On a stochastic particle model of the Keller-Segel equation and its macroscopic limit

Dissertation

an der Fakultät für Mathematik, Informatik und Statistik
der Ludwig-Maximilians-Universität
München

eingereicht von
ANA CAÑIZARES GARCÍA
2017

Erstgutachter: Prof. Dr. Peter Pickl

Zweitgutachter: Prof. Dr. Tomasz Cieslak

Drittgutachter: Prof. Dr. Jian-Guo Liu

Tag der Einreichung: 19. Mai 2017

Tag der mündlichen Prüfung: 29. Juni 2017

Eidesstattliche Versicherung

(Nach der Promotionsordnung vom 12.07.11, § 8, Abs. 2 Pkt. .5.)

Hiermit erkläre ich an Eidesstatt, dass die Dissertation von mir selbstständig, ohne unerlaubte Beihilfe angefertigt ist.

München, den 19. Mai 2015 Ana Cañizares García

Abstract

The aim of this thesis is to derive the two-dimensional Keller-Segel equation for *chemotaxis* from a stochastic system of N interacting particles in the situation in which bounded solutions are guaranteed to exist globally in time, that is in the case of *subcritical chemosensitivity* $\chi < 8\pi$. To this end we regularise the singular (Coulomb) interaction force by a cutoff of size $N^{-\alpha}$ for arbitrary $\alpha \in (0, 1/2)$. Our proof adapts a method originally developed for the derivation of the Vlasov-Poisson equation from an N-particle Coulomb system for typical initial conditions [8, 51]. In addition we discuss about recent results in the literature on the nature of the particle collisions [15, 33] that we obtained in an independent way.

Zusammenfassung

Ziel dieser Arbeit ist die Herleitung der zwei-dimensionalen Keller-Segel Gleichung für *Chemotaxis* aus einem wechselwirkenden, stochastischen N-Teilchen System, wenn die Existenz von beschränkten, für alle Zeiten definierten Lösungen vorgegeben ist. Dies entspricht dem *unterkritischen Fall* $\chi < 8\pi$. Hierfür regularisieren wir die singuläre (Coulomb) Wechselwirkung durch einem Cutoff der Ordnung $N^{-\alpha}$, für beliebiges $\alpha \in (0, 1/2)$. Der Beweis erweitert eine Methode, die ursprünglich für die Herleitung der Vlassov-Poisson Gleichung aus einem N-Teilchen Coulomb-System für typische Anfangspositionen entwickelt wurde [8, 51]. Des Weiteren besprechen wir neulich erschienene Ergebnisse über die Teilchenkollisionen [15, 33], die auch wir unabhängig erhielten.

Acknowledgements

I am profoundly grateful to Peter Pickl for guiding me through this work with enthusiasm and genuine insights, to Martin Kolb for suggesting the topic in the first place and for supporting the project, together with Detlef Dürr, with many valuable discussions, as well as to Tomasz Cieslak and Christian Stinner for kindly answering questions about the macroscopic solutions. I am especially thankful to Miguel de Benito for his mathematical, technical and editorial advice and to my family for their unconditional trust. The financial support of the Studienstiftung des Deutschen Volkes is also greatly acknowledged.

Table of Contents

Introduction

This dissertation is concerned with the mathematical analysis of the biological process of *chemotaxis*, the movement of an organism in response to a chemical stimulus, which is observed in some amoebae and bacteria, as well as in some other living beings of more complex structure.

This phenomenon is of great importance from both a theoretical and an applied point of view. Besides the intrinsic interest of understanding complex patterns and behaviours in biology, there are applications of profound social relevance e.g. in medicine, where pharmacological alteration of the chemotactic ability of microorganisms is a powerful tool to control disease spreading or morbidity [31, 45, 11]; another most interesting application is microbial biodegradation of polluted environments, e.g. due to oil spills, discarded pharmaceutical substances or residual radioactive isotopes and heavy metals [55, 60, 47].

Chemotaxis is mathematically modeled either phenomenologically from a macroscopic, global perspective by considering the population as a continuum, or from a microscopic one by describing the behaviour of a finite number of individual organisms. It is intuitively clear that both models must be related to each other since the movement of single organisms is a change of population density and indeed we rigorously prove that the macroscopic model may be derived from the microscopic, what is known in the field as *propagation of chaos*.

More generally, this idea of passing from the discrete to the continuous is pervasive in science as it provides ground for the belief in the validity of, and is often considered rigorous justification for many phenomenological models. This can be seen by the surge in related works in the recent applied mathematics, physics and engineering literature. Notorious examples include atomistic derivations of models in continuum mechanics like linear and non linear elasticity, analysis of fracture mechanics [9, 5, 4] or models in population dynamics and biological evolution [68]. More in the spirit of our passage from the stochastic to the deterministic, the field of *stochastic homogenisation* attempts to provide *effective* macroscopic models for heterogenous media whose microscopic properties display random behaviour, e.g. porous media or composite materials [17, 25]. The development of mathematical tools to tackle (some of) these problems is thus undeniably of great theoretical interest.

In this introductory chapter we walk through the main aspects of our object of study. We begin with a short description of the biological context of the Keller-Segel equation, then we introduce the equations corresponding to the aforementioned macroscopic and microscopic perspectives, and we close the chapter with the precise formulation of the *microscopic derivation*, which is the main result of this dissertation. Chapter 1 gathers some recent results in the literature concerning the microscopic system [15, 33], which we partially obtained in an independent way. Finally, in Chapter 2, we present our microscopic derivation of the Keller-Segel equation, jointly developed with Peter Pickl [16].

1 The underlying process: chemotaxis

Taxis (from the Ancient Greek τάξις: "arrangement") in biology refers to the movement of organisms in response to an external stimulus. In the case of *chemotaxis* it is an external chemical substance that guides the movement, but taxis occurs in relation to many other kinds of stimuli, like in *phototaxis*, *gravitaxis* or *electrotaxis*. The chemotactical movement is of vital importance for a great variety of organisms in processes such as the search for food (an example of *positive* chemotaxis, e.g. towards food) or in the protection from danger (*negative* chemotaxis, e.g. away from poison). For instance, the bacteria *Escherichia coli* is known to direct its movement towards an existing source of sugar [1]. Other examples of the many chemotactical processes presented in [29] are the migration of white blood cells or the growth of axons in the nervous system.

The classical model for chemotaxis is the Keller-Segel equation [49], initially motivated by the extraordinary behaviour of a unicellular organism: *Dictyostelium discoidium* (Dd). This organism is a myxamoeba which grows by cell division as long as the food resources are sufficient. When the nutrients are depleted the cells will first tend to spread out over the available region. After a while, starvation triggers an aggregation phase: some cells start emitting a chemical substance which attracts the other cells leading to the formation of aggregation centers. At each center a *slug* is formed out of several thousands of cells, which migrate together towards new food sources. At the end of migration a *fruiting body* is formed, spores are released, these become myxamoeba and the life cycle starts again.

The transition of Dd from unicellular to a more complex structure is a phenomenon observed in many other higher organisms. Because of its simple lifecycle, Dd has been chosen as model for biomedical research that could help understand the process of cell differentiation [41]. Related models have also been used to describe other chemotactical processes relevant in the development of deseases: *angiogenesis* [23] or the process of inducing new vasculature, related with tumor growth, *atherosclerosis* [44], a chronic desease which causes lipid cells to accumulate in the arterial wall [64], and Alzheimer [54]. We refer to [40] for a nice review on the modelling of chemotaxis and its numerous applications in biology and medicine. For some interpretations of the Keller-Segel model in astrophysics and statistical mechanics see [3, 73, 18, 19].

2 Macroscopic and microscopic approaches

The modelling of chemotaxis (and many other natural processes) may be approached from two different perspectives corresponding to micro- and macro- scales. In the macroscopic approach the whole population is considered and a description of the dynamics of the population density is given, whereas the microscopic point of view is concerned with the dynamics of the single individuals of the system. Microscopic approximations usually arise in a more intuitive way, may be founded on elementary governing laws instead of complicated phenomenological descriptions, and are useful for numerical simulations which can provide a deeper understanding of the problem. However, they quickly become impossible to treat analytically as the number of individuals increases. For this reason a macroscopic, or *effective*, description of the population is necessary that explains the global movement when the number of individuals is very large.

Historically, the first rigorous mathematical model for chemotaxis was given by Keller and Segel in 1970 [49] following the macroscopic approach. This is known as the classical chemotaxis model, although it had been previously derived heuristically by Patlak [61] using a microscopic approach. In the following Sections 2.1 and 2.2 we describe in detail the macroscopic and microscopic equations that we will be working with.

2.1 Macroscopic model

We are concerned with the Keller-Segel model in its parabolic-elliptic form

$$\partial_t \rho + \nabla\cdot(\chi\,\rho\,\nabla S - \nabla\rho) \;=\; 0, \tag{1}$$
$$-\Delta S \;=\; \rho, \tag{2}$$

for the density of cells $\rho\colon [0,\infty)\times\mathbb{R}^2\to\mathbb{R}$ and the concentration of the *chemoattractant* S: $[0,\infty)\times\mathbb{R}^2\to\mathbb{R}$. The constant $\chi>0$ denotes the *chemosensitivity* or response of the cells to the chemical substance. This form of the model has been studied for instance in [46, 10, 26, 7, 37], and can be derived from the classical model [49] when the chemoattractant diffuses much faster than the cells [46].

As described above, during the chemotaxis process of Dd cells are spreading out over the region looking for nutrients when some of them start producing the attracting substance. Therefore, the movement results in a competition between diffusion and aggregation which is represented in equation (1): the flux of cells is a combination of the diffusion term $-\nabla\rho$ and the drift term $\chi\ \rho\ \nabla S$. Equation (2) arises from the fact that the chemoattractant is produced by the cells and diffuses instantaneously.

From a mathematical point of view this equation displays many interesting effects and it has become a topic of intense mathematical research. One important aspect is that in some cases there exist global smooth solutions, while in other situations solutions *blow up* in finite time[1] (corresponding to the clustering of cells). Furthermore, the existence of global solutions or the presence of blow-up events strongly depends 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 $M:=\int_{\mathbb{R}^2}\rho_0(x)\,\mathrm{d}x$ and the chemosensitivity χ [20, 24, 7, 26, 6]. We name some of the many other questions that have been asked: on the steady state solutions [57, 20, 72], on the blow-up profile [39], what happens after a blow-up event [69, 70], or on some generalisations of the model (e.g. with a nonlinear diffusion) [42, 14, 50, 21]. A comprehensive survey on the known results related with the Keller-Segel model from 1970 to 2000 can be found in [41]. We also refer to the more recent reviews [40] and [62].

1. 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.

Blow-up solutions describe precisely a clumping event in the biological process, the creation of point-like aggregates. Experiments show that the process of aggregation requires a high number of individuals; there is a threshold under which no aggregation occurs and above which cells do aggregate [22]. The role of the mass should therefore show up in the model. This role for the 2-dimensional description was completely understood for the first time a decade ago: if $\chi M < 8\pi$, a global and 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 $t \to \infty$ [7, 26, 6]. 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 $\chi = 8\pi$.

In the two dimensional case the system (1)-(2) is often reduced to a single non-linear equation for the population density ρ by taking the Newtonian potential

$$S[\rho](t) := -\frac{1}{2\pi}\int_{\mathbb{R}^2} \log(|x-y|)\,\rho(y)\,\mathrm{d}y = \phi * \rho$$

as solution of (2). Substitution in (1) of the concentration of chemical substance S by this particular solution yields the *McKean-Vlasov equation* with Newtonian interaction potential $\phi := -\frac{1}{2\pi}\log(|x|)$. If we denote the corresponding force field kernel by $k\colon \mathbb{R}^2 \to \mathbb{R}^2$, $k(x) := -\nabla\phi(x) = \frac{x}{2\pi|x|^2}$, (1) becomes (assuming ρ is regular enough)

$$\partial_t \rho = \Delta\rho + \chi\,\nabla\cdot((k*\rho)\,\rho), \qquad \rho(0,\cdot) = \rho_0. \tag{3}$$

We will refer to this equation as the macroscopic model. In Chapter 2, Section 2.3.1 we include results on the existence of solutions of (3) in the subcritical case $\chi < 8\pi$ together with some boundedness and regularity properties.

2.2 Microscopic model

This approach is concerned with the displacements of single particles. The stochastic N-particle system[2] we consider is

$$\mathrm{d}X_t^i = -\frac{\chi}{N}\sum_{j\neq i}^{N} k(X_t^i - X_t^j)\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}B_t^i, \quad i = 1,\ldots,N, \quad X_0 \sim \bigotimes_{i=1}^{N} \rho_0, \tag{4}$$

where the process $X^i\colon [0,\infty) \to \mathbb{R}^2$ denotes the trajectory of the i-th particle, $(B^i)_{i\in\mathbb{N}}$ is a family of 2-dimensional independent Brownian motions, $X_t \in \mathbb{R}^{2N}$ denotes the vector $X_t := (X_t^1, \ldots, X_t^N)$, and at the initial time $t = 0$ the particles are independently distributed according to the initial density ρ_0. As before:

$$k(x) = \frac{x}{2\pi|x|^2}.$$

2. On the topic of stochastic differential equations we refer to [59] for an introduction with many examples and to [65] for a comprehensive exposition including more advanced material.

