Asymptotic behavior of a system of stochastic particles subject to nonlocal interactions.

Vincenzo Capasso, Daniela Morale

Department of Mathematics, University of Milan, 20133 Milan, Italy

vincenzo.capasso@unimi.it, daniela.morale@mat.unimi.it

Abstract: In this paper we present a rigorous mathematical derivation of a macroscopic model of aggregation, scaling up from a microscopic description of a family of individuals subject to aggregation/repulsion, described by a system of Itô type stochastic differential equations. We analyze the asymptotics of the system for both a large number of particles on a bounded time interval, and its long time behavior, for a fixed number of particles. As far as this second part is concerned, we show that a suitable localizing potential is required, in order that the system may admit a non trivial invariant distribution.

Keywords: Stochastic differential equations, measure-valued processes, empirical measure, law of large numbers, invariant measure.

Mathematics Subject Classification: 60H20, 60H10, 60F99

1 Introduction

In biology and medicine there is a wide spectrum of examples which exhibit collective behavior, leading to the formation of patterns and clustering. Indeed animals may form *swarms*, characterized by a cohesive but unorganized aggregation (midges), or *schools* with a cohesive and synchronized organization (in fish schooling, individuals are oriented so that distances are uniform), or *shoals* and *flocks* in which animals are gathered together for social aims, in a synchronized or asynchronized way, or *herds, congregation*, and so on. The strong biological attention to such phenomena has stimulated a lively interest in the last decades with respect to their mathematical modelling, and simulation, aimed at grasping the basic features that lead to the observed behaviors (see e.g. [1, 2, 3, 4, 5], and literature therein). In particular, the aim of the modelling is to catch the main features of the interaction at the lower scale of single individuals that are responsible, at a larger scale, for a more complex behavior that leads to the formation of the observed aggregating patterns. However, the development of a general and coherent framework for modelling collective behavior, built from the basic stochastic processes acting at the individual level, is far from complete (see e.g. [6], and literature therein); indeed the complexity of biological organization raises nontrivial mathematical problems.

Here we report on a rigorous mathematical derivation of a macroscopic model of aggregation, scaling up from a microscopic description of a family of individuals subject to aggregation/repulsion, described by a system of Itô type stochastic differential equations. We refer to a model proposed and partially analyzed by the authors in previous papers [7, 8, 9, 10]. The basic model describes the interaction of a system of a finite number of particles subject on one hand to a "force" of aggregation depending upon a "long-ranged" nonlocal gradient of the spatial distribution of the total population; on the other hand individuals are supposed to be subject to a "force" of repulsion depending upon a "short-ranged" local gradient of the population. Both components are assumed to depend upon the empirical spatial distribution of the total population. The stochasticity is modelled by a family of independent standard Brownian motions, so that the *Lagrangian description* of the movement of a population of N individuals is given via a system of N Itô type stochastic differential equations . The state of the k-th particle, out of N, is denoted by $\{X_N^k(t)\}_{t\in\mathbb{R}_+}$, a stochastic process, defined on a suitable probability space $(\Omega, \mathcal{F}, P)$ and valued in $\left(\mathbb{R}^d, \mathcal{B}_{\mathbb{R}^d}\right)$, where $\mathcal{B}_{\mathbb{R}^d}$ is the usual Borel σ-algebra generated by intervals in $\mathbb{R}^d$. Finally, the system of SDE's is taken of the type

$$dX_N^k(t) = \mathcal{H}_N^k(X_N^1(t), \dots, X_N^N(t), t)\, dt + \sigma\, dW^k(t), \quad k = 1, \dots, N.$$

For the *Eulerian description* we refer to the time evolution of the spatial distribution of the total population, i.e. the empirical measure associated with the system of N particles

$$X_N(t) = \frac{1}{N}\sum_{k=1}^{N} \epsilon_{X_N^k(t)} \in \mathcal{M}_P(\mathbb{R}^d).$$

For a finite and small number N of individuals the empirical measure suffers significant stochastic fluctuations. But a "law of large numbers" shows how for N tending to infinity the stochastic fluctuations tend to disappear. In cited papers [7, 8, 9, 10] the authors have already investigated some aspects of the model with

$$\mathcal{H}_N^k(X_N^1(t), \ldots, X_N^N(t), t) = \nabla G * X_N(X_N^k(t) - \nabla V_N * X_N(X_N^k(t),$$

where, as we see later, G is a Vlasov long ranged kernel, and V_N is a moderate short ranged kernel. In [10] an heuristic derivation of the dynamics of a limit measure whose density ρ is a solution of a deterministic integro-differential equation describing the evolution of the mean-field spatial density of the population. In particular the derivation is performed in the case of diffusion coefficients, depending upon N, and vanishing for N tending to infinity. As a consequence the limiting PDE would be degenerate

$$\frac{\partial}{\partial t}\rho(x,t) = \nabla \cdot (\rho(x,t)\nabla\rho(x,t)) - \ \nabla \cdot [\rho(x,t)(\nabla G * \rho(\cdot,t))(x)], \quad x \in \mathbb{R}^d, t \geq 0.$$

This causes problems for the uniqueness of the solution, as shown in [11], and discussed in [12], where the authors provide conditions for the existence and uniqueness of an entropy solution. In the present paper the diffusion coefficient is not allowed to vanish, so that the regularization due to diffusion guarantees the existence and uniqueness of the solution of the limiting PDE, a fact that helps the rigorous derivation of such limit. Furthermore it becomes essential in the many steps of the proof. As discussed also in [13, 14], the drift is not bounded uniformly in N, so the standard arguments used for example in [15, 16] cannot be adopted, since they assume that their drift functions are always bounded. So that a more careful investigation is required [10, 17, 13, 14]. In the model discussed here, an additional term has been included in the drift term, describing possible intrinsic dynamics of each individual particle ; as we will see later, the particular choice of this term is fundamental in the second part of this study, where the long time behavior of the system for a fixed number N of stochastic differential equations. We first discuss the behavior of the purely interacting and diffusive system, and show that in this case the system cannot admit a nontrivial invariant distribution [18]. On the other hand, under suitable conditions on a "localizing" potential U it does admit a nontrivial invariant distribution to which the system converges; we notice that, by applying recent results by Veretennikov [19, 20], the requirements on U about its convexity are less restrictive with respect to previous literature [21, 18, 22, 19].

The paper is organized as follows. In Sections 2 and 3 the system of stochastic differential equations is presented, and an equation for the empirical measure-valued process X_N is derived. In Section 4 a main result on the relative compactness of the sequence of laws $\mathcal{L}(X_N)$ of $\{X_N(t), t \in [0, T]\}$, $N \in \mathbb{N}$, for any given $T \in \mathbb{R}_+$, is shown, which is needed for the asymptotics, with respect to N, of the evolution equation of the measure-valued process $\{X_N(t), t \in [0, T]\}$. In Section 5 the regularity of the limit measure is studied, and in particular it is shown that, for any t in a finite time interval, the limit measure is absolutely continuous with respect to the relevant Lebesgue measure; the density is characterized as the weak solution of an advection-diffusion equation. In Section 6 the long time behavior of the system for a finite given number of particle is discussed. Finally in Appendix 8 proofs of some auxiliary lemmas, and notations used throughout the paper are presented.

2 The system of interacting particles

We consider a population of constant size $N \in \mathbb{N} - \{0\}$. From the Lagrangian point of view, we assume that the "state" of the k-th particle is described by a random vector

$X_N^k(t) \in \mathbb{R}^d, t \geq 0, d \in \mathbb{N} \setminus \{0\}$. Hence, for each $k \in \{1, .., N\}$, $\{X_N^k(t), t \in \mathbb{R}_+\}$ is a stochastic process in the state space $(\mathbb{R}^d, \mathcal{B}_{\mathbb{R}^d})$, on a common probability space $(\Omega, \mathcal{F}, P)$. Notice that X_N^k may describe the spatial position, but may also describe any state of the k-th particle.

In an equivalent way we may describe the k-th particle as the (random) measure

$$\epsilon_{X_N^k(t)} \in \mathcal{M}_{\mathcal{P}}(\mathbb{R}^d),$$

where $\mathcal{M}_{\mathcal{P}}(\mathbb{R}^d)$ is the space of all probability measures on $\mathbb{R}^d$. Note that for any (sufficiently smooth) function $\varphi : \mathbb{R}^d \to \mathbb{R}$, we have $\int_{\mathbb{R}^d} \varphi(y)\, \epsilon_{X_N^k(t)}(dy) \; = \; \varphi(X_N^k(t))$.

As a consequence, the spatial distribution of the system of N particles at time t is described by the random measure on $\mathbb{R}^d$

$$X_N(t) = \frac{1}{N} \sum_{k=1}^{N} \epsilon_{X_N^k(t)} \in \mathcal{M}_P(\mathbb{R}^d). \tag{1}$$

This measure may be regarded as the empirical distribution of the location of a single particle of the system in $\mathbb{R}^d$ at time $t \in \mathbb{R}_+$.

We consider the case of a continuous time evolution, in which the time change of the individuals, apart from an advection term, is due to a stochastic individual component. The stochastic component is modelled by a family of independent standard Wiener processes $\{W^k, k = 1, \ldots\}$. Advection may be due to both an interaction dynamics among the particles (aggregation, repulsion), modelled by a functional $F_N : \mathcal{M}_{\mathcal{P}}(\mathbb{R}^d) \to C(\mathbb{R}^d, \mathbb{R}^d)$ and an individual dynamics modelled by $\Phi : \mathbb{R}^d \to \mathbb{R}$. As a consequence the dynamics of the N stochastic particles is described by a system of stochastic differential equations subject to additive noise

$$dX_N^k(t) \;\; = \;\; \left(\Phi(X_N^k(t)) + F_N\left[X_N(t)\right]\left(X_N^k(t)\right)\right) dt + \sigma dW^k(t), \quad k = 1, \ldots, N. \tag{2}$$

The measure (1) describes the system according to an Eulerian approach: the collective behavior of the discrete (in the number of particles) system is given in terms of the spatial distribution of particles at time t.

Here we specify the advection components on the basis of possible assumptions inducing self-organization of biological populations. “Social” forces are responsible for interaction of each individual with other individuals in the population within suitable neighborhoods. Additionally, the movement of each individual particle might be driven by an external information coming from the the environment, expressed via suitable potentials. These systems have been already discussed by the authors in several papers [7, 8, 10, 9]. Based on these modelling assumptions, we consider the following system of SDEs as follows

$$
\begin{aligned}
dX_N^k(t) &= \left[\gamma_1 \nabla U(X_N^k(t)) + \gamma_2 \left(\nabla \left(G - V_N\right) * X_N\right)\left(X_N^k(t)\right)\right] dt \\
&\quad + \sigma dW^k(t), \qquad k = 1, \ldots, N, \qquad (3)
\end{aligned}
$$

where $\gamma_1, \gamma_2, \sigma \in \mathbb{R}_+$.

