Microscopic derivation of the Keller-Segel equation in the sub-critical regime

ANA CAÑIZARES GARCÍA
LMU-Munich
Email: `caniz@math.lmu.de`

PETER PICKL
LMU-Munich
Email: `pickl@math.lmu.de`

May, 12th 2017

ABSTRACT. We derive the two-dimensional Keller-Segel equation from a stochastic system of N interacting particles in the case of sub-critical chemosensitivity $\chi < 8\pi$. The Coulomb interaction force is regularised with a cutoff of size $N^{-\alpha}$, with arbitrary $\alpha \in (0, 1/2)$. In particular we obtain a quantitative result for the maximal distance between the real and mean-field N-particle trajectories.

1. INTRODUCTION

The Keller-Segel equation [15] is known as the classical model of *chemotaxis*, which in Biology refers to the movement of organisms guided by an external chemical substance and has been observed in some species of bacteria or amoeba. The Keller-Segel equation, concretely motivated by the behaviour of the unicellular organism *Dictyostelium discoidium*, models a situation in which cells naturally spread out but under starvation circumstances also attract other cells by segregating an attractive chemical substance. We consider the two-dimensional Keller-Segel equation:

$$\partial_t \rho = \Delta \rho + \chi \nabla \cdot ((k * \rho)\, \rho), \qquad \rho(0, \cdot) = \rho_0. \tag{1}$$

Here ρ: $[0, \infty) \times \mathbb{R}^2 \to [0, \infty)$ is the evolution of the cell population density for an initial value $\rho_0 \colon \mathbb{R}^2 \to [0, \infty)$, the interaction force kernel $k \colon \mathbb{R}^2 \to \mathbb{R}^2$ is given by $k(x) := \frac{x}{2\pi |x|^2}$ and the constant $\chi > 0$ denotes the *chemosensitivity* or response of the cells to the chemical substance[1]. This model reflects the characteristic competition between diffusion and aggregation in such a chemotactical process. Mathematically this results in the interesting effect that in some cases smooth solutions exist for all times, while in others solutions *blow up* in finite time[2] (corresponding to clustering of the cells). Furthermore, the existence of global solutions or the presence of blow-up events strongly depend on the dimension, mass and chemosensitivity of the system: in one dimension the solution exists globally, but in higher dimensions blow-up events in finite time may or may not occur depending on the initial mass $M := \int_{\mathbb{R}^2} \rho_0(x)$ and the chemosensitivity χ. This role for the 2-dimensional description was completely understood for the first time less than a decade ago: if $\chi M < 8\pi$, a global bounded solution exists, while for $\chi M > 8\pi$ blow-up in finite time always takes place. Finally, if $\chi M = 8\pi$ a global solution exists which possibly becomes unbounded as $t \to \infty$ [3],[8],[2]. Here we work in a probabilistic setting and for convenience assume an initial mass $M = 1$. The threshold condition for the existence of global solutions is therefore at $\chi = 8\pi$.

1. This form of the equation results from the Keller-Segel parabolic-elliptic system if the concentration of chemical substance is taken to be the Newtonian potential of the density of cells [12].

2. A solution $\rho(t, x)$ is said to blow up in finite time if $\lim_{t \to T} \|\rho(t, \cdot)\|_{L^\infty} = \infty$ for some finite time T.

Our purpose in this paper is to derive the deterministic *macroscopic* equation (1) in the sub-critical regime $\chi \in (0, 8\,\pi)$ as the mean-field limit of the following *microscopic* stochastic N-particle system as $N \to \infty$:

$$\mathrm{d}X_t^{i(N)} = -\frac{\chi}{N}\sum_{j\neq i}^{N} k\left(X_t^{i(N)} - X_t^{j(N)}\right)\mathrm{d}t + \sqrt{2}\,\mathrm{d}B_t^i, \quad i = 1, \ldots, N, \quad X_0^{(N)} \sim \bigotimes_{i=1}^{N} \rho_0, \tag{2}$$

where the process $X^{i(N)}\colon [0,\infty) \to \mathbb{R}^2$ denotes the trajectory of the i-th particle, $(B^i)_{i\in\mathbb{N}}$ is a family of 2-dimensional independent Brownian motions, $X_t^{(N)} \in \mathbb{R}^{2N}$ denotes the vector $X_t^{(N)} := (X_t^{1(N)}, \ldots, X_t^{N(N)})$[3], and at the initial time $t = 0$ the particles are independently distributed according to the initial density ρ_0. To this end we prove the property of *propagation of chaos* for regularised versions (with a cutoff depending on N) of these equations in Corollary 2. We obtain the propagation of chaos as a consequence of Theorem 1, where the real trajectories $X^{i(N)}$ are shown to remain close to the mean-field trajectories, defined by (3) below, if both started at the same point. The mean-field trajectories are given by the following equation:

$$\mathrm{d}Y_t^{i(N)} = -\chi\,(k * \rho_t)\,(Y_t^{i(N)})\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}B_t^i, \quad i = 1, \ldots, N, \quad Y_0^{(N)} = X_0^{(N)}, \tag{3}$$

where $\rho_t = \mathscr{L}(Y_t^{i(N)})$ is the probability distribution of any of the i.i.d. $Y_t^{i(N)}$. We remark that the Keller-Segel equation (1) is Kolmogorov's forward equation for any solution of (3), and in particular their probability distribution ρ_t solves (1).

The work of Cattiaux and Pédèches [6] is relevant for the existence of solutions of the stochastic particle system (2) and their properties. Furthermore, the derivation of the macroscopic equation (1) from the many-particle system (2) or propagation of chaos has been addressed in the past years by several mathematicians for modified problems: for a regularised interaction force $k_\varepsilon(x) := \frac{x}{|x|\,(|x| + \varepsilon)}$ in [12] and for a sub-Keller-Segel equation with a less singular force $k_\alpha(x) := \frac{x}{|x|^{\alpha+1}}$, $0 < \alpha < 1$, in [11]. More recently, great progress has been made for the purely Coulomb case ($\alpha = 1$): Fournier and Jourdain [10] proved the convergence of a subsequence for the particle system (2) by a tightness argument in the *very sub-critical* case $\chi < 2\,\pi$ using no cutoff at all; the convergence of the whole sequence (and therefore propagation of chaos) was nevertheless not achieved. Liu et al. published in the past year several results on propagation of chaos of (2) [18], [14], [19], the last of them containing the strongest result available to date to our knowledge. We improve their result in two aspects. On the one hand our conditions (4) on the initial density ρ_0 are weaker: Liu and Zhang assume ρ_0 is compactly supported, Lipschitz continuous and $\rho_0 \in H^4(\mathbb{R}^2)$. On the other hand our initial configuration for the N particles is less restrictive: ours are i.i.d. random variables on $\mathbb{R}^2$, while their particles are distributed on a grid. Our approach adapts a method that seems to be powerful for deriving the mean-field limit of some N-particle systems with Coulomb interactions, which was presented by Boers, Pickl [4] and Lazarovizi, Pickl [17] for the derivation of the Vlasov-Poisson equation from an N-particle Coulomb system for typical initial conditions.

3. We introduce the notation (N) for the number of particles in order to differentiate between these trajectories and the regularised ones. We nevertheless just use this notation during the introduction, since in the following sections we only work with the regularised equations.

Conditions on the chemosensitivity and the initial density.

We assume throughout this note a sub-critical chemosensitivity $\chi \in (0, 8\pi)$ and the following conditions on the initial density ρ_0:

$$\begin{aligned} \rho_0 &\in L^1(\mathbb{R}^2, (1+|x|^2)\,\mathrm{d}x) \cap L^\infty(\mathbb{R}^2) \cap H^2(\mathbb{R}^2),\\ \rho_0 &\geqslant 0,\\ \int_{\mathbb{R}^2} \rho_0(x)\,\mathrm{d}x &= 1,\\ \rho_0 \log \rho_0 &\in L^1(\mathbb{R}^2). \end{aligned} \tag{4}$$

These conditions guarantee global existence, uniqueness and further good properties of the solution of the macroscopic equation (1). Section 3 reviews these results and the corresponding ones for the solutions of the microscopic system.

Regularisation of the interaction force.

We introduce the following N-dependent regularisation of the Coulomb interaction force. Let $\phi^1 : \mathbb{R}^2 \to [0, \infty)$ be a radially symmetric, smooth function with the following properties:

$$\phi^1(x) := \begin{cases} -\frac{1}{2\pi} \log|x|, & |x| \geqslant 2,\\ 0, & |x| \leqslant 1, \end{cases}$$

$|\nabla \phi^1(x)| \leqslant (2\pi |x|)^{-1}$, $-\Delta \phi^1(x) \geqslant 0$ and $|\partial_{ij}^2 \phi^1(x)| \leqslant (\pi |x|^2)^{-1}$ for any $x \in \mathbb{R}^2$, $i, j = 1, 2$. For each $N \in \mathbb{N}$ and $\alpha \in (0, 1/2)$, let $\phi^N(x) = \phi^1(N^\alpha x)$ and consider the regularised interaction force $k^N = -\nabla \phi^N$, which by construction satisfies

$$k^N(x) := \begin{cases} \frac{x}{2\pi |x|^2}, & |x| \geqslant 2N^{-\alpha}\\ 0, & |x| \leqslant N^{-\alpha} \end{cases}$$

and

$$|\partial_i k^N(x)| \leqslant \begin{cases} \frac{1}{\pi |x|^2}, & |x| > N^{-\alpha}\\ 0, & |x| \leqslant N^{-\alpha} \end{cases}, \quad i = 1, 2.$$

For an initial density ρ_0 satisfying the above conditions (4) and each $N \in \mathbb{N}$ we consider the regularised Keller-Segel equation

$$\partial_t \rho^N = \Delta \rho^N + \chi \nabla((k^N * \rho^N)\rho^N), \qquad \rho^N(0, \cdot) = \rho_0, \tag{5}$$

the regularised microscopic N-particle system, for $i = 1, \ldots, N$,

$$\mathrm{d}X_t^{i(N),N} = -\frac{\chi}{N} \sum_{j \neq i} k^N\big(X_t^{i(N),N} - X_t^{j(N),N}\big)\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}B_t^i, \quad i = 1, \ldots, N, \quad X_0^{(N),N} \sim \bigotimes_{i=1}^N \rho_0, \tag{6}$$

and the regularised mean-trajectories

$$\mathrm{d}Y_t^{i(N),N} = -\chi\,(k^N * \rho_t^N)(Y_t^{i(N),N})\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}B_t^i, \quad i = 1, \ldots, N, \quad Y_0^{(N),N} = X_0^{(N),N} \tag{7}$$

where ρ_t^N denotes the probability distribution of $Y_t^{i(N),N}$, for any $i = 1, \ldots, N$. As in the non-regularised version this implies that ρ^N solves the regularised Keller-Segel equation (5). For simplicity of notation, and since the number of particles N already becomes apparent by the dependency of N of the cutoff, we will just write $X^{i,N}$ and $Y^{i,N}$ instead of $X^{i(N),N}$ and $Y^{i(N),N}$, as well as X^N and Y^N for the vectors $X^{(N),N}$ and $Y^{(N),N}$. It is also convenient to denote the regularised interaction force as

$$K_i^N(x_1, \ldots, x_N) := -\frac{\chi}{N} \sum_{j \neq i} k^N(x_i - x_j) \tag{8}$$

and the mean interaction force as

$$\overline{K}_{t,i}^N(x_1,\ldots,x_N) := -\chi(k^N * \rho_t^N)(x_i),$$

where $\rho_t^N = \mathscr{L}(Y_t^{i,N})$. We need to introduce one last process: For times $0 \leqslant s \leqslant t$ and a random variable $X \in \mathbb{R}^{2N}$, independent of the filtration generated by B_r, $r \geqslant s$, let $Z_{t,s}^{X,N}$ be the process starting at time s at the position X and evolving from time s up to time t with the mean force $\overline{K}^N$. Put in another way, the process $Z_{t,s}^{X,N} = (Z_{t,s}^{X,1,N},\ldots,Z_{t,s}^{X,N,N})$ is given by the solution of

$$\mathrm{d}Z_{t,s}^{X,i,N} = \overline{K}_{t,i}^N(Z_{t,s}^{X,N})\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}B_t^i, \quad i = 1,\ldots,N, \quad Z_{s,s}^{X,N} = X. \tag{9}$$

This paper is organised as follows. In the next section we state our main result and the ensuing propagation of chaos. We comment on the existence and properties of solutions of equations (1)-(9) in Section 3. Section 4 is devoted to some preliminary results that we need for the proof of the main result, Theorem 1, which is then proven in Section 5. We conclude with the proofs of Propositions 4 and 5 introduced in Section 3.

Notation. For simplicity we write single bars $|\cdot|$ for norms in $\mathbb{R}^n$ and $\|\cdot\|$ for norms in L^p spaces.

2. Main result

Let the chemosensitivity χ and the initial density ρ_0 satisfy condition (4), and for $N \in \mathbb{N}$ let X^N and Y^N be the real and mean-field trajectories solving the regularised microscopic equations (6) and (7), respectively. Our main result is that the N-particle trajectory X^N starting from a chaotic (product-distributed) initial condition $X_0^N \sim \otimes_{i=1}^N \rho_0$ typically remains close to the purely chaotic mean-field trajectory Y^N with same initial configuration $Y_0^N = X_0^N$ during any finite time interval $[0,T]$. More precisely, we prove that the measure of the set where the maximal distance $|X_t^N - Y_t^N|_\infty$ on $[0,T]$ excedes $N^{-\alpha}$ decreases exponentially with the number of particles N, as the number of particles grows to infinity.

Theorem 1. *Let $T > 0$ and $\alpha \in (0,1/2)$. For each $\gamma > 0$, there exist a positive constant C_γ and a natural number N_0 such that*

$$\mathbb{P}\Big(\sup_{0\leqslant t\leqslant T} |X_t^N - Y_t^N|_\infty \geqslant N^{-\alpha}\Big) \leqslant C_\gamma N^{-\gamma}, \quad \textit{for each } N \geqslant N_0.$$

C_γ depends on the coefficient χ, the initial density ρ_0, the final time T, α and γ, and N_0 depends on ρ_0, T and α.

Note that if the interaction force were Lipschitz continuous the statement would easily follow from a Grönwall-type argument. In our case we do not have this good property, but we can prove that the regularised force K^N is locally Lipschitz with a bound of order $\log N$, which follows from Lemma 7 and the Law of large numbers as presented in Proposition 8. This Lipschitz bound is good enough to prove the statement for short times but for larger times we need to introduce a new intermediate process. This process is proved to be close to X_t^N by the same argument as before for short times and close to Y_t^N by a new argument introduced in Lemma 9 which compares the densities of the processes instead of comparing the trajectories.

We remark that Theorem 1 directly implies the propagation of chaos, or the weak convergence of the k-particle marginals for X_t^N and Y_t^N. In order to show this, let us briefly introduce the first Wasserstein distance for measures: For $k \geqslant 1$ we denote by $\mathscr{P}(\mathbb{R}^{2k})$ the set of probability measures on $\mathbb{R}^{2k}$ and by $\mathscr{P}_1(\mathbb{R}^{2k})$ the subset of probability measures with finite expectation, $\mathscr{P}_1(\mathbb{R}^{2k}) := \{\mu \in \mathscr{P}(\mathbb{R}^{2k}) : \int |x|\,\mathrm{d}\mu < \infty\}$. We consider on $\mathscr{P}_1(\mathbb{R}^{2k})$ the first Wasserstein metric W_1 with respect to the normalised Euclidean distance on $\mathbb{R}^{2k}$

$$W_1(\mu,\nu) := \inf_{\pi\in\Pi(\mu,\nu)} \int_{\mathbb{R}^{2k}\times\mathbb{R}^{2k}} \frac{1}{k}\sum_{i=1}^k |x^i - y^i|\,\mathrm{d}\pi(x,y), \tag{10}$$

where $\Pi(\mu,\nu)$ is the set of all probability measures on $\mathbb{R}^{2k}\times\mathbb{R}^{2k}$ with first marginal μ and second marginal ν. It is a well known result (see, for instance, [22, Theorem 7.12]) that convergence with respect to this metric W_1 implies the weak convergence of measures in $\mathscr{P}_1(\mathbb{R}^{2k})$.

Corollary 2. *Consider the probability density $\otimes_{i=1}^N\rho_t^N$ of Y_t^N, denote by Ψ_t^N the probability density of X_t^N and by ${}^{(k)}\Psi_t^N$ its k-particle marginal*

$$ {}^{(k)}\Psi_t^N(x_1,\ldots,x_k) := \int_{\mathbb{R}^{2(N-k)}} \Psi_t^N(x_1,\ldots,x_N)\,\mathrm{d}x_{k+1}\cdots\mathrm{d}x_N, \qquad k\geqslant 1. $$

Then ${}^{(k)}\Psi_t^N$ converges weakly to $\otimes_{i=1}^k\rho_t$ as $N\to\infty$ for each fixed $k\geqslant 1$ and the full density Ψ_t^N converges weakly to $\otimes_{i=1}^N\rho_t$ as $N\to\infty$. More precisely, there exist a positive constant C and a natural number N_0 such that

$$ \sup_{0\leqslant t\leqslant T} W_1\big({}^{(k)}\Psi_t^N,\otimes_{i=1}^k\rho_t^N\big),\ \sup_{0\leqslant t\leqslant T} W_1(\Psi_t^N,\otimes_{i=1}^N\rho_t^N)\leqslant CN^{-\alpha} \tag{11} $$

holds for each $k\geqslant 1$ and $N\geqslant N_0$. W_1 denotes the first Wassertein distance (10), C depends on the coefficient χ, the initial density ρ_0, the final time T and α, and N_0 depends on ρ_0, T and α.