Equation (4) models a system of N stochastic interacting particles with identical masses $\frac{1}{N}$ and Coulomb interaction k. The stochastic character of the particle system in contrast to the deterministic character of the macroscopic equation should not be surprising. In fact this agrees with what is observed: irregular movement of single members results in a regular movement of the whole population. The competition between diffusion and aggregation of particles is also present at the microscopic level. The interaction (drift) term describes the guided movement towards a higher concentration of chemoattractant (which by assumption is produced by the particles themselves and diffuses infinitely fast, so it decays with the inverse of the distance to the particles for $d=2$) while the Brownian motion (diffusion) term describes the random spread of the particles. In this approach the chemosensitivity χ plays an important role in the clustering of particles too. This matter, together with the existence of solutions, is exposed in Chapter 1.

The microscopic system of equations (4) has been considered by several authors as a basis for numerical methods to simulate solutions of the Keller-Segel equation [37, 30], as well as for deriving from this microscopic model the macroscopic one [38, 34, 33]. On the issue of existence of solutions we refer to [15] and [33].

3 The problem: microscopic derivation

Since a change in population density is necessarily consequence of the movement of the single members, the following question arises naturally: can the macroscopic equation (3) be derived from the microscopic many-particle system (4)? As we mentioned before, finding reasonable microscopic equations whose limit, as the number of particles goes to infinity, agrees with the macroscopic equation supports the validity of the macroscopic model and is therefore an important question to answer. Our goal then is to rigorously derive (3) from the N-particle system (4) in the limit $N \to \infty$.

Let us precise what is meant by *microscopic derivation*. The result should be of the kind "*the positions of the N particles are well represented by the population density ρ_t if N is large enough*" or, in a more mathematical language, "*the empirical measure*

$$\mu_t^{X,N} := \frac{1}{N} \sum_{i=1}^{N} \delta_{X_t^i} \tag{5}$$

for the particle system converges in some sense to ρ_t as $N \to \infty$", always under the initial assumption that $X_0^1, \dots, X_0^N$ are independent and distributed according to the initial density ρ_0.

Let us first informally discuss why such a result should hold by introducing a new element to our problem: the *mean-field particles*

$$\mathrm{d}Y_t^i = -\chi\,(k * \rho_t)\,(Y_t^i)\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}B_t^i, \quad i = 1, \dots, N, \quad Y_0 = X_0, \tag{6}$$

where $\rho_t = \mathscr{L}(Y_t^i)$ is the probability distribution of any of the i.i.d. Y_t^i. Note that Y^i and X^i start at $t = 0$ at the same position and have a common diffusion term $\sqrt{2}\ \mathrm{d}B_t^i$. The Keller-Segel equation (3) is *Kolmogorov's forward equation* for any solution of (6) and consequently their probability distribution ρ_t solves indeed (3). Moreover, by the strong law of large numbers

$$\mu_t^{Y,N} \longrightarrow \rho_t \text{ a.s. for } N \to \infty,$$

where the empirical measure $\mu_t^{Y,N}$ for the independent particles Y_t is as in (5). One then would hope that the interacting particles X_t^N given by (4) behave asymptotically like Y_t, in particular that the empirical measure for the real particle system $\mu_t^{X,N}$ converges in some way to ρ_t. And in fact this is likely to be true since equation (4) is a linearisation of (6), in the sense that substituting ρ_t in (6) by its approximation $\mu_t^{Y,N}$ yields (4).

This is known as *propagation of chaos*, which refers to the propagation in time of the independence (chaoticity) for a system of N indistinguishable interacting particles. This concept was first introduced by Kac [48] for the derivation of the Boltzmann equation and since then it has become a popular method for showing the derivation of deterministic mean-field equations from systems of interacting stochastic particles [67, 58, 32, 36, 34]. The property of propagation of chaos can be expressed in terms of convergence of the empirical measure or convergence of the k-particle marginals. The following three statements are in fact equivalent:

Proposition 1. *Let $X=(X^1,\ldots,X^N)$ be an exchangeable*[3] *$\mathbb{R}^{2N}$-valued random variable. We denote by $\Psi^N\in\mathscr{P}(\mathbb{R}^{2N})$*[4] *the law of X, by ${}^{(k)}\Psi^N$ its k-th marginal*

$$ {}^{(k)}\Psi^N := \int_{\mathbb{R}^{2(N-k)}} \Psi^N \,\mathrm{d}x_{k+1}\ldots\mathrm{d}x_N, \quad k\geqslant 1, $$

and by $\mu^{X,N} := \frac{1}{N}\sum_{i=1}^N \delta_{X^i}$ the associated empirical measure. For a given probability measure $\rho\in\mathscr{P}(\mathbb{R}^2)$ there are equivalent:

i. For all $k\geqslant 1$, ${}^{(k)}\Psi^N$ converges weakly to $\otimes_{i=1}^k\rho$, as $N\to\infty$.
ii. ${}^{(2)}\Psi^N$ converges weakly to $\rho\otimes\rho$, as $N\to\infty$.
iii. The $\mathscr{P}(\mathbb{R}^2)$-valued random variable $\mu^{X,N}$ converges in law to the constant ρ, as $N\to\infty$.

We refer to [67, Prop. 2.2] and to [36, Theorem 1.2] for a quantitative version of the equivalence.

Ideally one would like to derive the Keller-Segel equation (3) directly from the particle system (4) in the case where global solutions of (3) exist. If we recall the dichotomy mentioned above, this corresponds to the sub-critical regime $\chi\in(0,8\pi)$. However, this remains an open problem and we are just able to prove the propagation of chaos for a regularised version of the particle system. The method we present in Chapter 2 needs the particle interaction to be bounded, although the bound is allowed to explode as $N\to\infty$. For this reason we introduce a regularisation of the interaction force k^N, a cutoff of order $N^{-\alpha}$ for an arbitrary $\alpha\in(0,1/2)$, and derive the Keller-Segel equation (3) from the corresponding regularised particle system, defined later in (2.6).

More precisely, we prove propagation of chaos in terms of the k-th marginals for the regularised particle system (Corollary 2.2):

3. The random variables $(X^1,\ldots,X^N)$ are exchangeable if the law of $(X^1,\ldots,X^N)$ is invariant under permutations.

4. We denote by $\mathscr{P}(\mathbb{R}^d)$ the space of Borel probability measures on $\mathbb{R}^d$.

Let X^N be the solution of the regularised particle system (2.6) starting at independent and identically distributed positions according to a given density ρ_0 (under some assumptions) and let ρ_t solve the Keller-Segel equation (3) with initial density ρ_0. Then, for each $t \geqslant 0$, $k \geqslant 1$,

$$^{(k)}\Psi_t^N \rightharpoonup \otimes_{i=1}^k \rho_t \text{ weakly, as } N \to \infty,$$

where $^{(k)}\Psi_t^N$ denotes the k-th marginal of X_t^N.

1

The microscopic system

Is the critical value $\chi=8\pi$ for the existence of global solutions of the macroscopic Keller-Segel equation also encoded in the microscopic system? In which way does the value of χ affect the behaviour of the single particles and the existence or non-existence of global solutions? These questions arise naturally for the microscopic equations in view of the known results for the macroscopic model. In this short chapter we give an overview of what has been done in this direction which provides some answers and a deeper understanding of the microscopic setting. The results presented here were first published by Fournier and Jourdain [33], and Cattiaux and Pédèches [15], who made great progress in the study of the microscopic equations. By the time these papers appeared we had arrived independently at the same results on the nature of collisions (Lemma 1.3 and Remark 1.4). The short answer to our opening question is of course yes, $\chi=8\pi$ is also critical for the microscopic system.

Recall that the microscopic stochastic N-particle system is described by a system of stochastic differential equations

$$\mathrm{d}X_t^i=-\frac{\chi}{N}\sum_{j\neq i}^{N}k(X_t^i-X_t^j)\,\mathrm{d}t+\sqrt{2}\,\mathrm{d}B_t^i,\quad i=1,\ldots,N,\quad X_0\sim\bigotimes_{i=1}^{N}\rho_0, \tag{1.1}$$

where the process X^i: $[0,\infty)\to\mathbb{R}^2$ denotes the trajectory of the i-th particle, $(B^i)_{i\in\mathbb{N}}$ is a family of 2-dimensional independent Brownian motions, $X_t\in\mathbb{R}^{2N}$ denotes the vector $X_t:=(X_t^1,\ldots,X_t^N)$, and at the initial time $t=0$ the particles are independent and identically distributed according to the initial probability measure ρ_0. The interaction force kernel k: $\mathbb{R}^2\to\mathbb{R}^2$ is given by $k(x):=\frac{x}{2\pi|x|^2}$ and the constant $\chi>0$ denotes the chemosensitivity. Any solution of (1.1) is a priori only defined up to the time of the first collision, where the interaction force becomes singular. We shall see below that in some cases, when the collisions between particles are not *too strong* in some precise sense, there exist solutions which are defined globally in time. In other situations this is not possible and a new description of the process after a too strong collision is necessary.

This chapter is structured as follows: In section 1.1 we discuss in an informal way what the expected nature of collisions is, based on comparisons with squared Bessel processes. In section 1.2 we collect some of the results in [15] and [33] on existence of solutions of the particle system (1.1). Finally we briefly present the description of the process in the supercritical case $\chi\geqslant 8\pi$ that was suggested in [33].

1.1 About collisions

By an m-particle collision for $2 \leqslant m \leqslant N$, we mean a collision where exactly m particles are involved. We say that an m-particle collision is *reflecting* if the m particles come appart from each other immediately after the collision. Reciprocally, we say that it is *glueing* if the particles remain together for all future times, forming an m-particle cluster.

In order to illustrate the main idea behind the results of this section we first discuss the 2-particle collisions. The general case $2 \leqslant m \leqslant N$ is presented at the end of the section.

2-particle collisions

Let us for the moment assume the simplest situation $N = 2$. We look at two new processes $U_t^1 \coloneqq \frac{1}{2}(X_t^1 - X_t^2)$, $U_t^2 \coloneqq \frac{1}{2}(X_t^1 + X_t^2)$ and the corresponding equations

$$\begin{aligned} \mathrm{d}U_t^1 &= -\frac{\chi}{2} k(X_t^1 - X_t^2)\,\mathrm{d}t + \mathrm{d}\tilde{B}_t^1, \\ \mathrm{d}U_t^2 &= \mathrm{d}\tilde{B}_t^2 \end{aligned} \tag{1.2}$$

where $\tilde{B}_t^1 \coloneqq \frac{\sqrt{2}}{2}(B_t^1 - B_t^2)$ and $\tilde{B}_t^2 \coloneqq \frac{\sqrt{2}}{2}(B_t^1 + B_t^2)$ are again two independent two-dimensional Brownian motions[1.1]. We end up with a system of two decoupled variables, and the center of mass U_t^2 is a Brownian motion. Most interestingly, the squared norm of U_t^1 is a squared Bessel process of order $\nu = 2 - \frac{\chi}{4\pi}$, since by Itô's chain rule:

$$\begin{aligned} \mathrm{d}|U_t^1|^2 &= 2\,U_t^1 \cdot \mathrm{d}U_t^1 + 2\,\mathrm{d}t \\ &= \left(2 - \frac{\chi}{4\pi}\right)\mathrm{d}t + 2\,U_t^1 \cdot \mathrm{d}\tilde{B}_t^1. \end{aligned} \tag{1.3}$$

Squared Bessel processes have well known properties, and the nature of their collisions with the origin, which depends on their order, is particularly relevant for us. More precisely:

Definition 1.1. *Let $\nu \geqslant 0$. The unique strong solution $Y_t^\nu \geqslant 0$ of the SDE*

$$\mathrm{d}Y_t^\nu = \nu\,\mathrm{d}t + 2\sqrt{Y_t^\nu}\,\mathrm{d}B_t, \qquad Y_0^\nu = y \geqslant 0, \tag{1.4}$$

is called a squared Bessel process of oder $\nu \geqslant 0$*. For $\nu < 0$ the above equation has no global solution, nevertheless we define the* squared Bessel process of oder $\nu < 0$ *to be the strong solution of (1.4) up to the first hitting time of the origin.*

1.1. Two independent two-dimensional Brownian motions are nothing else than a 4-dimensional Brownian motion, and it is a basic result in the theory of stochastic processes that any orthogonal transformation of a d-dimensional Brownian motion results in a new d-dimensional Brownian motion. See, for instance, [59, Exercise 2.5].

Lemma 1.2. *Let Y_t^ν be a squared Bessel process of order $\nu \in \mathbb{R}$ with $Y_0^\nu = y \geqslant 0$ and let τ_0 be the first hitting time of the origin.*

i. If $\nu \geqslant 2$ and $y \neq 0$, then $\tau_0 = \infty$ a.s.
ii. If $0 < \nu < 2$, then $\tau_0 < \infty$ a.s. and 0 is reflecting[1.2].
iii. If $\nu = 0$, then $\tau_0 < \infty$ and $Y_t^\nu = 0$ for $t \geqslant \tau_0$ a.s.
iv. If $\nu < 0$, then $\tau_0 < \infty$ a.s. and Y_t^ν terminates at τ_0.

We refer to [12, Proposition 24.7] or [65, Proposition XI.1.5] for the proof of this lemma.