The potential

$$U \in C_b^2(\mathbb{R}^d, \mathbb{R}_+) \qquad (4)$$

is taken as a non negative smooth even function. It satisfies the following condition [23, 19, 20]: there exist constants $M_0 \geq 0$ and $r > 0$ such that

$$\left(\nabla U(x), \frac{x}{|x|}\right) \leq -\frac{r}{|x|}, \quad |x| \geq M_0, \qquad (5)$$

where $(\cdot, \cdot)$ denotes the usual scalar product in $\mathbb{R}^d$.

We wish to assume that the interaction is composed of two components describing aggregation and repulsion, respectively. These two different ”forces”, modelled via two symmetric positive kernels

$$G \in C_b^2(\mathbb{R}^d, \mathbb{R}_+) \cap W^{1,1}(\mathbb{R}^d), \quad \text{and} \quad V_1 \in C_b^2(\mathbb{R}^d, \mathbb{R}_+), \qquad (6)$$

respectively, compete, but act at different scales. As discussed in [12, 10], aggregation, modelled by a McKean-Vlasov interaction, acts at the macroscale, via a "generalized" gradient operator

$$\begin{aligned} \left(\nabla G * X_N(t)\right)\left(X_N^k(t)\right) &= \int_{\mathbb{R}^d} \nabla G(X_N^k(t)-y)X_N(t)(dy) \\ &= \frac{1}{N}\sum_{i=1}^{N} \nabla G\left(X_N^k(t)-X_N^i(t)\right). \end{aligned}$$

We emphasize the great generality included in this definition. By using particular shapes of G, one may include possible angular ranges of sensitivity, asymmetries, etc. [2].

Repulsion acts at the mesoscale; the mesoscale is introduced as in [8, 10, 13] by rescaling a kernel V_1, chosen as a symmetric (with respect to zero) probability density

$$V_N(z) = N^{\beta} V_1(N^{\beta/d} z), \quad \beta \in (0,1). \tag{7}$$

In particular, we consider $V_1 = W_1 * W_1$, where W_1 is a function with compact support, such that

$$W_1(x) = W_1(-x); \quad W_1 \in W^{1,2}(\mathbb{R}^d), \tag{8}$$

where

$$W^{1,2}(\mathbb{R}^d) = \{f \in L^2(\mathbb{R}^d) \ : \ \int_{\mathbb{R}^d} (1+|\lambda|^2)|\tilde{f}(\lambda)|^2 d\lambda = ||f||_2^2 + ||\nabla f||_2^2 < \infty\}.$$

Here $\tilde{f}$ denotes the Fourier transform of the function f. From (8) we obtain that V_1 is Hölder continuous with exponent 2 and symmetry is required for the required symmetry of V_1. By taking $W_N(x) = N^{\beta} W_1(N^{\beta/d} x)$, we have $V_N = W_N * W_N$.

By standard arguments [24], we can prove the following

Proposition 1 *If U, G, V_1 satisfy assumptions* (4) *and* (6), *then system* (3) *admits a unique solution $X(t) = (X_N^1(t), \cdots, X_N^N(t))$ for all $t \in [0,T]$ with almost surely continuous trajectories.*

Note that the sufficient conditions for the existence and uniqueness of a solution, i.e. the Lipschitz condition and then the restriction on the growth of the coefficients of (3), derive from the fact that equation (3) is autonomous and from the assumptions on continuity and boundedness of the gradients of the kernels [25].

Different scales

The repelling force exterted on the k-th (out of N) particle located at $X_N^k(t)$ is then given by

$$\begin{aligned}(\nabla V_N * X_N(t))(X_N^k(t)) &= \int_{\mathbb{R}^d} \nabla V_N(X_N^k(t)-y)X_N(t)(dy) \\ &= \sum_{i=1}^{N} N^{\beta-1}\nabla V_1\left(N^{\beta/d}\left(X_N^k(t)-X_N^i(t)\right)\right) \\ &= \frac{N^\beta}{N}\nabla\sum_{i=1}^{N}\int_{\mathbb{R}^d} W_1\left(N^{\beta/d}X_N^k(t)-y\right)W_1\left(N^{\beta/d}X_N^i(t)-y\right)\end{aligned} \tag{9}$$

In (9) it is clear how the choice of β may determine the range and the strength of the influence of neighboring particles; indeed, any particle interacts (repelling) with $O\left(N^{1-\beta}\right)$ other particles in a volume of order $O\left(N^{-\beta}\right)$.

The idea of a spatial bound of the range over which interaction among individuals may occur has been used in previous Lagrangian models of schooling and swarming [1, 26]. On the other hand, a short range repulsion among individuals prevents their accumulation at a single point in space; this means that each individual feels the gradient of the population within a small range which decreases to zero as the size N of the population increases towards infinity.

For simulation results and comparison with experimental data, the interested reader can refer to [7, 10, 9].

In the case $\gamma_1 = 0$, the system is a purely diffusive interacting particle system. A continuum dynamics as the size of the system increases to infinite has been derived heuristically in [10]. There, the authors have considered the diffusion coefficient σ depending upon N, and have studied the cases of σ_∞ both positive (viscous case) and equal to zero (non viscous case). Here we want to carry out a rigorous mathematical derivation for the viscous case; the non viscous case requires further investigation due to regularity problems.

3 An equation for the empirical measure

First of all note that from system (3) we may get the evolution equation of the empirical measure (1). A fundamental tool for the limiting procedure is Itô's formula for the time evolution of a function $f \in C_b^{2,1}(\mathbb{R}^d \times \mathbb{R}_+)$, of the trajectory $\{X_N^k(t), t \in \mathbb{R}_+\}$ of the k-th individual subject to (3)

$$\begin{aligned} f(X_N^k(t),t) &= f(X_N^k(0),0) + \int_0^t \left[\gamma_1 \nabla U + \gamma_2 \left(\nabla\left(G - V_N\right) * X_N\right)\right](X_N^k(s)) \nabla f(X_N^k(s),s) ds \\ &+ \int_0^t \left[\frac{\partial}{\partial s} f(X_N^k(s),s) + \frac{\sigma^2}{2} \triangle f(X_N^k(s),s)\right] ds + \sigma \int_0^t \nabla f(X_N^k(s),s) dW_s. \end{aligned}$$

Hence, the evolution equation of the empirical measure (1) is, for any $f \in C_b^{2,1}(\mathbb{R}^d \times \mathbb{R}_+)$,

$$\begin{aligned} \langle X_N(t), f(\cdot,t)\rangle &= \langle X_N(0), f(\cdot,0)\rangle \\ & \int_0^t \langle X_N(s), \left[\gamma_1 \nabla U + \gamma_2 \left(\nabla\left(G - V_N\right) * X_N\right)\right](\cdot) \nabla f(\cdot,s)\rangle\, ds \\ & + \int_0^t \left\langle X_N(s), \frac{\sigma^2}{2} \triangle f(\cdot,s) + \frac{\partial}{\partial s} f(\cdot,s)\right\rangle ds \\ & + \sigma \int_0^t \langle X_N(s), \nabla f(\cdot,s)\rangle\, dW^k(s). \end{aligned} \tag{10}$$

Let us consider last term in (10),

$$M_N(f,t) = \frac{\sigma}{N} \int_0^t \sum_k \nabla f(X_N^k(s),s) dW^k(s);$$

it is a martingale with respect to the natural filtration of the process $\{X_N(t), t \in \mathbb{R}_+\}$. Hence we may apply Doob's inequality [27] to get

$$
\begin{aligned}
\mathbb{E}\left[\sup_{t\le T}|M_N(f,t)|\right]^2 &\le \mathbb{E}\left[\sup_{t\le T}|M_N(f,t)|^2\right] \qquad (11)\\
&\le 4\mathbb{E}\left[|M_N(f,T)|^2\right]\\
&= \frac{4\sigma_N^2}{N^2}\sum_{k=1}^{N}\mathbb{E}\left[\int_0^t |\nabla f(X_N^k(s),s)|^2 ds\right]\\
&\le \frac{4\sigma_N^2\|\nabla f\|_\infty^2\, T}{N}.
\end{aligned}
$$

So we have an averaged equation

$$
\begin{aligned}
\mathbb{E}\left[\langle X_N(t), f(\cdot,t)\rangle\right] &= \mathbb{E}\left[\langle X_N(0), f(\cdot,0)\rangle\right] \qquad (12)\\
&+\mathbb{E}\left[\int_0^t \langle X_N(s), \left[\gamma_1\nabla U + \gamma_2\left(\nabla\left(G - V_N\right) * X_N\right)\right](\cdot)\nabla f(\cdot,s)\rangle\, ds\right.\\
&\left.+\int_0^t \left\langle X_N(s), \frac{\sigma^2}{2}\triangle f(\cdot,s) + \frac{\partial}{\partial s}f(\cdot,s)\right\rangle ds\right]
\end{aligned}
$$

with a variance vanishing as N tends to infinity.

4 A relative compactness result

First of all, let us define the following two mollified measures

$$
\begin{aligned}
h_N(x,t) &= (W_N * X_N(t))(x); \qquad (13)\\
g_N(x,t) &= (V_N * X_N(t))(x). \qquad (14)
\end{aligned}
$$

We assume the following regularity conditions for the initial empirical measure $X_N(0)$,

$$
\sup_{N\in\mathbb{N}}\mathbb{E}\left[\int_{\mathbb{R}^d}|x|X_N(0)(dx)\right] < \infty, \qquad (15)
$$

$$\sup_{N\in\mathbb{N}} \mathbb{E}\left[\int_{\mathbb{R}^d} |h_N(x,0)|^2 dx\right] = \sup_{N\in\mathbb{N}} \mathbb{E}\left[||h_N(\cdot,0)||_2^2\right] < \infty, \tag{16}$$

For technical reasons, we impose the following restriction on β in the definition of the scaled kernel (7)

$$\beta \in \left(0, \frac{d}{d+2}\right) \tag{17}$$

Furthermore, for sake of simplicity, let $\gamma_1 = \gamma_2 = 1$ and denote by $A_N(x,t)$ the "attractor" term, i.e.

$$A_N(x,t) = U(x) + G * X_N(t)(x).$$

A main result needed for the asymptotics of the evolution equation of the measure-valued process $\{X_N(t), t \in \mathbb{R}_+\}$ is the following theorem on the properties of the sequence of laws $\mathcal{L}(X_N(t))$ of $X_N(t)$, for any t and N.

Theorem 1 *Under the hypotheses listed above, the sequence* $\{\mathcal{L}(X_N)\}_{N\in\mathbb{N}}$ *is relatively compact in the space* $\mathcal{M}_{\mathcal{P}}(C([0,T], \mathcal{M}_{\mathcal{P}}(\mathbb{R}^d)))$.

