{Y}$ defined in (2.7). This will be achieved by proving boundedness in the following space

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

which is compactly embedded in $\mathcal{Y}$, as proven in [7] (see also Section 2.3 of [9]), if $\eta \theta \geq \frac{1-p}{p} + \frac{q}{q}$, where $\theta := \frac{\beta-\lambda}{2+\beta}$.

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

**Proposition 4.1.** Let the Assumption (C0) hold. Let $p \geq 2$. Then there exists a constant $C_{\beta,T,A,p} > 0$ such that, for all $t \in (0,T]$ and $N \in \mathbb{N}$, it holds:

$$
\mathbb{E} \left[ \left\| (I - \Delta)^{\beta/2} g_t^N \right\|_{L^2(\mathbb{R}^2)}^p \right] \leq C_{\beta,T,A,p}.
$$

**Proposition 4.2.** Let the Assumption (C0) hold. Let $\eta \in (0,\frac{1}{4})$ and $q \geq 1$. There exists a constant $C_{\beta,T,A,q} > 0$ such that, for all $N \in \mathbb{N}$, it holds:

$$
\mathbb{E} \left[ \int_0^T \int_0^T \frac{\|g_t^N - g_s^N\|^q_{L^2(\mathbb{R}^2)}}{|t - s|^{1+q\eta}} ds dt \right] \leq C_{\eta,T,A,q}.
$$

The proofs of these two results are similar to the proofs of Propositions 6 and 7 in [9] (the kernel plays no role here), but we reproduce them in Appendix for the sake of completeness.

**Remark 4.3.** We recall the following classical inequality for $\beta > 1$, based on the isometry property of $\mathcal{F}^{-1}$,

$$
\| (I - \Delta)^{\beta/2} f \|^2_{L^2(\mathbb{R}^2)} = \| (1 + |x|^2)^{\beta/2} \mathcal{F}f \|^2_{L^2(\mathbb{R}^2)} \leq \int_{\mathbb{R}^2} (1 + |x|^2)^{\beta/2} |\mathcal{F}f(x)|^2 dx = \| (I - \Delta)^{\beta/2} f \|^2_{L^2(\mathbb{R}^2)}.
$$

Hence, Proposition 4.1 implies that $\mathbb{E} \left[ \| (I - \Delta)^{\beta/2} g_t^N \|^p_{L^2(\mathbb{R}^2)} \right] \leq C_{\beta,T,A,p}$. Observing that $\| \nabla \xi \|^2_{L^2(\mathbb{R}^2)} \leq C \| (I - \Delta)^{\beta/2} \xi \|^2_{L^2(\mathbb{R}^2)}$ and by applying Fatou’s lemma, it follows that

$$
\forall t \in [0,T], \quad \| \nabla \xi \|^2_{L^2(\mathbb{R}^2)} \leq C \liminf_{n \to \infty} \mathbb{E} \left[ \| (I - \Delta)^{\beta/2} g_t^N \|^2_{L^2(\mathbb{R}^2)} \right] \leq C_{\beta,T,A,2}, \quad (4.6)
$$

since $\xi$ is deterministic. Similarly, one gets that

$$
\forall t \in [0,T], \quad \| \xi_t \|^2_{L^2(\mathbb{R}^2)} \leq C_{\beta,T,A,2}.
$$

The Chebyshev inequality ensures that

$$
\mathbb{P}(\| g_t^N \|^2_{\mathcal{Y}_0} > R) \leq \frac{\mathbb{E}[\| g_t^N \|^2_{\mathcal{Y}_0}]}{R}, \quad \text{for any } R > 0.
$$
Thus by Proposition 4.1 and Proposition 4.2 we obtain
\[ \mathbb{P}(\|g^N\|^2_{\mathcal{Y}_0} > R) \leq \frac{C}{R^2} \quad \text{for any } R > 0, N \in \mathbb{N}. \]

Let \( \mathbb{P}_N \) be the law of \( g^N \) in \( \mathcal{Y} \). The last inequality implies that there exists a bounded set \( B_\epsilon \in \mathcal{Y}_0 \) such that \( \mathbb{P}_N(B_\epsilon) < 1 - \epsilon \) for all \( N \), and therefore there exists a compact set \( K_\epsilon \in \mathcal{Y} \) such that \( \mathbb{P}_N(K_\epsilon) < 1 - \epsilon \). That is, the sequence of random variables \( \{g^N\} \) is tight in \( \mathcal{Y} \). Therefore we deduce that Proposition 2.4 holds.