In view of Lemma 1.2 and equation (1.3), the following dichotomy should hold for the system of two particles: if $0 < \nu < 2$ (which translates into $0 < \chi < 8\pi$) then the two particles collide but come apart again immediately; if $\nu \leqslant 0$ (or $\chi \geqslant 8\pi$) then the particles collide building a cluster with double mass, which evolves further as the Brownian motion $\tilde{B}_t^2$. It is of course not that clear how to describe the system in the case $0 < \chi < 8\pi$ after a reflecting collision, since in principle we only have information about the norm of the distance $|U_t^1|$ and not about the direction in which the particles separate from each other. In particular, the existence of global solutions of (1.3) does not imply the existence of global solutions of (1.1), but we will come back to this issue in section 1.2.

Squared Bessel processes can in fact also be used for studying the nature of the 2-particle collisions in a system of an arbitrary number of particles $N > 2$ for the following simple reason: Imagine particles X_t^1 and X_t^2 (and only those) are about to collide. Then, at least during a short period of time before the collision, the system is in a spatial configuration where the distance between X_t^1 and X_t^2 is significantly smaller than their distance to the other particles $|X_t^i - X_t^j|$ for $i \in \{1, 2\}$, $j \in \{3, \ldots, N\}$. Intuitively, since the influence of the particles $X_t^3, \ldots, X_t^N$ on the dynamics of X_t^1, X_t^2 is minimal during this period, X_t^1, X_t^2 should behave almost as if no other particles were present. In the hypothetical setting of two particles of mass $1/N$ each $|U_t^1|^2$ would be a squared Bessel process of order $\nu_2 := 2 - \frac{\chi}{2\pi N}$. And it is in fact true that $|U_t^1|^2$ is a perturbation of such a process in the spatial configuration where the distance ratios $|X_t^1 - X_t^2|\,|X_t^i - X_t^j|^{-1}$ are small for $i \in \{1, 2\}$, $j \in \{3, \ldots, N\}$:

$$\begin{aligned}
\mathrm{d}|U_t^1|^2 &= (X_t^1 - X_t^2) \cdot \mathrm{d}U_t^1 + 2\,\mathrm{d}t\\
&= -\frac{\chi}{N}(X_t^1 - X_t^2) \cdot \left(k(X_t^1 - X_t^2) + \frac{1}{2} \sum_{j>2}^{N} [k(X_t^1 - X_t^j) - k(X_t^2 - X_t^j)] \right) \mathrm{d}t + 2\,\mathrm{d}t\\
&\quad + 2\,U_t^1 \cdot \mathrm{d}\tilde{B}_t^1\\
&= \left(2 - \frac{\chi}{2\pi N}\right) \mathrm{d}t + R_t^2\,\mathrm{d}t + 2\,U_t^1 \cdot \mathrm{d}\tilde{B}_t^1,
\end{aligned}$$

1.2. 0 is *reflecting* if the process "spends no time" at 0:

$$\int_0^\infty \mathbb{1}_{\{Y_t = 0\}}\,\mathrm{d}t = 0 \quad a.s.$$

See [13] for a classification of boundary points.

where the small perturbation R_t^2 arises from the interaction with the distant particles

$$R_t^2 := -\frac{\chi}{2N}(X_t^1 - X_t^2)\cdot\sum_{j>2}^{N}[k(X_t^1 - X_t^j) - k(X_t^2 - X_t^j)].$$

In this case we cannot apply Lemma 1.2 to the process $|U_t^1|^2$ directly, but we know from the comparison theorem [65, Theorem IX.3.7] that $|U_t^1|^2$ evolves between two squared Bessel processes in the neighbourhood of a 2-particle collision: if $|R_t^2| \leqslant \varepsilon$, then

$$Y_t^{\nu_2-\varepsilon} \leqslant |U_t^1|^2 \leqslant Y_t^{\nu_2+\varepsilon}.$$

Since $\varepsilon > 0$ can be chosen to be arbitrarily small the following situation is expected: if $\chi < 4\pi N$ then 2-particle collisions are reflecting; if $\chi \geqslant 4\pi N$ then 2-particle collisions are glueing.
The argument we presented for the 2-particle collisions can be generalised to the study of m-particle collisions in a system of $N \geqslant m$ particles, as we describe next.

m-particle collisions

Let $N \geqslant m \geqslant 2$. For simplicity we assume that the colliding particles are those labelled as $X_t^1, \ldots, X_t^m$. This does not affect the conclusion, since each particle is indistinguishable from each other. We define the processes

$$\begin{aligned} U_t^m &:= \frac{1}{\sqrt{2m(m+1)}}\sum_{i=1}^{m}(X_t^i - X_t^{m+1}), \qquad m = 1, \ldots, N-1, \\ U_t^N &:= \frac{1}{\sqrt{2N}}\sum_{i=1}^{N} X_t^i, \end{aligned}$$

which extend the above definition of U_t^1, U_t^2 for $N = 2$. If we denote by A the matrix corresponding to this change of variables then $\sqrt{2}\,A$ is again an orthogonal matrix. Therefore, $\{\sqrt{2}\,AB_t^m\}_{m=1}^N = \{\tilde{B}_t^m\}_{m=1}^N$ is a new family of independent Brownian motions. It is clear that a collision between the particles $X_1, \ldots, X_m$ takes place if and only if $\sum_{l=1}^{m-1}|U_t^l|^2$ hits the origin. As one could expect in view of the previous section, $\sum_{l=1}^{m-1}|U_t^l|^2$ is a perturbation of a squared Bessel process in the neighbourhood of such a collision and, if $m = N$, then $\sum_{l=1}^{N-1}|U_t^l|^2$ is itself a squared Bessel process:

Lemma 1.3. *Let* $\nu_m := (m-1)\left(2 - \frac{\chi m}{4\pi N}\right)$, $m = 2, \ldots, N$. *Then* $\sum_{l=1}^{N-1}\mathrm{d}|U_t^l|^2$ *is a squared Bessel process of order* ν_N *and for* $2 \leqslant m < N$

$$\sum_{l=1}^{m-1}\mathrm{d}|U_t^l|^2 = \nu_m\,\mathrm{d}t + R_t^m\,\mathrm{d}t + \sum_{l=1}^{m-1} 2\,U_t^l\cdot\mathrm{d}\tilde{B}_t^l$$

for a one-dimensional process R_t^m *such that* $|R_t^m| \leqslant CN\varepsilon$ *if* $\frac{|X_t^i - X_t^j|}{|X_t^i - X_t^r|} < \varepsilon$ *for* $i \neq j \in \{1, \ldots, m\}$ *and* $r \in \{m+1, \ldots, N\}$.

We refer to [33] or [15] for the proof. In [33] they work with the quantity $\frac{1}{2}\sum_{l=1}^{m}|X_t^l - \overline{X}_t^m|^2$, where $\overline{X}_t^m := \frac{1}{m}\sum_{l=1}^{m}X_t^l$, and in [15] with $\frac{1}{4m}\sum_{i,j=1}^{m}|X_t^i - X_t^j|^2$, but notice that both are in fact equal to $\sum_{l=1}^{m-1}|U_t^l|^2$.

Remark 1.4. Lemmas 1.2, 1.3 and the comparison theorem lead to the following expected behaviour for m-particle collisions:

Let $a_m := 8\pi \frac{N(m-2)}{m(m-1)}$, $b_m := 8\pi \frac{N}{m}$, for $m = 1, \ldots, N$.

i. If $\chi \geqslant b_m$ then m-particle collisions are glueing.
ii. If $\chi \leqslant a_m$ then there are no m-particle collisions.
iii. If $a_m < \chi < b_m$ then m-particle collisions are reflecting.

We write *expected* because in order to prove i and iii rigorously, one should first ensure the existence of the process after such collisions. However ii is always true, as well as i in the case $m = N$ since the continuation of the process in this case is clear: a single N-particle cluster which evolves as a Brownian motion.

1.2 Existence of solutions

The nature of the collisions is clearly related to the existence or non-existence of global solutions. For instance, the non-existence of solutions of (1.1) in the supercritical case $\chi \geqslant 8\pi$ follows already from the previous results on the N-particle collisions: Since $\sum_{l=1}^{N-1} |U_t^l|^2$ is a squared Bessel process of order $\nu_N = (N-1)\left(2 - \frac{\chi}{4\pi}\right)$, Lemma 1.2 proves that if $\chi \geqslant 8\pi$ (i.e. if $\nu_N \leqslant 0$) then the N particles collide (assuming the solution exists long enough) and after the collision either $\sum_{l=1}^{N-1} |U_t^l|^2$ is no longer defined, or the N particles stick together forming a cluster. In any case a solution of the original microscopic system (1.1) cannot be defined globally in time. The existence of solutions in the subcritical case $0 < \chi < 8\pi$ is however a more complicated issue. Cattiaux and Pédèches prove using the theory of Dirichlet forms the existence and uniqueness in law of (weak) solutions[1.3] of the particle system (1.1) for $0 < \chi < 8\pi$ if N is big enough. Their assumption on N ensures that no more than two particles collide at the same time. We collect these results in the next theorem.

Theorem 1.5. [15, Theorem 1.5]

i. *For $N \geqslant 2$, $\chi \geqslant 8\pi$, the system (1.1) does not have any global solution.*
ii. *For $N \geqslant 3$, $\chi < 8\pi\left(1 - \frac{1}{N-1}\right)$, there exists a unique (in law) solution of (1.1) starting from any $x \in M := \{X \in \mathbb{R}^{2N} : X^i = X^j$ for at most one pair $i \neq j\}$*[1.4].

The previous theorem does not cover the existence of solutions for $N = 2$ and $0 < \chi < 8\pi$. Fournier and Jourdain prove in [33] that the system (1.1) with $N = 2$ has a global weak solution which is unique in law if $\chi < 4\pi$, but that there is no global solution if $\chi \geqslant 4\pi$ [33, Remark 16]. They overcome this problem by looking at the equation corresponding to the process $Z_t := |U_t^1|^2 U_t^1$ instead of just U_t^1. Consider the equation which is formally satisfied by $Z_t := |U_t^1|^2 U_t^1$

$$\mathrm{d}Z_t = b(Z_t)\,\mathrm{d}t + \sigma(Z_t)\,\mathrm{d}B_t, \qquad Z_0 = |U_0^1|^2 U_0^1, \tag{1.5}$$

1.3. Otherwise stated, solutions of an SDE are to be understood in the *weak* sense.

1.4. Note that, even though the original assumption is $N \geqslant 4$, $\chi < 8\pi\left(1 - \frac{1}{N-1}\right)$, only the restrictions $\chi < 2\pi N$ and $\chi < 8\pi\left(1 - \frac{1}{N-1}\right)$ are actually needed for the proof. These are also true if $N = 3$, $\chi < 8\pi\left(1 - \frac{1}{N-1}\right)$.

for $b(z) := (16 - 3\,\chi/(2\,\pi))\,|z|^{-2/3}\,z$, $\sigma(z) = 2\,|z|^{-4/3}\,(|z|^2\,I_2 + z\,z^\top)$, where I_2 is the 2×2 identity matrix and $z^\top$ is the transpose of z. For this equation they prove the existence and uniqueness in law of solutions for the whole range $0<\chi<8\,\pi$ under the condition that the process spends no time at zero:

Theorem 1.6. [33, Theorem 17] *Let $N=2$. If $0<\chi<8\,\pi$, then (1.5) has a unique (in law) solution such that a.s. $\int_0^\infty \mathbb{1}_{\{Z_t=0\}}\,\mathrm{d}t=0$. Moreover, $U_t^1=|Z_t|^{-2/3}\,Z_t\,\mathbb{1}_{\{Z_t\neq 0\}}$ solves (1.2) when $0<\chi<4\,\pi$.*

1.3 Dynamics for heavier particles

As we have seen, although (1.1) is a priori only defined up to the time of the first collision, the particle system is still described by this equation for all times if the collisions are not too strong. However, it is clear that (1.1) cannot be fulfilled by the particle system after a glueing collision, where two or more particles remain sticked together. In this case a new description is necessary where heavier particles are allowed. Initially all particles have the same mass $1/N$. After a cluster of m-particles is formed, the cluster should be described as a heavy particle with mass equal to the sum of the m single masses, and the number of total particles should be reduced accordingly. Fournier and Jourdain propose in [33] the following description for the supercritical case $\chi\geqslant 8\,\pi$:

$$\mathrm{d}X_t^i=-\chi\sum_{j\neq i}^{N_t}\mu_t^j\,k(X_t^i-X_t^j)\,\mathrm{d}t+\sqrt{\frac{2}{N\,\mu_t^i}}\,\mathrm{d}B_t^i,\quad i=1,\dots,N_t, \tag{1.6}$$

where N_t denotes the number of particles at time $t\geqslant 0$ and the masses μ_t^i are such that $\sum_{i=1}^{N_t}\mu_t^i=1$. If the sum of the masses of the particles involved in a collision is greater or equal than $8\,\pi/\chi$, then the colliding particles form a cluster and the equations need to be rewritten for the new situation. Otherwise, the particles are instantaneously reflected and continue evolving according to the current equations. However, the existence of solutions for such a system remains an open problem.

2

Microscopic derivation

Abstract

We present a new derivation of 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 $N^{-\alpha}$ for 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.