Proof. For the distance on $\mathscr{P}(\mathbb{R}^{2N})$ between the full density Ψ_t^N and $\otimes_{i=1}^N\rho_t^N$ we find

$$ \begin{aligned} W_1(\Psi_t^N,\otimes_{i=1}^N\rho_t^N) &= \inf_{\pi\in\Pi(\Psi_t^N,\otimes_{i=1}^N\rho_t^N)} \int_{\mathbb{R}^{2N}\times\mathbb{R}^{2N}} \frac{1}{N}\sum_{i=1}^N |x^i-y^i|\,\pi(\mathrm{d}x,\mathrm{d}y) \\ &\leqslant \inf_{\pi\in\Pi(\Psi_t^N,\otimes_{i=1}^N\rho_t^N)} \int_{\mathbb{R}^{2N}\times\mathbb{R}^{2N}} \sqrt{2}\,|x-y|_\infty\,\pi(\mathrm{d}x,\mathrm{d}y) \\ &\leqslant \sqrt{2}\,\mathbb{E}(|X_t^N-Y_t^N|_\infty). \end{aligned} $$

Analogously, if we take a fix $k\geqslant 1$, the same bound holds for the corresponding Wasserstein distance between the k-particle marginal ${}^{(k)}\Psi_t^N$ and the product $\otimes_{i=1}^k\rho_t^N$. Let us consider the expectation $\mathbb{E}(|X_t^N-Y_t^N|_\infty)$ on the set $A := \{\sup_{0\leqslant t\leqslant T}\|X_t^N-Y_t^N\|_\infty\geqslant N^{-\alpha}\}$ and its complementary separately. On A^c the expectation is simply bounded by $N^{-\alpha}$; on A, according to Theorem 1, it is

$$ \begin{aligned} \int_A |X_t^N-Y_t^N|_\infty\,\mathrm{d}\mathbb{P} &= \int_A \left|\int_0^t K^N(X_s^N)-\overline{K}_s^N(Y_s^N)\,\mathrm{d}s\right|_\infty \mathrm{d}\mathbb{P} \\ &\leqslant t\left(\|K^N\|_\infty + \sup_{0\leqslant s\leqslant t}\|\overline{K}_s^N\|_\infty\right)\mathbb{P}(A) \\ &\leqslant T\,(2\,\pi)^{-1}N^{\alpha}\,C_{2\alpha}N^{-2\alpha} \\ &\leqslant CN^{-\alpha}, \end{aligned} $$

for a constant C depending on χ, ρ_0, T and α and all N greater than some N_0 depending on ρ_0, T and α. We conclude that

$$ W_1\big({}^{(k)}\Psi_t^N,\otimes_{i=1}^k\rho_t^N\big), W_1(\Psi_t^N,\otimes_{i=1}^N\rho_t^N)\leqslant CN^{-\alpha}, \quad k\geqslant 1, $$

holds for each $t\in[0,T]$ and $N\geqslant N_0$, where $C=C(\chi,\rho_0,T,\alpha)$ and $N_0=N_0(\rho_0,T,\alpha)$. After taking the supremum over $0\leqslant t\leqslant T$ we obtain the desired result. □

The above result implies also the weak convergence of the k-particle marginal ${}^{(k)}\Psi_t^N$, for $k\geqslant 1$ to the tensor product $\otimes_{i=1}^k\rho_t$ as $N\to\infty$, where ρ_t is the solution of the (non-regularised) Keller-Segel equation (1). Indeed, since ρ_t^N converges weakly to ρ_t (Proposition 3) it is also true that $\otimes_{i=1}^k\rho_t^N$ converges weakly to $\otimes_{i=1}^k\rho_t$ for any fix $k\geqslant 1$, $N\to\infty$. Here we do not include a quantitative version of this convergence, but it should not be difficult to prove.

3. PROPERTIES OF SOLUTIONS

3.1. Macroscopic equations.

Following [9] we say that ρ is a *weak solution* of (1) for an initial condition ρ_0 satisfying (4) if

$$0 \leqslant \rho \in L^\infty(0,T;L^1(\mathbb{R}^2)) \cap C([0,T);\mathscr{D}'(\mathbb{R}^2)), \quad T>0,$$

ρ satisfies the *conservation of mass*

$$\int_{\mathbb{R}^2} \rho \,\mathrm{d}x = \int_{\mathbb{R}^2} \rho_0 \,\mathrm{d}x \quad (=1),$$

the *second moment equation*

$$\int_{\mathbb{R}^2} \rho(t,x)\,|x|^2 \,\mathrm{d}x = 4\left(1-\frac{\chi}{8\pi}\right)t + \int_{\mathbb{R}^2} \rho_0(x)\,|x|^2 \,\mathrm{d}x,$$

the *free energy inequality*

$$\mathscr{F}[\rho(t)] + \int_0^t \int_{\mathbb{R}^2} \rho\,|\nabla(\log \rho) + \chi\,(k*\rho)|^2 \,\mathrm{d}x\mathrm{d}s \leqslant \mathscr{F}[\rho_0],$$

and the Keller-Segel equation in the following sense: for each $\varphi \in C_c^2([0,T)\times\mathbb{R}^2)$

$$\int_{\mathbb{R}^2} \rho_0(x)\,\varphi(0,x)\,\mathrm{d}x = \int_0^\infty \int_{\mathbb{R}^2} \rho\,[(\nabla(\log\rho) + \chi\,(k*\rho))\cdot\nabla\varphi - \partial_t\varphi]\,\mathrm{d}x\mathrm{d}t.$$

Here the *free energy* $\mathscr{F}$ is given by

$$\mathscr{F}[\rho] := \int_{\mathbb{R}^2} \rho\log\rho\,\mathrm{d}x - \frac{\chi}{2}\int_{\mathbb{R}^2} \rho\,(\phi*\rho)\,\mathrm{d}x.$$

Proposition 3. (Existence and convergence) *Under assumption (4) for the chemosensitivity χ and the initial density ρ_0 the following holds:*

i. For any $N\in\mathbb{N}$ and any $T>0$, there exists $\rho^N \in L^2(0,T;H^1(\mathbb{R}^2)) \cap C(0,T;L^2(\mathbb{R}^2))$ which solves (5) in the sense of distributions.

ii. The Keller-Segel equation (1) has a unique weak solution $\rho \in L^\infty(\mathbb{R}_+;L^1(\mathbb{R}^2))$.

iii. The sequence (ρ^N) of solutions of (5) converges weakly to the solution ρ of the Keller-Segel equation (1).

We refer to [3] and [9] for the proof. More precisely, the existence of the sequence ρ^N and the weak convergence of a subsequence of ρ^N to a weak solution of the Keller-Segel equation (1) were proved in [3]. Together with the uniqueness of the weak solution ρ of (1), which was proved in [9], it follows the weak convergence of the whole sequence ρ^N (and not just a subsequence) to this unique solution ρ.

For the proof of Proposition 3 only $\rho_0 \in L^1(\mathbb{R}^2,(1+|x|^2)\,\mathrm{d}x)$, and not $\rho_0\in L^1(\mathbb{R}^2,(1+|x|^2)\,\mathrm{d}x)\cap L^\infty(\mathbb{R}^2)\cap H^2(\mathbb{R}^2)$ as required in condition (4), is necessary. If moreover the initial density is bounded in L^∞ we find in Proposition 4 that the solutions of the Keller-Segel and the regularised Keller-Segel equations are uniformly bounded in L^∞ as well. Finally with the full condition $\rho_0 \in L^1(\mathbb{R}^2, (1+|x|^2)\,\mathrm{d}x)\cap L^\infty(\mathbb{R}^2)\cap H^2(\mathbb{R}^2)$ we prove some Hölder estimates in Proposition 5. The proofs of these two last propositions are contained in Section 6.

Proposition 4. (L^∞ estimates) *Assume χ and ρ_0 satisfy condition (4). Then for each $T>0$ there exists a positive constant C such that*

$$\sup_{t\in[0,T]} \|\rho_t^N\|_\infty,\ \sup_{t\in[0,T]} \|\rho_t\|_\infty \leqslant C$$

holds for the solution ρ^N of (5) and the solution ρ of (1).

Proposition 5. (Hölder estimates) *Assume χ and ρ_0 satisfy condition (4). Then for each $T>0$ there exist positive constants C_1 and C_2 depending on ρ_0 and T, such that for each $N\in\mathbb{N}$ and $t\in[0,T]$ the following estimates hold for the solution ρ^N of (5) and the solution ρ of (1):*

i. $[\rho^N(t)]_{0,\alpha}, [\rho(t)]_{0,\alpha} \leqslant C_1$,

ii. $[k^N*\rho^N(t)]_{0,1}, [k*\rho(t)]_{0,1} \leqslant C_2$.

3.2. Microscopic equations.

We focus first on the interacting N-particle system (2) and its regularised version (6). Since for each $N>0$ the kernel of (6) is globally Lipschitz continous, the solution of (6) is strongly and uniquely well-defined. For the original singular situation (2) it is much more delicate. Cattiaux and Pédèches [6, Theorem 1.5] proved the following existence and uniqueness result:

Proposition 6. (Existence and uniqueness) *Let $\mathcal{M}:=\{$there exists at most one pair $i\neq j$ such that $X^i=X^j\}$. Then, for $N\geqslant 4$ and $\chi<8\,\pi\left(1-\frac{1}{N-1}\right)$ there exists a unique (in distribution) non explosive solution of (2) starting from any $x\in\mathcal{M}$.*

We continue with the mean-field N-particle system (3), its regularised version (7) and its regularised and linearised version (9). According to Proposition 5 the mean-field force $\overline{K}^N$ is Lipschitz in the space variable, uniformly in $t\in[0,T]$ and $N\in\mathbb{N}$. Therefore, the linear equation (9) has a unique strong solution. For the existence and uniqueness of solutions of the non-linear equations (3) and (7) we refer to [18, Theorem 2.6].

4. PRELIMINARY RESULTS

4.1. Local Lipschitz bound for the regularised interaction force.

The regularised interaction force K^N defined in (8) is locally Lipschitz, with a local Lipschitz bound depending on N. The proof of this statement is conducted in the following Lemma, which is formulated to include more general cutoffs that we will need to consider in this paper.

Lemma 7. *Let $\nu=\nu(N)$ be a monotone increasing function of N with $\lim_{N\to\infty}\nu(N)=\infty$, and consider the force k^ν with cutoff at $\nu(N)^{-1}$, $k^\nu(x):=-\nabla(\phi^1(\nu x))$ for the bump function ϕ^1 defined in Section 1, meaning in particular that $k^\nu(x)\leqslant(2\,\pi\,|x|)^{-1}$ and*

$$k^\nu(x)=\begin{cases}\frac{x}{2\,\pi|x|^2}, & |x|\geqslant 2\,\nu^{-1},\\ 0, & |x|\leqslant\nu^{-1}\end{cases}.$$

i. For each $x,y\in\mathbb{R}^2$ with $|x-y|\leqslant 2\,\nu^{-1}$ it holds

$$|k^\nu(x)-k^\nu(y)|\leqslant l^\nu(y)\,|x-y|,$$

where

$$l^\nu(y):=\begin{cases}\frac{16}{|y|^2}, & |y|\geqslant 4\,\nu^{-1},\\ \nu^2, & |y|\leqslant 4\,\nu^{-1}\end{cases}.$$

ii. Let the resulting force be $K_i^\nu(x_1,\ldots,x_N):=-\frac{\chi}{N}\sum_{j\neq i}k^\nu(x_i-x_j)$ and define

$$L_i^\nu(y_1,\ldots,y_N):=\frac{\chi}{N}\sum_{j\neq i}l^\nu(y_i-y_j).$$

Then, for each $x, y \in \mathbb{R}^{2N}$ *with* $|x-y|_\infty \leqslant \nu^{-1}$ *it holds*

$$|K_i^\nu(x) - K_i^\nu(y)| \leqslant 2\, L_i^\nu(y)\, |x-y|_\infty.$$

Proof. (i) By the Mean Value Theorem the bound

$$|k^\nu(x) - k^\nu(y)| \leqslant |Dk^\nu(z)|\, |x-y|$$

holds for some point z in the segment which joins x and y. We distinguish between the following two cases:

Case 1: $|y| \leqslant 4\, \nu^{-1}$.

Since the derivative of k^ν is globally bounded by ν^2/π, and consequently by ν^2 as well, it follows that

$$|k^\nu(x) - k^\nu(y)| \leqslant \|Dk^\nu\|\, |x-y| \leqslant l^\nu(y)\, |x-y|.$$

Case 2: $|y| \geqslant 4\, \nu^{-1}$.

Since $|z-y| \leqslant |x-y| \leqslant 2\, \nu^{-1}$, it follows that $|z| \geqslant 2\, \nu^{-1}$. This means in particular that the derivative of k^ν at z is bounded by $|z|^{-2}/\pi$ and also that $|z-y| \leqslant |z|$, so

$$|y|^2 \leqslant (|y-z| + |z|)^2 \leqslant (2\, |z|)^2 = 4\, |z|^2.$$

Therefore,

$$\begin{aligned} |k^\nu(x) - k^\nu(y)| &\leqslant |Dk^\nu(z)|\, |x-y| \\ &\leqslant 2^{-1}\, |z|^{-2}\, |x-y| \\ &\leqslant 2\, |y|^{-2}\, |x-y| \\ &\leqslant l^\nu(y)\, |x-y|. \end{aligned}$$

Finally, (ii) follows directly from (i). □

4.2. Law of large numbers.

In the proof of the main theorem we define several "exceptional" sets and rely on the fact that the measure of these sets is exponentially small. This fact is proven in the next Proposition, a *law of large numbers* for our setting, for all these sets are events where the sample mean and expected values of some family of independent variables are not close. The steps we follow for this version of the law of large numbers are the standard ones, the only issue being that the k-th moments of the variables we consider are not bounded but instead grow with N to infinity. We'll see that their growth is nevertheless slow enough and we still obtain a rate of convergence which is faster than $C_\gamma N^{-\gamma}$ for any $\gamma > 0$, where $C_\gamma > 0$ is a constant depending on the choice of γ but not on N.

Proposition 8. (Law of large numbers) *Let* $\alpha, \delta > 0$ *be such that* $\alpha + \delta < 1/2$. *For* $N \in \mathbb{N}$ *let* $Z^1, \ldots, Z^N$ *be N independent random variables in* $\mathbb{R}^2$ *and assume that* Z^i *has a probability density that we denote by* u^i, $i = 1, \ldots, N$. *Let* $h = (h^1, h^2)\colon \mathbb{R}^2 \to \mathbb{R}^2$ *be a continuous function satisfying* $|h(x)| \leqslant C_h \min\{N^\alpha, |x|^{-1}\}$. *Define* $H_i(Z) = (H_i^1(Z), H_i^2(Z)) := \frac{1}{N}\sum_{j \neq i} h(Z^i - Z^j)$ *and the following sets*

$$S := \left\{ \sup_{1 \leqslant i \leqslant N} |H_i(Z) - \mathbb{E}(H_i(Z))| \geqslant N^{-(\alpha+\delta)} \right\},$$

$$\tilde{S} := \left\{ \sup_{1 \leqslant i \leqslant N} |H_i(Z) - \mathbb{E}_{(-i)}(H_i(Z))| \geqslant N^{-(\alpha+\delta)} \right\},$$

where $\mathbb{E}_{(-i)}$ stands for the expectation with respect to every variable but Z^i, that is, $\mathbb{E}_{(-i)}(H_i(Z)) = \frac{1}{N}\sum_{j\neq i}(h*u^j)(Z^i)$.

Define $\varepsilon := 1-2(\alpha+\delta)$ (strictly positive by assumption) and assume that, for each i,

$$\log N \,\|u^i\|_\infty + \|u^i\|_\infty^2 \leqslant C_0 N^{\varepsilon/2} \tag{12}$$

holds for some constant C_0 independent of N and i. Then, for each $\gamma>0$ there exists a constant C_γ (depending on γ, ε, C_0 and C_h) such that

$$\mathbb{P}(S), \mathbb{P}(\tilde{S}) \leqslant C_\gamma N^{-\gamma}.$$

Proof. Because we can replace $\mathbb{E}(H_i(Z))$ by $\mathbb{E}_{(-i)}(H_i(Z))$ in the proof, it is enough to prove the statement for the first set S. Also notice that since

$$\mathbb{P}\left(\sup_{1\leqslant i\leqslant N}|H_i(Z)-\mathbb{E}(H_i(Z))|\geqslant N^{-(\alpha+\delta)}\right)\leqslant \sum_{i=1,\nu=1}^{N,2}\mathbb{P}(|H_i^\nu(Z)-\mathbb{E}(H_i^\nu(Z))|\geqslant N^{-(\alpha+\delta)})$$

holds, it suffices to prove that

$$\mathbb{P}(|H_i^\nu(Z)-\mathbb{E}(H_i^\nu(Z))|\geqslant N^{-(\alpha+\delta)})\leqslant C_\gamma N^{-\gamma}$$

for each $\gamma>0$, each $i=1,\ldots N$ and $\nu=1,2$. Let then $\gamma>0$, $\nu\in\{1,2\}$ and let us for simplicity take $i=1$.

We use Markov's inequality of order 2 m and determine later the right choice of m for the given γ and the quantity $(\alpha+\delta)$ in the exponent of the allowed error $N^{-(\alpha+\delta)}$. For $j=2,\ldots,N$ let us denote by Θ_j the (independent) random variables $\Theta_j := h^\nu(Z^1-Z^j)$ and by μ_j its expected value

$$\mu_j := \int h^\nu(z_1-z_j)\,u^1(z_1)\,u^j(z_j)\,\mathrm{d}z_1\,\mathrm{d}z_j.$$

Now by Markov's inequality

$$\begin{aligned}\mathbb{P}(|H_1(Z)-\mathbb{E}(H_1(Z))|\geqslant N^{-(\alpha+\delta)}) &= \mathbb{P}\left(\frac{1}{N}\left|\sum_{j\neq 1}(\Theta_j-\mu_j)\right|\geqslant N^{-(\alpha+\delta)}\right)\\ &\leqslant N^{2(\alpha+\delta)m}\,\mathbb{E}\left(\left(\frac{1}{N}\sum_{j\neq 1}(\Theta_j-\mu_j)\right)^{2m}\right).\end{aligned}$$

The expectation on the right hand side can be estimated by using the multinomial formula

$$(x_2+\ldots+x_N)^{2m} = \sum_{a_2+\ldots+a_n=2m} C_a\prod_{j=2}^N x_j^{a_j},$$

where $a=(a_2,\ldots,a_N)$ is a multiindex and $C_a=\binom{2m}{a_2,\ldots,a_N}=\frac{(2m)!}{a_2!\ldots a_N!}$. Consequently

$$\mathbb{E}\left(\left(\frac{1}{N}\sum_{j\neq 1}(\Theta_j-\mu_j)\right)^{2m}\right)=N^{-2m}\sum_{a_2+\ldots+a_N=2m}C_a\prod_{j\neq 1}\mathbb{E}((\Theta_j-\mu_j)^{a_j}).$$

Here note that if $a_j=1$ for some j then the whole term is zero, since $\mathbb{E}((\Theta_j-\mu_j))=0$. Therefore we are left only with terms with at most m non-zero entries. If we denote by $|a|$ the number of non-zero entries of the multiindex a, the sum above simplifies to

$$\mathbb{E}\left(\left(\frac{1}{N}\sum_{j\neq 1}(\Theta_j-\mu_j)\right)^{2m}\right)=N^{-2m}\sum_{\substack{a_2+\ldots+a_N=2m\\|a|\leqslant m}}C_a\prod_{j\neq 1}\mathbb{E}((\Theta_j-\mu_j)^{a_j}).$$

Next we estimate the a_j-th order moment of Θ_j, for $a_j \leqslant 2m$: specifically we prove that

$$\mathbb{E}((\Theta_j - \mu_j)^{a_j}) \leqslant C_h^{a_j} C_0 N^{\alpha(a_j-2)+\varepsilon/2}.$$

The a_j-th order moment of Θ_j equals

$$\int_{\mathbb{R}^2} (h^\nu(z_1 - z_j) - \mu_j)^{a_j} u^1(z_1) u^j(z_j) \, \mathrm{d}z_1 \, \mathrm{d}z_j.$$

We factor the power in the integrand as

$$(h^\nu(z_1 - z_j) - \mu_j)^{a_j} = (h^\nu(z_1 - z_j) - \mu_j)^{a_j-2} (h^\nu(z_1 - z_j) - \mu_j)^2,$$

then estimate the term to the power $a_j - 2$ by its supremum norm and integrate only the second factor. It holds that

$$\begin{aligned} \|h^\nu(z_1 - z_j) - \mu_j\|_\infty &\leqslant C \|h^\nu\|_\infty + \|(h^\nu * u^j)\|_\infty \\ &\leqslant C \|h\|_\infty \leqslant C_h N^\alpha. \end{aligned}$$