### 4.3 Weak convergence to the PDE solution: Proof of Equality (2.9)

First it comes from Assumption (C0) on the initial condition that
\[ \langle g_0^N, \varphi \rangle \rightarrow \langle \rho_0, \varphi \rangle. \]

In view of Proposition 2.4, recall that \( g^N \rightarrow \xi \) in the space \( \mathcal{Y} \) which was defined in Equation (2.7).

First, it is clear that this result implies that we can pass to the limit in (2.8):
\[ \int_0^t \langle g_s^N, \Delta \varphi \rangle \, ds \rightarrow \int_0^t \langle \xi_s, \Delta \varphi \rangle \, ds, \]
and
\[ \mathbb{E} \left[ \left( \frac{1}{N} \sum_{i=1}^N \int_0^t \nabla(V^N \ast \varphi)(X_s^{i,N}) \cdot dW_s \right)^2 \right] \leq \frac{1}{N^2} \sum_{i=1}^N \int_0^t \mathbb{E} \left[ (\nabla(V^N \ast \varphi)(X_s^{i,N}))^2 \right] \, ds \rightarrow 0. \]

To conclude that \( \xi \) satisfies Equation (2.9), it remains to prove

**Lemma 4.4.** For any \( t \in [0, T] \), the following convergence happens a.s.
\[ \int_0^t \langle S_s^N, \nabla(V^N \ast \varphi) \cdot F(\nabla g^N) \rangle \, ds \rightarrow \int_0^t \int_{\mathbb{R}^2} \xi_s(x) \nabla \varphi(x) \cdot F(\nabla g \ast \xi_s)(x) \, dx \, ds. \quad (4.7) \]

**Proof.** First, let \( \epsilon > 0 \) and let \( B_\epsilon \) be a ball centred in 0, with a sufficiently large radius to ensure that
\[ \int_{\mathbb{R}^2} 1_{B_\epsilon}(y) |\nabla \varphi(y)|^2 \, dy \leq \epsilon^2. \quad (4.8) \]

In view of Proposition 2.4, one has that for all \( x \in \mathbb{R}^2 \), there is \( N \) large enough such that \( \sup_{t \in [0, T], y \in B_\epsilon} |g_s^N(x - y) - \xi_t(x - y)| \leq \epsilon \). It follows, using the Cauchy-Schwarz inequality in the second inequality and the bound (4.8) in the third, that
\[
|\nabla \ast (g_s^N - \xi_s)(x)| = |\int_{\mathbb{R}^2} 1_{B_\epsilon}(y) \nabla G(y)(g_s^N - \xi_s)(x - y) \, dy + \int_{\mathbb{R}^2} 1_{B_\epsilon}(y) \nabla G(y)(g_s^N - \xi_s)(x - y) \, dy| \\
\leq \epsilon \int_{\mathbb{R}^2} 1_{B_\epsilon}(y) |\nabla G(y)| \, dy + \left( \int_{\mathbb{R}^2} 1_{B_\epsilon}(y) |\nabla G(y)|^2 \, dy \int_{\mathbb{R}^2} (g_s^N - \xi_s)(x - y)^2 \, dy \right)^{\frac{1}{2}} \\
\leq \epsilon \left( \int_{B_\epsilon} |\nabla G(y)| \, dy + \|g_s^N - \xi_s\|_{L^2(\mathbb{R}^2)} \right).
\]

Hence it follows that for any \( s \in [0, T] \) and any \( x \in \mathbb{R}^2 \),
\[ \left( \nabla \ast g_s^N \right)(x) \rightarrow (\nabla \ast \xi_s)(x) \quad \text{a.s.} \]