The order and rate of convergence of our cutoff are comparable to those in [53], but our initial assumptions are more general. Moreover, our method takes explicit advantage of the diffusive character of the Brownian motion. This strategy seems to be new and it could help improve existent results.

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 initially presented by Boers and Pickl [8] and further developed by Lazarovizi and Pickl [51] for the derivation of the Vlasov-Poisson equation from an N-particle Coulomb system for typical initial conditions.

This chapter gathers the content of joint work with P. Pickl [16]. The results are the same as in [16], although here we include some minor corrections such as the use of Dini derivatives in Section 2.5.

2.1 Introduction

We consider the macroscopic and microscopic models presented in the introduction. The two-dimensional Keller-Segel equation

$$\partial_t \rho = \Delta \rho + \chi \nabla \cdot ((k * \rho)\,\rho), \qquad \rho(0,\cdot) = \rho_0, \tag{2.1}$$

where $\rho\colon [0,\infty)\times\mathbb{R}^2 \to [0,\infty)$ 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, and the microscopic stochastic N-particle system

$$\mathrm{d}X_t^i = -\frac{\chi}{N}\sum_{j\neq i}^{N} k(X_t^i - X_t^j)\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}B_t^i, \quad i = 1,\ldots,N, \quad X_0 \sim \bigotimes_{i=1}^{N} \rho_0, \tag{2.2}$$

where the process X^i: $[0, \infty) \to \mathbb{R}^2$ denotes the trajectory of the i-th particle, $(B^i)_{i\in\mathbb{N}}$ is a family of 2-dimensional independent Brownian motions, $X_t \in \mathbb{R}^{2N}$ denotes the vector $X_t := (X_t^1, \ldots, X_t^N)$, and at the initial time $t = 0$ the particles are independently distributed according to the initial density ρ_0.

Our purpose in this chapter is to derive the deterministic macroscopic equation (2.1) in the sub-critical regime $\chi \in (0, 8\pi)$ as the mean-field limit of (2.2) as $N \to \infty$.

To this end we prove the property of propagation of chaos, or weak convergence of the k-th marginals, for a regularised version (with a cutoff depending on N) of this equation in Corollary 2.2. Our method compares the trajectories of the interacting particles to the trajectories of the independent mean-field particles, which are given by the following equation:

$$\mathrm{d}Y_t^i = -\chi\,(k * \rho_t)(Y_t^i)\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}B_t^i, \quad i = 1, \ldots, N, \quad Y_0 = X_0, \tag{2.3}$$

where $\rho_t = \mathscr{L}(Y_t^i)$ is the probability distribution of any of the i.i.d. Y_t^i. We remark that by Itô's formula the Keller-Segel equation (2.1) is Kolmogorov's forward equation for any solution of (2.3) and in particular their probability distribution ρ_t solves (2.1).

Let us next specify our initial assumptions and introduce the announced regularisation of the interaction term.

Conditions on the chemosensitivity and the initial density

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

$$\begin{aligned} \rho_0 &\in L^1(\mathbb{R}^2, (1+|x|^2)\,\mathrm{d}x) \cap L^\infty(\mathbb{R}^2) \cap H^2(\mathbb{R}^2), \\ \rho_0 &\geqslant 0, \\ \int_{\mathbb{R}^2} \rho_0(x)\,\mathrm{d}x &= 1, \\ \rho_0 \log \rho_0 &\in L^1(\mathbb{R}^2). \end{aligned} \tag{2.4}$$

These conditions guarantee global existence, uniqueness and further good properties of the solution of the macroscopic equation (2.1). Section 2.3 reviews these results and the corresponding ones for the solutions of the microscopic systems.

Regularisation of the interaction force

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

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

as well as

$$|\nabla \phi^1(x)| \leqslant (2\pi|x|)^{-1}, \ -\Delta\phi^1(x) \geqslant 0 \text{ and } |\partial_{ij}^2 \phi^1(x)| \leqslant (\pi|x|^2)^{-1}$$

for all $x \in \mathbb{R}^2$ and $i, j \in \{1, 2\}$. For each $N \in \mathbb{N}$ and $\alpha \in (0, 1/2)$, let $\phi^N(x) = \phi^1(N^\alpha x)$ and define the regularised interaction force kernel $k^N\colon \mathbb{R}^2 \to \mathbb{R}^2$ as $k^N \coloneqq -\nabla\phi^N$, which by construction satisfies

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

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 (2.4) and each $N \in \mathbb{N}$ we consider the *regularised Keller-Segel equation*

$$\partial_t \rho^N = \Delta\rho^N + \chi \nabla((k^N * \rho^N)\rho^N), \qquad \rho^N(0, \cdot) = \rho_0, \tag{2.5}$$

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

$$\mathrm{d}X_t^{i,N} = -\frac{\chi}{N}\sum_{j \neq i} k^N\left(X_t^{i,N} - X_t^{j,N}\right)\mathrm{d}t + \sqrt{2}\,\mathrm{d}B_t^i, \quad i = 1, \ldots, N, \quad X_0^N \sim \bigotimes_{i=1}^N \rho_0, \tag{2.6}$$

and the *regularised mean-field trajectories*

$$\mathrm{d}Y_t^{i,N} = -\chi(k^N * \rho_t^N)(Y_t^{i,N})\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}B_t^i, \quad i = 1, \ldots, N, \quad Y_0^N = X_0^N, \tag{2.7}$$

where ρ_t^N denotes the probability distribution of $Y_t^{i,N}$ for each $i = 1, \ldots, N$. As in the non-regularised version this implies that ρ^N solves the regularised Keller-Segel equation (2.5). For $i = 1, \ldots, N$, it is also convenient to denote the regularised interaction force as

$$K_i^N(x_1, \ldots, x_N) \coloneqq -\frac{\chi}{N}\sum_{j \neq i} k^N(x_i - x_j), \quad (x_1, \ldots, x_N) \in \mathbb{R}^{2N} \tag{2.8}$$

and the mean interaction force as

$$\overline{K}_{t,i}^N(x_1, \ldots, x_N) \coloneqq -\chi(k^N * \rho_t^N)(x_i), \quad (x_1, \ldots, x_N) \in \mathbb{R}^{2N}$$

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 any random variable $X \in \mathbb{R}^{2N}$ which is independent of the filtration generated by B_r, $r \geqslant s$, we let $Z_{t,s}^{X,N}$ be the process starting at time s and position X and evolving from time s up to time t with the mean force $\overline{K}^N$, which is given by the solution of

$$\mathrm{d}Z_{t,s}^{X,i,N} = \overline{K}_{t,i}^N(Z_{t,s}^{X,N})\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}B_t^i, \quad i = 1, \ldots, N, \quad Z_{s,s}^{X,N} = X. \tag{2.9}$$

Previous results and overview of the chapter

The question of the microscopic derivation for modified problems has been addressed by several authors: Stevens [66] proved the first rigorous derivation of the Keller-Segel equation in its parabolic-parabolic setting from a stochastic system of *moderately interacting* cell and chemical particles, Haškovec and Schmeiser [37] derived a regularised equation from a regularised particle system with interaction force $k_\varepsilon(x) := \frac{x}{|x|\,(|x|+\varepsilon)}$ (in the limit $N \to \infty$ for fix $\varepsilon > 0$), and Godinho and Quiñinao [34] considered a sub-Keller-Segel equation with less singular force $k_\alpha(x) := \frac{x}{|x|^{\alpha+1}}$, $0 < \alpha < 1$. More recently, great progress has been made for the purely Coulomb case ($\alpha = 1$): Fournier and Jourdain [33] proved the convergence of a subsequence for the particle system (2.2) by a tightness argument in the *very sub-critical* case $\chi < 2\,\pi$ using no regularisation 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 for a regularised version of (2.2) of the same kind as ours [52, 43, 53], 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 (2.4) on the initial density ρ_0 are weaker: Liu and Zhang assume that ρ_0 is compactly supported, Lipschitz continuous and 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 in $\mathbb{R}^2$, while their particles are distributed on a grid. Moreover, in contrast to other similar methods, ours makes use of the diffusive character of the Brownian motion explicitly: It is intuitively clear that the Brownian motion has a "smearing effect" that should be an important ingredient in the propagation of chaos. Here we include a formalisation of this idea that hopefully contributes to the further improvement of the available results by reducing the cutoff, or ideally by getting rid of it.

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 (2.1)-(2.9) in Section 2.3. Section 2.4 is devoted to some preliminary results that we need for the proof of the main result, Theorem 2.1, which is then proven in Section 2.5. Section 2.6 contains the proofs of Propositions 2.4 and 2.5 and is followed by some final remarks.

Notation

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

2.2 Main results

Let the chemosensitivity χ and the initial density ρ_0 satisfy condition (2.4), and for $N \in \mathbb{N}$ let X^N and Y^N be the real and mean-field trajectories solving the regularised microscopic equations (2.6) and (2.7), 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 2.1. *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 initial density ρ_0, the final time T, α and γ, and N_0 depends on ρ_0, T and α.

We remark that Theorem 2.1 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 $\mathscr{P}_1(\mathbb{R}^{2k}):=\big\{\mu\in\mathscr{P}(\mathbb{R}^{2k}):\int|x|\,\mathrm{d}\mu<\infty\big\}$ the subset of probability measures with finite expectation. We define in the latter the first Wasserstein metric W_1 with respect to the normalised Euclidean distance on $\mathbb{R}^{2k}$

$$W_1(\mu,\nu):=\inf_{\pi\in\Pi(\mu,\nu)}\int_{\mathbb{R}^{2k}\times\mathbb{R}^{2k}}\frac{1}{k}\sum_{i=1}^{k}|x^i-y^i|\,\mathrm{d}\pi(x,y),\tag{2.10}$$

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, [71, Theorem 7.12]) that convergence with respect to this metric W_1 implies weak convergence of measures in $\mathscr{P}_1(\mathbb{R}^{2k})$.

Corollary 2.2. *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*

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

Then ${}^{(k)}\Psi_t^N$ converges weakly to $\otimes_{i=1}^k\rho_t^N$ as $N\to\infty$ for each fixed $k\geqslant1$ and the full density Ψ_t^N converges weakly to $\otimes_{i=1}^N\rho_t^N$ as $N\to\infty$. More precisely, there exist a positive constant C and a natural number N_0 such that

$$\sup_{0\leqslant t\leqslant T}W_1\big({}^{(k)}\Psi_t^N,\otimes_{i=1}^k\rho_t^N\big),\ \sup_{0\leqslant t\leqslant T}W_1(\Psi_t^N,\otimes_{i=1}^N\rho_t^N)\leqslant CN^{-\alpha}\tag{2.11}$$

holds for each $k\geqslant1$ and $N\geqslant N_0$. W_1 denotes the first Wasserstein distance (2.10), C and N_0 depend on the initial density ρ_0, the final time 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 some fixed $k \geqslant 1$ 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

$$A := \left\{ \sup_{0 \leqslant t \leqslant T} \|X_t^N - Y_t^N\|_\infty \geqslant N^{-\alpha} \right\}$$

and its complementary separately. On A^c the expectation is simply bounded by $N^{-\alpha}$; on A, according to Theorem 2.1, it is

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

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

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

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 also implies the weak convergence of the k-particle marginal ${}^{(k)}\Psi_t^N$, for $k \geqslant 1$ to the product of measures $\otimes_{i=1}^k \rho_t$ as $N \to \infty$, where ρ_t is the solution of the (non-regularised) Keller-Segel equation (2.1). Indeed since ρ_t^N converges weakly to ρ_t (Proposition 2.3) 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.

2.3 Properties of solutions

2.3.1 Macroscopic equations

Following [28] we say that ρ is a *weak solution* of (2.1) for an initial condition ρ_0 satisfying (2.4) 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 2.3. (Existence and convergence) *Under assumption (2.4) 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 (2.5) in the sense of distributions.
ii. The Keller-Segel equation (2.1) has a unique weak solution $\rho\in L^\infty(\mathbb{R}_+;L^1(\mathbb{R}^2))$.
iii. The sequence (ρ^N) of solutions of (2.5) converges weakly to the solution ρ of the Keller-Segel equation (2.1).

We refer to [7] and [28] 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 (2.1) were proved in [7]. Together with the uniqueness of the weak solution ρ of (2.1), which was proved in [28], 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 2.3 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 (2.4), is necessary. If in addition the initial density is bounded in L^∞ we find that the solutions of the Keller-Segel and the regularised Keller-Segel equations are uniformly bounded in L^∞ as well (Proposition 2.4). 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 2.5. The proofs of these two last propositions are contained in Section 2.6.

Proposition 2.4. (L^∞ estimates) *Assume that χ and ρ_0 satisfy condition (2.4). 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 solutions $(\rho^N)_{N\in\mathbb{N}}$ *of* (2.5) *and the solution* ρ *of* (2.1).

Proposition 2.5. (Hölder estimates) *Assume that* χ *and* ρ_0 *satisfy condition* (2.4). *Then for each* $T>0$ *there exist positive constants* C_1 *and* C_2 *depending on* ρ_0 *and* T*, such that*

i. $\sup_{t\in[0,T]}[\rho_t^N]_{0,\alpha}, \sup_{t\in[0,T]}[\rho_t]_{0,\alpha}\leqslant C_1$, *for any* $\alpha\in(0,1/4]$,
ii. $\sup_{t\in[0,T]}[k^N*\rho_t^N]_{0,1}, \sup_{t\in[0,T]}[k*\rho_t]_{0,1}\leqslant C_2$,

holds for the solutions $(\rho^N)_{N\in\mathbb{N}}$ *of* (2.5) *and the solution* ρ *of* (2.1).