For the proof of Theorem 1, some preliminary results are needed. The main problem in the derivation of the compactness properties is due to the unboundedness of the drift term; so we have to take care of the possible explosion of the system. In particular, we need to be sure that by controlling the initial conditions (15)-(16), we could control the mollified measures $\|h_N(\cdot,t)\|_2^2$ and the space variation of the drift term. In order to deal with this problem, we define the following stopping time

$$\tau_{N,\zeta} = \inf\{t \geq 0 \ : \ |S_N(t)| > \zeta\},$$

for $\zeta > 0$, where

$$
\begin{aligned}
S_N(t) &= ||h_N(\cdot,t)||_2^2 + \int_0^t du \left(\sigma^2 ||\nabla h_N(\cdot,u)||_2^2 \right. \\
&\quad \left. + \langle X_N(s), 2(|\nabla g_N(\cdot,u)|^2 - \nabla g_N(\cdot,u)\nabla A_N(\cdot,u)\rangle \right) \\
&= ||h_N(\cdot,t)||_2^2 + D_N(t) - \int_0^t \langle X_N(u), 2|\nabla A_N(\cdot,u)|^2\rangle du. \qquad (18)
\end{aligned}
$$

In the expression of $S_N(t)$, the positive process $D_N(t)$ is defined as

$$
\begin{aligned}
D_N(t) &= \int_0^t du \left(\sigma^2 ||\nabla h_N(\cdot,u)||_2^2 + \right. \qquad (19) \\
&+ \left. \langle X_N(s), 2\left(|\nabla g_N(\cdot,u)|^2 - \nabla g_N(\cdot,u)\nabla A_N(\cdot,u) + |\nabla A_N(\cdot,u)|^2\right)\rangle \right). \qquad (20)
\end{aligned}
$$

The positivity of $D_N(t)$ derives from the positivity of the term (19).

Lemma 1 *The process*

$$
M_N(t) = S_N(t) - C\sigma^2 t N^{\beta(d+2)/d-1}
$$

is a martingale.

Proof. See section 7.

□

As a consequence, it is possible to show the non explosion of S_N in a finite time, i.e.

Proposition 2 *For any T, such that $0 < T < \infty$,*

$$
\lim_{\zeta\to\infty} \inf_{N\in\mathbb{N}} \mathbb{P}\{\tau_{N,\zeta} > T\} = 1. \qquad (21)
$$

Proof. From the martingale property of the process $M_N(t)$, it derives that the process $S_N(t)$ is a submartingale. By Doob's inequality, the almost surely continuity of the solutions of (3), assumptions (16) and(17), we have

$$
\begin{aligned}
\mathbb{P}\left\{\sup_{t\le T}|S_N(t)|>\zeta\right\} &\le \frac{1}{k}\mathbb{E}[|S_N(T)|] \le \frac{1}{k}\mathbb{E}[|M_N(T)| + C\,T\sigma^2 N^{\beta(d+2)/d-1}] \\
&= \frac{1}{k}\mathbb{E}\left[\mathbb{E}[|M_N(T)|\;|\mathcal{F}_0]\right] + C\,T\,\sigma^2 N^{\beta(d+2)/d-1} \\
&= \frac{1}{k}\mathbb{E}[|M_N(0)|] + C\,T\,\sigma^2 N^{\beta(d+2)/d-1} \\
&= \frac{1}{k}\left(\mathbb{E}[||h_N(\cdot,0)||_2^2] + C\,T\,\sigma^2 N^{\beta(d+2)/d-1}\right) \\
&\le C(T)/\zeta. \qquad (22)
\end{aligned}
$$

From inequality (22), uniform in N, we get the limit (21).

□

Note that, from inequality (22), we also obtain the control

$$
\mathbb{E}\left[\left|\,||h_N(\cdot,t)||_2^2 + \sigma^2\int_0^t ||\nabla h_N(\cdot,u)||_2^2 du \right.\right.
$$
$$
\left.\left. + \int_0^t du\langle X_N(s), 2(|\nabla g_N(\cdot,u)|^2 - \nabla g_N(\cdot,u)\nabla A_N(\cdot,u)\rangle\right|\right] < \infty.
$$

In the sequel the following lemma will be useful. For the proof see section 7.

Lemma 2 *Let ϕ be a positive function in $C^2(\mathbb{R}^d)$ such that*

$$
\phi(x) = |x| \quad \text{for } |x|\ge 1 \quad \text{and} \quad ||\nabla\phi||_\infty + ||\Delta\phi||_\infty < \infty. \qquad (23)
$$

Then for any $0\le s<t\le T$ we have

$$
\langle X_N(t),\phi\rangle + C\,D_N(t) + C\,t \quad \textit{is a submartingale} \qquad (24)
$$

and

$$
\langle X_N(t),\phi\rangle - C\,D_N(t) - C\,t \quad \textit{is a supermartingale,} \qquad (25)
$$

with $C \in \mathbb{R}_+$.

Proof of Theorem 1.

We first prove the relative compactness of the stopped process

$$X_{N,\zeta}(t) = X_N(t \wedge \tau_{N,\zeta}), \tag{26}$$

for a constant $\zeta \in \mathbb{N}$ sufficiently large. We follow the characterization of the relative compactness by Ethier-Kurtz [28]. We prove first the tightness and then the boundedness of small variations of the process, in the bounded Lipschitz metric (60), as defined in Section 8.

Tightness: For any $\epsilon > 0$ *there exists a compact* K_ϵ^ζ *in* $(\mathcal{M}_\mathcal{P}(\mathbb{R}^d), d_{BL})$ *such that*

$$\inf_{N \in \mathbb{N}} \mathbb{P}\{X_{N,\zeta}(t) \in K_\epsilon^k, \ \forall t \in [0,T]\} \geq 1 - \epsilon.$$

Let $B_\lambda^c = \{x \in \mathbb{R}^d \ : \ |x| > \lambda, \ \lambda > 1\}$; for any function ϕ which satisfies condition (23), one obtains $\langle X_{N,\zeta}(t), \phi \rangle \geq \lambda \langle X_{N,\zeta}(t), \mathbf{1}_{B_\lambda^c} \rangle$; therefore, from Lemmas 1 and 2, and conditions (15), (16)

$$\begin{aligned}
\mathbb{P}\left\{\sup_{t\le T}\langle X_{N,\zeta}(t),\mathbf{1}_{B^c_\lambda}\rangle>\delta\right\} &\le \mathbb{P}\left\{\sup_{t\le T}\langle X_{N,\zeta}(t),\phi\rangle>\lambda\delta\right\}\\
&\le \mathbb{P}\left\{\sup_{t\le T}\left(\langle X_{N,\zeta}(t),\phi\rangle+C\,D_N(t\wedge\tau_N^k)+C\,(t\wedge\tau_N^k)\right)>\lambda\delta\right\}\\
&\le \frac{1}{\lambda\delta}\mathbb{E}\left[\langle X_{N,\zeta}(T),\phi\rangle+C\,D_N(T\wedge\tau_N^k)+C\,T\wedge\tau_N^k\right]\\
&\le \frac{1}{\lambda\delta}\Big(\mathbb{E}\left[\langle X_{N,\zeta}(0),\phi\rangle\right]+2C\,\mathbb{E}\left[\mathbb{E}[D_N(T\wedge\tau_N^k)|\mathcal{F}_0]\right]\\
&\qquad +2C\,(T\wedge\tau_N^k)\Big)\\
&\le \frac{1}{\lambda\delta}\Big(\mathbb{E}\left[\langle X_{N,\zeta}(0),\phi\rangle\right]+2C\,\mathbb{E}\left[\mathbb{E}[M_N(T\wedge\tau_N^k)|\mathcal{F}_0]\right]\\
&\qquad +C\,(T\wedge\tau_N^k)-\sigma^2(T\wedge\tau_N^k)N^{\beta(d+2)/d-1}\Big)\\
&\le \frac{1}{\lambda\delta}\left(\mathbb{E}\left[\langle X_N(0),\phi\rangle\right]+\mathbb{E}\left[\|h_N(\cdot,0)\|_2^2\right]+CT\right)\\
&\le \frac{C(T)}{\lambda\delta} \qquad (27)
\end{aligned}$$

Let us now take $\epsilon>0$ and two sequences μ_i and δ_i of positive numbers such that $\sum_{i=1}^\infty\mu_i=\epsilon$ and $\delta_i\searrow 0$. Let $\lambda_i=\frac{c_5(k,T)}{\mu_i\delta_i}\to\infty$. Then (27) yields

$$\begin{aligned}
\mathbb{P}\left\{\sup_{t\le T}\langle X_{N,\zeta}(t),\mathbf{1}_{B^c_{\lambda_i}}\rangle>\delta_i,\quad\forall i\in\mathbb{N}\right\} &\le \sum_{i=1}^\infty\mathbb{P}\left\{\sup_{t\le T}\langle X_{N,\zeta}(t),\mathbf{1}_{B^c_{\lambda_i}}\rangle>\delta_i\right\}\\
&\le \sum_{i=1}^\infty\frac{c_5(\zeta,T)}{\lambda_i\delta_i}=\sum_{i=1}^\infty\mu_i=\epsilon. \qquad (28)
\end{aligned}$$

By Prohorov's Theorem [24], the set

$$K_\epsilon^k=\{\mu\in\mathcal{M}_{\mathcal{P}}(\mathbb{R}^d)\ :\ \langle\mu,\mathbf{1}_{B^c_{\lambda_i}}\rangle\le\delta_i,\quad\forall i\in\mathbb{N}\}$$

is compact in $\mathcal{M}_{\mathcal{P}}(\mathbb{R}^d)$; since

$$\mathbb{P}\left\{\sup_{t\le T}\langle X_{N,\zeta}(t),\mathbf{1}_{B^c_{\lambda_i}}\rangle>\delta_i,\forall i\in\mathbb{N}\right\}=1-\mathbb{P}\left\{\langle X_{N,\zeta}(t),\mathbf{1}_{B^c_{\lambda_i}}\rangle\le\delta_i,\forall i\in\mathbb{N},\forall t\in[0,T]\right\},$$

by (28), $\forall \epsilon > 0$ there exists a compact set $K_\epsilon^\zeta \subset \mathcal{M}_{\mathcal{P}}(\mathbb{R}^d)$ such that

$$\inf_{N\in\mathbb{N}} \mathbb{P}\{X_{N,\zeta}(t) \in K_\epsilon^\zeta,\ \forall t \in [0,T]\} \geq 1-\epsilon.$$