Next, observe that
\[
\left| \left\langle S_s^N, \nabla(V^N \ast \varphi) \cdot F(\nabla g^N) \right\rangle - \left\langle g_s^N, \nabla(V^N \ast \varphi) \cdot F(\nabla g^N) \right\rangle \right| \\
\leq \sup_{x \in \mathbb{R}^2} \left| \nabla(V^N \ast \varphi)(x) \cdot F(\nabla g^N) - (\nabla(V^N \ast \varphi) \cdot F(\nabla g^N)) \ast V^N(x) \right|. 
\]
Using the fact that $\int_{\mathbb{R}^2} V = 1$ and $V \geq 0$, one first gets that
\begin{align*}
&\left| \nabla (V^N * \phi)(x) \cdot F(V^N * g_s^N)(x) - (\nabla (V^N * \phi) \cdot F(V^N * g_s^N)) * V^N(x) \right| \\
&\leq \int_{\mathbb{R}^2} V(y) \left| \nabla (V^N * \phi)(x) \right| \left| F(V^N * g_s^N)(x) - F(V^N * g_s^N) (x - \frac{y}{N^\alpha}) \right| \, dy \\
&\quad + \int_{\mathbb{R}^2} V(y) \left| \nabla (V^N * \phi)(x) - \nabla (V^N * \phi)(x - \frac{y}{N^\alpha}) \right| \left| F(V^N * g_s^N)(x) - F(V^N * g_s^N) (x - \frac{y}{N^\alpha}) \right| \, dy \\
&\leq C \int_{\mathbb{R}^2} V(y) \left| \nabla (V^N * \phi)(x) \right| \left| \nabla g_s^N(x) - \nabla g_s^N (x - \frac{y}{N^\alpha}) \right| \, dy \\
&\quad + \frac{C}{N^\alpha} \int_{\mathbb{R}^2} V(y) |y| \, dy
\end{align*}
where the second inequality comes using the Lipschitz continuity and boundedness of $F$. Now in view of (A.21), for some $p \in (2, \infty)$ and $\eta = 1 - \frac{2}{p}$, one has
\begin{align*}
\left| \nabla g_s^N(x) - \nabla g_s^N (x - \frac{y}{N^\alpha}) \right| &\leq \left\| \nabla g_s^N \right\|_{C^0} \left( \frac{|y|}{N^\alpha} \right)^\eta \\
&\leq C_p \left\| g_s^N \right\|_{L^p(\mathbb{R}^2)} \frac{|y|^\eta}{N^\eta \alpha^\eta}.
\end{align*}
Therefore,
\begin{align*}
&\left| \nabla (V^N * \phi)(x) \cdot F(V^N * g_s^N)(x) - (\nabla (V^N * \phi) \cdot F(V^N * g_s^N)) * V^N(x) \right| \\
&\leq C \int_{\mathbb{R}^2} V(y) \left| \nabla (V^N * \phi)(x) \right| \left\| g_s^N \right\|_{L^p(\mathbb{R}^2)} \frac{|y|^\eta}{N^\eta \alpha^\eta} \, dy + \frac{C}{N^\alpha} \int_{\mathbb{R}^2} V(y) |y| \, dy.
\end{align*}
Thus we have obtained
\begin{align*}
&\left| \nabla (V^N * \varphi)(x) \cdot F(V^N * g_s^N)(x) - (\nabla (V^N * \varphi) \cdot F(V^N * g_s^N)) * V^N(x) \right| \\
&\leq C \left( \frac{1}{N^\alpha} + \frac{\left\| g_s^N \right\|_{L^p(\mathbb{R}^2)}}{N^\eta \alpha^\eta} \right).
\end{align*}

Recall that $\{g^N\}_{N \in \mathbb{N}}$ converges almost surely in $L^2([0, T], H^\beta)$ for the weak topology, hence it is bounded in this space (by the uniform boundedness principle). Thus, $\sup_N \int_0^T \left\| g_s^N \right\|_{H^\beta}^2 \, ds < \infty$, and by interpolation inequality and Sobolev embedding, $\sup_N \int_0^T \left\| g_s^N \right\|_{L^p(\mathbb{R}^2)}^\theta \, ds < \infty$ for $\theta = 1 - \frac{1}{p}$.