After integrating the term to the second power we find

$$\begin{aligned} \int_{\mathbb{R}^2} (h^\nu(z_1 - z_j) - \mu_j)^2 u^1(z_1) u^j(z_j) \, \mathrm{d}z_1 \, \mathrm{d}z_j &= \mu_j^2 + 2\mu_j \int h^\nu(z_1 - z_j) u^1(z_1) u^j(z_j) \, \mathrm{d}z_1 \, \mathrm{d}z_j \\ &\quad + \int_{\mathbb{R}^2} h^\nu(z_1 - z_j)^2 u^1(z_1) u^j(z_j) \, \mathrm{d}z_1 \, \mathrm{d}z_j \\ &\leqslant 3 \|h * u^j\|_\infty^2 + \|h^2 * u^j\|_\infty \\ &\leqslant C_h (\|u^j\|_\infty^2 + \log N \|u^j\|_\infty) \\ &\leqslant C_h C_0 N^{\varepsilon/2}. \end{aligned}$$

Altogether

$$\begin{aligned} \mathbb{E}((\Theta_j - \mu_j)^{a_j}) &= \int_{\mathbb{R}^2} (h^\nu(z_1 - z_j) - \mu_j)^{a_j} u^1(z_1) u^j(z_j) \, \mathrm{d}z_1 \, \mathrm{d}z_j \\ &\leqslant \|h^\nu(z_1 - z_j) - \mu_j\|_\infty^{a_j-2} \int_{\mathbb{R}^2} (h^\nu(z_1 - z_j) - \mu_j)^2 u^1(z_1) u^j(z_j) \, \mathrm{d}z_1 \, \mathrm{d}z_j \\ &\leqslant C_h^{a_j} C_0 N^{\alpha(a_j-2)+\varepsilon/2}. \end{aligned}$$

Let now $k \leqslant m$ and consider only the multiindices a with k non-zero entries, that is with $|a| = k$. It holds

$$\begin{aligned} \sum_{\substack{a_2 + \ldots + a_N = 2m \\ |a| = k}} C_a \prod_{j \neq 1} \mathbb{E}((\Theta_j - \mu_j)^{a_j}) &\leqslant \sum_{\substack{a_2 + \ldots + a_N = 2m \\ |a| = k}} C_a C_h^{2m} C_0^k N^{\alpha(2m-2k)+\varepsilon k/2} \\ &\leqslant \sum_{\substack{a_2 + \ldots + a_N = 2m \\ |a| = k}} (2m)^{2m} C_h^{2m} C_0^m N^{\alpha(2m-2k)+\varepsilon k/2}, \end{aligned}$$

where we used that $C_a = \binom{2m}{a_2, a_3, \ldots, a_N} \leqslant (2m)^{2m}$. Since the number of terms in the sum, i.e. the number of ways of choosing k numbers that add up $2m$ counting all permutations, is bounded by $N^k (2m)^k$, we find that

$$\begin{aligned} \sum_{\substack{a_2 + \ldots + a_N = 2m \\ |a| = k}} C_a \prod_{j \neq 1} \mathbb{E}((\Theta_j - \mu_j)^{a_j}) &\leqslant (2m)^{3m} C_h^{2m} C_0^m N^{\alpha(2m-2k)+\varepsilon k/2} N^k \\ &\leqslant C_m N^{2m\alpha} N^{k(1-2\alpha+\varepsilon/2)}, \end{aligned} \tag{13}$$

for a constant $C_m > 0$ only depending on m, C_h and C_0. At this point we can estimate the desired expected value

$$\begin{aligned}
\mathbb{E}\left(\left(\frac{1}{N}\sum_{j\neq 1}(\Theta_j-\mu_j)\right)^{2m}\right) &= N^{-2m}\sum_{\substack{a_2+\ldots+a_N=2m\\ |a|\leqslant m}} C_a\prod_{j\neq 1}\mathbb{E}((\Theta_j-\mu_j)^{a_j})\\
&= N^{-2m}\sum_{k=1}^{m}\sum_{\substack{a_2+\ldots+a_N=2m\\ |a|=k}} C_a\prod_{j\neq 1}\mathbb{E}((\Theta_j-\mu_j)^{a_j})\\
&\leqslant C_m N^{-2m}\sum_{k=1}^{m} N^{2m\alpha}N^{k(1-2\alpha+\varepsilon/2)}\\
&\leqslant C_m N^{-2m}N^{m(2\alpha+1-2\alpha+\varepsilon/2)}\\
&\leqslant C_m N^{-m(1-\varepsilon/2)},
\end{aligned}$$

where we used (13) and the positivity of $1-2\alpha+\varepsilon/2$. Finally we find that

$$\begin{aligned}
\mathbb{P}(|H_1(Z)-\mathbb{E}(H_1(Z))|\geqslant N^{-(\alpha+\delta)}) &\leqslant N^{2(\alpha+\delta)m}\,\mathbb{E}\left(\left(\frac{1}{N}\sum_{j\neq 1}(\Theta_j-\mu_j)\right)^{2m}\right)\\
&\leqslant C_m N^{2(\alpha+\delta)m}N^{-m(1-\varepsilon/2)}\\
&= C_m N^{-m(1-2(\alpha+\delta)-\varepsilon/2)}\\
&\leqslant C_m N^{-m\varepsilon/2} = \tilde{C}_\gamma N^{-\gamma}
\end{aligned}$$

holds for $m=2\gamma/\varepsilon$, where $\tilde{C}_\gamma := C_{2\gamma/\varepsilon}$ depends on γ, ε, C_0 and C_h. □

4.3. Comparison of solutions of (9) starting at different points.

In this section we address the following question: how different is the action of the force K^N on two solutions of (9) that start at different points? An estimate of this difference will be very useful in the second case (for large times) of the proof of the main theorem and innovates the methods presented in [4] and [17]. The estimate is provided in Corollary 10. Recall that for each $x\in\mathbb{R}^{2N}$, $Z_{t,s}^{x,N}\in\mathbb{R}^{2N}$ denotes the process starting at point x at time s and evolving for times greater that s according to the mean-field force $\overline{K}^N$. That is, $Z_{t,s}^{x,N}$ solves (9) with constant initial condition x and initial time s. Furthermore $Z_{t,s}^{x,N}$ has a transition probability density for $t>s$ since $\overline{K}^N$ is bounded [21, Chapter 1, Theorem 10.2]. Since the processes $Z_{t,s}^{x,1}, \ldots, Z_{t,s}^{x,N}$ are independent, the joint transition probability density $u_{t,s}^{x,N}(z_1,\ldots,z_N)$ is given by the product $u_{t,s}^{x,N}(z_1,\ldots,z_N) := \prod u_{t,s}^{x,i,N}(z_i)$. Here each term $u_{t,s}^{x,i,N}$ is the transition probability density of $Z_{t,s}^{x,i,N}$ and also the solution of the *linearised Keller-Segel equation*

$$\partial_t u_{t,s}^{x,i,N} = \Delta u_{t,s}^{x,i,N} - \nabla\cdot(f_t^N u_{t,s}^{x,i,N}), \qquad u_{s,s}^{x,i,N} = \delta_{x_i}, \tag{14}$$

where $f_t^N := \chi\, k^N * \rho_t^N$ and ρ_t^N solves the regularised Keller-Segel equation (5) with initial condition ρ_0. Consider now the processes $Z_{t,s}^{x,N}$ and $Z_{t,s}^{y,N}$ for two different starting points $x, y\in\mathbb{R}^{2N}$. It is intuitively clear that the probability densities $u_{t,s}^{x,N}$ and $u_{t,s}^{y,N}$ are just a shift of each other. The next lemma gives an estimate for the L^∞ norm of each $u_{t,s}^{x,N}$ as well as for the distance in L^∞ between any two densities $u_{t,s}^{x,N}$ and $u_{t,s}^{y,N}$ in terms of the distance between the starting points x and y and the elapsed time $t-s$.

Lemma 9. *There exists a positive constant C depending on ρ_0 and T such that for each $N\in\mathbb{N}$, any starting points $x,y\in\mathbb{R}^{2N}$ and any times $0\leqslant s<t\leqslant T$ the following estimates for the transition probability densities $u_{t,s}^{x,N}$ resp. $u_{t,s}^{y,N}$ of the processes $Z_{t,s}^{x,N}$ resp. $Z_{t,s}^{y,N}$ given by (9) hold:*

i. $\|u_{t,s}^{x,N}\|_\infty\leqslant C\,((t-s)^{-1}+1)$,

ii. $\|u_{t,s}^{x,N}-u_{t,s}^{y,N}\|_\infty\leqslant C\,((t-s)^{-3/2}+1)\,|x-y|_\infty$.

Proof. Both estimates can be proved in the same way. We just give the proof for part (ii), which can be easily adapted for part (i). For simplicity of notation we assume $s=0$ and write simply $u_t^{x_i}$ instead of $u_{t,0}^{x,i,N}$. What we need to show is then that

$$\|u_t^{x_i}-u_t^{y_i}\|_\infty\leqslant C\,(t^{-3/2}+1)\,|x_i-y_i|$$

holds for each $i=1,\ldots,N$ and for a constant $C>0$ depending only on ρ_0 and T. We show this inductively.

Let us then fix $i\in\{1,\ldots,N\}$ and define $v_t:=u_t^{x_i}-u_t^{y_i}$. For a solution of (14) we see that

$$\begin{aligned}
u_t^{x_i} &= G(t)*\delta_{x_i}-\int_0^t G(t-s)*\mathrm{div}\,(u_s^{x_i}f_s^N)\,\mathrm{d}s\\
&= G(t)*\delta_{x_i}-\int_0^t \nabla G(t-s)*(u_s^{x_i}f_s^N)\,\mathrm{d}s, \qquad (15)
\end{aligned}$$

where $G(t,x):=\frac{1}{2\pi t}\exp\!\left(-\frac{|x|^2}{2t}\right)$ denotes the heat kernel. By substracting the corresponding equations for $u_t^{x_i}$ and $u_t^{y_i}$ it follows

$$v_t=G(t)*(\delta_{x_i}-\delta_{y_i})-\int_0^t \nabla G(t-s)*(v_sf_s^N)\,\mathrm{d}s$$

and consequently, for $p\in[1,\infty]$,

$$\|v_t\|_p\leqslant\|G(t)*(\delta_{x_i}-\delta_{y_i})\|_p+\int_0^t\|\nabla G(t-s)*(v_sf_s^N)\|_p\,\mathrm{d}s$$

holds due to Bochner's Theorem. Next we use Young's inequality for convolutions[4]. We split the last integral into two parts and use Young's inequality with different exponents for each part

$$\begin{aligned}
\int_0^t\|\nabla G(t-s)*(v_sf_s^N)\|_p\,\mathrm{d}s &= \int_0^{t/2}\|\nabla G(t-s)*(v_sf_s^N)\|_p\,\mathrm{d}s\\
&\quad+\int_{t/2}^t\|\nabla G(t-s)*(v_sf_s^N)\|_p\,\mathrm{d}s\\
&\leqslant C\int_0^{t/2}\|\nabla G(t-s)\|_p\,\|v_s\|_1\,\mathrm{d}s
\end{aligned}$$

4. For two functions $a,b:\mathbb{R}^n\to\mathbb{R}$ and exponents $p,q,r\in[1,\infty]$ satisfying $1+\frac{1}{p}=\frac{1}{r}+\frac{1}{q}$ it holds

$$\|a*b\|_p\leqslant\|a\|_r\,\|b\|_q.$$

$$+C\int_{t/2}^{t}\|\nabla G(t-s)\|_{3/2}\,\|v_s\|_q\,\mathrm{d}s, \tag{16}$$

where $C:=\sup_{0\leqslant t\leqslant T}\|f_t^N\|_\infty$ is finite since $\|\rho_t^N\|_1$ is equal to $\|\rho_0\|_1$ and by Proposition 4 $\|\rho_t^N\|_\infty$ is also uniformly bounded in $t\in[0,T]$ and $N\in\mathbb{N}$. The choice of the exponent $r=3/2$ for the norm of ∇G in the second integral is as good as any other choice $r\in(1,2)$ since we just need the term $\|\nabla G\|_r$ to be integrable in $[0,t]$. Observe that with the previous bound for $\|v_t\|_p$ and taking $p_n:=q$ and $p_{n+1}:=p$ in (16) we find

$$\begin{aligned}\|v_t\|_{p_{n+1}} &\leqslant \|G(t)*(\delta_{x_i}-\delta_{y_i})\|_{p_{n+1}}+C\int_0^{t/2}\|\nabla G(t-s)\|_{p_{n+1}}\,\|v_s\|_1\,\mathrm{d}s\\ &\quad +C\int_{t/2}^{t}\|\nabla G(t-s)\|_{3/2}\,\|v_s\|_{p_n}\,\mathrm{d}s,\end{aligned} \tag{17}$$

where the relation between the exponents $p_{n+1}=3\,\frac{p_n}{3-p_n}$ follows from Young's inequality. Therefore, if we are able to estimate $\|v_t\|_1$ we can then iteratively estimate the L^p norms of v_t for higher exponents. Since the function $x\mapsto 3\,\frac{x}{3-x}$ on $[0,3)$ is strictly monotone increasing, grows to infinity as x approaches 3 and its first derivative is non-decreasing, it is already clear that starting at $p_1=1$ the exponent $p_k=\infty$ must be attained after a finite number k of steps. Specifically, if we take $p_1=1$, we reach the desired norm $\|v_t\|_\infty$ after $k=4$ steps. Below we go through the first two steps in detail, the last two can be completed analogously. We will need some estimates for the L^p norms of the heat kernel G and its derivative, which are given in Lemma 11.

Step $k=1$, $p_1=1$: We compute the first norm directly using a Grönwall-type inequality.

$$\begin{aligned}\|v_t\|_1 &\leqslant \|G(t,\cdot-x_0)-G(t,\cdot-y_0)\|_1+\int_0^t\|\nabla G(t-s)*(v_sf_s^N)\|_p\,\mathrm{d}s\\ &\leqslant \|G(t,\cdot-x_0)-G(t,\cdot-y_0)\|_1+\int_0^t\|\nabla G(t-s)\|_1\,\|v_s\|_1\,\|f_s^N\|_\infty\,\mathrm{d}s\\ &\leqslant C\,\frac{|x_0-y_0|}{t^{1/2}}+C\int_0^t(t-s)^{-1/2}\,\|v_s\|_1\,\mathrm{d}s.\end{aligned}$$

By Grönwall's inequality we find

$$\begin{aligned}\|v_t\|_1 &\leqslant C\,\frac{|x_0-y_0|}{t^{1/2}}+C\,|x_0-y_0|\int_0^t s^{-1/2}\,(t-s)^{-1/2}\,\mathrm{e}^{C\int_s^t(t-\sigma)^{-1/2}\mathrm{d}\sigma}\,\mathrm{d}s\\ &\leqslant C\,\frac{|x_0-y_0|}{t^{1/2}}+C\,\mathrm{e}^{Ct^{1/2}}\,|x_0-y_0|.\end{aligned}$$

Here we used that the integral $\int_0^t s^{-1/2}\,(t-s)^{-1/2}$ is finite since it can be split into

$$\int_0^t s^{-1/2}\,(t-s)^{-1/2}\,\mathrm{d}s=\int_0^{t/2}s^{-1/2}\,(t-s)^{-1/2}\,\mathrm{d}s+\int_{t/2}^{t}s^{-1/2}\,(t-s)^{-1/2}\,\mathrm{d}s,$$

and both terms are finite. Consequently

$$\|v_t\|_1\leqslant C\,(t^{-1/2}+1)\,|x_0-y_0|$$

holds for a constant C depending only on $\sup_{0\leqslant t\leqslant T}\|f_t^N\|_\infty$.

Step $k=2$, $p_2=\frac{3}{2}$: Recall that the next exponent is computed via the relationship $p_{n+1}=3\,\frac{p_n}{3-p_n}$. In this and the following steps we just need to substitute the found estimates into (17):

$$\begin{aligned}
\|v_t\|_{3/2} &\leqslant \|G(t,\cdot-x_0)-G(t,\cdot-y_0)\|_{3/2}+C\int_0^t \|\nabla G(t-s)\|_{3/2}\,\|v_s\|_1\,\mathrm{d}s\\
&\leqslant C\,\frac{|x_0-y_0|}{t^{5/6}}+C\int_0^t (t-s)^{-5/6}\,\|v_s\|_1\,\mathrm{d}s\\
&\leqslant C\,\frac{|x_0-y_0|}{t^{5/6}}+C\,|x_0-y_0|\int_0^t (t-s)^{-5/6}\,(s^{-1/2}+1)\,\mathrm{d}s\\
&\leqslant C\,\frac{|x_0-y_0|}{t^{5/6}}+C\,|x_0-y_0|\left(\int_0^{t/2} (t-s)^{-5/6}\,s^{-1/2}\,\mathrm{d}s+\int_{t/2}^t (t-s)^{-5/6}\,s^{-1/2}\,\mathrm{d}s\right)\\
&\quad +C\,|x_0-y_0|\,t^{1/6}\\
&\leqslant C\,(t^{-5/6}+t^{-1/3}+t^{1/6})\,|x_0-y_0| = C \leqslant (t^{-5/6}+1)\,|x_0-y_0|.
\end{aligned}$$

The last two steps with $k=3$, $p_3=3$ and $k=4$, $p_4=\infty$ are analogous.

□

As a consequence we find the following estimate:

Corollary 10. *Let $f\in L^1(\mathbb{R}^2)$ and define $F\colon\mathbb{R}^{2N}\to\mathbb{R}^{2N}$ by $F_i(z):=\frac{1}{N}\sum_{j\neq i} f(z_i-z_j)$, for $i=1,\ldots,N$. Then,*

$$\left|\mathbb{E}(F(Z_{t,s}^{x,N}))-\mathbb{E}\big(F\big(Z_{t,s}^{y,N}\big)\big)\right|_\infty \leqslant C\,((t-s)^{-3/2}+1)\,\|f\|_1\,|x-y|_\infty$$

holds for $x,y\in\mathbb{R}^{2N}$, $t\in[0,T]$ and $Z_{t,s}^{x,N},Z_{t,s}^{y,N}$ given by (9).

Note that the interaction force K^N is a function of this kind.