Small Variations: For any $0 < \delta < 1$, there exists a sequence $\{\gamma_n^T(\delta)\}_{n\in\mathbb{N}}$ of non negative random variables such that

$$\mathbb{E}\left[d_{BL}(X_{N,\zeta}(t+\delta), X_{N,\zeta}(t))^4\right] \leq \mathbb{E}\left[\gamma_n^T(\delta)\right] \quad 0 \leq t \leq T, \tag{29}$$

and

$$\lim_{\delta\to 0} \limsup_{n\to\infty} \mathbb{E}[\gamma_n^T(\delta)] = 0. \tag{30}$$

Indeed, for $0 \leq s \leq t \leq T$, we have

$$d_{BL}(X_{N,\zeta}(t), X_{N,\zeta}(s)) = \sup_{f\in\mathcal{H}_1} \frac{1}{N}\sum_{i=1}^N \left(f(X_{N,\zeta}^i(t)) - f(X_{N,\zeta}^i(s))\right) \tag{31}$$

$$\leq \frac{1}{N}\sum_{i=1}^N |X_{N,\zeta}^i(t) - X_{N,\zeta}^i(s)|$$

$$\leq \frac{1}{N}\sum_{i=1}^N \int_{s\wedge\tau_{N,\zeta}}^{t\wedge\tau_{N,\zeta}} \left|(\nabla A_N(X_N^i(u),u) - \nabla g_N(X_N^i(u),u)\right| du$$

$$+\frac{\sigma}{N}\sum_{i=1}^N \left|W^i(t\wedge\tau_{N,\zeta}) - W^i(s\wedge\tau_{N,\zeta})\right|$$

$$= \int_{s\wedge\tau_{N,\zeta}}^{t\wedge\tau_{N,\zeta}} \langle X_N(u), |\nabla A_N(\cdot,u) - \nabla g_N(\cdot,u)|\rangle du + \frac{\sigma}{N}\sum_{i=1}^N \left|W^i(t\wedge\tau_{N,\zeta}) - W^i(s\wedge\tau_{N,\zeta})\right|.$$

By the Cauchy-Schwartz and Jensen inequalities,

$$\int_{s\wedge\tau_{N,\zeta}}^{t\wedge\tau_{N,\zeta}} \langle X_N(u), |\nabla A_N(\cdot,u) - \nabla g_N(\cdot,u)|\rangle du \tag{32}$$

$$\leq (t-s)^{1/2}\left(\int_{s\wedge\tau_{N,\zeta}}^{t\wedge\tau_{N,\zeta}} \langle X_N(u), |\nabla A_N(\cdot,u) - \nabla g_N(\cdot,u)|^2\rangle du\right)^{1/2}$$

Therefore, by (18), (22), (31), and (32), we have

$$
\begin{aligned}
d_{BL}(X_{N,\zeta}(t), X_{N,\zeta}(s)) &\leq (t-s)^{1/2}\left(\int_0^{T\wedge\tau_{N,\zeta}} \langle X_N(u), |\nabla A_N(\cdot,u) - \nabla g_N(\cdot,u)|^2\rangle du\right)^{1/2} \\
&\quad + \frac{\sigma}{N}\sum_{i=1}^N \left|W^i(t\wedge\tau_{N,\zeta}) - W^i(s\wedge\tau_{N,\zeta})\right| \\
&\leq (t-s)^{1/2}\left(\int_0^{T\wedge\tau_{N,\zeta}} du\left(\sigma^2 \|\nabla h_N(\cdot, t\wedge\tau_{N,\zeta})\|_2^2\right.\right. \\
&\qquad\qquad + \langle X_N(u), 2\left(|\nabla A_N(\cdot,u)|^2\right. \\
&\qquad\qquad \left.\left. - \nabla A_N(\cdot,u)\nabla g_N(\cdot,u) + |\nabla g_N(\cdot,u)|^2\rangle\right)\right) \\
&\qquad \left. + \|h_N(\cdot, T\wedge\tau_{N,\zeta})\|_2^2\right)^{1/2} \\
&\quad + \frac{\sigma}{N}\sum_{i=1}^N \left|W^i(t\wedge\tau_{N,\zeta}) - W^i(s\wedge\tau_{N,\zeta})\right| \\
&= (t-s)^{1/2}\left(D_N(T\wedge\tau_{N,\zeta}) + \|h_N(\cdot, T\wedge\tau_{N,\zeta})\|_2^2\right)^{1/2} \\
&\quad + \frac{\sigma}{N}\sum_{i=1}^N \left|W^i(t\wedge\tau_{N,\zeta}) - W^i(s\wedge\tau_{N,\zeta})\right| \\
&\leq (t-s)^{1/2}\left(S_N(T\wedge\tau_{N,\zeta}) + \int_0^{T\wedge\tau_{N,\zeta}} \langle X_N(s), 2|\nabla A_N(\cdot,u)|^2\rangle du\right)^{1/2} \\
&\quad + \frac{\sigma}{N}\sum_{i=1}^N \left|W^i(t\wedge\tau_{N,\zeta}) - W^i(s\wedge\tau_{N,\zeta})\right| \\
&\leq \left((t-s)\left(\zeta + CT\right)\right)^{1/2} + \frac{\sigma}{N}\sum_{i=1}^N \left|W^i(t\wedge\tau_{N,\zeta}) - W^i(s\wedge\tau_{N,\zeta})\right|
\end{aligned}
$$

As a consequence, for $0 \leq s < t \leq T$,

$$
\mathbb{E}\left[d_{BL}(X_{N,\zeta}(t+\delta), X_{N,\zeta}(t))^4\right] \leq C(T)(t-s)^2.
$$

By considering the process $\gamma_N^T(\delta) = C(T)\delta^2$, we get the thesis.

Thanks to a characterization of relative compactness [28], we may then state that $\{\mathcal{L}(X_N(\cdot\wedge\tau_{N,\zeta}))\}_{N\in\mathbb{N}}$, the sequence of probability laws of the processes $\{X_N((t\wedge\tau_{N,\zeta})),\ 0 \leq t \leq T\}$ is relatively compact in $\mathcal{M}_{\mathcal{P}}(C([0,T], \mathcal{M}_{\mathcal{P}}(\mathbb{R}^d)))$, for any $\zeta > 0$. This fact, together with Proposition 2 imply the relative compactness of the full process [17], so that

Theorem 1 is proven.

□

5 About the existence and regularity of the weak limit X_∞

Theorem 1 implies the existence of a subsequence $N_k \subset \mathbb{N}$, $N_1 < N_2 < \ldots$, such that the sequence $\{\mathcal{L}(X_{N_k})\}_{k\in\mathbb{N}}$ converges in $\mathcal{M}_{\mathcal{P}}(C([0,T],\mathcal{M}_{\mathcal{P}}(\mathbb{R}^d)))$ to some limit $\mathcal{L}(X)$, which is the distribution of some process $\{X(t), t \in [0,T]\}$, with trajectories in $C([0,T],\mathcal{M}_{\mathcal{P}}(\mathbb{R}^d))$. We discuss the uniqueness of the limit later on. By now we assume the uniqueness, so that we may take $\{N_k\} = \mathbb{N}$; by Skorokhod Theorem ([29], p.9), we may assert that, corresponding to the possible unique limit law, we can also have an almost sure convergence, i.e.

$$\lim_{N\to\infty} \sup_{t\le T} d_{BL}(X_N(t), X(t)) = 0 \quad \mathbb{P}-a.s. \tag{33}$$

The next step is to study the *regularity properties of the limit measure.* Due to the bound (22), it is possible to show [13] that

$$\lim_{N,N'\to\infty} \mathbb{E}\left[\int_0^T \int_{\mathbb{R}^d} |h_N(x,t) - h_{N'}(x,t)|^2 dxdt\right] = 0.$$

Then there exits a positive (random) function h_∞ defined on $[0,T]\times\mathbb{R}^d$ such that

$$\lim_{N\to\infty} \mathbb{E}\left[\int_0^T \int_{\mathbb{R}^d} |h_N(x,t) - h_\infty(x,t)|^2 dxdt\right] = 0. \tag{34}$$

Equation (34) shows that the limit measure $X_\infty \in \mathcal{M}_P([0,T]\times\mathbb{R}^d)$ has P-a.s. a density

$$h_\infty \in L^2\left([0,T]\times\mathbb{R}^d\right) \tag{35}$$

with respect to the Lebesgue measure on $[0,T]\times\mathbb{R}^d$, i.e. for any $f\in C_b([0,T]\times\mathbb{R}^d)$

$$\int_0^T\int_{\mathbb{R}^d} f(t,x)X_\infty(dx,dt)=\int_0^T\int_{\mathbb{R}^d} f(t,x)h_\infty(t,x)(dx,dt). \tag{36}$$

By now, we do not know neither whether the measure $X_\infty(t)$ has a density for any fixed $t\in[0,T]$ or the density is deterministic.

Let us try to identify the limit by acquiring information on the *limit dynamics*. We prove the following

Proposition 3 *Let us suppose that a law of large number holds at initial time*

$$\lim_{N\to\infty}\mathcal{L}(X_N(0))=\delta_{\mu_0}\quad\text{in}\quad\mathcal{M}_{\mathcal{P}}(\mathcal{M}_{\mathcal{P}}(\mathbb{R}^d)), \tag{37}$$

where μ_0 has a density p_0 in $L^2(\mathbb{R}^d)$. Then, almost surely, for any $f\in C_b^{2,1}(\mathbb{R}^d,\mathbb{R}_+), 0\leq t\leq, T$,

$$\begin{aligned}\langle X_\infty(t),f(\cdot,t)\rangle &= \langle\mu_0,f(\cdot,0)\rangle+\int_0^t\langle h_\infty(\cdot,s),\frac{1}{2}\sigma_\infty^2\Delta f(\cdot,s)+\frac{\partial}{\partial s}f(\cdot,s)\\ &\quad+[(\nabla G_a*h_\infty(\cdot,s))(\cdot)+\nabla U(\cdot)-\nabla h_\infty(\cdot,s)]\cdot\nabla f(\cdot,s)\rangle ds.\end{aligned} \tag{38}$$

Proof.