It follows that $\sup_N \int_0^T \left\| g_s^N \right\|_{L^p(\mathbb{R}^2)} \, ds < \infty$ and therefore
\begin{align*}
&\lim_{N \to \infty} \int_0^t \left\langle S_s^N, \nabla (V^N * \varphi) \cdot F(V^N * g_s^N) \right\rangle \, ds = \lim_{N \to \infty} \int_0^t \left\langle g_s^N, \nabla (V^N * \varphi) \cdot F(V^N * g_s^N) \right\rangle \, ds \\
&= \lim_{N \to \infty} \int_0^t \int_{\mathbb{R}^2} g_s^N(x) \nabla (V^N * \varphi)(x) \cdot F(V^N * g_s^N)(x) \, dx ds \\
&= \int_0^t \int_{\mathbb{R}^2} \xi_s(x) \nabla \varphi(x) \cdot F(\nabla g_s^N)(x) \, dx ds
\end{align*}
where in the last equality we used that $g^N \xrightarrow{a.s.} \xi$ strongly in $L^2([0, T]; C(D))$ for $D$ the compact support of $\varphi$ (recall that $g^N$ converges a.s. in $C([0, T], H^1_{loc})$, hence by Sobolev embedding and dominated convergence, the convergence in $L^2([0, T]; C(D))$ holds). \hfill \Box

### 4.4 Time and space regularity of $\xi$: Proof of Proposition 2.6

As $\xi \in {\mathcal{Y}}$, we know that for any $t \in [0, T]$, $\xi_t \in L^1 \cap H^1(\mathbb{R}^2)$. Observe that for $p = 1, 2$ and any $t \in [0, T]$, we have $\left\| g_t^N \right\|_{L^p(\mathbb{R}^2)} \leq C_{T, p}$; indeed, for $p = 1$, this is because $g_t^N$ is a probability density function; for $p = 2$, this follows from Proposition 4.1 and is explained in Remark 4.3. Hence with Fatou’s lemma, this implies that
\begin{equation}
\sup_{t \leq T} \left( \left\| \xi_t \right\|_{L^1(\mathbb{R}^2)} + \left\| \xi_t \right\|_{L^2(\mathbb{R}^2)} \right) \leq C_T.
\end{equation}
In addition, in view of Remark 4.3, one has
\[ \sup_{t \leq T} \| \nabla \xi_t \|_{L^2(\mathbb{R}^2)} \leq C_T. \] (4.10)

It only remains to prove that for any \( t \in [0, T) \), one has
\[ \lim_{s \to t} \| \xi_t - \xi_s \|_{L^1 \cap H^1(\mathbb{R}^2)} = 0. \] (4.11)

This follows from the above properties of \( \xi \) and the mild form satisfied by \( \xi \). Namely, almost everywhere in \( \mathbb{R}^2 \), one has
\[ \xi_t = e^{(t-s)\Delta} \xi_s + \int_s^t e^{(t-r)\Delta} (\xi_r F(\nabla G * \xi_r)) \, dr. \]

To check (4.11), we need to ensure that
\[ \lim_{s \to t} \int_s^t \| e^{(t-r)\Delta} (\xi_r F(\nabla G * \xi_r)) \|_{L^1 \cap H^1(\mathbb{R}^2)} \, dr = 0. \] (4.12)

This will follow from the continuity of the integral if the integral is well-defined. For \( p = 1, 2 \), we have that
\[ \int_s^t \| e^{(t-r)\Delta} (\xi_r F(\nabla G * \xi_r)) \|_{L^p(\mathbb{R}^2)} \, dr \leq \int_s^t \frac{C_A}{\sqrt{t-s}} \| \xi_r \|_{L^p(\mathbb{R}^2)} \, dr. \]