Proof. Let $i\in\{1,\ldots,N\}$.

$$\mathbb{E}(F(Z_t^x))_i=\mathbb{E}(F_i(Z_t^x))=\frac{1}{N}\sum_{j\neq i}\int f(z_i-z_j)\,u_t^{x_i}(z_i)\,u_t^{x_j}(z_j)\,\mathrm{d}z_i\,\mathrm{d}z_j.$$

Therefore

$$\begin{aligned}
|\mathbb{E}(F(Z_t^x))_i-\mathbb{E}(F(Z_t^y))_i| &= \frac{1}{N}\left|\sum_{j\neq i}\int f(z_i-z_j)\,(u_t^{x_i}(z_i)\,u_t^{x_j}(z_j)-u_t^{y_i}(z_i)\,u_t^{y_j}(z_j))\,\mathrm{d}z_i\,\mathrm{d}z_j\right|\\
&\leqslant \frac{1}{N}\sum_{j\neq i}\left|\int f(z_i-z_j)\,u_t^{x_i}(z_i)\,(u_t^{x_j}(z_j)-u_t^{y_j}(z_j))\,\mathrm{d}z_i\,\mathrm{d}z_j\right.\\
&\qquad\left.+\int f(z_i-z_j)\,u_t^{y_j}(z_j)\,(u_t^{x_i}(z_i)-u_t^{y_i}(z_i))\,\mathrm{d}z_i\,\mathrm{d}z_j\right|\\
&\leqslant \frac{1}{N}\sum_{j\neq i}(\|u_t^{x_j}-u_t^{y_j}\|_\infty\,\|f*u_t^{x_i}\|_1+\|u_t^{x_i}-u_t^{y_i}\|_\infty\,\|f*u_t^{x_j}\|_1)\\
&\leqslant \frac{1}{N}\sum_{j\neq i}C\,(t^{-3/2}+1)\,|x-y|_\infty\,(\|f\|_1\,\|u_t^{x_i}\|_1+\|f\|_1\,\|u_t^{x_i}\|_1)\\
&\leqslant C\,(t^{-3/2}+1)\,\|f\|_1\,|x-y|_\infty,
\end{aligned}$$

by Lemma 9. □

We finally collect some standard estimates for the heat kernel which we required in the proof of Lemma 9.

Lemma 11. (p-norm estimates of the heat kernel) *Let* $G(t,x) := \frac{1}{2\pi t}\exp\left(-\frac{|x|^2}{2t}\right)$ *and* $p \in [1,\infty]$. *Then there exists a constant* $C>0$ *such that the following holds:*

i. $\|G(t)\|_p \leqslant C\frac{1}{t^{1-1/p}}$ *and* $\|\nabla_x G(t)\|_p \leqslant C\frac{1}{t^{3/2-1/p}}$,

ii. $\|G(t,\cdot-x_0)-G(t,\cdot-y_0)\|_p \leqslant C\frac{|x_0-y_0|}{t^{3/2-1/p}}$.

Proof. i. $\|G(t)\|_p \leqslant C\frac{1}{t^{1-1/p}}$ for $p\in[1,\infty]$.

For $p=\infty$ the statement is clearly true.

For $1\leqslant p<\infty$

$$\begin{aligned}\|G(t)\|_p &= \frac{1}{2\pi t}\left(\int \exp\left(-\frac{p|x|^2}{2t}\right)\mathrm{d}^2x\right)^{1/p}\\ &= \frac{C}{t^{1-1/p}}\left(\int \exp(-p|y|^2)\,\mathrm{d}^2y\right)^{1/p}\\ &\leqslant \frac{C}{t^{1-1/p}}\left(\int \exp(-|y|^2)\,\mathrm{d}^2y\right)^{1/p}\\ &\leqslant \frac{C}{t^{1-1/p}}.\end{aligned}$$

Next we show that $\|\nabla_x G(t)\|_p \leqslant C\frac{1}{t^{3/2-1/p}}$ for $p\in[1,\infty]$. For $p=\infty$, since $a\exp(-a)$ is bounded, one has

$$|\nabla_x G(t,x)| = \left|\frac{x}{2\pi t^2}\exp\left(-\frac{|x|^2}{2t}\right)\right| = \frac{C}{t^{3/2}}\frac{|x|}{t^{1/2}}\exp\left(-\frac{|x|^2}{2t}\right) \leqslant \frac{C}{t^{3/2}},$$

for $(t,x)\in\mathbb{R}_0^+\times\mathbb{R}^2$. For $1\leqslant p<\infty$:

$$\begin{aligned}\|\nabla_x G(t)\|_p &= \frac{1}{2\pi t^2}\left(\int |x|^p\exp\left(-\frac{p|x|^2}{2t}\right)\mathrm{d}^2x\right)^{1/p}\\ &\leqslant \frac{C}{t^{3/2-1/p}}\left(\int |y|^p\exp(-p|y|^2)\,\mathrm{d}^2y\right)^{1/p}\\ &\leqslant \frac{C}{t^{3/2-1/p}}\left(\left\||\cdot|^p\exp\left(-\frac{p|\cdot|^2}{2}\right)\right\|_\infty\right)^{1/p}\left(\int \exp\left(-\frac{p|y|^2}{2}\right)\mathrm{d}^2y\right)^{1/p}\\ &\leqslant \frac{C}{t^{3/2-1/p}}\left\||\cdot|\exp\left(-\frac{|\cdot|^2}{2}\right)\right\|_\infty\left(\int \exp\left(-\frac{|y|^2}{2}\right)\mathrm{d}^2y\right)\\ &\leqslant \frac{C}{t^{3/2-1/p}}.\end{aligned}$$

ii. Let $V(t,x) := G(t,x-x_0)-G(t,x-y_0)$. For $p=\infty$, it follows from part i that

$$|V(t,x)| \leqslant \|\nabla_x G(t)\|_\infty |x_0-y_0| \leqslant C\frac{|x_0-y_0|}{t^{3/2}}.$$

For $p=1$ one can directly compute

$$\|V(t,\cdot)\|_1 \leqslant C\frac{|x_0-y_0|}{t^{1/2}}.$$

For $1<p<\infty$ then

$$\begin{aligned}\|V(t,\cdot)\|_p &\leqslant \|V(t,\cdot)\|_\infty^{(p-1)/p}\|V(t,\cdot)\|_1^{1/p}\\ &\leqslant C\left(\frac{|x_0-y_0|}{t^{3/2}}\right)^{(p-1)/p}\left(\frac{|x_0-y_0|}{t^{1/2}}\right)^{1/p}\\ &= C\frac{|x_0-y_0|}{t^{3/2-1/p}}.\end{aligned}$$

□

5. Proof of the main theorem

In this section we prove Theorem 1. We show that if the regularised real trajectory X^N given by (6) and the regularised mean-field trajectory Y^N solving (7) start at the same point then for given $T>0$, $\alpha\in(0,1/2)$ and $\gamma>0$

$$\mathbb{P}\left(\sup_{0\leqslant t\leqslant T}|X_t^N-Y_t^N|_\infty\geqslant N^{-\alpha}\right)\leqslant C_\gamma N^{-\gamma}$$

holds for an appropriate constant $C_\gamma=C_\gamma(\chi,\rho_0,T,\alpha,\gamma)$ and for all N larger than some $N_0=N_0(\rho_0,T,\alpha)$. This is done by two slightly different methods, depending on how big the elapsed time is. For short times the proof follows quite directly from a local Lipschitz bound of oder $\log N$ for K^N. For large times it gets more involved and we need to introduce the new process $Z_{t,s}^{X_s^N,N}$ starting at an intermediate time $s\in(0,t)$, which we show to be close to X_t^N and to Y_t^N. Recall that $Z_{t,s}^{X_s^N,N}$ is given by (9) with initial condition $Z_{s,s}^{X_s^N,N}=X_s^N$. In order to simplify the notation we will omit the superindex in $Z_{t,s}^{X_s^N,N}$ referring to to the initial condition X_s^N and denote just by $Z_{t,s}^N$ the solution of (9) with initial condition $Z_{s,s}^N=X_s^N$. In particular, the identities $Z_{t,0}^N=Y_t^N$ and $Z_{t,t}^N=X_t^N$ hold.

It can clarify things to first present the main ideas in a simple, intuitive way before we start with the actual proof. Let us already fix $T>0$, $\alpha\in(0,1/2)$ and $\delta:=\frac{1}{2}\left(\frac{1}{2}-\alpha\right)>0$.

Sketch of the proof.

We show that the real and the mean-field trajectories remain close by controlling the growth of their distance $|Z_{t,t}-Z_{t,0}|_\infty$. The derivative of $|Z_{t,t}-Z_{t,0}|_\infty$ is bounded by the difference of the respective forces $|K^N(Z_{t,t}^N)-\overline{K}_t^N(Z_{t,0}^N)|_\infty$, which we can prove to be sufficiently small if we previously assume that the particles are close enough, more specifically if $|Z_{t,t}-Z_{t,0}|_\infty\leqslant N^{-\alpha}$. Since under this assumption we are able to show that the difference $|Z_{t,t}-Z_{t,0}|_\infty$ goes to zero as $N\to\infty$ faster than $N^{-\alpha}$, or equivalently that $N^\alpha|Z_{t,t}-Z_{t,0}|_\infty\to0$, the initial condition $|Z_{t,t}-Z_{t,0}|_\infty\leqslant N^{-\alpha}$ is actually harmless. This idea can be formalised by considering the following object:

$$J_t^N:=\min\{1,N^\alpha|Z_{t,t}-Z_{t,0}|_\infty\},\qquad 0\leqslant t\leqslant T.$$

The corresponding object in the real proof looks somewhat more complicated, but the underlying arguments are the same. The time derivative of J_t^N is less or equal than 0 on the event $\{|Z_{t,t}-Z_{t,0}|_\infty>N^{-\alpha}\}$ (since in this case J_t^N attains its maximum value 1). Therefore, we just need to control the derivative of $N^\alpha|Z_{t,t}-Z_{t,0}|_\infty$ when $|Z_{t,t}-Z_{t,0}|_\infty\leqslant N^{-\alpha}$. Let us then look at the growth of the scaled difference $N^\alpha|Z_{t,t}^N-Z_{t,0}^N|_\infty$:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}N^\alpha|Z_{t,t}^N-Z_{t,0}^N| &\leqslant N^\alpha|K^N(Z_{t,t}^N)-\overline{K}_t^N(Z_{t,0}^N)|\\ &\leqslant N^\alpha|K^N(Z_{t,t}^N)-K^N(Z_{t,0}^N)|+N^\alpha|K^N(Z_{t,0}^N)-\overline{K}_t^N(Z_{t,0}^N)|.\end{aligned}$$

From the law of large numbers (Proposition 8) for the i.i.d. random variables $Z_{t,0}^{1,N}$, …, $Z_{t,0}^{N,N}$ it follows that the last quantity is small with high probability. Moreover, thanks to the the local Lipschitz bound for the interaction force K^N (Lemma 7), together with the law of large numbers we can show that K^N is locally Lipschitz with bound $L(N) \sim \log N$. With this we can prove that, if $|Z_{t,t} - Z_{t,0}|_\infty \leqslant N^{-\alpha}$, then

$$\frac{\mathrm{d}}{\mathrm{d}t} N^\alpha |Z_{t,t}^N - Z_{t,0}^N| \leqslant N^\alpha \log N \, |Z_{t,t}^N - Z_{t,0}^N| + N^{-\delta}$$

holds with high probability. Then Grönwall's inequality yields the desired result, but only if the time t is small enough, since we obtain

$$N^\alpha |Z_{t,t}^N - Z_{t,0}^N| \leqslant t N^{-\delta} \mathrm{e}^{t \log N},$$

which goes to zero as N grows to infinity if $t \in [0, \tau_0]$, with $\tau_0 \sim (\log N)^{-\varepsilon}$ (or $\tau_0 < \delta$, independently of N).

If t is bigger than τ_0 we proceed in a different way, combining the above idea for small times with the diffusive effect of the Brownian motion. We introduce the new process $Z_{t,s}^N$ starting at position X_s^N at an intermediate time $s \in (0, t)$ and show that it is close to X_t^N and to Y_t^N. The intermediate time $s \in (0, t)$ is chosen such that $t - s \leqslant \tau_0$, so that $Z_{t,s}^N$ is close to $Z_{t,t}^N$ by the same argument as before (under the additional assumption $|Z_{t,t}^N - Z_{t,s}^N| \leqslant N^{-\alpha}$), but still not too small so we can control the distance between $Z_{t,s}^N$ and $Z_{t,0}^N$. For the latter we make use of an entirely different argument: we do not compare the trajectories but the probability densities of $Z_{t,s}^N$ and $Z_{t,0}^N$. Both processes evolved during a period of time $(t - s)$ according to the mean-field dynamics and starting at time s from their respective positions $Z_{s,s}^N$ and $Z_{s,0}^N$. If we also assume that $Z_{s,s}^N$ and $Z_{s,0}^N$ are close, i.e. if $|Z_{s,s}^N - Z_{s,0}^N| \leqslant N^{-\alpha}$, then their probability densities are close in L^∞ as a consequence of the diffusion (Lemma 9). The additional conditions $|Z_{t,t}^N - Z_{t,s}^N| \leqslant N^{-\alpha}$ and $|Z_{s,s}^N - Z_{s,0}^N| \leqslant N^{-\alpha}$ are included into J_t^N as

$$J_t^N := \min\left\{1, \sup_{0 \leqslant \tau \leqslant s \leqslant t} N^\alpha |Z_{s,s}^N - Z_{s,\tau}^N|_\infty\right\}, \qquad 0 \leqslant t \leqslant T.$$

This object still differs from the one in the real proof by some extra factors, which have the only purpose of improving the rate of convergence.

Proof of Theorem 1.

As we previously argued, instead of directly considering the evolution of the difference $|Z_{t,t}^N - Z_{t,0}^N|_\infty$ we work with a more complicated but technically convenient stochastic process. For $T > 0$, $\alpha \in (0, 1/2)$ and $\delta = \frac{1}{2}\left(\frac{1}{2} - \alpha\right) > 0$ fixed at the beginning of the section we consider the auxiliary process

$$J_t^N := \min\left\{1, \sup_{0 \leqslant s \leqslant t} \mathrm{e}^{C_N(T-s)} \sup_{0 \leqslant \tau \leqslant s} (N^\alpha f_N(s - \tau) \, |Z_{s,s}^N - Z_{s,\tau}^N|_\infty + N^{-\delta})\right\}, \quad 0 \leqslant t \leqslant T, \tag{18}$$

where $C_N := 18 (\log N)^{3/4}$ and $f_N : [0, \infty) \to \mathbb{R}$ is a positive, monotone decreasing, smooth function satisfying

$$f_N(t) = \begin{cases} \frac{4}{t \log N + (\log N)^{-1/4}}, & 0 \leqslant t \leqslant 2 (\log N)^{-1}, \\ 1, & t \geqslant 4 (\log N)^{-1}, \end{cases}$$

and $-C \log N f_N^2(t) \leqslant f_N'(t) \leqslant 0$, for some positive constant C, $t \in [0, \infty)$.

As we shall see, the process J_t^N helps us control the maximal distance $|Z_{s,s}^N - Z_{s,\tau}^N|_\infty$ for all intermediate times and the parameters in J_t^N are optimised for the desired rate of convergence. We now explain how to express our problem in terms of this new process. For $s \geqslant \tau \geqslant 0$ let

$$a(\tau,s) := N^\alpha f_N(s-\tau)\,|Z_{s,s}^N - Z_{s,\tau}^N|_\infty + N^{-\delta}. \tag{19}$$

Since for each t the bound

$$\sup_{0\leqslant s\leqslant t} N^\alpha |Z_{s,s}^N - Z_{s,0}^N|_\infty \leqslant \sup_{0\leqslant s\leqslant t} \mathrm{e}^{C_N(T-s)} \sup_{0\leqslant \tau\leqslant s} a(\tau,s)$$

holds true, $J_t^N < 1$ implies that $\sup\limits_{0\leqslant s\leqslant t} \mathrm{e}^{C_N(T-s)} \sup\limits_{0\leqslant \tau\leqslant s} a(\tau,s) = J_t^N < 1$, and $\sup_{0\leqslant s\leqslant t} |Z_{s,s}^N - Z_{s,0}^N|_\infty < N^{-\alpha}$ follows. Moreover, since $\mathrm{e}^{C_N T}$ grows slower than N^ε for any $\varepsilon > 0$, there exists $N_0 \in \mathbb{N}$ depending on T and α such that if $N \geqslant N_0$ then $J_0^N = \mathrm{e}^{C_N T}\, N^{-\delta}$ is bounded by some constant, say $1/2$. Therefore, we can estimate

$$\begin{aligned}
\mathbb{P}\Bigg(\sup_{0\leqslant t\leqslant T} |Z_{t,t}^N - Z_{t,0}^N|_\infty \geqslant N^{-\alpha}\Bigg) &\leqslant \mathbb{P}(J_T^N \geqslant 1)\\
&\leqslant \mathbb{P}(J_T^N - J_0^N \geqslant 1/2)\\
&\leqslant 2\,\mathbb{E}(J_T^N - J_0^N)\\
&= 2\int_0^T \mathbb{E}(\overline{D}_t^+ J_t^N)\,\mathrm{d}t,
\end{aligned}$$

where $\overline{D}_t^+$ denotes the upper right Dini derivative given in Definition 12 at the end of this proof. This definition is followed by Proposition 13, where the Fundamental Theorem of Calculus for this notion of derivative is stated, together with a few other relevant properties. Finally, we provide in Lemma 14 a formula for computing the Dini derivative of functions like J_t defined over suprema of other functions.