$[\cdot]_{0.\alpha}$ in the previous proposition denotes for $\alpha\in(0,1]$ the Hölder seminorm of a Hölder continuous function $f:\mathbb{R}^n\to\mathbb{R}^m$

$$[f]_{0,\alpha}:=\sup_{x\neq y\in\mathbb{R}^n}\frac{|f(x)-f(y)|}{|x-y|^{\alpha}}.$$

2.3.2 Microscopic equations

We first focus on the interacting N-particle system (2.2) and its regularised version (2.6). Since for each $N>0$ the interaction kernel k^N is globally Lipschitz continuous, the solution of (2.6) is strongly and uniquely well-defined [63, Theorem 1.7.1]. For the original singular situation (2.2) it is much more delicate as we discussed in Chapter 1. Theorem 1.5 states the result by Cattiaux and Pédèches [15, Theorem 1.5] on the existence and uniqueness in law of the particle system (2.2) for $\chi<8\,\pi$ and a big enough N, starting from any $x\in M:=\{X\in\mathbb{R}^{2N}: X^i=X^j$ for at most one pair $i\neq j\}$.

We continue with the mean-field N-particle system (2.3), its regularised version (2.7) and its regularised and linearised version (2.9). According to Proposition 2.5 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 (2.9) has a unique strong solution. For the existence and uniqueness of strong solutions of the non-linear equations (2.3) and (2.7) we refer to [52, Theorem 2.6].

2.4 Preliminary results

Here we provide the results our proof of the main theorem relies on. 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 convenient property, but one can still prove that the regularised force K^N is locally Lipschitz with a bound of order $\log N$, which follows from Lemma 2.6 and the Law of large numbers as presented in Proposition 2.7. This Lipschitz bound is good enough to prove the statement for short times but for larger ones we need to introduce a new intermediate process. This process is proved to be close to X_t^N by the same argument as for short times and close to Y_t^N by a new argument introduced in Lemma 2.8 which compares the densities of the processes instead of comparing the trajectories.

2.4.1 Local Lipschitz bound for the regularised interaction force

The regularised interaction force K^N defined in (2.8) is locally Lipschitz with a 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 later.

Lemma 2.6. *Let $\nu = \nu(N)$ be a finite, unbounded, monotone increasing function of N, and consider the force kernel $k^\nu \colon \mathbb{R}^2 \to \mathbb{R}^2$ with cutoff at $\nu(N)^{-1}$, $k^\nu(x) := -\nabla(\phi^1(\nu\, x))$ for the bump function ϕ^1 defined in Section 2.1, meaning in particular that $k^\nu(x) \leqslant (2\,\pi\,|x|)^{-1}$ and*

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

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

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

where

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

ii. For $i = 1, \ldots, N$, let the i-th component of the resulting force be

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

and define

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

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

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

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

From $|z-y| \leqslant |x-y| \leqslant 2\,\nu^{-1}$ follows $|z| \geqslant 2\,\nu^{-1}$. In particular the derivative of k^ν at z is bounded by $\pi^{-1}\,|z|^{-2} < 2^{-1}\,|z|^{-2}$. Also, since $|z-y| \leqslant |z|$,

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

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). □

2.4.2 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, since all these sets are events where the sample mean and expectation 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 will 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 2.7. (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,

$$\mathbb{E}_{(-i)}(H_i(Z)) := \frac{1}{N}\sum_{j\neq i} (h * u^j)(Z^i). \tag{2.12}$$

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

$$\log N \, \|u^i\|_\infty + \|u^i\|_\infty^2 \leqslant C_0 N^{\varepsilon/2} \tag{2.13}$$

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

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

it suffices to prove

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

for each $\gamma > 0$, $i = 1, \ldots N$ 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 $j = 2, \ldots, N$ we denote by Θ_j the (independent) random variables $\Theta_j := h^\nu(Z^1 - Z^j)$ and by μ_j its expectation

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

Now by Markov's inequality

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

The expectation on the right hand side can be estimated 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 $a := (a_2, \ldots, a_N) \in \mathbb{N}_0^{N-1}$ is a multiindex and $C_a := \binom{2\,m}{a_2, \ldots, a_N} = \frac{(2\,m)!}{a_2! \ldots a_N!}$. Consequently

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

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 $|a|_0$ the number of non-zero entries of the multiindex a, the sum above simplifies to

$$\mathbb{E}\left(\left(\frac{1}{N}\sum_{j\neq 1}^{N}(\Theta_j-\mu_j)\right)^{2m}\right)=N^{-2m}\sum_{\substack{a_2+\ldots+a_N=2m\\|a|_0\leqslant m}} C_a\prod_{j\neq 1}^{N}\mathbb{E}((\Theta_j-\mu_j)^{a_j}). \tag{2.14}$$

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

$$\mathbb{E}((\Theta_j-\mu_j)^{a_j})\leqslant C_h^{a_j}C_0 N^{\alpha(a_j-2)+\varepsilon/2}. \tag{2.15}$$

The a_j-th central 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

$$\begin{aligned}\|h^\nu(z_1-z_j)-\mu_j\|_\infty &\leqslant \|h^\nu\|_\infty+\|(h^\nu*u^j)\|_\infty\\ &\leqslant 2\,\|h\|_\infty\leqslant C_h N^\alpha.\end{aligned}$$

By integrating the term to the second power we find

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

Altogether

$$\begin{aligned}\mathbb{E}((\Theta_j-\mu_j)^{a_j}) &= \int_{\mathbb{R}^2}(h^\nu(z_1-z_j)-\mu_j)^{a_j}u^1(z_1)\,u^j(z_j)\,\mathrm{d}z_1\,\mathrm{d}z_j\\ &\leqslant \|h^\nu(z_1-z_j)-\mu_j\|_\infty^{a_j-2}\int_{\mathbb{R}^2}(h^\nu(z_1-z_j)-\mu_j)^2u^1(z_1)\,u^j(z_j)\,\mathrm{d}z_1\,\mathrm{d}z_j\\ &\leqslant C_h^{a_j}C_0N^{\alpha(a_j-2)+\varepsilon/2},\end{aligned}$$

which proves (2.15).

Let now $k \leqslant m$ and consider only the terms in (2.14) corresponding to multiindices a with k non-zero entries, that is with $|a|_0 = k$. It holds

$$\begin{aligned}\sum_{\substack{a_2+\ldots+a_N=2m\\|a|_0=k}} C_a \prod_{j\neq 1}^{N} \mathbb{E}((\Theta_j-\mu_j)^{a_j}) &\leqslant \sum_{\substack{a_2+\ldots+a_N=2m\\|a|_0=k}} C_a C_h^{2m} C_0^k N^{\alpha(2m-2k)+\varepsilon k/2}\\ &\leqslant \sum_{\substack{a_2+\ldots+a_N=2m\\|a|_0=k}} (2m)^{2m} C_h^{2m} C_0^m N^{\alpha(2m-2k)+\varepsilon k/2},\end{aligned}$$

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 previous sum, i.e. the number of ways of choosing k natural numbers that add up $2m$ and placing them in k positions out of $N-1$, is bounded by $N^k (2m)^k$, we find

$$\begin{aligned}\sum_{\substack{a_2+\ldots+a_N=2m\\|a|_0=k}} C_a \prod_{j\neq 1}^{N} \mathbb{E}((\Theta_j-\mu_j)^{a_j}) &\leqslant (2m)^{3m} C_h^{2m} C_0^m N^{\alpha(2m-2k)+\varepsilon k/2} N^k\\ &\leqslant C_m N^{2m\alpha} N^{k(1-2\alpha+\varepsilon/2)},\end{aligned} \tag{2.16}$$

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 with (2.14) and (2.16)

$$\begin{aligned}\mathbb{E}\left(\left(\frac{1}{N}\sum_{j\neq 1}^{N}(\Theta_j-\mu_j)\right)^{2m}\right) &= N^{-2m} \sum_{\substack{a_2+\ldots+a_N=2m\\|a|_0\leqslant m}} C_a \prod_{j\neq 1}^{N} \mathbb{E}((\Theta_j-\mu_j)^{a_j})\\ &= N^{-2m} \sum_{k=1}^{m} \sum_{\substack{a_2+\ldots+a_N=2m\\|a|_0=k}} C_a \prod_{j\neq 1}^{N} \mathbb{E}((\Theta_j-\mu_j)^{a_j})\\ &\leqslant C_m N^{-2m} \sum_{k=1}^{m} N^{2m\alpha} N^{k(1-2\alpha+\varepsilon/2)}\\ &\leqslant m C_m N^{-2m} N^{m(2\alpha+1-2\alpha+\varepsilon/2)}\\ &\leqslant C_m N^{-m(1-\varepsilon/2)},\end{aligned}$$

where we used the positivity of $1-2\alpha+\varepsilon/2$ and redefined C_m to $m C_m$. Finally we find

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

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

2.4.3 Comparison of solutions of (2.9) 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 (2.9) that start at different points? Corollary 2.9 provides a quantitative answer using the closeness in L^∞ of the probability densities of those solutions, a property that we prove in Lemma 2.8. Although this lemma seems to be an elementary result we were not able to find it explicitly in the literature. Moreover, its application in the proof of the main theorem is an innovation with respect to [8] and [51].

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 $t\geqslant s$ according to the mean-field force $\overline{K}^N$. That is, $Z_{t,s}^{x,N}$ solves (2.9) 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 [63, Theorem 1.10.2]. 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 $u_{t,s}^{x,N}(z_1, \ldots, z_N) \coloneqq \prod u_{t,s}^{x,i,N}(z_i)$. 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*

$$\partial_t u_{t,s}^{x,i,N} = \Delta u_{t,s}^{x,i,N} - \nabla\cdot(f_t^N u_{t,s}^{x,i,N}), \qquad u_{s,s}^{x,i,N} = \delta_{x_i}, \tag{2.17}$$

where $f_t^N \coloneqq \chi k^N * \rho_t^N$ and ρ_t^N solves the regularised Keller-Segel equation (2.5) 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 2.8. *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 hold 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 (2.9):*

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 are 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 $u_t^{x_i}$ instead of $u_{t,0}^{x,i,N}$. What we need to show then is

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

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 fix $i\in\{1,\ldots,N\}$ and define $v_t \coloneqq u_t^{x_i} - u_t^{y_i}$. For a solution of (2.17) it holds

$$\begin{aligned} u_t^{x_i} &= G(t)*\delta_{x_i} - \int_0^t G(t-s)*\mathrm{div}\,(u_s^{x_i} f_s^N)\,\mathrm{d}s \\ &= G(t)*\delta_{x_i} - \int_0^t \nabla G(t-s)*(u_s^{x_i} f_s^N)\,\mathrm{d}s, \end{aligned} \tag{2.18}$$

where $G(t,x) \coloneqq \frac{1}{2\pi t}\exp\left(\frac{-|x|^2}{2t}\right)$ denotes the heat kernel in $\mathbb{R}^2$. By subtracting the corresponding equations for $u_t^{x_i}$ and $u_t^{y_i}$ it follows that

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

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

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

holds due to Bochner's Theorem. Next we split the last integral into two parts and use Young's inequality for convolutions with different exponents for each part:[2.1]

$$\begin{aligned}\int_0^t \|\nabla G(t-s) * (v_s f_s^N)\|_p\,\mathrm{d}s &= \int_0^{t/2} \|\nabla G(t-s) * (v_s f_s^N)\|_p\,\mathrm{d}s \\ &\quad + \int_{t/2}^t \|\nabla G(t-s) * (v_s f_s^N)\|_p\,\mathrm{d}s \\ &\leqslant C\int_0^{t/2} \|\nabla G(t-s)\|_p\,\|v_s\|_1\,\mathrm{d}s \\ &\quad + C\int_{t/2}^t \|\nabla G(t-s)\|_{3/2}\,\|v_s\|_q\,\mathrm{d}s, \qquad (2.19)\end{aligned}$$

where the constant $C \coloneqq \sup_{0\leqslant t\leqslant T}\|f_t^N\|_\infty$ is finite since $\|\rho_t^N\|_1$ is equal to $\|\rho_0\|_1$ and by Proposition 2.4 $\|\rho_t^N\|_\infty$ is also uniformly bounded in $t \in [0,T]$ and $N \in \mathbb{N}$. $p, q \in [1,\infty]$ satisfy $p = 3\frac{q}{3-q}$, which follows from the required relationship $1 + \frac{1}{p} = \frac{1}{r} + \frac{1}{q}$ for the choice $r = 3/2$ in the second integral. The choice of the exponent $r = 3/2$ for the norm of ∇G 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 by taking $p_n \coloneqq q$ and $p_{n+1} \coloneqq p$ in (2.19) we can construct a recursive sequence of inequalities

$$\begin{aligned}\|v_t\|_{p_{n+1}} &\leqslant \|G(t) * (\delta_{x_i} - \delta_{y_i})\|_{p_{n+1}} + C\int_0^{t/2} \|\nabla G(t-s)\|_{p_{n+1}}\,\|v_s\|_1\,\mathrm{d}s \\ &\quad + C\int_{t/2}^t \|\nabla G(t-s)\|_{3/2}\,\|v_s\|_{p_n}\,\mathrm{d}s, \qquad (2.20)\end{aligned}$$

where the exponents satisfy $p_{n+1} = 3\frac{p_n}{3-p_n}$. 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, it 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. In fact, one can check that $k = 4$. Below we go through the first two steps in detail, the last two can be completed analogously. We will need some well-known estimates for the L^p norms of the heat kernel G and its derivative, which are provided in Lemma 2.10.