In view of (4.9), the integral is well-defined. Now we turn to the \( H^1 \)-norm. Notice that
\[ \int_s^t \| e^{(t-r)\Delta} (\xi_r F(\nabla G * \xi_r)) \|_{H^1(\mathbb{R}^2)} \, dr \leq \int_s^t \| e^{(t-r)\Delta} (\nabla \xi_r F(\nabla G * \xi_r)) \|_{L^2(\mathbb{R}^2)} \, dr \]
\[ + \int_s^t \| \nabla e^{(t-r)\Delta} (\xi_r f' (\nabla_1 G * \xi_r)) \nabla_2 G * \xi_r \|_{L^2(\mathbb{R}^2)} \, dr =: I + II. \]

Now let us use the boundedness of \( F \), a convolution inequality and (4.10). Then it comes
\[ I \leq C_A \int_s^t \frac{1}{\sqrt{t-s}} \| \nabla \xi_r \|_2 \, dr \leq C_{T,A} \sqrt{t-s}. \]

For \( II \), we use the properties of \( f_A \), the property (A.20) of \( G \) and (4.9). It comes
\[ II \leq C_A \int_s^t \frac{1}{\sqrt{t-s}} \| \nabla_2 G * \xi_r \|_{L^2(\mathbb{R}^2)} \, dr \leq C \int_s^t \frac{1}{\sqrt{t-s}} \| \xi_r \|_{L^2(\mathbb{R}^2)} \, dr \leq C \sqrt{t-s}. \]

Hence the proof is complete.

Appendix

A.1 Proofs of technical results

Proof of Proposition 4.1. \textbf{Step 1}. Let \( H = L^2(\mathbb{R}^2) \) and let \( F \) stand for the function \( F_A \) defined in (2.1). From (4.5) after applying \( (I - \Delta)^{\beta/2} \) and by the triangular inequality we have
\[ \left\| (I - \Delta)^{\beta/2} \dot{g}^N_t \right\|_{L^p(\Omega; H)} \leq \left\| (I - \Delta)^{\beta/2} e^{t\Delta} \dot{g}^N_0 \right\|_{L^p(\Omega; H)} \]
\[ + \int_0^t \left\| (I - \Delta)^{\beta/2} \nabla e^{(t-s)\Delta} \left( V^N * (F(\nabla G * g^N_s)S^N_s) \right) \right\|_{L^p(\Omega; H)} \, ds \] (A.13)
\[ + \left\| \frac{1}{N} \sum_{i=1}^N \int_0^t (I - \Delta)^{\beta/2} \nabla e^{(t-s)\Delta} \left( V^N \left( \cdot - X^N_s \right) \right) dW^i_s \right\|_{L^p(\Omega; H)}. \] (A.14)
Step 2. Noticing that by a convolution inequality \( \| (I - \Delta)^{\beta/2} e^{t \Delta} g_0^N \|_{L^2(\mathbb{R}^2)} \leq \| e^{t \Delta} \|_{L^1 \rightarrow L^1} \| (I - \Delta)^{\beta/2} g_0^N \|_{L^2(\mathbb{R}^2)} \), one gets that the first term (A.13) can be estimated by

\[
\left\| (I - \Delta)^{\beta/2} e^{t \Delta} g_0^N \right\|_{L^p(\Omega; H)} \leq \left\| (I - \Delta)^{\beta/2} g_0^N \right\|_{L^p(\Omega; H)} \leq C_\beta,
\]

with \( C_\beta > 0 \). The boundedness of the norm of \( g_0^N \) follows from Assumption (C0).

Step 3. Let us come to the second term (A.14):

\[
\int_0^t \left\| (I - \Delta)^{\beta/2} \nabla e^{(t-s)\Delta} \left( V^N \ast (F(\nabla G \ast g_s^N)S_s^N) \right) \right\|_{L^p(\Omega; H)} ds \leq C \int_0^t \left\| e^{(t-s)\Delta} \left( V^N \ast (F(\nabla G \ast g_s^N)S_s^N) \right) \right\|_{L^p(\Omega; H)} ds.
\]

We have

\[
\left\| (I - \Delta)^{1/2} e^{((t-s)/2)\Delta} \right\|_{L^2 \rightarrow L^2} \leq \frac{C}{(t-s)^{1/2}}.
\]