For a given $\gamma > 0$ the problem then reduces to finding a constant C_γ such that

$$\mathbb{E}(\overline{D}_t^+ J_t^N) \leqslant C_\gamma N^{-\gamma}, \qquad t \in [0,T].$$

Let us next compute the Dini derivative of J_t^N (18) using Lemma 14. Note that we can write it as

$$J_t^N = \min\Big\{1, \sup_{0\leqslant \tau\leqslant s\leqslant t} g(\tau,s)\Big\},$$

where

$$g(\tau,s) := \mathrm{e}^{C_N(T-s)}\,(N^\alpha f_N(s-\tau)\,|Z_{s,s}^N - Z_{s,\tau}^N|_\infty + N^{-\delta}). \tag{20}$$

It is clear that $\overline{D}_t^+ J_t^N \leqslant \max\{0, \overline{D}_t^+ \sup_{0\leqslant\tau\leqslant s\leqslant t} g(\tau,s)\}$. Moreover, the function g satisfies the conditions of Lemma 14 below. Indeed, the diagonal points are minimal for g so the supremum cannot be attained there or g would be constant. We now show that g has finite right upper Dini derivatives in both variables and that $\overline{D}_s^+ g$ is continuous. The function

$$\varphi(\tau,s) := Z_{s,s}^N - Z_{s,\tau}^N = \int_\tau^s (K^N(Z_{u,u}^N) - \overline{K}_u^N(Z_{u,\tau}^N))\,\mathrm{d}u$$

is continuous, Lipschitz in τ and in s and has a partial derivative wrt. s. In particular, $|Z_{s,s}^N - Z_{s,\tau}^N|_\infty$ is Lipschitz in τ and s, so its right upper Dini derivatives in τ and s are finite. Since $\partial^+ \max\{f, g\} = \max\{\partial^+ f, \partial^+ g\}$ if f and g are right-differentiable, it follows that

$$\overline{D}_s^+ |Z_{s,s}^N - Z_{s,\tau}^N|_\infty = \partial_s^+ |Z_{s,s}^N - Z_{s,\tau}^N|_\infty = |\partial_s^+(Z_{s,s}^N - Z_{s,\tau}^N)|_\infty.$$

Finally, from the sum and product rules (Proposition 13) it follows that $D_\tau^+ g$ and $D_s^+ g$ are finite that $\overline{D}_s^+ g$ is a continuous function, namely

$$\begin{aligned} \overline{D}_s^+ g(\tau,s) &= \partial_s^+ g(\tau,s) \\ &= -\mathrm{e}^{C_N(T-s)}(C_N a(\tau,s) - N^\alpha f_N'(s-\tau)\,|Z_{s,s}^N - Z_{s,\tau}^N|_\infty) \\ &\quad + \mathrm{e}^{C_N(T-s)} N^\alpha f_N(s-\tau)\,\big|K^N(Z_{s,s}^N) - \overline{K}_s^N(Z_{s,\tau}^N)\big|_\infty, \end{aligned} \tag{21}$$

where $a(\tau,s)$ is defined in (19). We can then apply Lemma 14 to the function $\sup_{0\leqslant\tau\leqslant s\leqslant t} g(\tau,s)$ with the set of boundary maximal points

$$\hat{M}_t := \Big\{(\tau_t, t) : 0 \leqslant \tau_t \leqslant t,\ \sup_{0\leqslant\tau\leqslant s\leqslant t} g(\tau,s) = g(\tau_t,t)\Big\}$$

and find the following estimate:

$$\overline{D}_t^+ J_t^N \leqslant \max\Big\{0, \sup_{(\tau,t)\in\hat{M}_t} \overline{D}_s^+ g(\tau,t)\Big\}.$$

If $\hat{M}_t = \varnothing$ then $\overline{D}_t^+ J_t^N \leqslant 0$ and there is nothing to prove, so we assume this set is not empty. Then there exists $(\tau_t,\ t) \in \hat{M}_t$ where the supremum of $\overline{D}_s^+ g$ over $\hat{M}_t$ is attained (by the continuity of $\overline{D}_s^+ g$ and compactness of $\hat{M}_t$). Let us continue by trivially reducing the problem to a smaller set where $|Z_{s,s}^N - Z_{s,\tau}^N|_\infty \leqslant N^{-\alpha}$ holds for each $0\leqslant\tau\leqslant s\leqslant t$. Consider the event $\mathcal{A}_t := \{\overline{D}_t^+ J_t^N > 0\}$. Since $\mathcal{A}_t \subseteq \{\overline{D}_s^+ g(\tau_t,t) \geqslant \overline{D}_t^+ J_t^N\}$ it holds that

$$\mathbb{E}(\overline{D}_t^+ J_t^N) = \mathbb{E}(\overline{D}_t^+ J_t^N|\mathcal{A}_t^c) + \mathbb{E}(\overline{D}_t^+ J_t^N|\mathcal{A}_t) \leqslant 0 + \mathbb{E}(\overline{D}_t^+ J_t^N|\mathcal{A}_t) \leqslant \mathbb{E}(\overline{D}_s^+ g(\tau_t,t)|\mathcal{A}_t), \tag{22}$$

where by $\mathbb{E}(X|A)$ we denote the *restricted expectation* of X to the set A, i.e. $\mathbb{E}(X\,\mathbb{1}_A)$.

We shall prove that the latter is bounded by $C_\gamma N^{-\gamma}$ for some constant $C_\gamma \geqslant 0$. Note that in $\mathcal{A}_t$ one has $J_t^N \leqslant 1$ and in particular $\sup_{0\leqslant\tau\leqslant s\leqslant t} |Z_{s,s}^N - Z_{s,\tau}^N|_\infty \leqslant N^{-\alpha}$ holds. As a first estimate we can prove that in this set the bound $\overline{D}_s^+ g(\tau_t,t)$ of the derivative $\overline{D}_t^+ J_t^N$ grows slower than N^2: Using that

$$|f_N'(t-\tau)| = C \log N f_N^2(t-\tau) \leqslant C(\log N)^{3/2}$$

and

$$\big|K^N(Z_{t,t}^N) - \overline{K}_t^N(Z_{t,\tau}^N)\big| \leqslant C N^\alpha,$$

we find that in $\mathcal{A}_t$ and for each $0\leqslant\tau\leqslant t$

$$\begin{aligned} \overline{D}_s^+ g(\tau,t) &\leqslant \mathrm{e}^{C_N(T-t)}(C_N a(\tau,t) + N^\alpha |f_N'(t-\tau)|\,|Z_{t,t}^N - Z_{t,\tau}^N|) \\ &\quad + \mathrm{e}^{C_N(T-t)} N^\alpha f_N(t-\tau)\,\big|K^N(Z_{t,t}^N) - \overline{K}_t^N(Z_{t,\tau}^N)\big| \\ &\leqslant C\,\mathrm{e}^{C_N(T-t)}((\log N)^{3/4} + N^\alpha(\log N)^{3/2} N^{-\alpha} + N^\alpha(\log N)^{1/4} N^\alpha) \\ &\leqslant C\,\mathrm{e}^{C_N T} N^{3/2} < C N^2. \end{aligned} \tag{23}$$

With bounds (22) and (23) in mind we proceed to show $\mathbb{E}(\overline{D}_t^+ J_t^N) \leqslant C_\gamma N^{-\gamma}$. For $t\in[0,T]$ and $(\tau_t,\ t) \in \hat{M}_t$ where the supremum of $\overline{D}_s^+ g$ is attained we need to distinguish between two cases depending on the difference $t - \tau_t$.

Case 1: $t - \tau_t \leqslant 2(\log N)^{-1}$.

Here we show that $\overline{D}_s^+ g(\tau_t,\ t) \leqslant 0$ holds with high enough probability. The most important ingredient is that with high probability the regularised force K^N is locally Lipschitz with constant of order $\log N$, and this bound is good enough for short elapsed times. In order to develop this idea let us first recall the result of Lemma 7 applied to K^N. In the notation of Lemma 7, K^N is equal to $K^{v(N)}$ for $v(N) := N^\alpha$ and so it is locally Lipschitz with bound $L^{v(N)}$, which was defined as

$$L_i^{v(N)}(y_1,\dots,y_N) = \frac{\chi}{N}\sum_{j\neq i} l^{v(N)}(y_i - y_j), \quad (y_1,\dots,y_N) \in \mathbb{R}^{2N}$$

for

$$l^\nu(y) = \begin{cases} \frac{16}{|y|^2}, & |y| \geqslant 4\,\nu^{-1}, \\ \nu^2, & |y| \leqslant 4\,\nu^{-1}, \end{cases} \quad y \in \mathbb{R}^2.$$

Let us just write L^N instead of $L^{\nu(N)}$ and denote by $\overline{L}_t^N$ the averaged version of L^N given by

$$\overline{L}_{t,i}^N(y_1,\ldots,y_N) := \chi(l^{\nu(N)} * \rho_t^N)(y_i).$$

Then, from Lemma 7 it follows that if $x,y \in \mathbb{R}^{2N}$, $|x-y|_\infty \leqslant N^{-\alpha}$, then

$$|K^N(x) - K^N(y)| \leqslant 2\,|L^N(y)|\,|x-y|. \tag{24}$$

Notice that $\|\overline{L}_t^N\|_\infty$ is of order $O(\log N)$: Indeed, $l^{\nu(N)} = l_1^{\nu(N)} + l_\infty^{\nu(N)} \in L^1(\mathbb{R}^2) + L^\infty(\mathbb{R}^2)$ with integrable part satisfying $\|l_1^{\nu(N)}\|_1 = O(\log N)$ and ρ_t^N is bounded in $L^1(\mathbb{R}^2) \cap L^\infty(\mathbb{R}^2)$ uniformly in N and $t \in [0,T]$. If L^N is close to $\overline{L}_t^N$, then K^N is locally Lipschitz with a constant of the appropriate order. We also need that K^N be well approximated by the mean-field force $\overline{K}_t^N$ acting on the i.i.d. particles Y_t^N. In this spirit we introduce the set

$$\mathscr{B}_t^1 := \{|K^N(Y_t^N) - \overline{K}_t^N(Y_t^N)| \leqslant N^{-(\alpha+\delta)}\} \cap \{|L^N(Y_t^N) - \overline{L}_t^N(Y_t^N)| \leqslant C\}, \tag{25}$$

where the two desired properties hold. Moreover, as a consequence of the law of large numbers (Proposition 8) the measure of the event $\Omega \backslash \mathscr{B}_t^1$ decays to zero as N grows to infinity faster than any polynomial in N (see Proposition 15 at the end of this section).

Recalling (22) we write

$$\mathbb{E}(\overline{D}_t^+ J_t^N) \leqslant \mathbb{E}(\overline{D}_s^+ g(\tau_t,t)|\mathscr{A}_t) = \mathbb{E}(\overline{D}_s^+ g(\tau_t,t)|\mathscr{A}_t \backslash \mathscr{B}_t^1) + \mathbb{E}(\overline{D}_s^+ g(\tau_t,t)|\mathscr{A}_t \cap \mathscr{B}_t^1). \tag{26}$$

Since $\overline{D}_s^+ g(\tau_t,t)$ grows in the set $\mathscr{A}_t$ only polynomially in N by estimate (23), by Proposition 15 we can find a positive constant C_γ such that the first term in (26) satisfies

$$\mathbb{E}(\overline{D}_s^+ g(\tau_t,t)|\mathscr{A}_t \backslash \mathscr{B}_t^1) \leqslant C_\gamma N^{-\gamma}. \tag{27}$$

It is therefore enough to prove that $\overline{D}_s^+ g(\tau_t,t) \leqslant 0$ holds in $\mathscr{A}_t \cap \mathscr{B}_t^1$.

Note that $\overline{D}_s^+ g(\tau_t,t) \leqslant 0$ holds if the following inequality is true:

$$\begin{aligned} f_N(t-\tau_t)\,|K^N(Z_{t,t}^N) - \overline{K}_t^N(Z_{t,\tau_t}^N)| \;&\leqslant\; -f_N'(t-\tau_t)\,|Z_{t,t}^N - Z_{t,\tau_t}^N| \\ &\quad + C_N\,(f_N(t-\tau_t)\,|Z_{t,t}^N - Z_{t,\tau_t}^N| + N^{-(\alpha+\delta)}). \end{aligned} \tag{28}$$

We next estimate the term $|K^N(Z_{t,t}^N) - \overline{K}_t^N(Z_{t,\tau_t}^N)|$ in the set in $\mathscr{A}_t \cap \mathscr{B}_t^1$ by splitting in three:

$$\begin{aligned} |K^N(Z_{t,t}^N) - \overline{K}_t^N(Z_{t,\tau_t}^N)| \;&\leqslant\; |K^N(Z_{t,t}^N) - K^N(Z_{t,0}^N)| + |K^N(Z_{t,0}^N) - \overline{K}_t^N(Z_{t,0}^N)| \\ &\quad + |\overline{K}_t^N(Z_{t,0}^N) - \overline{K}_t^N(Z_{t,\tau_t}^N)|. \end{aligned}$$

The last term is the least problematic, since the function $\overline{K}_t^N$ is globally Lipschitz. The middle term is small in the event $\mathscr{B}_t^1$ by definition (recall that $Z_{t,0}^N = Y_t^N$). For the first one we use that in this event the force K^N is locally Lipschitz with bound of order $\log N$: with (24), $|Z_{t,t}^N - Z_{t,0}^N| \leqslant N^{-\alpha}$ in $\mathscr{A}_t$ and $|L^N(Z_{t,0}^N) - \overline{L}_t^N(Z_{t,0}^N)| \leqslant C$ in $\mathscr{B}_t^1$ we find

$$\begin{aligned} |K^N(Z_{t,t}^N) - K^N(Z_{t,0}^N)| \;&\leqslant\; 2\,|L^N(Z_{t,0}^N)|\,|Z_{t,t}^N - Z_{t,0}^N| \leqslant 2\,(C + \|\overline{L}_t^N\|_\infty)\,|Z_{t,t}^N - Z_{t,0}^N| \\ &\leqslant\; 2\,(C + \log N)\,|Z_{t,t}^N - Z_{t,0}^N|. \end{aligned}$$

Consequently,

$$\begin{aligned}|K^N(Z^N_{t,t})-\overline{K}^N_t(Z^N_{t,\tau_t})| &\leqslant |K^N(Z^N_{t,t})-K^N(Z^N_{t,0})|+|K^N(Z^N_{t,0})-\overline{K}^N_t(Z^N_{t,0})|\\ &\quad +|\overline{K}^N_t(Z^N_{t,0})-\overline{K}^N_t(Z^N_{t,\tau_t})|\\ &\leqslant 2\,(C+\log N)\,|Z^N_{t,t}-Z^N_{t,0}|+N^{-(\alpha+\delta)}+L\,|Z^N_{t,t}-Z^N_{t,0}|\\ &\leqslant (2\log N+2\,C+L)\,|Z^N_{t,t}-Z^N_{t,0}|+L\,|Z^N_{t,t}-Z^N_{t,\tau_t}|+N^{-(\alpha+\delta)},\end{aligned}$$

where L is the Lipschitz constant of $\overline{K}^N_t$ (uniform in $t\in[0,T]$). Now observe that, by the definition of J^N_t, $f_N(t-s)\,|Z^N_{t,t}-Z^N_{t,s}|\leqslant f_N(t-\tau_t)\,|Z^N_{t,t}-Z^N_{t,\tau_t}|$ holds for each $0\leqslant s\leqslant t$. Therefore, we can choose a maybe greater N_0, depending now also on the Lipschitz constant L, such that for $N\geqslant N_0$ we have

$$\begin{aligned}|K^N(Z^N_{t,t})-\overline{K}^N_t(Z^N_{t,\tau_t})| &\leqslant 2\,(C+\log N)\frac{f_N(t-\tau_t)}{f_N(t)}\,|Z^N_{t,t}-Z^N_{t,\tau_t}|+L|Z^N_{t,t}-Z^N_{t,\tau_t}|+N^{-(\alpha+\delta)}\\ &\leqslant 3\log N f_N(t-\tau_t)\,|Z^N_{t,t}-Z^N_{t,\tau_t}|+N^{-(\alpha+\delta)}\\ &\leqslant -\frac{f_N'(t-\tau_t)}{f_N(t-\tau_t)}\,|Z^N_{t,t}-Z^N_{t,\tau_t}|+\frac{C_N}{f_N(t-\tau_t)}N^{-(\alpha+\delta)},\end{aligned}$$

which proves (28). Here we used that $1\leqslant f\leqslant C_N$ and $3\log N f_N^2(t-\tau_t)\leqslant -f_N'(t-\tau_t)$. Consequently $\overline{D}^+_s g(\tau_t,t)\leqslant 0$ holds in the set $\mathscr{A}_t\cap\mathscr{B}^1_t$ and with (27)

$$\mathbb{E}(\overline{D}^+_t J^N_t)\leqslant\mathbb{E}(\overline{D}^+_s g(\tau_t,t)|\mathscr{A}_t\backslash\mathscr{B}^1_t)+\mathbb{E}(\overline{D}^+_s g(\tau_t,t)|\mathscr{A}_t\cap\mathscr{B}^1_t)\leqslant C_\gamma N^{-\gamma}$$

as required.

Case 2: $t-\tau_t>2\,(\log N)^{-1}$.

The key now is to introduce the process $Z^N_{t,s}$ starting at an appropriate intermediate time $s\in[0,t]$ and show that it is close to both the real trajectory X^N_t and the mean-field trajectory Y^N_t. That it is close to the real trajectory is proven by the same argument as in the previous case, using that the elapsed time $t-s$ is small enough. We compare $Z^N_{t,s}$ to the mean-field trajectory by proving that their densities are close in L^∞ thanks to the diffusive effect of the Brownian motion (Lemma 9 and Corollary 10). We also need to split the interaction force K^N into $K^N=K^N_1+K^N_2$, where K^N_2 is the result of choosing a wider cutoff of order $(\log N)^{-3/2}$ in the force kernel k and the most singular part is “isolated” in $K^N_1:=K^N-K^N_2$. More precisely, in the notation of Lemma 7 let $k^N_2:=k^{v_2(N)}$ for $v_2(N):=(\log N)^{-3/2}$ and define $k^N_1:=k^N-k^N_2$. For $i=1,\ldots,N$, the i-th components of K^N_1 and K^N_2 are then given by

$$(K^N_1)_i(x_1,\ldots,x_N)\;:=\;-\frac{\chi}{N}\sum_{j\neq i}k^N_1(x_i-x_j),\quad (x_1,\ldots,x_N)\in\mathbb{R}^{2N},$$

and

$$(K^N_2)_i(x_1,\ldots,x_N)\;:=\;-\frac{\chi}{N}\sum_{j\neq i}k^N_2(x_i-x_j),\quad (x_1,\ldots,x_N)\in\mathbb{R}^{2N}. \tag{29}$$

We denote the local Lipschitz bound for K^N_2 given by Lemma 7 as $L^N_2:=L^{v_2(N)}$ and its averaged version as $\overline{L}^N_{2,t}$, defined analogously to $\overline{L}^N_t$. Let us denote by $\mathscr{B}^2_t$ the intersection of the set $\mathscr{B}^1_t$ from the previous case and the set $\{|L^N_2(Y^N_t)-\overline{L}^N_2(Y^N_t)|\leqslant C\}$ concerning the Lipschitz bound of the second part K^N_2 of K^N:

$$\mathscr{B}^2_t:=\mathscr{B}^1_t\cap\{|L^N_2(Y^N_t)-\overline{L}^N_{2,t}(Y^N_t)|\leqslant C\}. \tag{30}$$

We write again

$$\mathbb{E}(\overline{D}^+_t J^N_t)\leqslant\mathbb{E}(\overline{D}^+_s g(\tau_t,t)|\mathscr{A}_t)=\mathbb{E}(\overline{D}^+_s g(\tau_t,t)|\mathscr{A}_t\backslash\mathscr{B}^2_t)+\mathbb{E}(\overline{D}^+_s g(\tau_t,t)|\mathscr{A}_t\cap\mathscr{B}^2_t).$$

The first term is bounded as in the previous section: due to the exponential decay of the measure of $\mathscr{A}_t \backslash \mathscr{B}_t^2$ (proven in Proposition 15 below) in contrast to the milder polynomial growth of $\overline{D}_s^+ g(\tau_t, t)$, we find a constant $C_\gamma \geqslant 0$ such that

$$\mathbb{E}(\overline{D}_s^+ g(\tau_t, t) | \mathscr{A}_t \backslash \mathscr{B}_t^2) \leqslant C_\gamma N^{-\gamma}. \tag{31}$$

It remains to show that also $\mathbb{E}(\overline{D}_s^+ g(\tau_t, t) | \mathscr{A}_t \cap \mathscr{B}_t^2) \leqslant C_\gamma N^{-\gamma}$ holds (for a possibly different constant C_γ, which we do not rename for simplicity of notation).