For fixed $f\in C_b^{2,1}(\mathbb{R}^d,\mathbb{R}_+)$ and $t\in[0,T]$

$$
\begin{aligned}
\mathbb{E} \quad & \Bigg[\Bigg|\langle X_\infty(t), f(\cdot,t)\rangle - \langle \mu_0, f(\cdot,0)\rangle - \int_0^t \langle h_\infty(\cdot,s), \frac{1}{2}\sigma_\infty^2 \Delta f(\cdot,s) + \frac{\partial}{\partial s} f(\cdot,s) \\
& +[(\nabla G_a * h_\infty(\cdot,s))(\cdot) + \nabla U(\cdot) - \nabla h_\infty(\cdot,s)] \cdot \nabla f(\cdot,s)\rangle ds \ \Big|\Big] \\
\leq \quad & \mathbb{E}\left[|\langle X_\infty(t), f(\cdot,t)\rangle - \langle X_N(t), f(\cdot,t)\rangle|\right] + \mathbb{E}\left[|\langle \mu_0, f(\cdot,0)\rangle - \langle X_N(0), f(\cdot,0)\rangle|\right] \\
+ \quad & \frac{\sigma_\infty^2}{2}\mathbb{E}\left[\int_0^t |\langle -h_\infty(\cdot,s), \Delta f(\cdot,s)\rangle + \langle X_N(s), \Delta f(\cdot,s)\rangle| ds\right] \\
+ \quad & \mathbb{E}\left[\int_0^t |-\langle h_\infty(\cdot,s), \frac{\partial}{\partial s} f(\cdot,s)\rangle + \langle X_N(s), \frac{\partial}{\partial s} f(\cdot,s)\rangle| ds\right] \\
+ \quad & \mathbb{E}\left[\int_0^t |\langle h_\infty(\cdot,s), \nabla h_\infty(\cdot,s)\cdot \nabla f(\cdot,s)\rangle - \langle h_N(\cdot,s), \nabla h_N(\cdot,s)\cdot\nabla f(\cdot,s)\rangle| ds\right] \\
+ \quad & \mathbb{E}\left[\left|\int_0^t \langle h_N(\cdot,s), \nabla h_N(\cdot,s)\cdot\nabla f(\cdot,s)\rangle - \langle X_N(s), \nabla g_N \cdot \nabla f(\cdot,s)\rangle ds\right|\right] \\
+ \quad & \mathbb{E}\Bigg[\int_0^t |-\langle h_\infty(\cdot,s), [(\nabla G_a * h_\infty(\cdot,s))(\cdot) + \nabla U(\cdot)]\cdot\nabla f(\cdot,s)\rangle \\
& + \langle X_N(s), [(\nabla G_a * X_N(s))(\cdot) + \nabla U(\cdot)]\cdot \nabla f(\cdot,s)\rangle|\, ds\Big] \\
+ \quad & \mathbb{E}\left[\left|\frac{\sigma_N}{N}\int_0^t \sum_{k=1}^N \nabla f(X_N^k(s),s) dW_k(s)\right|\right] \\
+ \quad & \mathbb{E}\Bigg[\Bigg|\langle X_N(t), f(\cdot,t)\rangle - \langle X_N(0), f(\cdot,0)\rangle \\
& - \int_0^t \langle X_N(s), (\nabla G_a * X_N(s))\cdot\nabla f(\cdot,s)\rangle ds + \int_0^t \langle X_N(s), \nabla g_N(\cdot,s)\cdot\nabla f(\cdot,s)\rangle ds \\
& - \int_0^t \langle X_N(s), \nabla U(\cdot)\cdot\nabla f(\cdot,s)\rangle ds - \int_0^t \langle X_N(s), \frac{1}{2}\sigma_N^2 \Delta f(\cdot,s) + \frac{\partial}{\partial s} f(\cdot,s)\rangle ds \\
& - \frac{\sigma_N}{N}\int_0^t \sum_{k=1}^N \nabla f(X_N^k(s),s) dW_k(s)\Bigg|\Bigg] \\
:= \quad & \sum_{i=1}^9 I_N^i(t). \qquad (39)
\end{aligned}
$$

Clearly, by (34) and assumption (37) $\lim_N \sum_{i=1}^4 I_N^i(t) = 0$, by (12) $\lim_N I_N^9(t) = 0$, and by (11) $\lim_N I_N^8(t) = 0$. It remains to estimate the terms $I_N^5(t)$, $I_N^6(t)$ and $I_N^7(t)$.

$$\begin{aligned} I_N^5(t) &= \mathbb{E}\left[\int_0^t |\langle h_\infty(\cdot,s), h_\infty(\cdot,s)\Delta f(\cdot,s)\rangle - \langle h_N(\cdot,s), h_N(\cdot,s)\Delta f(\cdot,s)\rangle| ds\right] \\ &\leq ||\Delta f||_\infty \int_0^T \mathbb{E}\left[\int_{\mathbb{R}^d} |h_N(x,t) - h_\infty(x,t)||h_N(x,t) + h_\infty(x,t)| dx\right] dt \\ &\leq ||\Delta f||_\infty \left(\mathbb{E}\left[\int_0^T \int_{\mathbb{R}^d} |h_N(x,t) - h_\infty(x,t)|^2 dx dt\right]\right)^{1/2} \\ &\cdot \left(\mathbb{E}\left[\int_0^T \int_{\mathbb{R}^d} |h_N(x,t) + h_\infty(x,t)|^2 dx dt\right]\right)^{1/2}; \end{aligned}$$

by (22) and (34) we obtain

$$\lim_{N\to\infty} I_N^5(t) = 0. \tag{40}$$

By the symmetry of W_1,

$$\begin{aligned} I_N^6(t) &= \mathbb{E}\left[\left|\int_0^t \langle X_N(s), W_N * (\nabla h_N(\cdot,s) \cdot \nabla f(\cdot,s))\right.\right. \\ &\quad -(W_N * \nabla h_N(\cdot,s)) \cdot \nabla f(\cdot,s)\rangle ds|] \\ &= \mathbb{E}\left[\left|\int_0^t \left(\int_{\mathbb{R}^d} X_N(s)(dx) \int_{\mathbb{R}^d} W_N(x-y)\nabla h_N(y,s)\right.\right.\right. \\ &\quad \cdot(\nabla f(y) - \nabla f(x))dy)\, ds|]\,. \end{aligned} \tag{41}$$

By the definition of W_N and since W_1 has compact support, with $c = \mathrm{diam}(\mathrm{supp} W_1(\cdot))$ and $||D^2 f||_\infty = \sup_{i,j\leq d} ||\partial_{ij}^2||_\infty$, (41) is less than or equal to

$$\begin{aligned} & c\chi_N^{-1}||D^2 f||_\infty \mathbb{E}\left[\int_0^t \langle X_N(s) * W_N, |\nabla h_N(\cdot,s)|\rangle ds\right] \\ \leq\; & c\chi_N^{-1}||D^2 f||_\infty \left(\mathbb{E}\left[\int_0^T ||h_N(\cdot,s)||_2^2 ds\right]\right)^{1/2} \left(\mathbb{E}\left[\int_0^T ||\nabla h_N(\cdot,s)||_2^2 ds\right]\right)^{1/2} \\ \leq\; & c\chi_N^{-1}||D^2 f||_\infty. \end{aligned}$$

It follows that

$$\lim_{N\to\infty} I_N^6(t) = 0. \tag{42}$$

$$
\begin{aligned}
I_N^7(t) &= \mathbb{E}\left[\int_0^t |-\langle h_\infty(\cdot,s), [(\nabla G_a * h_\infty(\cdot,s))(\cdot) + \nabla U(\cdot)]\cdot\nabla f(\cdot,s)\rangle\right.\\
&\quad + \langle X_N(s), [(\nabla G_a * X_N(s))(\cdot) + \nabla U(\cdot)]\cdot\nabla f(\cdot,s)\rangle\\
&\quad + \langle X_N(s), [(\nabla G_a * h_\infty(\cdot,s))(\cdot) + \nabla U(\cdot)]\cdot\nabla f(\cdot,s)\rangle\\
&\quad - \langle X_N(s), [(\nabla G_a * h_\infty(\cdot,s))(\cdot) + \nabla U(\cdot)]\cdot\nabla f(\cdot,s)\rangle|\, ds]\\
&\le \mathbb{E}\left[\int_0^t |\langle X_N(s) - h_\infty(\cdot,s), [(\nabla G_a * h_\infty(\cdot,s))(\cdot) + \nabla U(\cdot)]\cdot\nabla f(\cdot,s)\rangle\,|\right.\\
&\quad + |\langle X_N(s), [(\nabla G_a * h_\infty(\cdot,s))(\cdot) - (\nabla G_a * X_N(s))(\cdot)]\cdot\nabla f(\cdot,s)\rangle|\, ds]\,.
\end{aligned}
$$

By (34) and (36) we may control also last term

$$\lim_{N\to\infty} I_N^7(t) = 0;$$

hence

$$\lim_{N\to\infty}\sum_{i=1}^{9} I_N^i(t) = 0.$$

Thus we have proven that equation (38) is true almost surely, for any $f \in C_b^{2,1}(\mathbb{R}^d, \mathbb{R}_+)$ and $t \in [0,T]$. Since $X_\infty \in \mathcal{C}([0,T], \mathcal{M}_P(\mathbb{R}^d))$ and the map

$$
\begin{aligned}
(f,t) \;\to\; & \int_0^t \langle h_\infty(\cdot,s), \frac{1}{2}\sigma_\infty^2 \Delta f(\cdot,s) + \frac{\partial}{\partial s} f(\cdot,s)\\
& + [(\nabla G_a * h_\infty(\cdot,s))(\cdot) + \nabla U(\cdot) - \nabla h_\infty(\cdot,s)]\cdot\nabla f(\cdot,s)\rangle ds
\end{aligned}
$$

is continuous, so, because of (35), the thesis follows.

□

Note that we can write (38) also in the form [17], i.e.

for any $f \in C_b^2(\mathbb{R}^d), 0 \le t \le T$,

$$\langle X_\infty(t), f\rangle = \langle \mu_0, f\rangle + \int_0^t \langle h_\infty(\cdot,s), \frac{1}{2}\left(\sigma_\infty^2 + h_\infty(\cdot,s)\right)\Delta f + A_\infty(\cdot,s)\nabla f\rangle ds, \quad (43)$$

where we have denoted by $A_\infty(x,t) = \nabla G * h_\infty(x,t) - \nabla U(x)$.

So far we have shown that any limit measure $X_\infty \in \mathcal{C}([0,T], \mathcal{M}_P(\mathbb{R}^d))$ is a solution of the equation (38), with $h_\infty \in L^2\left([0,T] \times \mathbb{R}^d\right)$, satisfying the relation (36).

We should prove that for any $t \in [0,T]$, the measure $X_N(t)$ is absolutely continuous with respect to the Lebesgue measure, so it admits a density for each t. We prove that by showing that the Fourier transform of the measure $X_N(t)$ is in L^2 for any $t \in [0,T]$, so that a density exists and the latter is also in $L^2(\mathbb{R}^d)$ and we prove that it is also L^2 unformely bounded.