On the other hand, for any \( x \in \mathbb{R}^2 \),

\[
\left| \left( V^N \ast (F(\nabla G \ast g_s^N)S_s^N) \right)(x) \right| \leq \| F(\nabla G \ast g_s^N) \|_\infty \left| V^N \ast S_s^N(x) \right| \leq A |g_s^N(x)|.
\]

By Lemma 16 in [8] we have

\[
\left\| (I - \Delta)^{\beta/2} e^{((t-s)/2)\Delta} \left[ V^N \ast (F(\nabla G \ast g_s^N)S_s^N) \right] \right\|_{L^p(\Omega; H)} \leq C_A \left\| e^{((t-s)/2)\Delta} (I - \Delta)^{\beta/2} g_s^N \right\|_{L^p(\Omega; H)} \leq C_A \left\| (I - \Delta)^{\beta/2} g_s^N \right\|_{L^p(\Omega; H)}.
\]

To summarize, we have

\[
\int_0^t \left\| (I - \Delta)^{\beta/2} \nabla e^{(t-s)\Delta} \left( V^N \ast (F(\nabla G \ast g_s^N)S_s^N) \right) \right\|_{L^p(\Omega; H)} ds \leq C_{\beta,A} \int_0^t (t-s)^{-\frac{1}{2}} \left\| (I - \Delta)^{\beta/2} g_s^N \right\|_{L^p(\Omega; H)} ds.
\]

This bounds the second term.

Step 4. For the third term (A.15), we have by Lemma 10 in [9] that for any \( \delta > 0 \), there exists \( C_{\beta,T,p,\delta} > 0 \) such that

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

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

Collecting the three bounds together, we get

\[
\left\| (I - \Delta)^{\beta/2} g_t^N \right\|_{L^p(\Omega; H)} \leq C_{\beta,T,p} + C_{\beta,A} \int_0^t (t-s)^{-\frac{1}{2}} \left\| (I - \Delta)^{\beta/2} g_s^N \right\|_{L^p(\Omega; H)} ds.
\]

We can now apply Gronwall’s Lemma to deduce

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

\[\square\]
Proof of Proposition 4.2. Let us now prove the second estimate on \( g^N \) given in Proposition 4.2. In this proof we use the fact that \( L^2(\mathbb{R}^d) \subset H_2^{-2} \) with continuous embedding, and that the linear operator \( \Delta \) is bounded from \( L^2(\mathbb{R}^d) \) to \( H_2^{-2} \).

We first observe that

\[
g^N_t(x) - g^N_s(x) = \int_s^t \langle S^N_t, (\nabla G \ast F(g^N_t)) \nabla V^N(x - \cdot) \rangle \, dr
+ \nu \int_s^t \Delta g^N_t(x) \, dr
+ \frac{1}{N} \sum_{i=1}^N \int_s^t \nabla (V^N)(x - X^i_t) \, dW^i_t.
\]

Thus we obtain

\[
\mathbb{E} \left[ \| g^N_t(x) - g^N_s(x) \|_{-2,2}^q \right] \leq (t-s)^{q-1} \int_s^t \mathbb{E} \left[ \| S^N_t, F(\nabla G \ast g^N_t) \nabla V^N(x - \cdot) \|_{-2,2}^q \right] \, dr \tag{A.16}
\]

\[
+ (t-s)^{q-1} \frac{1}{2} \int_s^t \mathbb{E} \left[ \| \Delta g^N_t(x) \|_{-2,2}^q \right] \, dr \tag{A.17}
\]

\[
+ \mathbb{E} \left[ \left\| \frac{1}{N} \sum_{i=1}^N \int_s^t \nabla (V^N)(x - X^i_t) \, dW^i_t \right\|_{-2,2}^q \right]. \tag{A.18}
\]