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

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

Step $k=1$, $p_1=1$: We bound the first norm directly:

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

By a Grönwall-type argument one can show that

$$\|v_t\|_1 \leqslant C\,(t^{-1/2}+1)\,|x_0-y_0|, \qquad t\in[0,T],$$

for some positive constant depending on $\sup_{0\leqslant t\leqslant T}\|f_t^N\|_\infty$ and T.
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 in (2.20) the estimates that we already found:

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

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

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

holds for $x,y\in\mathbb{R}^{2N}$, $t\in(s,T]$ and $Z_{t,s}^{x,N}, Z_{t,s}^{y,N}$ given by (2.9).

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

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

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

Therefore

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

by Lemma 2.8. □

We finally collect some standard estimates for the heat kernel which were required for the proof of Lemma 2.8.

Lemma 2.10. (p-norm estimates of the heat kernel) *Define the heat kernel*

$$
G(t,x) := \frac{1}{2\pi t}\exp\left(\frac{-|x|^2}{2t}\right), \quad (t,x)\in(0,\infty)\times\mathbb{R}^2.
$$

There exists a constant $C>0$ such that for each $p\in[1,\infty]$ 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. We start by showing that $\|G(t)\|_p \leqslant C\frac{1}{t^{1-1/p}}$ 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 $\|\nabla_x G(t)\|_p \leqslant C\,\frac{1}{t^{3/2-1/p}}$, $p\in[1,\infty]$. For $p=\infty$, since $a\exp(-a^2)$ is bounded, one has

$$
|\nabla_x G(t,x)| = \left|\frac{x}{2\pi t^2}\exp\left(-\frac{|x|^2}{2t}\right)\right| = \frac{C}{t^{3/2}}\frac{|x|}{t^{1/2}}\exp\left(-\frac{|x|^2}{2t}\right) \leqslant \frac{C}{t^{3/2}}, \quad (t,x)\in[0,\infty)\times\mathbb{R}^2
$$

For $1 \leqslant p < \infty$:

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

again by the boundedness of $a\exp(-a^2)$ together with the integrability in $\mathbb{R}^2$ of $\exp(-|x|^2)$.
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 check that

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

Finally, for $1 < p < \infty$,

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

□

2.5 Proof of the main theorem

In this section we prove Theorem 2.1. We show that if the regularised real trajectory X^N given by (2.6) and the regularised mean-field trajectory Y^N solving (2.7) 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 $C_\gamma = C_\gamma(\rho_0, T, \alpha, \gamma)$ 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 (2.9) 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 (2.9) 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 (see Figure 2.1).

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 $|Z_{t,t}^N-Z_{t,0}^N|_\infty$. The derivative of $|Z_{t,t}^N-Z_{t,0}^N|_\infty$ 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 $|Z_{t,t}^N-Z_{t,0}^N|_\infty\leqslant N^{-\alpha}$. Since under this assumption we are able to show that the difference $|Z_{t,t}^N-Z_{t,0}^N|_\infty$ goes to zero as $N\to\infty$ faster than $N^{-\alpha}$, or equivalently that $N^\alpha\,|Z_{t,t}^N-Z_{t,0}^N|_\infty\to 0$, the initial condition $|Z_{t,t}^N-Z_{t,0}^N|_\infty\leqslant N^{-\alpha}$ is actually harmless. This idea can be formalised by considering the following object:

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

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 $\{|Z_{t,t}^N-Z_{t,0}^N|_\infty>N^{-\alpha}\}$ (since in this case J_t^N attains its maximum value 1). Therefore, we just need to control the derivative of $N^\alpha\,|Z_{t,t}^N-Z_{t,0}^N|_\infty$ when $|Z_{t,t}^N-Z_{t,0}^N|_\infty\leqslant N^{-\alpha}$. 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 2.7) for the i.i.d. random variables $Z_{t,0}^{1,N},\ldots,Z_{t,0}^{N,N}$ 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 2.6), 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 $|Z_{t,t}^N-Z_{t,0}^N|_\infty\leqslant N^{-\alpha}$, 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

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

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 2.8).

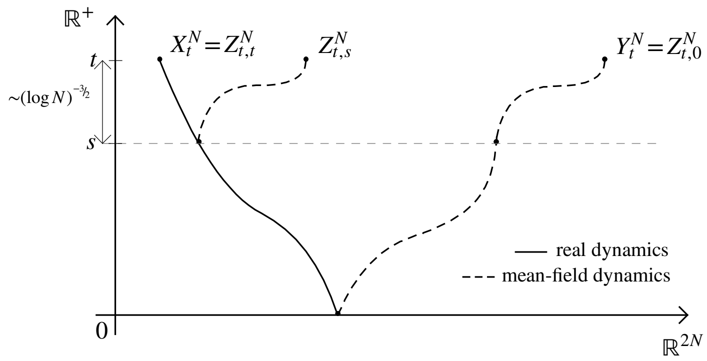

Fig. 2.1. Time splitting

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 2.1

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

$$J_t^N := \min\left\{1, \sup_{0\leqslant s\leqslant t} \mathrm{e}^{C_N(T-s)} \sup_{0\leqslant\tau\leqslant s} \left(N^\alpha f_N(s-\tau)\, |Z_{s,s}^N - Z_{s,\tau}^N|_\infty + N^{-\delta}\right)\right\}, \quad 0 \leqslant t \leqslant T, \tag{2.21}$$

where $C_N := 32\,(\log N)^{3/4}$ and f_N: $[0,\infty)\to\mathbb{R}$ is a positive, monotone decreasing, smooth function satisfying

$$f_N(t)=\begin{cases}\frac{1}{4\,t\log N+(\log N)^{-1/4}}, & 0\leqslant t\leqslant 2^{-3}(\log N)^{-1},\\ 1, & t\geqslant 2^{-2}(\log N)^{-1},\end{cases}$$

and $f_N'(t)\leqslant -C\log N f_N^2(t)\leqslant 0$, for some positive constant C, $0\leqslant t\leqslant 2^{-3}(\log N)^{-1}$.

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

$$a(\tau,s) := N^\alpha f_N(s-\tau)\,|Z_{s,s}^N - Z_{s,\tau}^N|_\infty + N^{-\delta}. \tag{2.22}$$

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

$$\sup_{0\leqslant s\leqslant t} \mathrm{e}^{C_N(T-s)} \sup_{0\leqslant\tau\leqslant s} a(\tau,s) = J_t^N < 1,$$

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

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

where $\overline{D}_t^+$ denotes the upper right Dini derivative given in Definition 2.11 at the end of this proof. This definition is followed by Proposition 2.12, 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 2.13 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 (2.21) using Lemma 2.13. Note that we can write it as

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

where

$$g(\tau,s) := e^{C_N(T-s)}\left(N^{\alpha} f_N(s-\tau)\,|Z^N_{s,s} - Z^N_{s,\tau}|_\infty + N^{-\delta}\right). \tag{2.23}$$

It is clear that $\overline{D}^+_t J^N_t \leqslant \max\{0, \overline{D}^+_t \sup_{0\leqslant\tau\leqslant s\leqslant t} g(\tau,s)\}$. Moreover, the function g satisfies the conditions of Lemma 2.13 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

$$\varphi(\tau,s) := Z^N_{s,s} - Z^N_{s,\tau} = \int_\tau^s \left(K^N(Z^N_{u,u}) - \overline{K}^N_u(Z^N_{u,\tau})\right) \mathrm{d}u$$

is continuous, Lipschitz in τ and in s and has a partial derivative wrt. s. In particular, $|Z^N_{s,s} - Z^N_{s,\tau}|_\infty$ 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

$$\overline{D}^+_s |Z^N_{s,s} - Z^N_{s,\tau}|_\infty = \partial^+_s |Z^N_{s,s} - Z^N_{s,\tau}|_\infty = |\partial^+_s (Z^N_{s,s} - Z^N_{s,\tau})|_\infty.$$

Finally, from the sum and product rules (Proposition 2.12) it follows that $\overline{D}^+_\tau g$ and $\overline{D}^+_s g$ are finite and that $\overline{D}^+_s g$ is a continuous function, namely

$$\begin{aligned} \overline{D}^+_s g(\tau,s) &= \partial^+_s g(\tau,s) \\ &= -e^{C_N(T-s)}\left(C_N a(\tau,s) - N^{\alpha} f'_N(s-\tau)\,|Z^N_{s,s} - Z^N_{s,\tau}|_\infty\right) \\ &\quad + e^{C_N(T-s)} N^{\alpha} f_N(s-\tau)\left|K^N(Z^N_{s,s}) - \overline{K}^N_s(Z^N_{s,\tau})\right|_\infty, \end{aligned} \tag{2.24}$$

where $a(\tau,\, s)$ is defined in (2.22). We can then apply Lemma 2.13 to the function $\sup_{0\leqslant\tau\leqslant s\leqslant t} g(\tau,s)$ with the set of boundary maximal points

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

and find the following estimate:

$$\overline{D}^+_t J^N_t \;\leqslant\; \max\left\{0,\ \sup_{(\tau,t)\in\hat{M}_t} \overline{D}^+_s g(\tau,t)\right\}.$$

If $\hat{M}_t = \varnothing$ then $\overline{D}^+_t J^N_t \leqslant 0$ 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 $|Z^N_{s,s} - Z^N_{s,\tau}|_\infty \leqslant N^{-\alpha}$ holds for each $0 \leqslant \tau \leqslant s \leqslant t$. Consider the event $\mathscr{A}_t := \{\overline{D}^+_t J^N_t > 0\}$. Since $\mathscr{A}_t \subseteq \{\overline{D}^+_s g(\tau_t, t) \geqslant \overline{D}^+_t J^N_t\}$ it holds that

$$\mathbb{E}(\overline{D}^+_t J^N_t) = \mathbb{E}(\overline{D}^+_t J^N_t | \mathscr{A}^c_t) + \mathbb{E}(\overline{D}^+_t J^N_t | \mathscr{A}_t) \leqslant 0 + \mathbb{E}(\overline{D}^+_t J^N_t | \mathscr{A}_t) \leqslant \mathbb{E}(\overline{D}^+_s g(\tau_t, t) | \mathscr{A}_t), \tag{2.25}$$

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 $\mathscr{A}_t$ one has $J^N_t \leqslant 1$ and in particular $\sup_{0\leqslant\tau\leqslant s\leqslant t} |Z^N_{s,s} - Z^N_{s,\tau}|_\infty \leqslant N^{-\alpha}$ holds. As a first estimate we can prove that in this set the bound $\overline{D}^+_s g(\tau_t,\, t)$ of the derivative $\overline{D}^+_t J^N_t$ grows slower than N^2: Using that

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

and

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

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

$$\begin{aligned}\overline{D}_s^+ g(\tau,t) &\leqslant \mathrm{e}^{C_N(T-t)}(C_N a(\tau,t) + N^\alpha |f_N'(t-\tau)|\,|Z_{t,t}^N - Z_{t,\tau}^N|) \\ &\quad + \mathrm{e}^{C_N(T-t)} N^\alpha f_N(t-\tau) \left|K^N(Z_{t,t}^N) - \overline{K}_t^N(Z_{t,\tau}^N)\right| \\ &\leqslant C\,\mathrm{e}^{C_N(T-t)}((\log N)^{3/4} + N^\alpha (\log N)^{3/2} N^{-\alpha} + N^\alpha (\log N)^{1/4} N^\alpha) \\ &\leqslant C\,\mathrm{e}^{C_N T} N^{3/2} < C N^2. \end{aligned} \tag{2.26}$$

With bounds (2.25) and (2.26) 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 2.6 applied to K^N. In the notation of Lemma 2.6, 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

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

for

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

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

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

Then, from Lemma 2.6 it follows that if $x, y \in \mathbb{R}^{2N}$, $|x - y|_\infty \leqslant N^{-\alpha}$, then

$$|K^N(x) - K^N(y)| \leqslant 2\,|L^N(y)|\,|x - y|. \tag{2.27}$$

Notice that $\|\overline{L}_t^N\|_\infty$ is of order $O(\log N)$: Indeed, $l^{v(N)} = l_1^{v(N)} + l_\infty^{v(N)} \in L^1(\mathbb{R}^2) + L^\infty(\mathbb{R}^2)$ with integrable part satisfying $\|l_1^{v(N)}\|_1 = O(\log N)$ 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

$$\mathscr{B}_t^1 := \{|K^N(Y_t^N) - \overline{K}_t^N(Y_t^N)| \leqslant N^{-(\alpha+\delta)}\} \cap \{|L^N(Y_t^N) - \overline{L}_t^N(Y_t^N)| \leqslant C\}, \tag{2.28}$$

where the two desired properties hold. Moreover, as a consequence of the law of large numbers (Proposition 2.7) 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 2.14 at the end of this section).