Let be $f \in C_b^2(\mathbb{R}^d \times \mathbb{R}^d)$; from (43) we have

$$\begin{aligned}
&\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} X_\infty(t)(dx)X_\infty(t)(dy)f(x,y) = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} p_0(x)p_0(y)f(x,y)dxdy \qquad (44)\\
&-\frac{1}{2}\int_0^t\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} h_\infty(x,s)h_\infty(y,s)\left(\sigma^2 + h_\infty(x,s)\right)\Delta_x f(x,y)dxdyds\\
&-\frac{1}{2}\int_0^t\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} h_\infty(x,s)h_\infty(y,s)\left(\sigma^2 + h_\infty(y,s)\right)\Delta_y f(x,y)dxdyds\\
&+\int_0^t\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} h_\infty(x,s)h_\infty(y,s)\left[A_\infty(x,s)\nabla_x f(x,y) + A_\infty(y,s)\nabla_y f(x,y)\right]dxdyds
\end{aligned}$$

As proposed in [17], let us consider as test function a rescaled and smoothed version of the Green function G_d of the Laplace operator in $\mathbb{R}^d$, i.e. $f(x,y) = q_{r,\delta,\epsilon}(x-y)$, where

$$\begin{aligned}
q_{r,\delta,\epsilon}(x) &= \frac{1}{2\delta}\int_{r-\delta}^{r+\delta}\left(\int_{\mathbb{R}^d} q_\eta(x-y)\sigma_\epsilon(y)dy\right)d\eta\\
q_r(x) &= \frac{2d}{r^2}((G_d(|x|) - G_d(r)) \vee 0) = q_1(x/r)/r^d
\end{aligned}$$

Among the properties of q_r, the following are useful in the sequel

$$\begin{aligned} \widetilde{q_r}(\lambda) &\geq 0 \\ 0 &\leq \widetilde{q_{r,\delta,\epsilon}}(\lambda) \leq (2\pi)^{-d/2} \\ \lim_{\epsilon,\delta\to 0} \widetilde{q_{r,\,\delta,\,\epsilon}}(\lambda) &= \widetilde{q_r}(\lambda) \\ \lim_{r\to 0} \widetilde{q_r}(\lambda) &= (2\pi)^{-d/2} \end{aligned}$$

For a derivation of eqauation (45)-(45), please refer to [17]. Then equation (44) becomes

$$\begin{aligned} & \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} X_\infty(t)(dx)X_\infty(t)(dy)q_{r,\delta,\epsilon}(x-y) = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} p_0(x)p_0(y)q_{r,\delta,\epsilon}(x-y)dxdy \\ & +\int_0^t\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} h_\infty(x,s)h_\infty(y,s)\left(\sigma^2+h_\infty(x,s)\right)\Delta q_{r,\delta,\epsilon}(x-y)dxdyds \\ = & +\int_0^t ds\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} h_\infty(y,s)\left(\sigma^2 h_\infty(x,s)+h_\infty(x,s)^2\right) \\ & \qquad \left(\frac{1}{2\delta}\int_{r-\delta}^{r+\delta}\int_{S^{d-1}}\left(h_\infty(x+\theta\eta-y,s)-h_\infty(x-y,s)\right)\sigma_\epsilon(y)d\theta d\eta\right)dxdy \end{aligned}$$

Hence

$$\begin{aligned} & \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} X_\infty(t)(dx)X_\infty(t)(dy)q_{r,\delta,\epsilon}(x-y) \\ & -\int_0^t ds\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} h_\infty(y,s)\left(\sigma^2 h_\infty(x,s)+h_\infty(x,s)^2\right) \\ & \qquad \left(\frac{1}{2\delta}\int_{r-\delta}^{r+\delta}\int_{S^{d-1}}\left(h_\infty(x+\theta\eta-y,s)-h_\infty(x-y,s)\right)\sigma_\epsilon(y)d\theta d\eta\right)dxdy \\ = & \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} p_0(x)p_0(y)q_{r,\delta,\epsilon}(x-y)dxdy. \end{aligned} \tag{45}$$

Note that in the last equation the only contribution comes from the diffusion part of the dynamics. So we could not proceed in the estimation in the present setting, in the case the limit equation would be degenerate (not viscous case).

Let us now suppose that the density $p_0 \in C_b^2(\mathbb{R}^d)$, then as in [17], pag. 315-318, one can prove that $h_\infty \in L^3(\mathbb{R}_d \times [0,T])$,

$$\sup_{r,\delta,\epsilon>0} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} p_0(x)p_0(y)q_{r,\delta,\epsilon}(x-y)dxdy \leq \|p_0\|_2^2,$$

and that suitable subsequences $\delta_k, \epsilon_k \downarrow 0$ exist such that from (45)

$$\begin{aligned} \|p_0\|_2^2 &\geq \liminf_{r\to 0} \limsup_{k\to\infty} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} X_\infty(t)(dx)X_\infty(t)(dy)q_{r,\delta_k,\epsilon_k}(x-y) \\ &= (2\pi)^{d/2} \liminf_{r\to 0} \limsup_{k\to\infty} \int_{\mathbb{R}^d} \|\widetilde{X_\infty(\lambda)}\|^2 \widetilde{q_{r,\delta_k,\epsilon_k}}(\lambda)d\lambda \\ &\geq \int_{\mathbb{R}^d} \|\widetilde{X_\infty(\lambda)}\|^2 d\lambda. \end{aligned} \tag{46}$$

From the boundedness (46) and classical result in Fourier analysis [30], we may state that for any fixed $t \in [0,T]$ the measure $X_\infty(t)$ has a density with respect to the Lebesgue measure, and because of relation (36), we have

$$X_\infty(t) = h_\infty(\cdot,t)\nu^d, \tag{47}$$

where ν^d denote the Lebesgue measure in $\mathbb{R}^d$. Furthermore, again from (46) and (47) the density in bounded in L^2

$$\|h_\infty(\cdot,t)\|_2 \leq \|p_0\|_2.$$

So we have shown the following result

Theorem 2 *Under the hypotheses of the theorem 1, let us suppose that a law of large number exists at initial time*

$$\lim_{N\to\infty} \mathcal{L}(X_N(0)) = \delta_{\mu_0} \quad \text{in} \quad \mathcal{M}_\mathcal{P}(\mathcal{M}_\mathcal{P}(\mathbb{R}^d)), \tag{48}$$

where μ_0 has a density p_0 in $L^2(\mathbb{R}^d) \cap C_b^2(\mathbb{R}^d)$. Then, almost surely, the sequence X_N converges in law to a determinist measure X_∞. For any $t \in [0,T]$ the measure $X_N(t)$ has *a density* $h_\infty(\cdot,t)$ *such that , for any* $f \in C_b^{2,1}(\mathbb{R}^d,\mathbb{R}_+), 0 \le t \le, T,$

$$\begin{aligned}\langle h_\infty(\cdot,t), f(\cdot,t)\rangle &= \langle \mu_0, f(\cdot,0)\rangle + \int_0^t \langle h_\infty(\cdot,s), \frac{1}{2}\sigma_\infty^2 \Delta f(\cdot,s) + \frac{\partial}{\partial s} f(\cdot,s) \\ &\quad + [(\nabla G_a * h_\infty(\cdot,s))(\cdot) + \nabla U(\cdot) - \nabla h_\infty(\cdot,s)] \cdot \nabla f(\cdot,s)\rangle ds.\end{aligned} \tag{49}$$

One can easily see that equation (49) is the weak form of the following partial differential equation

$$\begin{aligned}\frac{\partial}{\partial t}\rho(x,t) &= \frac{\sigma_\infty^2}{2}\triangle \rho(x,t) + \nabla\cdot(\rho(x,t)\nabla U(x)) \\ &\quad + \nabla\cdot[\rho(x,t)\nabla(\rho(x,t) - G*\rho(\cdot,t))(x)], \quad x\in\mathbb{R}^d, t\ge 0, \\ \rho(x,0) &= p_0(x), \quad x\in\mathbb{R}^d.\end{aligned} \tag{50}$$

The uniqueness of the limit h_∞ derives from the uniqueness of the weak solution of the viscous equation (50), which can be achieve with classical arguments [31, 32].

6 Long time behavior

In this section we investigate the long time behavior of the particle system, for a fixed number N of particles.

Interacting-Diffusing Particles

First of all, let us consider system (3) with $\gamma_1 = 0$, i.e. the case in which the advection is due only to interactions among particles. Following [18], from (3), it follows that the location of the center of mass $\bar{X}_N$ of the N particles,

$$\bar{X}_N(t) = \frac{1}{N}\sum_{k=1}^N X_N^k(t),$$

evolves according the following equation

$$d\bar{X}_N(t) = -\frac{1}{N^2}\sum_{k,j=1}^{N}\nabla\left(V_N - G\right)\left(X_N^k(t) - X_N^j(t)\right)dt + \sigma d\bar{W}(t), \tag{51}$$

where $\bar{W}(t) = \frac{1}{N}\sum_{k=1}^{N} W^k(t)$ is still a Brownian motion; by the symmetry of the kernels V_1 and G, the first term on the right hand side vanishes and we get

$$d\bar{X}_N(t) = \sigma d\bar{W}(t), \tag{52}$$

i.e. the stochastic process $\bar{X}_N$ is a Wiener process. Hence, its law, conditional upon the initial state, is

$$\mathcal{L}\left(\bar{X}_N(t)|\bar{X}_N(0)\right) = \mathcal{L}\left(\bar{X}_N(0), \sigma^2\bar{W}(t)\right) = \mathcal{N}\left(\bar{X}_N(0), \frac{\sigma^2}{N}t\right);$$

with variance $\frac{\sigma^2}{N}t$, which, for any fixed N, increases as t tends to infinity. Consequently we may claim that the probability law of the system does not converge to any non trivial probability law, since otherwise the same would happen for the law of the center of mass.

Complete System

Let us now consider the complete system of SDE System (3) with $\gamma_1 > 0$. This means that particles are also subject to a force due to the confining potential U. Equations of the type

$$dX_t = -\nabla P(X_t) + \sigma dW_t, \tag{53}$$

have been thoroughly analyzed in literature, under the sufficient condition of strict convexity of the symmetric potential U [21, 18, 22]; it has been shown that (53) does admit a nontrivial invariant distribution. From a biological point of view a strictly convex confining potential is difficult to explain; it would mean an infinite range of attraction of the force which becomes infinitely strong at the infinite, with an at least constant, drift even far from origin.

A weaker sufficient condition for the existence of a unique invariant measure has been more recently suggested by Veretennikov [19, 20], following Has'minski [33]. This condition states that there exist constants $M_0 \geq 0$ and $r > 0$ such that for $|x| \geq M_0$

$$\left(-\nabla P(\mu)(x), \frac{x}{|x|} \right) \leq -\frac{r}{|x|}. \tag{54}$$

It is ease to prove that without any further condition on the interaction kernels V_N and G, by considering condition (5) on U, we may apply the results by Veretennikov and prove the existence of an invariant measure for the joint law of the particles locations. Condition (5) means that ∇U may decay to zero as $|x|$ tends to infinity, provided that its tails are sufficiently "fat".