To estimate the first term (A.16) we observe first that

\[
\mathbb{E} \left[ \left\| \langle S^N_t, F(\nabla G \ast g^N_t) \nabla V^N(x - \cdot) \rangle \right\|_{-2,2}^q \right] = \mathbb{E} \left[ \left\| \nabla(S^N_t F(\nabla G \ast g^N_t) \ast V^N) \right\|_{-2,2}^q \right]
\leq \mathbb{E} \left[ \left\| (S^N_t F(\nabla G \ast g^N_t) \ast V^N) \right\|_{-1,2}^q \right]
\leq C_A \mathbb{E} \left[ \| g^N_t \|_{L^2(\mathbb{R}^2)}^q \right] \leq C_A.
\]

Moreover, for the second term (A.17) we have

\[
\mathbb{E} \left[ \| \Delta g^N_t \|_{-2,2}^q \right] \leq C \mathbb{E} \left[ \| g^N_t \|_{L^2(\mathbb{R}^2)}^q \right] \leq C. \tag{A.19}
\]

Now, we bound the last term (A.18):

\[
\mathbb{E} \left[ \left\| \frac{1}{N} \sum_{i=1}^N \int_s^t \nabla (V^N)(x - X^i_t) \, dW^i_t \right\|_{-2,2}^q \right]
\leq C_q \mathbb{E} \left[ \left\| \frac{1}{N^2} \sum_{i=1}^N \int_s^t \nabla (V^N)(x - X^i_t) \right\|_{-2,2} \right]^{q/2}
\]

Then we have

\[
\frac{1}{N^2} \int_{\mathbb{R}} \sum_{i=1}^N \int_s^t \left| (1 - \Delta)^{-1} \nabla (V^N)(x - X^i_t) \right|^2 \, dx \, dr
= (t-s) \frac{1}{N} \left\| V^N \right\|_{0,2}^2 \leq (t-s) \frac{1}{N} \left\| V^N \right\|_{0,2}^2 \leq CN^{2\alpha-1}(t-s) \leq C(t-s).
\]

In order to conclude the lemma, we need to divide (A.16)–(A.18) 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 \( \eta < \frac{1}{2} \)).
A.2 Properties of the Kernel $G$

Recall that $\nabla G(x) = -2(x_1/||x||^2, x_2/||x||^2)$. Then for any ball $D \subset \mathbb{R}^2$, it follows from a polar coordinate change of variables that $\nabla G \in L^p(D)$ if and only if $p \in [1, 2)$. Let $B$ denote the unit ball of $\mathbb{R}^2$. Let $p \in [1, \infty)$. Then it follows from Young’s inequality that for any $q_1 \in (\frac{2p}{p+1}, 2)$, any $q_2 \in (\frac{2p}{p+1}, \frac{2p}{p+2})$ and any $f \in L^{q_1} \cap L^{q_2}(\mathbb{R}^2)$,

$$\|\nabla G * f\|_{L^p(\mathbb{R}^2)} \leq \|(1_B \nabla G) * f\|_{L^p(\mathbb{R}^2)} + \|(1_B^c \nabla G) * f\|_{L^p(\mathbb{R}^2)}$$

$$\leq \|1_B \nabla G\|_{L^{q_1}(\mathbb{R}^2)} \|f\|_{L^{q_1}(\mathbb{R}^2)} + \|1_B^c \nabla G\|_{L^{q_2}(\mathbb{R}^2)} \|f\|_{L^{q_2}(\mathbb{R}^2)}$$

where $r_1 = 1 + \frac{1}{q_1} - \frac{1}{q_2} \in (\frac{1}{2}, 1)$ and $r_2 = 1 + \frac{1}{p} - \frac{1}{q_2} \in (0, \frac{1}{2})$, which ensures that $C_G$ is finite. Moreover,

$$(g_{ij}(x))_{i,j \in \{1, 2\}} := \nabla^2 G(x) = \left( \begin{array}{cc} -2|x|^2 + 4x_1^2 & 4x_1x_2 \\ 4x_1x_2 & -2|x|^2 + 4x_2^2 \end{array} \right)$$

is a Calderón-Zygmund operator, in the sense that it satisfies: for all $i, j \in \{1, 2\}$,

- $g_{ij}(x) = \frac{g_{ij}(x)}{|x|^2}$, where $g_{ij}^0 \in L^2(S_1)$;
- $g_{ij}^0$ is homogeneous of order 0, i.e. for any $a > 0$, $g_{ij}^0(ax) = g_{ij}^0(x), \forall x \in \mathbb{R}^2$;
- $\int_{S_1} g_{ij}^0 = 0$.