Recalling (2.25) we write

$$\mathbb{E}(\overline{D}_t^+ J_t^N) \leqslant \mathbb{E}(\overline{D}_s^+ g(\tau_t,t)|\mathscr{A}_t) = \mathbb{E}(\overline{D}_s^+ g(\tau_t,t)|\mathscr{A}_t \backslash \mathscr{B}_t^1) + \mathbb{E}(\overline{D}_s^+ g(\tau_t,t)|\mathscr{A}_t \cap \mathscr{B}_t^1). \tag{2.29}$$

Since $\overline{D}_s^+ g(\tau_t,t)$ grows in the set $\mathscr{A}_t$ only polynomially in N by estimate (2.26), by Proposition 2.14 we can find a positive constant C_γ such that the first term in (2.29) satisfies

$$\mathbb{E}(\overline{D}_s^+ g(\tau_t,t)|\mathscr{A}_t \backslash \mathscr{B}_t^1) \leqslant C_\gamma N^{-\gamma}. \tag{2.30}$$

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:

$$\begin{aligned} f_N(t-\tau_t)\,|K^N(Z_{t,t}^N) - \overline{K}_t^N(Z_{t,\tau_t}^N)| \;&\leqslant\; -f_N'(t-\tau_t)\,|Z_{t,t}^N - Z_{t,\tau_t}^N| \\ &\quad + C_N\,(f_N(t-\tau_t)\,|Z_{t,t}^N - Z_{t,\tau_t}^N| + N^{-(\alpha+\delta)}). \end{aligned} \tag{2.31}$$

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 (2.27), $|Z_{t,t}^N - Z_{t,0}^N| \leqslant N^{-\alpha}$ 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,

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

where L is the Lipschitz constant of $\overline{K}_t^N$ (uniform in $t \in [0, T]$). Now observe that, by the definition of J_t^N, $f_N(t-s)\,|Z_{t,t}^N - Z_{t,s}^N| \leqslant f_N(t-\tau_t)\,|Z_{t,t}^N - Z_{t,\tau_t}^N|$ 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

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

which proves (2.31). 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 $\mathscr{A}_t \cap \mathscr{B}_t^1$ and with (2.30)

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

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 2.8 and Corollary 2.9). 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 2.6 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

$$(K^N_2)_i(x_1,\ldots,x_N) \;:=\; -\frac{\chi}{N}\sum_{j\neq i} k^N_2(x_i-x_j), \quad (x_1,\ldots,x_N)\in\mathbb{R}^{2N}. \tag{2.32}$$

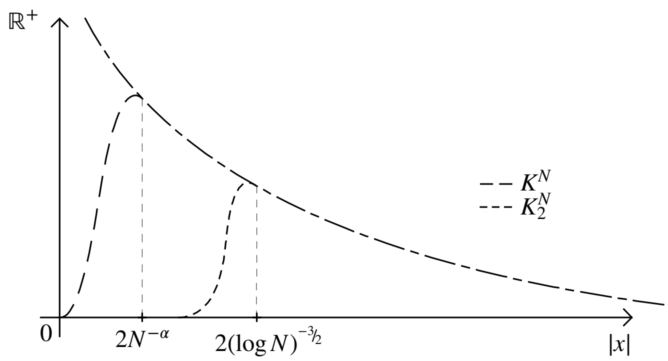

Fig. 2.2. Force splitting

We denote the local Lipschitz bound for K^N_2 given by Lemma 2.6 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:

$$\mathscr{B}^2_t := \mathscr{B}^1_t \cap \{|L^N_2(Y^N_t) - \overline{L}^N_{2,t}(Y^N_t)| \leqslant C\}. \tag{2.33}$$

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 2.14 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

$$\mathbb{E}(\overline{D}_s^+ g(\tau_t, t) | \mathscr{A}_t \backslash \mathscr{B}_t^2) \leqslant C_\gamma N^{-\gamma}. \tag{2.34}$$

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:

$$\begin{aligned} f_N(t-\tau_t)\, \mathbb{E}(|K^N(Z_{t,t}^N) - \overline{K}_t^N(Z_{t,\tau_t}^N)| | \mathscr{A}_t \cap \mathscr{B}_t^2) \;\leqslant\; & -f_N'(t-\tau_t)\, \mathbb{E}(|Z_{t,t}^N - Z_{t,\tau_t}^N| | \mathscr{A}_t \cap \mathscr{B}_t^2) \\ & + C_N f_N(t-\tau_t)\, \mathbb{E}(|Z_{t,t}^N - Z_{t,\tau_t}^N| | \mathscr{A}_t \cap \mathscr{B}_t^2) \\ & + C_N N^{-(\alpha+\delta)}\, \mathbb{P}(\mathscr{A}_t \cap \mathscr{B}_t^2) \\ & + C_\gamma N^{-\gamma}. \end{aligned} \tag{2.35}$$

To this end we write as before

$$\begin{aligned} |K^N(Z_{t,t}^N) - \overline{K}_t^N(Z_{t,\tau_t}^N)| \;\leqslant\; & |K^N(Z_{t,t}^N) - K^N(Z_{t,0}^N)| + |K^N(Z_{t,0}^N) - \overline{K}_t^N(Z_{t,0}^N)| \\ & + |\overline{K}_t^N(Z_{t,0}^N) - \overline{K}_t^N(Z_{t,\tau_t}^N)|. \end{aligned} \tag{2.36}$$

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 2.6 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 (2.32), which results in

$$\begin{aligned} |K^N(Z_{t,t}^N) - K^N(Z_{t,0}^N)| \;\leqslant\; & |K^N(Z_{t,t}^N) - K^N(Z_{t,s}^N)| + |K^N(Z_{t,s}^N) - K^N(Z_{t,0}^N)| \\ \leqslant\; & |K^N(Z_{t,t}^N) - K^N(Z_{t,s}^N)| + |K_1^N(Z_{t,s}^N) - K_1^N(Z_{t,0}^N)| \\ & + |K_2^N(Z_{t,s}^N) - K_2^N(Z_{t,0}^N)|. \end{aligned} \tag{2.37}$$

We can now use the Lipschitz bound for the first and third terms in (2.37): By (2.27), in $\mathscr{A}_t \cap \mathscr{B}_t^2$ it holds that

$$\begin{aligned} |K^N(Z_{t,t}^N) - K^N(Z_{t,s}^N)| \;\leqslant\; & 2\,|L^N(Z_{t,s}^N)|\, |Z_{t,t}^N - Z_{t,s}^N| \\ \leqslant\; & 6\,(C + \|\overline{L}_t^N\|_\infty)\, |Z_{t,t}^N - Z_{t,s}^N| \\ \leqslant\; & 7 \log N\, |Z_{t,t}^N - Z_{t,s}^N| \\ \leqslant\; & 7 \log N \frac{f_N(t-\tau_t)}{f_N(t-s)}\, |Z_{t,t}^N - Z_{t,\tau_t}^N| \\ \leqslant\; & 14\,(\log N)^{3/4} f_N(t-\tau_t)\, |Z_{t,t}^N - Z_{t,\tau_t}^N|, \end{aligned} \tag{2.38}$$

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 $f_N(t-s) = f_N((\log N)^{-3/2}) \geqslant 2^{-1} (\log N)^{1/4}$. We analogously obtain the following estimate for the third term in (2.37)

$$\begin{aligned} |K_2^N(Z_{t,s}^N) - K_2^N(Z_{t,0}^N)| \;\leqslant\; & 2\,|L_2^N(Z_{t,0}^N)|\, |Z_{t,s}^N - Z_{t,0}^N| \\ \leqslant\; & 2\,(\|\overline{L}_{2,t}^N\|_\infty + C)\, |Z_{t,s}^N - Z_{t,0}^N| \\ \leqslant\; & 3 \log\log N\, |Z_{t,s}^N - Z_{t,0}^N| \\ \leqslant\; & 3 \log\log N f_N(t-\tau_t) \left(\frac{1}{f_N(t-s)} + \frac{1}{f_N(t)} \right) |Z_{t,t}^N - Z_{t,\tau_t}^N| \\ \leqslant\; & 6 \log\log N f_N(t-\tau_t)\, |Z_{t,t}^N - Z_{t,\tau_t}^N|. \end{aligned} \tag{2.39}$$

The estimate provided by the local Lipschitz bound from Lemma 2.6 works for $|K^N(Z^N_{t,t}) - K^N(Z^N_{t,s})|$ and $|K^N_2(Z^N_{t,s}) - K^N_2(Z^N_{t,0})|$ 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^N_2 has a milder derivative which is only of order log log N. For the remaining term $|K^N_1(Z^N_{t,s}) - K^N_1(Z^N_{t,0})|$ in (2.37) we use that the probability densities of $Z^N_{t,s}$ and $Z^N_{t,0}$ are close in L^∞ by Lemma 2.8 and its Corollary 2.9. Note that in order to complete the last argument we need independence of the particles and, although the mean-field particles $Z^{1,N}_{t,0}, \ldots, Z^{N,N}_{t,0}$ are independent, this does not hold for the particles $Z^{1,N}_{t,s}, \ldots, Z^{N,N}_{t,s}$ (recall that by definition $Z^N_{t,s} = Z^{X^N_s}_{t,s}$ and that $Z^N_{t,0} = Z^{Y^N_s}_{t,s}$ for $t \geqslant s$). For this reason, instead of considering the processes starting at time s at the r.v. X^N_s and Y^N_s 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 $Z^{x,N}_{t,s}$ and $Z^{y,N}_{t,s}$. This being done, we can recover the original processes $Z^N_{t,s}$ and $Z^N_{t,0}$ by writing $\mathbb{E}(|K^N_1(Z^N_{t,s}) - K^N_1(Z^N_{t,0})||\mathscr{A}_t \cap \mathscr{B}^2_t)$ as

$$\int_{(x,y)\in(Z^N_{s,s}, Z^N_{s,0})(\mathscr{A}_t\cap\mathscr{B}^2_t)} \mathbb{E}\left(|K^N_1(Z^{x,N}_{t,s}) - K^N_1\left(Z^{y,N}_{t,s}\right)||\mathscr{A}_t \cap \mathscr{B}^2_t\right) \mathbb{P}(X^N_s \in \mathrm{d}x, Y^N_s \in \mathrm{d}y). \tag{2.40}$$

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

$$\begin{aligned} \mathbb{E}\left(|K^N_1(Z^{x,N}_{t,s}) - K^N_1\left(Z^{y,N}_{t,s}\right)||\mathscr{A}_t \cap \mathscr{B}^2_t\right) &= \mathbb{E}\left(|K^N_1(Z^{x,N}_{t,s}) - K^N_1\left(Z^{y,N}_{t,s}\right)||(\mathscr{A}_t \cap \mathscr{B}^2_t)\backslash\mathscr{C}^{x,y}_t\right) \\ &\quad + \mathbb{E}\left(|K^N_1(Z^{x,N}_{t,s}) - K^N_1\left(Z^{y,N}_{t,s}\right)||\mathscr{A}_t \cap \mathscr{B}^2_t \cap \mathscr{C}^{x,y}_t\right), \end{aligned}$$

where we introduced the new set

$$\begin{aligned} \mathscr{C}^{x,y}_t := &\{|K^N_1(Z^{x,N}_{t,s}) - \mathbb{E}(K^N_1(Z^{x,N}_{t,s}))| \leqslant N^{-(\alpha+\delta)}\} \\ &\cap \left\{|K^N_1\left(Z^{y,N}_{t,s}\right) - \mathbb{E}\left(K^N_1\left(Z^{y,N}_{t,s}\right)\right)| \leqslant N^{-(\alpha+\delta)}\right\}, \end{aligned} \tag{2.41}$$

for $s = t - (\log N)^{-3/2}$. By Proposition 2.14 below the measure of the set $\Omega\backslash\mathscr{C}^{x,y}_t$ is exponentially small. Also note that the bound given in Proposition 2.14 does not depend of the points x, y. Since K^N_1 is of order $O(N^\alpha)$ we can find a constant $C_\gamma > 0$ such that

$$\mathbb{E}\left(|K^N_1(Z^{x,N}_{t,s}) - K^N_1\left(Z^{y,N}_{t,s}\right)||(\mathscr{A}_t \cap \mathscr{B}^2_t)\backslash\mathscr{C}^{x,y}_t\right) \leqslant C_\gamma N^{-\gamma}.$$

Next we estimate $|K^N_1(Z^{x,N}_{t,s}) - K^N_1(Z^{y,N}_{t,s})|$ in the set $\mathscr{A}_t \cap \mathscr{B}^2_t \cap \mathscr{C}^{x,y}_t$. We write

$$\begin{aligned} \left|K^N_1(Z^{x,N}_{t,s}) - K^N_1\left(Z^{y,N}_{t,s}\right)\right| \leqslant\; &|K^N_1(Z^{x,N}_{t,s}) - \mathbb{E}(K^N_1(Z^{x,N}_{t,s}))| + \left|K^N_1\left(Z^{y,N}_{t,s}\right) - \mathbb{E}\left(K^N_1\left(Z^{y,N}_{t,s}\right)\right)\right| \\ &+ \left|\mathbb{E}(K^N_1(Z^{x,N}_{t,s})) - \mathbb{E}\left(K^N_1\left(Z^{y,N}_{t,s}\right)\right)\right|. \end{aligned}$$