Proposition 4 *Under the hypotheses for the existence and uniqueness stated in Proposition 1 and condition* (5)*, system* (3) *admits a unique invariant measure.*

Proof.

Let $\pi_i(\mathbf{x}) = x_i, i = 1, ..., N$ be the i-th projection of $\mathbf{x} \in (R^d)^N$, $\tilde{U}(\mathbf{x})$ and $\tilde{K}(\mathbf{x})$ the vector function defined by

$$\tilde{U}(\mathbf{x}) = (U \circ \pi_i(\mathbf{x}))_{1 \leq i \leq N}, \quad \tilde{K}(\mathbf{x}) = \left((G - V_N) * \frac{1}{N} \sum_i \epsilon_{\pi_i(\mathbf{x})} \circ \pi_i(\mathbf{x}) \right)_{1 \leq i \leq N}$$

In order to apply Theorem 2 in [23], we have to prove that there exist constants $M \geq 0$ and $\tilde{r} > (\frac{Nd}{2} + 1)$ such that for all $\mathbf{x} \in (R^d)^N : |x| \geq M$

$$\left(-\gamma_1 \nabla \tilde{U}(\mathbf{x}) + \gamma_2 \nabla \tilde{K}(\mathbf{x}), \frac{\mathbf{x}}{|\mathbf{x}|} \right) \leq -\frac{\tilde{r}}{|\mathbf{x}|}. \tag{55}$$

We have

$$\begin{aligned}
\left(-\gamma_1\nabla\tilde{U}(\mathbf{x})+\gamma_2\nabla\tilde{K}(\mathbf{x}),\frac{\mathbf{x}}{|\mathbf{x}|}\right) &= -\gamma_1\sum_{k=1}^{N}\nabla U(x_k)\frac{x_k}{|\mathbf{x}|}\\
&+ \gamma_2\frac{1}{N}\sum_{k=1}^{N}\sum_{i=1}^{N}\nabla(G-V_N)(x_i-x_k)\frac{x_k}{|\mathbf{x}|}\\
&\le -\gamma_1\sum_{k=1}^{N}\nabla U(x_k)\frac{x_k}{|\mathbf{x}|}\\
&+ \gamma_2\frac{1}{N}\sum_{k=1}^{N}\sum_{i=1}^{N}\nabla(G-V_N)(x_i-x_k)\\
&= -\gamma_1\sum_{k=1}^{N}\nabla U(x_k)\frac{x_k}{|\mathbf{x}|}\le-\frac{\gamma_1 rN}{|\mathbf{x}|}
\end{aligned}$$

The last two inequalities derive from the symmetry of the G and V_N and (5). So if for $\tilde{r}=\gamma_1 rN$ and condition on r in (5), we have condition (55).

□

Let now $P_N^{x_0}(t)$ denote the joint distribution of the N particles at time t, conditional upon a non random initial condition x_0, and let P_S denote the invariant distribution. As far as the convergence of $P_N^{x_0}(t)$ is concerned, for t tending to infinity, as in [19], one can prove the following result.

Proposition 5 *Under the same assumptions of Proposition 4, for any k, $0<k<\tilde{r}-\frac{Nd}{2}-1$ with $m\in(2k+2,2\tilde{r}-Nd)$ and $\tilde{r}=\gamma_1 Nr$, there exists a positive constant c such that*

$$\left|P_N^{x_0}(t)-P_N^S\right|\le c(1+|x_0|^m)(1+t)^{-(k+1)},$$

where $\left|P_N^{x_0}(t)-P_N^S\right|$ denotes the total variation distance of the two measures,i.e.

$$\left|P_N^{x_0}(t)-P_N^S\right|=\sup_{A\in\mathcal{B}_{\mathbb{R}^d}}\left[P_N^{x_0}(t)(A)-P_N^S(A)\right],$$

and x_0 the initial data.

So Proposition 4 states a polynomial convergence rate to invariant measure. To improve the rate of convergence, one has to consider more restricted assumptions on U [20].

7 Appendix 1:proofs of Lemma 1 and Lemma 2

Proof of Lemma 1

With the same procedure we have obtained equation (10), with $f = g_N$, since $||h_N(\cdot,t)||_2^2 = \langle X_N(t), g_N(\cdot,t)\rangle$, one get

$$\begin{aligned}
\mathbb{E}\left[||h_N(\cdot,t)||_2^2|\mathcal{F}_s\right] &= ||h_N(\cdot,s)||_2^2 \\
&- \mathbb{E}\left[2\int_s^t \langle X_N(u), |\nabla g_N(\cdot,u)|^2\rangle du\right. \\
&- 2\int_s^t \langle X_N(u), \nabla g_N(\cdot,u)\cdot(\nabla U(\cdot) + (\nabla G * X_N(u))(\cdot))\rangle du \\
&+ \left.\sigma^2\int_s^t ||\nabla h_N(\cdot,u)||_2^2 du|\mathcal{F}_s\right] + \frac{\sigma^2(t-s)}{N}\Delta V_N(0). \qquad (56)
\end{aligned}$$

Since $|\Delta V_N(x)| = N^{\beta(d+2)/d}|\Delta V_1(N^{\beta/d}x)|$ and by assumption (8),

$$\Delta V_1(0) = (\nabla W_1 * \nabla W_1)(0) = -\int_{\mathbb{R}^d} \nabla W_1(y)\nabla W_1(y)dy < \infty,$$

so that

$$\frac{\sigma^2(t-s)}{N}\Delta V_N(0) = C\,\sigma^2(t-s)N^{\beta(d+2)/d-1}. \qquad (57)$$

The thesis follows.

□

Proof of Lemma 2

By applying Ito's Formula to $\langle X_N(t), \phi\rangle$,

$$\begin{aligned}\mathbb{E}\left[\langle X_N(t), \phi\rangle | \mathcal{F}_s\right] &= \langle X_N(s), \phi\rangle + \mathbb{E}\left[\int_s^t \langle X_N(u), \nabla A_N(\cdot, u) - \nabla g_N(\cdot, u))\nabla\phi \right.\\ &\quad \left. + \frac{\sigma^2}{2}\Delta\phi\rangle du | \mathcal{F}_s\right] \\ &\geq \langle X_N(s), \phi\rangle - C\,\mathbb{E}\left[\int_s^t \langle X_N(u), \nabla A_N(\cdot, u) - \nabla g_N(\cdot, u) + 1\rangle du | \mathcal{F}_s\right].\end{aligned} \tag{58}$$

Since

$$\begin{aligned}0 &\leq \langle X_N(u), |\nabla A_N(\cdot, u) - \nabla g_N(\cdot, u) - 1|^2\rangle \\ &= \langle X_N(u), |\nabla A_N(\cdot, u) - \nabla g_N(\cdot, u)|^2 - 2(\nabla A_N(\cdot, u) - \nabla g_N(\cdot, u)) + 1\rangle,\end{aligned}$$

$$\langle X_N(u), \nabla A_N(\cdot, u) - \nabla g_N(\cdot, u)\rangle \leq \langle X_N(u), | + \nabla A_N(\cdot, u) - \nabla g_N(\cdot, u)|^2 + 1\rangle.$$

This implies that (58) is greater than or equal to

$$\begin{aligned}&\langle X_N(s), \phi\rangle - C\,\mathbb{E}\left[\int_s^t \langle X_N(u), |\nabla A_N(\cdot, u) - \nabla g_N(\cdot, u)|^2 + 2\rangle du | \mathcal{F}_s\right] \\ &\geq \langle X_N(s), \phi\rangle - C\,\mathbb{E}\left[\int_s^t \langle X_N(u), |\nabla A_N(\cdot, u) - \nabla g_N(\cdot, u)|^2 + 1\rangle du | \mathcal{F}_s\right] \\ &\geq \langle X_N(s), \phi\rangle - C\,\mathbb{E}\left[\int_s^t \langle X_N(u), 2(|\nabla A_N(\cdot, u)|^2 - \nabla A_N(\cdot, u)\nabla g_N(\cdot, u)\right. \\ &\quad \left. + |\nabla g_N(\cdot, u)|^2)\rangle + \sigma^2 ||\nabla h_N(\cdot, u)||_2^2 du + \int_s^t du | \mathcal{F}_s\right] \\ &= \langle X_N(s), \phi\rangle - C\,\mathbb{E}\left[D_N(t) - D_N(s) + t - s | \mathcal{F}_s\right] \\ &= \langle X_N(s), \phi\rangle - C\,\mathbb{E}\left[D_N(t) + t | \mathcal{F}_s\right] + C\,D_N(s) + C\,s.\end{aligned} \tag{59}$$

Hence,

$$\mathbb{E}\left[\langle X_N(t), \phi\rangle + C\,D_N(t) + C\,t | \mathcal{F}_s\right] \geq \langle X_N(s), \phi\rangle + C\,D_N(s) + C\,s$$

and (24) follows. In a completely analogous way, we obtain property (25).

□

8 Appendix 2: Notations.

Measure Spaces

$\mathcal{M}_{\mathcal{P}}(\mathbb{R}^d)$ is the space of probability measures on $\mathbb{R}^d$. This space is equipped with the usual weak topology. Denote by $Lip_b(\mathbb{R}^d)$ the set of bounded and Lipschitz real function on $\mathbb{R}^d$. For any $\mu, \nu \in \mathcal{M}(\mathbb{R}^d)$, define the *bounded Lipschitz metric* as follows

$$d_{BL}(\mu, \nu) = \sup_{f \in \mathcal{H}_1} (\langle \mu, f \rangle - \langle \nu, f \rangle), \tag{60}$$

where

$$\langle \mu, f \rangle = \int_{\mathbb{R}^d} f(x) \mu(dx) \quad f \in C_b(\mathbb{R}^d),$$

and

$$\mathcal{H}_1 = \left\{ f \in Lip_b(\mathbb{R}^d) \ : \ \|f\|_{Lip} = \|f\|_\infty + \sup_{x,y \in \mathbb{R}^d, x \neq y} \frac{|f(x) - f(y)|}{|x - y|} \leq 1 \right\}. \tag{61}$$

For some $T \in (0, \infty)$, $C([0,T],\ \mathcal{M}_{\mathcal{P}}(\mathbb{R}^d))$ is the space of all continuous functions $f = f(t),\ 0 \leq t \leq T$ from $[0,T]$ to $\mathcal{M}_{\mathcal{P}}(\mathbb{R}^d)$, equipped with the metric

$$\rho(f, g) = \sup_{0 \leq t \leq T} ||f(t) - g(t)||_1.$$

Given a metric space (S, ρ), for any S-valued random variable Y, we denote by $\mathcal{L}(Y) \in \mathcal{M}_{\mathcal{P}}(S)$ its distribution.