Therefore, for any $p \in (1, \infty)$, there exists $C_p > 0$ (see [22, Th. 3.1, p. 225]) such that for any $f \in L^p(\mathbb{R}^2)$,

$$\|\nabla (\nabla G * f)\|_{L^p(\mathbb{R}^2)} \leq C_p \|f\|_{L^p(\mathbb{R}^2)}.$$  \hspace{1cm} (A.20)

and one now deduces that if $f \in L^p \cap L^{q_1} \cap L^{q_2}(\mathbb{R}^2)$, then $\nabla G * f \in \mathbb{W}^{1,p}(\mathbb{R}^2)$. Hence it follows from Morrey’s inequality [5, Th. 9.12] that for $p > 2$, there exists $C_p > 0$ such that for any $f \in L^p \cap L^{q_1} \cap L^{q_2}(\mathbb{R}^2)$,

$$\|\nabla G * f\|_{C^\eta} \leq C_p \|f\|_{L^p(\mathbb{R}^2)},$$  \hspace{1cm} (A.21)

where $\eta = 1 - \frac{2}{p}$.

References

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

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

[3] A. Blanchet, J. Dolbeault, and B. Perthame. Two-dimensional Keller-Segel model: optimal critical mass and qualitative properties of the solutions. *Electron. J. Differential Equations*, 2006(44):1–33, 2006.

[4] 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.

[5] H. Brezis. *Functional analysis, Sobolev Spaces and Partial Differential Equations*. Springer, 2010.
[6] P. Cattiaux and L. Pédèches. The 2-D stochastic Keller-Segel particle model: existence and uniqueness. *ALEA*, 13:447–463, 2016.

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

[8] 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.

[9] F. Flandoli, C. Olivera, and M. Simon. Uniform approximation of 2d Navier-Stokes equation by stochastic interacting particle systems. *ArXiv preprint arXiv:2004.00458*, 2020.

[10] 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.

[11] M. A. Herrero and J. J. L. Velázquez. Singularity patterns in a chemotaxis model. *Math. Ann.*, 306(3):583–623, 1996.

[12] D. Horstmann. From 1970 until present: the Keller-Segel model in chemotaxis and its consequences. I. *Jahresber. Deutsch. Math.-Verein.*, 105(3):103–165, 2003.

[13] D. Horstmann. From 1970 until present: the Keller-Segel model in chemotaxis and its consequences. II. *Jahresber. Deutsch. Math.-Verein.*, 106(2):51–69, 2004.

[14] E. F. Keller and L. A. Segel. Initiation of slime mold aggregation viewed as an instability. *J. Theoret. Biol.*, 26(3):399–415, 1970.

[15] E. F. Keller and L. A. Segel. Model for chemotaxis. *J. Theoret. Biol.*, 30(2):225–234, 1971.

[16] S. Méleard 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.

[17] 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.

[18] T. Nagai and T. Ogawa. Global existence of solutions to a parabolic-elliptic system of drift-diffusion type in $\mathbb{R}^2$. *Funkcial. Ekvac.*, 59(1):67–112, 2016.

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

[20] B. Perthame. PDE models for chemotactic movements: parabolic, hyperbolic and kinetic. *Appl. Math.*, 49(6):539–564, 2004.

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

[22] E. M. Stein and G. Weiss. *Introduction to Fourier analysis on Euclidean spaces*, volume 32. Princeton University Press, 1971.

[23] A. Stevens. The derivation of chemotaxis equations as limit dynamics of moderately interacting stochastic many-particle systems. *SIAM J. Appl. Math.*, 61(1):183–212, 2000.

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

[25] M. Tomašević. A new McKean-Vlasov stochastic interpretation of the parabolic-parabolic Keller-Segel model: The two dimensional case. *ArXiv preprint arXiv:1902.08024*, 2020.

[26] H. Triebel. *Interpolation Theory, Function Spaces, Differential Operators*. North-Holland, 1978.