Notice that $\mathbb{E}(\overline{D}_s^+ g(\tau_t, t) | \mathscr{A}_t \cap \mathscr{B}_t^2) \leqslant C_\gamma N^{-\gamma}$ holds if the following inequality is true:

$$\begin{aligned} f_N(t-\tau_t)\, \mathbb{E}(|K^N(Z_{t,t}^N) - \overline{K}_t^N(Z_{t,\tau_t}^N)| | \mathscr{A}_t \cap \mathscr{B}_t^2) \;\leqslant\; & -f_N'(t-\tau_t)\, \mathbb{E}(|Z_{t,t}^N - Z_{t,\tau_t}^N| | \mathscr{A}_t \cap \mathscr{B}_t^2) \\ & + C_N f_N(t-\tau_t)\, \mathbb{E}(|Z_{t,t}^N - Z_{t,\tau_t}^N| | \mathscr{A}_t \cap \mathscr{B}_t^2) \\ & + C_N N^{-(\alpha+\delta)}\, \mathbb{P}(\mathscr{A}_t \cap \mathscr{B}_t^2) \\ & + C_\gamma N^{-\gamma}. \end{aligned} \tag{32}$$

To this end we write as before

$$\begin{aligned} |K^N(Z_{t,t}^N) - \overline{K}_t^N(Z_{t,\tau_t}^N)| \;\leqslant\; & |K^N(Z_{t,t}^N) - K^N(Z_{t,0}^N)| + |K^N(Z_{t,0}^N) - \overline{K}_t^N(Z_{t,0}^N)| \\ & + |\overline{K}_t^N(Z_{t,0}^N) - \overline{K}_t^N(Z_{t,\tau_t}^N)|. \end{aligned} \tag{33}$$

The last two terms can be bounded in the same way as in the previous section, but for $|K^N(Z_{t,t}^N) - K^N(Z_{t,0}^N)|$ we can no longer use the corresponding Lipschitz bound from Lemma 7 directly. Here we need to add the intermediate time $s = t - (\log N)^{-3/2}$ and to split the force into $K^N = K_1^N + K_2^N$ as described in (29), which results in

$$\begin{aligned} |K^N(Z_{t,t}^N) - K^N(Z_{t,0}^N)| \;\leqslant\; & |K^N(Z_{t,t}^N) - K^N(Z_{t,s}^N)| + |K^N(Z_{t,s}^N) - K^N(Z_{t,0}^N)| \\ \leqslant\; & |K^N(Z_{t,t}^N) - K^N(Z_{t,s}^N)| + |K_1^N(Z_{t,s}^N) - K_1^N(Z_{t,0}^N)| \\ & + |K_2^N(Z_{t,s}^N) - K_2^N(Z_{t,0}^N)|. \end{aligned} \tag{34}$$

We can now use the Lipschitz bound for the first and third terms in (34): By (24), in $\mathscr{A}_t \cap \mathscr{B}_t^2$ it holds that

$$\begin{aligned} |K^N(Z_{t,t}^N) - K^N(Z_{t,s}^N)| \;\leqslant\; & 2\, |L^N(Z_{t,s}^N)|\, |Z_{t,t}^N - Z_{t,s}^N| \\ \leqslant\; & 6\, (C + \|\overline{L}_t^N\|_\infty)\, |Z_{t,t}^N - Z_{t,s}^N| \\ \leqslant\; & 12 \log N\, |Z_{t,t}^N - Z_{t,s}^N| \\ \leqslant\; & 12 \log N \frac{f_N(t-\tau_t)}{f_N(t-s)}\, |Z_{t,t}^N - Z_{t,\tau_t}^N| \\ \leqslant\; & 6\, (\log N)^{3/4} f_N(t-\tau_t)\, |Z_{t,t}^N - Z_{t,\tau_t}^N|, \end{aligned} \tag{35}$$

since $f_N(s-r)\, |Z_{s,s} - Z_{s,r}| \leqslant f_N(t-\tau_t)\, |Z_{t,t}^N - Z_{t,\tau_t}^N|$ is true for each $0 \leqslant r \leqslant s \leqslant t$, and $f_N(t-s) \geqslant 2\, (\log N)^{1/4}$. We analogously obtain the following estimate for the third term in (34)

$$\begin{aligned} |K_2^N(Z_{t,s}^N) - K_2^N(Z_{t,0}^N)| \;\leqslant\; & 2\, |L_2^N(Z_{t,0}^N)|\, |Z_{t,s}^N - Z_{t,0}^N| \\ \leqslant\; & 2\, (\|\overline{L}_{2,t}^N\|_\infty + C)\, |Z_{t,s}^N - Z_{t,0}^N| \\ \leqslant\; & 4 \log\log N\, |Z_{t,s}^N - Z_{t,0}^N| \\ \leqslant\; & 4 \log\log N f_N(t-\tau_t) \left(\frac{1}{f_N(t-s)} + \frac{1}{f_N(t)} \right) |Z_{t,t}^N - Z_{t,\tau_t}^N| \\ \leqslant\; & 8 \log\log N f_N(t-\tau_t)\, |Z_{t,t}^N - Z_{t,\tau_t}^N|. \end{aligned} \tag{36}$$

The estimate provided by the local Lipschitz bound from Lemma 7 works for $|K^N(Z_{t,t}^N)-K^N(Z_{t,s}^N)|$ and $|K_2^N(Z_{t,s}^N)-K_2^N(Z_{t,0}^N)|$ because in the first term the elapsed time $t-s$ is small enough (so we can compensate the $\log N$ order coming from the derivative of K^N with $(f_N(t-s))^{-1}$) and in the other one the force K_2^N has a milder derivative which is only of order $\log\log N$. For the remaining term $|K_1^N(Z_{t,s}^N)-K_1^N(Z_{t,0}^N)|$ in (34) we use that the probability densities of $Z_{t,s}^N$ and $Z_{t,0}^N$ are close in L^∞ by Lemma 9 and its Corollary 10. Note that in order to complete the last argument we need independence of the particles and, although the mean-field particles $Z_{t,0}^{1,N},\dots,Z_{t,0}^{N,N}$ are independent, this does not hold for the particles $Z_{t,s}^{1,N},\dots,Z_{t,s}^{N,N}$ (recall that by definition $Z_{t,s}^N=Z_{t,s}^{X_s^N}$ and that $Z_{t,0}^N=Z_{t,s}^{Y_s^N}$ for $t\geqslant s$). For this reason, instead of considering the processes starting at time s at the r.v. X_s^N and Y_s^N respectively, it is convenient to first fix the starting points at time s to be some given points $x,y\in\mathbb{R}^{2N}$ and to compare the corresponding (product distributed) processes $Z_{t,s}^{x,N}$ and $Z_{t,s}^{y,N}$. This being done, we can recover the original processes $Z_{t,s}^N$ and $Z_{t,0}^N$ by writing $\mathbb{E}(|K_1^N(Z_{t,s}^N)-K_1^N(Z_{t,0}^N)||\mathscr{A}_t\cap\mathscr{B}_t^2)$ as

$$\int_{(x,y)\in(Z_{s,s}^N,Z_{s,0}^N)(\mathscr{A}_t\cap\mathscr{B}_t^2)}\mathbb{E}\left(\left|K_1^N(Z_{t,s}^{x,N})-K_1^N\left(Z_{t,s}^{y,N}\right)\right||\mathscr{A}_t\cap\mathscr{B}_t^2\right)\mathbb{P}(X_s^N\in\mathrm{d}x,Y_s^N\in\mathrm{d}y). \tag{37}$$

Let us then fix $x,y\in\mathbb{R}^{2N}$ and write

$$\begin{aligned}\mathbb{E}\left(\left|K_1^N(Z_{t,s}^{x,N})-K_1^N\left(Z_{t,s}^{y,N}\right)\right||\mathscr{A}_t\cap\mathscr{B}_t^2\right) &= \mathbb{E}\left(\left|K_1^N(Z_{t,s}^{x,N})-K_1^N\left(Z_{t,s}^{y,N}\right)\right||(\mathscr{A}_t\cap\mathscr{B}_t^2)\backslash\mathscr{C}_t^{x,y}\right)\\ &\quad+\mathbb{E}\left(\left|K_1^N(Z_{t,s}^{x,N})-K_1^N\left(Z_{t,s}^{y,N}\right)\right||\mathscr{A}_t\cap\mathscr{B}_t^2\cap\mathscr{C}_t^{x,y}\right),\end{aligned}$$

where we introduced the new set

$$\begin{aligned}\mathscr{C}_t^{x,y} &:= \{|K_1^N(Z_{t,s}^{x,N})-\mathbb{E}(K_1^N(Z_{t,s}^{x,N}))|\leqslant N^{-(\alpha+\delta)}\}\\ &\quad\cap\left\{\left|K_1^N\left(Z_{t,s}^{y,N}\right)-\mathbb{E}\left(K_1^N\left(Z_{t,s}^{y,N}\right)\right)\right|\leqslant N^{-(\alpha+\delta)}\right\},\end{aligned} \tag{38}$$

for $s=t-(\log N)^{-3/2}$. By Proposition 15 below the measure of the set $\Omega\backslash\mathscr{C}_t^{x,y}$ is exponentially small. Also note that the bound given in Proposition 15 does not depend of the points x,y. Since K_1^N is of order $O(N^\alpha)$ we can find a constant $C_\gamma>0$ such that

$$\mathbb{E}\left(\left|K_1^N(Z_{t,s}^{x,N})-K_1^N\left(Z_{t,s}^{y,N}\right)\right||(\mathscr{A}_t\cap\mathscr{B}_t^2)\backslash\mathscr{C}_t^{x,y}\right)\leqslant C_\gamma N^{-\gamma}.$$

Next we estimate $|K_1^N(Z_{t,s}^{x,N})-K_1^N(Z_{t,s}^{y,N})|$ in the set $\mathscr{A}_t\cap\mathscr{B}_t^2\cap\mathscr{C}_t^{x,y}$. We write

$$\begin{aligned}\left|K_1^N(Z_{t,s}^{x,N})-K_1^N\left(Z_{t,s}^{y,N}\right)\right| &\leqslant |K_1^N(Z_{t,s}^{x,N})-\mathbb{E}(K_1^N(Z_{t,s}^{x,N}))|+\left|K_1^N\left(Z_{t,s}^{y,N}\right)-\mathbb{E}\left(K_1^N\left(Z_{t,s}^{y,N}\right)\right)\right|\\ &\quad+\left|\mathbb{E}(K_1^N(Z_{t,s}^{x,N}))-\mathbb{E}\left(K_1^N\left(Z_{t,s}^{y,N}\right)\right)\right|.\end{aligned}$$

In $\mathscr{C}_t^{x,y}$ the first two terms are bounded. For the remaining term $|\mathbb{E}(K_1^N(Z_{t,s}^{x,N}))-\mathbb{E}(K_1^N(Z_{t,s}^{y,N}))|$ we use the following fact: both processes $Z_{t,s}^{x,N}$ and $Z_{t,s}^{y,N}$ evolved according to the mean-field dynamics during a period of time $t-s$, which is long enough to ensure that the densities of $Z_{t,s}^{x,N}$ and $Z_{t,s}^{y,N}$ are close if their starting positions x and y are close. It follows that the difference $|\mathbb{E}(K_1^N(Z_{t,s}^{x,N}))-\mathbb{E}(K_1^N(Z_{t,s}^{y,N}))|$ is also small in that case (Corollary 10). More precisely,

$$\begin{aligned}\left|K_1^N(Z_{t,s}^{x,N})-K_1^N\left(Z_{t,s}^{y,N}\right)\right| &\leqslant |K_1^N(Z_{t,s}^{x,N})-\mathbb{E}(K_1^N(Z_{t,s}^{x,N}))|+\left|K_1^N\left(Z_{t,s}^{y,N}\right)-\mathbb{E}\left(K_1^N\left(Z_{t,s}^{y,N}\right)\right)\right|\\ &\quad+\left|\mathbb{E}(K_1^N(Z_{t,s}^{x,N}))-\mathbb{E}\left(K_1^N\left(Z_{t,s}^{y,N}\right)\right)\right|\\ &\leqslant 2N^{-(\alpha+\delta)}+\frac{|x-y|}{(t-s)^{3/2}}\|k_1^N\|_1,\end{aligned}$$

is true in the event $\mathscr{A}_t\cap\mathscr{B}_t^2\cap\mathscr{C}_t^{x,y}$. Consequently the expected value in $\mathscr{A}_t\cap\mathscr{B}_t^2$ for fixed starting points x,y can be bounded as:

$$\mathbb{E}\left(\left|K_1^N(Z_{t,s}^{x,N})-K_1^N\left(Z_{t,s}^{y,N}\right)\right|\,|\mathscr{A}_t\cap\mathscr{B}_t^2\right)\leqslant\frac{|x-y|}{(t-s)^{3/2}}\,\|k_1^N\|_1+2\,N^{-(\alpha+\delta)}\,\mathbb{P}(\mathscr{A}_t\cap\mathscr{B}_t^2)+C_\gamma N^{-\gamma}.$$

Using this bound in (37) we find an estimate for the original processes

$$\begin{aligned}\mathbb{E}(|K_1^N(Z_{t,s}^N)-K_1^N(Z_{t,0}^N)|\,|\mathscr{A}_t\cap\mathscr{B}_t^2) &\leqslant \frac{\mathbb{E}(|Z_{s,s}^N-Z_{s,0}^N|\,|\mathscr{A}_t\cap\mathscr{B}_t^2)}{(t-s)^{3/2}}\,\|k_1^N\|_1\\ &\quad+2\,N^{-(\alpha+\delta)}\,\mathbb{P}(\mathscr{A}_t\cap\mathscr{B}_t^2)\\ &\quad+C_\gamma N^{-\gamma}\\ &\leqslant (\log N)^{3/4}\,\mathbb{E}(|Z_{s,s}^N-Z_{s,0}^N|\,|\mathscr{A}_t\cap\mathscr{B}_t^2)\\ &\quad+2\,N^{-(\alpha+\delta)}\,\mathbb{P}(\mathscr{A}_t\cap\mathscr{B}_t^2)\\ &\quad+C_\gamma N^{-\gamma},\end{aligned}$$

where for the last inequality we used that $t-s=(\log N)^{-3/2}$ and $\|k_1^N\|_1\leqslant(\log N)^{-3/2}$. Consequently,

$$\begin{aligned}\mathbb{E}(|K_1^N(Z_{t,s}^N)-K_1^N(Z_{t,0}^N)|\,|\mathscr{A}_t\cap\mathscr{B}_t^2) &\leqslant (\log N)^{3/4}\frac{f_N(t-\tau_t)}{f_N(s)}\,\mathbb{E}(|Z_{t,t}^N-Z_{t,\tau_t}^N|\,|\mathscr{A}_t\cap\mathscr{B}_t^2)\\ &\quad+2\,N^{-(\alpha+\delta)}\,\mathbb{P}(\mathscr{A}_t\cap\mathscr{B}_t^2)+C_\gamma N^{-\gamma}\\ &\leqslant (\log N)^{3/4}\,f_N(t-\tau_t)\,\mathbb{E}(|Z_{t,t}^N-Z_{t,\tau_t}^N|\,|\mathscr{A}_t\cap\mathscr{B}_t^2)\\ &\quad+2\,N^{-(\alpha+\delta)}\,\mathbb{P}(\mathscr{A}_t\cap\mathscr{B}_t^2)+C_\gamma N^{-\gamma}.\end{aligned}$$

Together with (35) and (36) this covers all three terms appearing in (34). We can adapt $N_0\in\mathbb{N}$ chosen at the beginning of the proof so that for $N\geqslant N_0$:

$$\begin{aligned}\mathbb{E}(|K^N(Z_{t,t}^N)-K^N(Z_{t,0}^N)|\,|\mathscr{A}_t\cap\mathscr{B}_t^2) &\leqslant 7\,(\log N)^{3/4}\,f_N(t-\tau_t)\,\mathbb{E}(|Z_{t,t}^N-Z_{t,\tau_t}^N|\,|\mathscr{A}_t\cap\mathscr{B}_t^2)\\ &\quad+8\log\log N f_N(t-\tau_t)\,\mathbb{E}(|Z_{t,t}^N-Z_{t,\tau_t}^N|\,|\mathscr{A}_t\cap\mathscr{B}_t^2)\\ &\quad+2\,N^{-(\alpha+\delta)}\,\mathbb{P}(\mathscr{A}_t\cap\mathscr{B}_t^2)+C_\gamma N^{-\gamma}\\ &\leqslant 8\,(\log N)^{3/4}\,f_N(t-\tau_t)\,\mathbb{E}(|Z_{t,t}^N-Z_{t,\tau_t}^N|\,|\mathscr{A}_t\cap\mathscr{B}_t^2)\\ &\quad+2\,N^{-(\alpha+\delta)}\,\mathbb{P}(\mathscr{A}_t\cap\mathscr{B}_t^2)+C_\gamma N^{-\gamma}.\end{aligned}$$

Going back to (33) we use this last estimate for the first term, the bound

$$|K^N(Z_{t,0}^N)-\overline{K}_t^N(Z_{t,0}^N)|\leqslant N^{-(\alpha+\delta)}$$

in $\mathscr{A}_t\cap\mathscr{B}_t^2$ for the second term and the Lipschitz continuity of $\overline{K}_t^N$

$$|\overline{K}_t^N(Z_{t,0}^N)-\overline{K}_t^N(Z_{t,\tau_t}^N)|\leqslant L\,|Z_{t,0}^N-Z_{t,\tau_t}^N|\leqslant L\left(1+\frac{f_N(t-\tau_t)}{f_N(t)}\right)|Z_{t,t}^N-Z_{t,\tau_t}^N|$$

for the third one. Bringing everything together, (33) becomes

$$\begin{aligned}\mathbb{E}(|K^N(Z_{t,t}^N)-\overline{K}_t^N(Z_{t,\tau_t}^N)|\,|\mathscr{A}_t\cap\mathscr{B}_t^2) &\leqslant 8\,(\log N)^{3/4}\,f_N(t-\tau_t)\,\mathbb{E}(|Z_{t,t}^N-Z_{t,\tau_t}^N|\,|\mathscr{A}_t\cap\mathscr{B}_t^2)\\ &\quad+2\,N^{-(\alpha+\delta)}\,\mathbb{P}(\mathscr{A}_t\cap\mathscr{B}_t^2)+C_\gamma N^{-\gamma}\\ &\quad+N^{-(\alpha+\delta)}\,\mathbb{P}(\mathscr{A}_t\cap\mathscr{B}_t^2)\\ &\quad+L\,(1+f_N(t-\tau_t))\,\mathbb{E}(|Z_{t,t}^N-Z_{t,\tau_t}^N|\,|\mathscr{A}_t\cap\mathscr{B}_t^2)\\ &\leqslant 9\,(\log N)^{3/4}\,f_N(t-\tau_t)\,\mathbb{E}(|Z_{t,t}^N-Z_{t,\tau_t}^N|\,|\mathscr{A}_t\cap\mathscr{B}_t^2)\\ &\quad+3\,N^{-(\alpha+\delta)}\,\mathbb{P}(\mathscr{A}_t\cap\mathscr{B}_t^2)+C_\gamma N^{-\gamma},\end{aligned}$$

which is true if N is greater than some new N_0 depending now also on the Lipschitz constant L. Finally, from $f_N(t-\tau_t)\leqslant 2$ it follows that

$$\begin{aligned}\mathbb{E}(|K^N(Z_{t,t}^N)-\overline{K}_t^N(Z_{t,\tau_t}^N)||\mathcal{A}_t\cap\mathcal{B}_t^2) &\leqslant 18(\log N)^{3/4}\,\mathbb{E}(|Z_{t,t}^N-Z_{t,\tau_t}^N||\mathcal{A}_t\cap\mathcal{B}_t^2)\\ &\quad +3N^{-(\alpha+\delta)}\,\mathbb{P}(\mathcal{A}_t\cap\mathcal{B}_t^2)+C_\gamma N^{-\gamma},\end{aligned}$$

proving (32). As a consequence of (31) and (32)

$$\mathbb{E}(\overline{D}_t^+J_t^N)\leqslant\mathbb{E}(\overline{D}_s^+g(\tau_t,t)|\mathcal{A}_t\backslash\mathcal{B}_t^2)+\mathbb{E}(\overline{D}_s^+g(\tau_t,t)|\mathcal{A}_t\cap\mathcal{B}_t^2)\leqslant 2\,C_\gamma N^{-\gamma}=:\tilde{C}_\gamma N^{-\gamma}.$$

Dini derivatives and the derivative of J_t.