Function Spaces

For an open set $D \subset \mathbb{R}^N$, we denote by $C(D)$ the space of continuous functions on D and by $C^k(D)$ the space of k-times continuously differentiable functions equipped with the usual supremum-norms. Moreover, we will use the Lebesgue spaces $L^p(D)$, $1 \leq q \leq \infty$, with

$$\|u\|_{L^p(D)} = \begin{cases} \left(\int_D |u(x)|^p \; dx\right)^{\frac{1}{p}} & \text{if } 1 \leq p < \infty \\ \text{ess } \sup_{x\in D} |u(x)| & \text{if } p = \infty \end{cases} \tag{62}$$

and the Sobolev spaces $W^{k,p}(D)$, $1 \leq p \leq \infty$, $0 \leq k$ of functions with distributional derivatives up to order k in $L^p(D)$. The Sobolev space norms are defined by

$$\|u\|_{W^{k,p}(D)} = \left(\|u\|^p_{L^p(D)} + \sum_{1\leq|\alpha|\leq k} \|\partial^\alpha u\|^p_{L^p(D)} \right)^{\frac{1}{p}} \tag{63}$$

for $1 \leq p < \infty$, and by

$$\|u\|_{W^{k,\infty}(D)} = \max \left\{ \|u\|_{L^\infty(D)}, \sup_{1\leq|\alpha|\leq k} \|\partial^\alpha u\|_{L^\infty(D)} \right\}. \tag{64}$$

Moreover, we will use the standard notations $H^k(D) = W^{k,2}(D)$ and $H^1_0(D)$ for the subspace of functions in $H^1(D)$ with vanishing trace on ∂D. For further details on the spaces $W^{k,p}(D)$ we refer to the monograph by Evans [31].

Finally, we need the Banach valued function spaces on a real interval $[0,T] \subset \mathbb{R}$; let $u : [0,T] \to X$ be a function defined almost everywhere in $[0,T]$ with values in some Banach space X. If u is continuous, then we say that $u \in C([0,T];X)$, and equip this space with the supremum norm

$$\|u\|_{C([0,T];X)} := \sup_{t\leq T} \|u(t)\|_X.$$

In an analogous way we define the spaces $C^k([0,T];X)$, $L^p([0,T];X)$ and $W^{k,p}([0,T];X)$ and their norms. For a detailed discussion on such spaces, we refer to [32].

Fourier Transform

For $f \in L^2(\mathbb{R}^d)$ we denote by

$$\tilde{f}(\lambda) = \lim_{a \to \infty} \left(\frac{1}{2\pi}\right)^{d/2} \int_{\{|x| \leq a\}} e^{i\lambda x} f(x) dx$$

its Fourier transform.

In connection with Fourier transforms we shall use the relations

$$\int_{\mathbb{R}^d} f(x)\overline{g(x)}dx = \int_{\mathbb{R}^d} \tilde{f}(\lambda)\overline{\tilde{g}(\lambda)}d\lambda \quad f, g \in L^2(\mathbb{R}^d), \tag{65}$$

$$\widetilde{f * g}(\lambda) = (2\pi)^{d/2}\tilde{f}(\lambda)\tilde{g}(\lambda) \quad f, g \in L^2(\mathbb{R}^d), \tag{66}$$

$$\widetilde{\nabla f}(\lambda) = -i\lambda\tilde{f}(\lambda) \quad f \in W_2^1(\mathbb{R}^d); \tag{67}$$

where

$$W^{1,2}(\mathbb{R}^d) = \{f \in L^2(\mathbb{R}^d) \ : \ \int_{\mathbb{R}^d} (1 + |\lambda|^2)|\tilde{f}(\lambda)|^2 d\lambda = ||f||_2^2 + ||\nabla f||_2^2 < \infty\}.$$

The Convolution Kernel G

In the above model for the aggregation kernel, it is usually assumed that G is a bounded function with finite support, which represents the fact that individuals interact only over some finite range. For our analysis, we can relax this assumption to

$$G \in C^1(\mathbb{R}^d) \cap W^{3,2}(\mathbb{R}^d) \cap W^{2,\infty}(\mathbb{R}^d) \tag{68}$$

which implies that $G \in W^{2,p}(\mathbb{R}^d)$ for any $p \in [2, \infty]$.

Note that due to $\frac{\partial^3 G}{\partial x_i x_j x_k} \in L^2(\mathbb{R}^d)$, the convolution $\frac{\partial^3 G}{\partial x_i x_j x_k} * u$ is well-defined as a function in $L^1(\mathbb{R}^d)$ due to Plancherel's Theorem and the corresponding convolution operator is continuous on $L^1(\Omega)$. Similarly, because of $\frac{\partial^2 G}{\partial x_i x_j} \in L^\infty(\mathbb{R}^d)$, the convolution $\frac{\partial^2 G}{\partial x_i x_j} * u$ is well-defined as a function in $L^1(\mathbb{R}^d)$ and the corresponding convolution operator is continuous on $L^1(\Omega)$, which can be seen from a straight-forward estimate. The interested reader may refer to Champerey [34].

References

[1] Grünbaum, D., Okubo, A. Modelling social animal aggregations. In *Frontiers of Theoretical Biology* (S.A. Levin, Ed.), Lectures Notes in Biomathematics, 100, Springer Verlag, New York, 1994; 296-325.

[2] Gueron, S., Levin, S.A., Rubenstein, D.I., The dynamics of herds: from individuals to aggregations, J. Theor. Biol., 1996, 182: 85-98 .

[3] Okubo, A. Dynamical aspects of animal grouping: swarms, school, flocks and herds, Adv. BioPhys., 1986, 22: 1-94. .

[4] Satsuma J., Animal aggregation and solitons: exact solutions for nonlinear diffusion equations, in Special Issue "Solitons, Mathematical Sciences, 1985: 108-113. (in Japanese)

[5] Warburton, K., Lazarus, J., Tendency-distance models of social cohesion in animal groups, J. Theor. Biol., 1991, 150: 473-488.

[6] Capasso V., D. Morale, Rescaling Stochastic Processes: Asymptotics. in *Multiscale Problems in the Life Sciences From Microscopic to Macroscopic*, Lecture Notes in Mathematics / Fondazione C.I.M.E., Firenze, 2008, 1940: 91-146.

[7] Boi S., Capasso V., Morale D., Modeling the aggregative behavior of ants of the species Polyergus rufescens. Spatial heterogeneity in ecological models. Nonlinear Anal. Real World Appl., 2000, 1: 163-176, .

[8] Morale D., Cellular automata and many-particles systems modeling aggregation behaviour among populations, Int. Journal of Applied Mathematics and Computer

Science, Special issue on "Mathematical Aspects of Population Dynamics", 2000, 10, (1): 157-173.

[9] Morale D. , Modeling and simulating animal grouping: individual-based models, Future Generation Computer Systems, 2001, 71, (7): 883-891.

[10] Morale D., Capasso V., Oelschläger K., An interacting particle system modelling aggregation behavior: from individuals to populations. J. Math. Biol., 2005, 50: 49-66.

[11] Diekmann O., M.Gyllenberg, H.R.Thieme, Lack of uniqueness in transport equations with a nonlocal nonlinearity, Math. Models and Meth. in Appl. Sciences, 2000, 10, 581-592.

[12] Burger M., Capasso V., Morale D., On an aggregation model with long and short range interactions, Nonlinear Analysis: Real World Applications, 2007, 3: 939-958.

[13] Oelschläger K., On the derivation of reaction-diffusion equations as limit dynamics of systems of moderately interacting stochastic Processes, Prob. Th. Rel. Fields, 1989, 82: 565-586.

[14] Oelschläger K., Large systems of interacting particles and porous medium equatio, J. Differential Equations, 1990, 88: 294-346.

[15] Meleard S., Asymptotic behaviour for interacting diffusion processes with space-time random birth, with B. Fernandez, Bernoulli, 2000, 6, (2): 1-21.

[16] Sznitman, A. S., *Topics in propagation of chaos.*, Ecole d'Ete de Probabilites de Saint-Flour XIX, 1989. Lecture Notes in Math., 1464, Springer, 1991: 165-251.

[17] Oelschläger K., A law of large numbers for moderately interacting diffusion processes. Z. Wahrscheinlichkeitstheorie verw. Gabiete ,1985, 69: 279-322.

[18] Malrieu, F., Convergence to equilibrium for granular media equations and their Euler schemes. The Annals of Applied Probability, 2003, 13: 540-560.

[19] Veretennikov A.Y., On polynomial mixing and convergence rate for for stochastic differential equations. Theory Probab. Appl., 1999, 44: 361-374.

[20] Veretennikov, A.Y., On subexponential mixing rate for Markov processes. Theory Probab. Appl., 2005, 49: 110-122.

[21] Carrillo J.A., McCann R.J., Villani C. Kinetic equilibration rates for granular media and related equations: entropy dissipation and mass transportation estimates. Rev. Mat. Iberoamericana, 2003, 19: 971-1018.

[22] Markowich P.A., Villani C. On the trend to equilibrium for the Fokker-Planck equation: an interplay between physics and functional analysis. Mathematica Contemporanea, 2000, **19**: 1-31.

[23] Veretennikov A.Y., On polynomial mixing bounds for stochastic differential equations. Stochastic Processes and their Applications, 1997, 70: 115-127.

[24] Capasso V., Bakstein D., *An Introduction to Continuous-Time Stochastic Processes - Theory, Models and Applications to Finance, Biology and Medicine.*, Birkhäuser, Boston, 2005.

[25] Arnold L., *Stochastic differential equations: theory and applications. Wiley-Interscience*, New York, 1974.

[26] Murray, J., *Mathematical Biology*, Springer-Verlag, Berlin, Heidelberg, 1989.

[27] Friedman A., *Stochastic Differential Equations and Applications*, Vols. I and II, Academic Press, London, 1975.

[28] Ethier S. N., Kurtz T. G., *Markov Processes, Characterization and Convergence*, Wiley, New York, 1986.

[29] Ikeda N., Watanabe S., *Stochastic Differential Equations and Diffusion Processes.* North-Holland Mathematical Library, Amsterdam, 1981.

[30] Rudin, W., *Fourier Analysis on Groups*, reprint, Wiley, New York, 1990.

[31] Evans L.C., *Partial Differential Equations*, AMS, Providence, 1998.

[32] R.E.Showalter, *Monotone Operators and Nonlinear Partial Differential Equations in Banach Space*, AMS Mathematical Surveys and Monographs, Providence, Rhode Island, 1997.

[33] Has'minski, R.Z. *Stochastic stability of differential equations.* Sijthoff & Noordhoff, Alphen aan den Rijn, The Netherlands and Rockville, Maryland, USA, 1980.

[34] D.C.Champerey, *A Handbook of Fourier Theorems* (Cambridge Univ. Press, Cambridge, New York, 1987).