Here we define the derivatives in the sense of Dini and provide a necessary result for computing the derivative of J_t.

Definition 12. (Dini derivatives) *For any function* $f\colon\mathbb{R}\to\mathbb{R}$ *the* right upper, right lower, left upper, left lower Dini derivatives *are defined, in that order, as*

$$\begin{aligned}\overline{D}^+f(x)&:=\limsup_{h\to 0^+}\frac{f(x+h)-f(x)}{h}, & \underline{D}^+f(x)&:=\liminf_{h\to 0^+}\frac{f(x+h)-f(x)}{h},\\ \overline{D}^-f(x)&:=\limsup_{h\to 0^-}\frac{f(x+h)-f(x)}{h}, & \underline{D}^-f(x)&:=\liminf_{h\to 0^-}\frac{f(x+h)-f(x)}{h}.\end{aligned}$$

For any function $g\colon\mathbb{R}^d\to\mathbb{R}$ *and any normed vector* $v\in\mathbb{R}^d$ *the* upper *resp.* lower directional Dini derivatives *of* g *in the direction* v *are given by*

$$\overline{D}_v g(x):=\limsup_{h\to 0^+}\frac{g(x+hv)-g(x)}{h},\quad \underline{D}_v g(x):=\liminf_{h\to 0^+}\frac{g(x+hv)-g(x)}{h}.$$

We denote by $\overline{D}_i^+,\underline{D}_i^+$ *(resp.* $\overline{D}_i^-,\underline{D}_i^-$*) the directional Dini derivatives in the direction* e_i *(resp.* $-e_i$*).*

Note that the Dini derivatives are always well defined (taking values in $[-\infty,\infty]$). Moreover, for locally Lipschitz functions they are finite at every point. For differentiable functions, all four Dini derivatives coincide and are equal to the derivative. Similarly, if a function f is right-differentiable, then $\overline{D}^+f=\underline{D}^+f=\partial^+f$.

Proposition 13. (Some properties of Dini derivatives) *Let* $f,g\colon\mathbb{R}\to\mathbb{R}$ *be continuous.*

i. $\overline{D}^+(f+g)\leqslant\overline{D}^+f+\overline{D}^+g$.

ii. If $f>0$*, then* $\overline{D}^+(fg)=f\overline{D}^+g+g\,\overline{D}^+f$.

iii. [13, Theorem 11] If f has a finite Dini derivative $\overline{D}^+$ *at every* $t\in\mathbb{R}$*, then*

$$f(b)-f(a)=\int_a^b\overline{D}^+f(t)\,\mathrm{d}t$$

for each interval $[a,b]$*, provided that* $\overline{D}^+f$ *is Lebesgue integrable over* $[a,b]$.[5]

The proof of the sum and product rules are an easy consequence of the properties of the limit superior. Some other properties of Dini derivatives can be found in [20].

5. The following result [13, Theorem 3] is also enough for our purpose and can be proven in a much simpler way: If f has a finite Dini derivative $\overline{D}^+$ at every $t\in\mathbb{R}$, then

$$f(b)-f(a)\leqslant\overline{\int_a^b}\overline{D}^+f(t)\,\mathrm{d}t$$

for each interval $[a,b]$, where $\overline{\int_a^b}$ denotes the upper Riemann integral.

Lemma 14. *Let $g\colon [0,T]\times[0,T]\to\mathbb{R}$ be a continuous function with finite Dini derivatives and such that the right upper Dini derivative in the second variable $\overline{D}_2^+$ is continuous in the first variable. We define the function $f(t) := \sup_{0\leqslant\tau\leqslant s\leqslant t} g(\tau,s)$ for $t\in[0,T]$. Let $M_t := \{(\tau,s): 0\leqslant\tau\leqslant s\leqslant t, f(t)=g(\tau,s)\}$ be the set of maximal points for g and assume that none of these points is on the diagonal, i.e. $M_t\cap\{(s,s): s\in[0,T]\}=\varnothing$. Then*

$$\overline{D}^+ f(t) \leqslant \max\left\{0, \sup_{(\tau,s)\in M_t\cap\{s=t\}} \overline{D}_2^+ g(\tau,s)\right\}.$$

Proof. Let $\Delta_t := \{(\tau,s): 0\leqslant\tau\leqslant s\leqslant t\}$, $0\leqslant t\leqslant T$, be the triangle where the supremum in the definition of $f(t)$ is taken. We consider two cases.

Assume first that there exists $(\tau,s)\in M_t$ with $s<t$. In this situation it is clear (since g is continuous) that $g(\tau,s)$ is also the supremum of g over Δ_{t+h} for small enough $h>0$. Therefore, $f(t+h)=f(t)$ for h in a small right neighbourhood of 0 and so is $\overline{D}^+ f(t)=0$.

Next assume that the previous situation does not hold, that is, that the supremum of g over Δ_t is only attained when $s=t$. Let $\hat{M}_t := \{(\tau,s)\in M_t: s=t\}$ be the set of such maximal points. For $(\tau,t)\in\hat{M}_t$ we also know that the coordinates must satisfy $\tau<t$, since we assumed that the supremum is not attained on the diagonal. Furthermore, it holds

$$\underline{D}_1^- g(\tau,t)\geqslant 0 \quad\text{and}\quad \overline{D}_1^+ g(\tau,t)\leqslant 0, \quad (\tau,t)\in\hat{M}_t. \tag{39}$$

We only prove the first inequality, the other being analogous. Assume to the contrary that $\underline{D}_1^- g(\tau,t)<0$. Then, by definition,

$$\inf_{h\leqslant\delta<0} \frac{g(\tau+\delta,t)-g(\tau,t)}{\delta}<0, \quad \text{for each } h<0.$$

In particular there exists a sequence $\delta_n\nearrow 0$ such that $g(\tau+\delta_n,t)>g(\tau,t)$, which contradicts (τ,t) being a maximal point. Next we prove that the direction of maximal growth for $(\tau,t)\in\hat{M}_t$ is the e_2 direction. It is already clear, since (τ,t) is a maximal point over Δ_t, that $\overline{D}_v g(\tau,s)\leqslant 0$ for any direction $v=(v_1,v_2)$ with $v_2\leqslant 0$. Let then v with $v_2\geqslant 0$ and $|v|=1$. It holds

$$\begin{aligned}\overline{D}_v g(\tau,s) &= \limsup_{h\to 0^+} \frac{g((\tau,s)+hv)-g(\tau,s)}{h}\\ &\leqslant \limsup_{h\to 0^+} \frac{g(\tau+hv_1,s+hv_2)-g(\tau+hv_1,s)}{h} + \limsup_{h\to 0^+} \frac{g(\tau+hv_1,s)-g(\tau,s)}{h}.\end{aligned}$$

The second term is equal to

$$v_1 \limsup_{h\to 0^+} \frac{g(\tau+h,s)-g(\tau,s)}{h} = v_1\overline{D}_1^+ g(\tau,s), \quad \text{if } v_1\geqslant 0$$

or to

$$v_1 \liminf_{h\to 0^-} \frac{g(\tau+h,s)-g(\tau,s)}{h} = v_1\underline{D}_1^- g(\tau,s), \quad \text{if } v_1<0.$$

The first term is smaller or equal than

$$\begin{aligned}\limsup_{h_1\to 0^+}\limsup_{h_2\to 0^+} \frac{g(\tau+h_1v_1,s+h_2v_2)-g(\tau+h_1v_1,s)}{h_2} &= v_2 \limsup_{h_1\to 0^+} \overline{D}_2^+ g(\tau+h_1v_1,s)\\ &= v_2\overline{D}_2^+ g(\tau,s),\end{aligned}$$

by the continuity of $\overline{D}_2^+$ in the τ direction. Therefore,

$$\overline{D}_v g(\tau,s)\leqslant \max\{-|v_1|\underline{D}_1^- g(\tau,s), |v_1|\overline{D}_1^+ g(\tau,s)\} + v_2\overline{D}_2^+ g(\tau,s), \quad v_2\geqslant 0.$$

With (39) and $0\leqslant v_2\leqslant 1$ we find for $(\tau,t)\in\hat{M}_t$

$$\overline{D}_v g(\tau,t)\leqslant \overline{D}_2^+ g(\tau,t), \quad v_2\geqslant 0.$$

Since $D_2^+ g(\tau, t)$ is continuous in τ and $\hat{M}_t$ is compact there exists $(\tau_t, t) \in \hat{M}_t$ such that $D_2^+ g(\tau_t, t) = \sup_{(\tau,t)\in\hat{M}_t} D_2^+ g(\tau, t)$. Then

$$\lim_{h\to 0^+} \sup_{0<\delta\leqslant h} \frac{f(t+\delta)-f(t)}{\delta} = \lim_{h\to 0^+} \sup_{0<\delta\leqslant h} \frac{g(\tau_t, t+\delta)-g(\tau_t,t)}{\delta} = D_2^+ g(\tau_t, t).$$

Note that although the set M_t is certainly non-empty, it might happen that $\hat{M}_t = \varnothing$. In this case we are in the first situation, $\overline{D}^+ f(t) = 0$, and with $\sup\varnothing = -\infty$ the bound is still valid. ☐

Measure of the exceptional sets.

It just remains to estimate the measure of the complementary sets of $\mathscr{B}_t^1$, $\mathscr{B}_t^2$ and $\mathscr{C}_t^{x,y}$ as defined in (25), (30) and (38). The constants $T>0$, $\alpha\in(0,1/2)$ and $\delta>0$ are the ones we fixed at the beginning of this section.

Proposition 15. (Measure of the exceptional sets) *For each $\gamma>0$ there exists a positive constant C_γ such that*

i. $\mathbb{P}(S_t^1 \cup S_t^2 \cup S_t^3) \leqslant C_\gamma N^{-\gamma}$ *for each* $0\leqslant t\leqslant T$, *where*

$$S_t^1 := \{|K^N(Y_t^N) - \overline{K}_t^N(Y_t^N)|_\infty \geqslant N^{-(\alpha+\delta)}\},$$
$$S_t^2 := \{|L^N(Y_t^N) - \overline{L}_t^N(Y_t^N)|_\infty \geqslant 1\}, \qquad S_t^3 := \{|L_2^N(Y_t^N) - \overline{L}_{2,t}^N(Y_t^N)|_\infty \geqslant 1\}.$$

Consequently $\mathbb{P}(\Omega\backslash\mathscr{B}_t^1) \leqslant C_\gamma N^{-\gamma}$ *and* $\mathbb{P}(\Omega\backslash\mathscr{B}_t^2) \leqslant C_\gamma N^{-\gamma}$ *hold for each* $0\leqslant t\leqslant T$.

ii. $\mathbb{P}(|K_1^N(Z_{t,s}^{x,N}) - \mathbb{E}(K_1^N(Z_{t,s}^{x,N}))|_\infty \geqslant N^{-(\alpha+\delta)}) \leqslant C_\gamma N^{-\gamma}$ *holds for any* $x\in\mathbb{R}^{2N}$ *and any* $T\geqslant t\geqslant s\geqslant 0$ *satisfying* $t-s \geqslant (\log N)^{-r}$ *for some* $r\geqslant 0$. *Consequently,* $\mathbb{P}(\Omega\backslash\mathscr{C}_t^{x,y}) \leqslant 2\, C_\gamma N^{-\gamma}$ *for any* $x,y\in\mathbb{R}^{2N}$ *and* $0\leqslant s\leqslant t\leqslant T$.

Proof. Fix $\gamma>0$.

i. First note that the mean-field force $\overline{K}_{t,i}^N(Y_t^N)$ can be written in terms of the expected value of K^N as $\overline{K}_{t,i}^N(Y_t^N) = \mathbb{E}_{(-i)}(K_i^N(Y_t^N))$ and therefore the first set S_t^1 is equal to the set $\{\sup_{1\leqslant i\leqslant N} |K_i^N(Y_t^N) - \mathbb{E}_{(-i)}(K_i^N(Y_t^N))| \geqslant N^{-(\alpha+\delta)}\}$. Moreover, $Y_t^1, \ldots, Y_t^N$ are already pairwise independent and the L^∞-norm of its probability density ρ_t^N is bounded uniformly in N and $t\in[0,T]$ by Proposition 4. Therefore, from Proposition 8 follows the existence of a constant $C_\gamma > 0$, independent of t, with

$$\mathbb{P}(S_t^1) = \mathbb{P}(|K^N(Y_t^N) - \overline{K}_t^N(Y_t^N)|_\infty \geqslant N^{-(\alpha+\delta)}) \leqslant C_\gamma N^{-\gamma},$$

for each $t\in[0,T]$.

The remaining sets S_t^2 and S_t^3 can be expressed in terms of the expected value of L^N resp. L_2^N in an analogous way. Also note that both $|N^{-\alpha} L_i^N(x)|$ and $|N^{-\alpha} L_{2i}^N(x)|$ are bounded by $C\,\chi \min\{N^\alpha, |x|^{-1}\}$. Proposition 8 then implies for S_t^2 that

$$\begin{aligned}\mathbb{P}(|L^N(Y_t^N) - \overline{L}_t^N(Y_t^N)|_\infty \geqslant 1) &= \mathbb{P}(N^{-\alpha}|L^N(Y_t^N) - \overline{L}_t^N(Y_t^N)|_\infty \geqslant N^{-\alpha})\\ &\leqslant \mathbb{P}(N^{-\alpha}|L^N(Y_t^N) - \overline{L}_t^N(Y_t^N)|_\infty \geqslant N^{-(\alpha+\delta)})\\ &\leqslant C_\gamma N^{-\gamma},\end{aligned}$$

and in the same manner that $\mathbb{P}(S_t^3) = \mathbb{P}(|L_2(Y_t^N) - \overline{L}_{2,t}^N(Y_t^N)|_\infty \geqslant 1) \leqslant C_\gamma N^{-\gamma}$ for each $t\in[0,T]$.

ii. Let $T\geqslant t\geqslant s\geqslant 0$ be such that $t-s\geqslant(\log N)^{-r}$ holds for some $r\geqslant 0$. First notice that for each fixed starting point $x\in\mathbb{R}^{2N}$ the processes $Z_{t,s}^{x,1,N}, \ldots, Z_{t,s}^{x,N,N}$ are pairwise independent. Furthermore, the probability density $u_{t,s}^{x,i,N}$ of $Z_{t,s}^{x,i,N}$ satisfies

$$\|u_{t,s}^{x,i,N}\|_\infty \leqslant C((t-s)^{-1}+1) \leqslant C(\log N)^r$$

for $i = 1, \ldots, N$, by Lemma 9, meaning that the growth of $\|u_{t,s}^{x,i,N}\|_\infty$ is only logarithmic in N and consequently condition (12) is fulfilled independently of the times t, s and the exponent r. Therefore there exists a constant $C_\gamma > 0$ such that, for any such t, s:

$$\mathbb{P}(|K_1^N(Z_{t,s}^{x,N}) - \mathbb{E}(K_1^N(Z_{t,s}^{x,N}))|_\infty \geqslant N^{-(\alpha+\delta)}) \leqslant C_\gamma N^{-\gamma}. \qquad \square$$

6. Proofs of Propositions 4 and 5

Proof of Proposition 4.

One first proves the boundedness of ρ in L^p for each $1 < p < \infty$. The L^∞ estimate follows from this fact and the boundedness of $\nabla c = -k * \rho$ by an iterative argument.

Step 1: Uniform bounds in L^p, $p < \infty$.

Notice that under our assumptions ρ_0 is in $L^p(\mathbb{R}^2)$ for each $p \in [1, \infty]$. Then $\rho \in L^\infty(0, T; L^p(\mathbb{R}^2))$ for any $T > 0$ and $1 \leqslant p < \infty$, and the same holds for ρ^N with bounds which are uniform in N. See either [3, Proposition 17] or [9, Lemma 2.7] for the proof for ρ and [3, Lemma 13] for ρ^N.

Step 2: Uniform bounds in L^∞.

For this step we follow [5, Lemma 3.2] and [16, Lemma 4.1]. The second reference is much more detailed but only handles bounded domains. The proof can nevertheless be adapted for the whole space $\mathbb{R}^2$ as described in the first paper.

For readability the following computations are performed only formally. One can write the same proof for the solutions ρ^N of the regularised equation (5) and then deduce the result for ρ by passing to the limit (recall that $\rho^N \rightharpoonup \rho$ weakly by Theorem 3).

Let $\rho_m := (\rho - m)_+$. First notice that $\nabla c = -k * \rho$ is in $L^\infty(0, T; L^\infty(\mathbb{R}^2))$: $\|k * \rho\|_\infty \leqslant C(\|\rho\|_3 + \|\rho\|_1)$ since $k \in L^{3/2} + L^\infty$, and the right hand side is uniformly bounded by the first step. We then prove the inequality

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} \int \rho_m^p \,\mathrm{d}x \;&\leqslant\; -p^2 \int \rho_m^p \,\mathrm{d}x \\ &\quad + C p^4 \left(\int \rho_m^{p/2} \,\mathrm{d}x \right)^2 + C p^2, \end{aligned} \tag{40}$$

for some constant C depending on $\|\nabla c\|_\infty$.

From this we will conclude that $\sup_{t \in [0,T]} \|\rho_m\|_p$ is bounded independently of p. The proof is then complete after taking the limit $p \to \infty$.

We first multiply on both sides of the Keller-Segel equation (1) by ρ_m^{p-1} and integrate to find

$$\frac{1}{p} \frac{\mathrm{d}}{\mathrm{d}t} \int \rho_m^p \,\mathrm{d}x = \int \nabla \cdot (\nabla \rho + \chi \nabla c\, \rho)\, \rho_m^{p-1}.$$

Let $\Omega_t := \{\rho(t) \geqslant m\}$ and notice that Ω_t is uniformly bounded: $1 = \|\rho(t)\|_1 \geqslant m\, |\Omega_t|$. Then the integral on the right hand side equals

$$\begin{aligned} \int_{\Omega_t} \nabla \cdot (\nabla \rho + \chi (k * \rho) \rho)\, \rho_m^{p-1} &= -\int_{\Omega_t} (\nabla \rho + \chi (k * \rho) \rho) \nabla \rho_m^{p-1} \\ &= -(p-1) \int \rho_m^{p-2} |\nabla \rho_m|^2 \\ &\quad + \chi (p-1) \int \rho\, \rho_m^{p-2} \nabla c \cdot \nabla \rho_m \\ &= -(p-1) \int \rho_m^{p-2} |\nabla \rho_m|^2 + \chi (p-1) \int \rho_m^{p-1} \nabla c \cdot \nabla \rho_m \\ &\quad + \chi m (p-1) \int \rho_m^{p-2} \nabla c \cdot \nabla \rho_m. \end{aligned}$$

Using that $\rho_m^{(p-k)/2}\nabla\rho_m^{p/2}=\frac{p}{2}\rho_m^{p-(k/2+1)}\nabla\rho_m$ for any $k\in\mathbb{R}$ the last expression equals

$$-\frac{4(p-1)}{p^2}\int|\nabla\rho_m^{p/2}|^2+\frac{2\chi(p-1)}{p}\int\rho_m^{p/2}\nabla c\cdot\nabla\rho_m^{p/2}+\frac{2\chi m(p-1)}{p}\int\rho_m^{(p-2)/2}\nabla c\cdot\nabla\rho_m^{p/2}.$$

For the last two terms we use the following Young's inequality, $|a\cdot b|\leqslant\frac{1}{4}|a|^2+|b|^2$ for $a,b\in\mathbb{R}^2$, and find

$$(p-1)\int\chi\rho_m^{p/2}\nabla c\cdot\frac{2}{p}\nabla\rho_m^{p/2}\leqslant\frac{(p-1)}{p^2}\int|\nabla\rho_m^{p/2}|^2+\chi^2(p-1)\|\nabla c\|_\infty^2\int\rho_m^p$$

and

$$\begin{aligned}(p-1)\int\chi m\rho_m^{(p-2)/2}\nabla c\cdot\frac{2}{p}\nabla\rho_m^{p/2} &\leqslant \frac{(p-1)}{p^2}\int|\nabla\rho_m^{p/2}|^2\\ &\quad+\chi^2m^2(p-1)\|\nabla c\|_\infty^2\int_{\Omega_t}\rho_m^{p-2}\\ &\leqslant \frac{(p-1)}{p^2}\int|\nabla\rho_m^{p/2}|^2\\ &\quad+C(p-1)\|\nabla c\|_\infty^2\int_{\Omega_t}(\rho_m^p+1)\\ &\leqslant \frac{(p-1)}{p^2}\int|\nabla\rho_m^{p/2}|^2+C(p-1)\|\nabla c\|_\infty^2\int\rho_m^p\\ &\quad+C(p-1)\|\nabla c\|_\infty^2|\Omega_t|.\end{aligned}$$

Altogether

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\int\rho_m^p\,\mathrm{d}x &\leqslant -\frac{2(p-1)}{p}\int|\nabla\rho_m^{p/2}|^2+Cp(p-1)\|\nabla c\|_\infty^2\int\rho_m^p+Cp(p-1)\|\nabla c\|_\infty^2\\ &\leqslant -\int|\nabla\rho_m^{p/2}|^2+Cp^2\int\rho_m^p+Cp^2,\end{aligned}$$

for $p\geqslant2$ and a constant C depending on $\|\nabla c\|_\infty^2$ (which is bounded uniformly in $t\in[0,T]$).

Now we use the Galiardo-Nirenberg-Sobolev inequality followed by Young's inequality

$$\lambda^2\|u\|_2^2\leqslant\lambda^2C_{\mathrm{GNS}}\|\nabla u\|_2\|u\|_1\leqslant\frac{\lambda^4C_{\mathrm{GNS}}^2}{4}\|u\|_1^2+\|\nabla u\|_2^2$$

for $u=\rho_m^{p/2}$ and $\lambda=\sqrt{C}\,p$:

$$Cp^2\int\rho_m^p\leqslant\frac{p^4C^2C_{\mathrm{GNS}}^2}{4}\left(\int\rho_m^{p/2}\right)^2+\int|\nabla\rho_m^{p/2}|^2.$$

Therefore

$$\frac{\mathrm{d}}{\mathrm{d}t}\int\rho_m^p\,\mathrm{d}x\leqslant-p^2\int\rho_m^p+Cp^4\left(\int\rho_m^{p/2}\right)^2+Cp^2,$$

which proves (40).

Let now $w_j=\int\rho_m^{2^j}$, $S_j:=\sup_{t\in[0,T]}\int\rho_m^{2^j}$ for $j\in\mathbb{N}$. Then

$$\frac{\mathrm{d}}{\mathrm{d}t}w_j\,\mathrm{d}x\leqslant-2^{2j}w_j+2^{2j}(C\,2^{2j}S_{j-1}^2+C).$$

The solution of

$$\frac{\mathrm{d}}{\mathrm{d}t}v=-\varepsilon v+\varepsilon C$$

is $v(t)=\mathrm{e}^{-\varepsilon t}v_0+C(1-\mathrm{e}^{-\varepsilon t})$. If we set $v_0=w_j(0)$ it holds

$$w_j\leqslant v\leqslant w_j(0)+C\,2^{2j}S_{j-1}^2+C\leqslant\|\rho_0\|_\infty^{2^j}|\Omega_0|+C\,2^{2j}S_{j-1}^2+C.$$

It follows that

$$S_j=\sup_{t\in[0,T]}w_j\leqslant C\max\left\{\|\rho_0\|_\infty^{2^j},2^{2j}S_{j-1}^2+1\right\}.$$

For $\tilde{S}_j := S_j \|\rho_0\|_\infty^{2^{-j}}$ is

$$\tilde{S}_j \leqslant C \max\left\{1, 2^{2j} \tilde{S}_{j-1}^2\right\}.$$

Hence

$$\begin{aligned}\log_+ \tilde{S}_j &\leqslant \max\left\{\log_+ C, \log_+ C\, 2^{2j} \tilde{S}_{j-1}^2\right\}\\ &\leqslant 2\log_+ \tilde{S}_{j-1} + j\log 4 + C,\end{aligned}$$

which implies $2^{-j} \log_+ \tilde{S}_j - 2^{-(j-1)} \log_+ \tilde{S}_{j-1} \leqslant j\, 2^{-j} \log 4 + C\, 2^{-j}$ for $j \in \mathbb{N}$. Adding up both sides over $j = 1, \ldots, J$ we find

$$\begin{aligned}2^{-J} \log_+ \tilde{S}_J - \log_+ \tilde{S}_0 &= \sum_{j=1}^{J} 2^{-j} \log_+ \tilde{S}_j - 2^{-(j-1)} \log_+ \tilde{S}_{j-1}\\ &\leqslant \sum_{j=1}^{\infty} j\, 2^{-j} \log 4 + C\, 2^{-j} \leqslant C,\end{aligned}$$

for a constant C independent of J. Since $\tilde{S}_0 \leqslant \sup_{t \in [0,T]} \frac{\|\rho(t)\|_1}{\|\rho_0\|_\infty}$ is also bounded, we conclude that $S_j^{2^{-j}} = \left(\sup_{t \in [0,T]} \int \rho_m^{2^j}\right)^{2^{-j}} = \sup_{t \in [0,T]} \left(\int \rho_m^{2^j}\right)^{2^{-j}} \leqslant C$ for some contant C not depending on j. We finally perform the limit $j \to \infty$ and conclude

$$\sup_{t \in [0,T]} \|\rho_m\|_\infty = \sup_{t \in [0,T]} \lim_{j \to \infty} \|\rho_m\|_{2^j} \leqslant \lim_{j \to \infty} \sup_{t \in [0,T]} \|\rho_m\|_{2^j} \leqslant C.$$

Proof of Proposition 5.

i. From the proof of [9, Lemma 2.8], with the additional assumption that ρ_0 is in $L^\infty(\mathbb{R}^2) \cap H^2(\mathbb{R}^2)$, it follows that ρ and ρ^N are in $W^{1,p}((0,T) \times \mathbb{R}^2)$ with

$$\|\rho\|_{W^{1,p}((0,T) \times \mathbb{R}^2)}, \|\rho^N\|_{W^{1,p}((0,T) \times \mathbb{R}^2)} \leqslant C\, \|\rho_0\|_{H^2(\mathbb{R}^2)}$$

for any $p \in (1, \infty)$ and $N \in \mathbb{N}$, where $C > 0$ is some constant depending on T. Then, by Morrey's inequality, $\rho, \rho^N \in C^{0,\alpha}((0,T) \times \mathbb{R}^2)$ for each $N \in \mathbb{N}$ and $0 < \alpha \leqslant 1/4$ and their norms in this space are also bounded by $C\, \|\rho_0\|_{H^2(\mathbb{R}^2)}$. This means in particular that, for $0 < \alpha \leqslant 1/4$, $\rho, \rho^N \in L^\infty(0,T; C^{0,\alpha}(\mathbb{R}^2))$ and

$$\sup_{t \in [0,T]} [\rho(t)]_{0,\alpha}, \sup_{t \in [0,T]} [\rho^N(t)]_{0,\alpha} \leqslant C_1, \quad N \in \mathbb{N},$$

where $C_1 := C\, \|\rho_0\|_{H^2(\mathbb{R}^2)}$.

ii. Let $w = \phi * \rho = -\log|\cdot| * \rho$. We need to prove that $-\nabla w^N = k^N * \rho^N$ and $-\nabla w = k * \rho$ are Lipschitz continuous in $\mathbb{R}^2$ uniformly in $N \in \mathbb{N}$ and $t \in [0,T]$. It is then enough to show that all second derivatives of w^N and w are uniformly bounded. More precisely, we find

$$\|\partial_{ij} w^N(t)\|_\infty \leqslant C\,(\|\rho^N(t)\|_1 + \|\rho^N(t)\|_\infty + [\rho^N(t)]_{0,\alpha}), \quad N \in \mathbb{N}$$

and

$$\|\partial_{ij} w(t)\|_\infty \leqslant C\,(\|\rho(t)\|_1 + \|\rho(t)\|_\infty + [\rho(t)]_{0,\alpha})$$

for some constant $C > 0$ and any $\alpha \in (0, 1/4]$. These are uniformly bounded on $[0,T]$ and $N \in \mathbb{N}$ by part i and Proposition 4. We just write down the proof for the limiting case $k * \rho$. For $k^N * \rho^N$ the steps are completely analogous.

We split the integral as follows:

$$\partial_{ij} w(t,x) = \int_{|x-y| \leqslant 1} \partial_{ij} \phi(x-y)\, \rho(y)\, \mathrm{d}y + \int_{|x-y| \geqslant 1} \partial_{ij} \phi(x-y)\, \rho(y)\, \mathrm{d}y. \tag{43}$$

Note that $|\partial_{ij}\phi(x-y)| \leqslant \frac{C}{|x-y|^2}$. Therefore, the second term is bounded by $C\,\|\rho\|_1$. For the first term we write

$$\begin{aligned}\int_{|x-y|\leqslant 1} \partial_{ij}\phi(x-y)\,\rho(y)\,\mathrm{d}y &= \int_{|x-y|\leqslant 1} \partial_{ij}\phi(x-y)\,(\rho(y)-\rho(x))\,\mathrm{d}y\\ &\quad +\rho(x)\int_{|x-y|\leqslant 1} \partial_{ij}\phi(x-y)\,\mathrm{d}y\\ &= \int_{|x-y|\leqslant 1} \partial_{ij}\phi(x-y)\,(\rho(y)-\rho(x))\,\mathrm{d}y\\ &\quad -\rho(x)\int_{|x-y|=1} \partial_{i}\phi(x-y)\,\nu_j(y)\,\mathrm{d}S(y).\end{aligned}$$

Consequently, for any $\alpha\in(0,1/4]$ part (i) implies

$$\begin{aligned}\left|\int_{|x-y|\leqslant 1} \partial_{ij}\phi(x-y)\,\rho(y)\,\mathrm{d}y\right| &\leqslant C\,[\rho(t)]_{0,\alpha}\int_{|x-y|\leqslant 1}\frac{1}{|x-y|^{2-\alpha}}\,\mathrm{d}y\\ &\quad +C\,\|\rho(t)\|_\infty\\ &\leqslant C\,([\rho(t)]_{0,\alpha}+\|\rho(t)\|_\infty).\end{aligned}$$

Using both estimates in (43) we find

$$\|\partial_{ij}w(t)\|_\infty \leqslant C\,(\|\rho(t)\|_1+\|\rho(t)\|_\infty+[\rho(t)]_{0,\alpha}),\quad \alpha\in(0,1/4].$$

Remark 16. Below we list the space embeddings we used in the proof.

i. $H^2(\mathbb{R}^2)\hookrightarrow C^{0,\alpha}(\mathbb{R}^2)$, for any $0<\alpha<1$, by the Sobolev embedding theorem [7, Theorem 2.31].

ii. If $f\in L^2(0,T;H^2(\mathbb{R}^2))\cap L^\infty(0,T;H^1(\mathbb{R}^2))$ then $\nabla_x f\in L^p((0,T)\times\mathbb{R}^2)$ for any $p\in(1,4)$. Since $W^{1,2}(\mathbb{R}^2)\subseteq L^q(\mathbb{R}^2)$ for any $2\leqslant q<\infty$, we have that

$$\nabla_x f\in L^2(0,T;L^q(\mathbb{R}^2))\cap L^\infty(0,T;L^2(\mathbb{R}^2)),\quad \text{for } 2\leqslant q<\infty.$$

We then use the interpolation inequality

$$\|u\|_{p_\theta}\leqslant \|u\|_{p_0}^{\theta}\,\|u\|_{p_1}^{1-\theta},\quad \text{for } \theta\in[0,1],\quad \frac{1}{p_\theta}=\frac{\theta}{p_0}+\frac{1-\theta}{p_1}$$

and find

$$\begin{aligned}\int_0^T \|\nabla_x f(t)\|_p^p\,\mathrm{d}t &\leqslant \int_0^T (\|\nabla_x f\|_2^{\theta}\,\|\nabla_x f\|_{p_1}^{1-\theta})^p\\ &\leqslant \sup_{t\in[0,T]}\|\nabla_x f(t)\|_2^{\theta p}\int_0^T \|\nabla_x f\|_{p_1}^{(1-\theta)p}.\end{aligned}$$

By choosing $\theta=1/2$ it holds for $p<4$ that $p\,(1-\theta)\leqslant 2$ and $p_1=\frac{2\,p\,(1-\theta)}{2-\theta\,p}=\frac{2\,p}{4-p}<\infty$, and so is the right hand side of the last inequality finite.

iii. $H^2(\mathbb{R}^2)\subseteq W^{2-2/p,p}(\mathbb{R}^2)$, for any $2<p\leqslant 4$.
By the Sobolev embedding theorem for fractional spaces [1, Theorem 7.58] it holds

$$H^2(\mathbb{R}^2)\subseteq W^{1+2/p,p}(\mathbb{R}^2),\quad \text{for any } 2<p<\infty.$$

Since $1+2/p\geqslant 2-2/p$ holds if $p\leqslant 4$, we conclude that

$$H^2(\mathbb{R}^2)\subseteq W^{1+2/p,p}(\mathbb{R}^2)\subseteq W^{2-2/p,p}(\mathbb{R}^2),\quad \text{for any } 2<p\leqslant 4.$$

ACKNOWLEDGEMENTS

We are greatful to Martin Kolb, Detlef Dürr, Christian Stinner and Tomasz Cieslak for many valuable discussions on this topic and to Miguel de Benito Delgado for reading and improving the manuscript. We further acknowledge the finantial support of the Studienstiftung des Deutschen Volkes.

BIBLIOGRAPHY

[1] Robert A. Adams. *Sobolev Spaces*. Academic Press, 1975.

[2] Adrien Blanchet, José A. Carrillo, and Nader Masmoudi. Infinite time aggregation for the critical Patlak-Keller-Segel model in ℝ2. *Communications on Pure and Applied Mathematics*, 61(10):1449–1481, 2008.

[3] Adrien Blanchet, Jean Dolbeault, Benoît Perthame, and others. Two-dimensional Keller-Segel model: optimal critical mass and qualitative properties of the solutions. *Electronic Journal of Differential Equations*, 44, 2006.

[4] Niklas Boers and Peter Pickl. On Mean Field Limits for Dynamical Systems. *J Stat Phys*, 164(1):1–16, jul 2016.

[5] Vincent Calvez and José A. Carrillo. Volume effects in the Keller–Segel model: energy estimates preventing blow-up. *Journal de mathématiques pures et appliquées*, 86(2):155–175, 2006.

[6] Patrick Cattiaux and Laure Pédèches. The 2-D stochastic Keller-Segel particle model: existence and uniqueness. *ALEA*, 13(1):447–463, 2016.

[7] Françoise Demengel and Gilbert Demengel. *Functional spaces for the theory of elliptic partial differential equations*. Springer, 2012.

[8] Jean Dolbeault and Benoît Perthame. Optimal critical mass in the two dimensional Keller–Segel model in R2. *Comptes Rendus Mathematique*, 339(9):611–616, 2004.

[9] Giani Egaña Fernández and Stéphane Mischler. Uniqueness and Long Time Asymptotic for the Keller–Segel Equation: The Parabolic–Elliptic Case. *Arch Rational Mech Anal*, 220(3):1159–1194, jun 2016.

[10] Nicolas Fournier and Benjamin Jourdain. Stochastic particle approximation of the Keller-Segel equation and two-dimensional generalization of Bessel processes. *ArXiv preprint arXiv:1507.01087*, jul 2015.

[11] David Godinho and Cristobal Quiñinao. Propagation of chaos for a subcritical Keller–Segel model. *Ann. Inst. H. Poincaré Probab. Statist.*, 51(3):965–992, aug 2015.

[12] Jan Haškovec and Christian Schmeiser. Stochastic Particle Approximation for Measure Valued Solutions of the 2d Keller-Segel System. *J Stat Phys*, 135(1):133–151, apr 2009.

[13] John W. Hagood and Brian S. Thomson. Recovering a function from a Dini derivative. *The American Mathematical Monthly*, 113(1):34–46, 2006.

[14] Hui Huang and Jian-Guo Liu. Error estimate of a random particle blob method for the Keller-Segel equation. *Mathematics of Computation*, may 2016.

[15] Evelyn F. Keller and Lee A. Segel. Initiation of slime mold aggregation viewed as an instability. *Journal of Theoretical Biology*, 26(3):399–415, mar 1970.

[16] R. Kowalczyk. Preventing blow-up in a chemotaxis model. *Journal of Mathematical Analysis and Applications*, 305(2):566–588, may 2005.

[17] Dustin Lazarovici and Peter Pickl. A mean-field limit for the Vlasov-Poisson system. *Arxiv preprint arXiv:1502.04608 [math-ph]*, feb 2015.

[18] Jian-Guo Liu and Rong Yang. A random particle blob method for the Keller-Segel equation and convergence analysis. *Mathematics of Computation*, 86(304):725–745, 2017.

[19] Jian-Guo Liu and Yuan Zhang. Convergence of stochastic interacting particle systems in probability under a Sobolev norm. *Annals of Mathematical Sciences and Applications*, 1(2):251–299, 2016.

[20] Edward J. McShane. *Integration*. Princeton University Press, 1944.

[21] Ross G. Pinsky. *Positive harmonic functions and diffusion*, volume 45. Cambridge University Press, 1995.

[22] Cédric Villani. *Topics in Optimal Transportation*. American Mathematical Soc., 2003.