In $\mathscr{C}^{x,y}_t$ the first two terms are bounded. For the remaining term $|\mathbb{E}(K^N_1(Z^{x,N}_{t,s})) - \mathbb{E}(K^N_1(Z^{y,N}_{t,s}))|$ we use the following fact: both processes $Z^{x,N}_{t,s}$ and $Z^{y,N}_{t,s}$ evolved according to the mean-field dynamics during a period of time $t - s$, which is long enough to ensure that the densities of $Z^{x,N}_{t,s}$ and $Z^{y,N}_{t,s}$ 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 2.9). 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:

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

Using this bound in (2.40) we find an estimate for the original processes

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

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,

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

Together with (2.38) and (2.39) this covers all three terms appearing in (2.37). We can adapt $N_0 \in \mathbb{N}$ chosen at the beginning of the proof so that for $N \geqslant N_0$:

$$\begin{aligned}\mathbb{E}(|K^N(Z_{t,t}^N) - K^N(Z_{t,0}^N)||\mathscr{A}_t \cap \mathscr{B}_t^2) &\leqslant 14(\log N)^{3/4} f_N(t-\tau_t)\,\mathbb{E}(|Z_{t,t}^N - Z_{t,\tau_t}^N||\mathscr{A}_t \cap \mathscr{B}_t^2) \\ &\quad + 6\log\log N f_N(t-\tau_t)\,\mathbb{E}(|Z_{t,t}^N - Z_{t,\tau_t}^N||\mathscr{A}_t \cap \mathscr{B}_t^2) \\ &\quad + 2N^{-(\alpha+\delta)}\mathbb{P}(\mathscr{A}_t \cap \mathscr{B}_t^2) + C_\gamma N^{-\gamma} \\ &\leqslant 15(\log N)^{3/4} f_N(t-\tau_t)\,\mathbb{E}(|Z_{t,t}^N - Z_{t,\tau_t}^N||\mathscr{A}_t \cap \mathscr{B}_t^2) \\ &\quad + 2N^{-(\alpha+\delta)}\mathbb{P}(\mathscr{A}_t \cap \mathscr{B}_t^2) + C_\gamma N^{-\gamma}.\end{aligned}$$

Going back to (2.36) 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, (2.36) becomes

$$\begin{aligned}\mathbb{E}(|K^N(Z^N_{t,t})-\overline{K}^N_t(Z^N_{t,\tau_t})||\mathscr{A}_t\cap\mathscr{B}^2_t) \leqslant\ & 15\,(\log N)^{3/4} f_N(t-\tau_t)\,\mathbb{E}(|Z^N_{t,t}-Z^N_{t,\tau_t}||\mathscr{A}_t\cap\mathscr{B}^2_t)\\ &+2\,N^{-(\alpha+\delta)}\,\mathbb{P}(\mathscr{A}_t\cap\mathscr{B}^2_t)+C_\gamma N^{-\gamma}\\ &+N^{-(\alpha+\delta)}\,\mathbb{P}(\mathscr{A}_t\cap\mathscr{B}^2_t)\\ &+L\,(1+f_N(t-\tau_t))\,\mathbb{E}(|Z^N_{t,t}-Z^N_{t,\tau_t}||\mathscr{A}_t\cap\mathscr{B}^2_t)\\ \leqslant\ & 16\,(\log N)^{3/4} f_N(t-\tau_t)\,\mathbb{E}(|Z^N_{t,t}-Z^N_{t,\tau_t}||\mathscr{A}_t\cap\mathscr{B}^2_t)\\ &+3\,N^{-(\alpha+\delta)}\,\mathbb{P}(\mathscr{A}_t\cap\mathscr{B}^2_t)+C_\gamma N^{-\gamma},\end{aligned}$$

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

$$\begin{aligned}\mathbb{E}(|K^N(Z^N_{t,t})-\overline{K}^N_t(Z^N_{t,\tau_t})||\mathscr{A}_t\cap\mathscr{B}^2_t) \leqslant\ & 32\,(\log N)^{3/4}\,\mathbb{E}(|Z^N_{t,t}-Z^N_{t,\tau_t}||\mathscr{A}_t\cap\mathscr{B}^2_t)\\ &+3\,N^{-(\alpha+\delta)}\,\mathbb{P}(\mathscr{A}_t\cap\mathscr{B}^2_t)+C_\gamma N^{-\gamma},\end{aligned}$$

proving (2.35). As a consequence of (2.34) and (2.35)

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

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

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

For any function $g:\mathbb{R}^d\to\mathbb{R}$ *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*

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

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 2.12. (Some properties of Dini derivatives) *Let* $f, g: \mathbb{R} \to \mathbb{R}$ *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. [35, 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* $[a, b]$.[2.2]

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 [56].

Lemma 2.13. *Let* $g: [0, T] \times [0, T] \to \mathbb{R}$ *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*

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

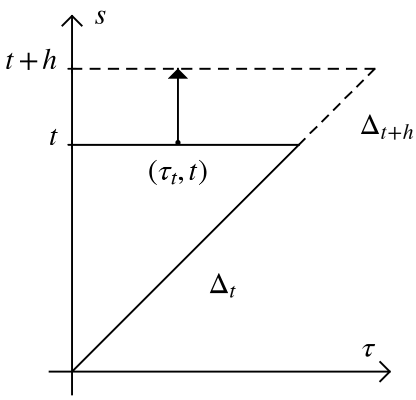

Fig. 2.3. Steepest ascent with Dini derivatives

2.2. The following result [35, 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.

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

$$\underline{D}_1^- g(\tau, t) \geqslant 0 \quad \text{and} \quad \overline{D}_1^+ g(\tau, t) \leqslant 0, \quad (\tau, t) \in \hat{M}_t. \tag{2.42}$$

We only prove the first inequality, the other one 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

$$\begin{aligned} \overline{D}_v g(\tau, s) &= \limsup_{h \to 0^+} \frac{g((\tau, s) + h v) - g(\tau, s)}{h} \\ &\leqslant \limsup_{h \to 0^+} \frac{g(\tau + h v_1, s + h v_2) - g(\tau + h v_1, s)}{h} + \limsup_{h \to 0^+} \frac{g(\tau + h v_1, s) - g(\tau, s)}{h}. \end{aligned}$$

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

$$\begin{aligned} \limsup_{h_1 \to 0^+} \limsup_{h_2 \to 0^+} \frac{g(\tau + h_1 v_1, s + h_2 v_2) - g(\tau + h_1 v_1, s)}{h_2} &= v_2 \limsup_{h_1 \to 0^+} \overline{D}_2^+ g(\tau + h_1 v_1, s) \\ &= v_2 \overline{D}_2^+ g(\tau, s), \end{aligned}$$

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 (2.42) 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 (2.28), (2.33) and (2.41). The constants $T>0$, $\alpha\in(0,1/2)$ and $\delta>0$ are the ones we fixed at the beginning of this section.

Proposition 2.14. (Measure of the exceptional sets) *For each $\gamma>0$ there exists a positive constant C_γ such that:*

i. For each $0\leqslant t\leqslant T$,

$$\mathbb{P}(S_t^1 \cup S_t^2 \cup S_t^3) \leqslant C_\gamma N^{-\gamma},$$

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

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

ii. For any $x\in\mathbb{R}^{2N}$ and any $0\leqslant s\leqslant t\leqslant T$ satisfying $t-s\geqslant(\log N)^{-r}$ for some $r\geqslant 0$,

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

Consequently,

$$\mathbb{P}(\Omega\backslash\mathscr{C}_t^{x,y}) \leqslant 2\, C_\gamma N^{-\gamma}, \text{ for } x,y\in\mathbb{R}^{2N} \text{ and } 0\leqslant t\leqslant T.$$

Proof. It is a direct consequence of the law of large numbers (Proposition 2.7). 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 $\{\sup_{1\leqslant i\leqslant N} |K_i^N(Y_t^N) - \mathbb{E}_{(-i)}(K_i^N(Y_t^N))| \geqslant N^{-(\alpha+\delta)}\}$. Recall that $\mathbb{E}_{(-i)}$ denotes the expectation with respect to every variable but the i-th, as defined in (2.12). Moreover, $Y_t^1, \ldots, Y_t^N$ are already independent and the L^∞-norm of its probability density ρ_t^N is bounded uniformly in N and $t\in[0,T]$ by Proposition 2.4. Therefore, from Proposition 2.7 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 $|N^{-\alpha}(L_2^N)_i(x)|$ are bounded by $C\chi \min\{N^\alpha, |x|^{-1}\}$. Proposition 2.7 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 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 2.8, meaning that the growth of $\|u_{t,s}^{x,i,N}\|_\infty$ is only logarithmic in N and consequently condition (2.13) is fulfilled independently of the times t, s and the exponent r. Therefore, by Proposition 2.7, there exists a constant $C_\gamma > 0$ such that, for any such t, s:

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

2.6 Proofs of Propositions 2.4 and 2.5

Proof of Proposition 2.4

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 [7, Proposition 17] or [28, Lemma 2.7] for the proof for ρ and [7, Lemma 13] for ρ^N.

Step 2: Uniform bounds in L^∞.

For this step we follow [14, Lemma 3.2] and [50, 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 (2.5) and then deduce the result for ρ by passing to the limit (recall that $\rho^N \rightharpoonup \rho$ weakly by Theorem 2.3).

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

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\int \rho_m^p \,\mathrm{d}x \;\leqslant\; & -p^2 \int \rho_m^p \,\mathrm{d}x \\ & + C p^4 \left(\int \rho_m^{p/2} \,\mathrm{d}x\right)^2 + C p^2,\end{aligned} \tag{2.43}$$

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 (2.1) 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

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

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 $p\geqslant 2$ 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

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

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

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

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 (2.43).

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 2.5

i. From the proof of [28, 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 $p\in(2,4]$ 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 2.4. 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:

$$\partial_{ij} w(t,x) = \int_{|x-y|\leqslant 1} \partial_{ij}\phi(x-y)\,\rho(y)\,\mathrm{d}y + \int_{|x-y|\geqslant 1} \partial_{ij}\phi(x-y)\,\rho(y)\,\mathrm{d}y. \tag{2.46}$$

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 (2.46) 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 2.15. 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 [27, 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 [2, 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

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

2.7 Final remarks

There are several ways of improving or extending the present microscopic derivation of the Keller-Segel equation. In the ideal case one would like to prove the propagation of chaos for the original equations and not the regularised ones. An intermediate step towards this result would consist on reducing (narrowing) the cutoff, i.e. regularising the interaction force with a cutoff of order $N^{-\alpha}$, $\alpha \geqslant 1/2$. We believe that our method, specially the way we use the results in Section 2.4.3, could be helpful in this direction. Regarding the supercritical case $\chi \geqslant 8\pi$, where solutions of the macroscopic and microscopic equations are known to blow-up or explode in finite time, it is not clear to us whether some sort of derivation could be possible. However, our method relies on the boundedness and regularity of the solution ρ and is therefore not suitable for $\chi \geqslant 8\pi$.

References

1 ADLER, J. Chemotaxis in Bacteria. *Science 153*, 3737 (Aug 1966), 708–716.

2 ADAMS, R. A. *Sobolev Spaces*. Academic Press, 1975.

3 BILER, P. and NADZIEJA, T. A class of nonlocal parabolic problems occurring in statistical mechanics. *Colloquium Mathematicae 66*, 1 (1993), 131–145.

4 BITZEK, E., KERMODE, J. R., and GUMBSCH, P. Atomistic aspects of fracture. *Int J Fract 191*, 1-2 (Feb 2015), 13–30.

5 BLANC, X., BRIS, C. L., and LIONS, P.-L. Atomistic to Continuum limits for computational materials science. *ESAIM: M2AN 41*, 2 (Mar 2007), 391–426.

6 BLANCHET, A., CARRILLO, J. A., and MASMOUDI, N. Infinite time aggregation for the critical Patlak-Keller-Segel model in $\mathbb{R}2$. *Communications on Pure and Applied Mathematics 61*, 10 (2008), 1449–1481.

7 BLANCHET, A., DOLBEAULT, J., PERTHAME, B., and OTHERS. Two-dimensional Keller-Segel model: optimal critical mass and qualitative properties of the solutions. *Electronic Journal of Differential Equations 44* (2006).

8 BOERS, N. and PICKL, P. On Mean Field Limits for Dynamical Systems. *J Stat Phys 164*, 1 (Jul 2016), 1–16.

9 BRAUN, J. and SCHMIDT, B. On the passage from atomistic systems to nonlinear elasticity theory for general multi-body potentials with p-growth. *Networks and Heterogeneous Media 8*, 4 (Nov 2013), 879–912.

10 BRENNER, M. P., CONSTANTIN, P., KADANOFF, L. P., SCHENKEL, A., and VENKATARAMANI, S. C. Diffusion, attraction and collapse. *Nonlinearity 12*, 4 (1999), 1071.

11 BUTLER, S. M. and CAMILLI, A. Going against the grain: Chemotaxis and infection in Vibrio Colerae. *Nat Rev Microbiol 3*, 8 (Aug 2005), 611–620.

12 BASS, R. F. *Stochastic Processes*. Cambridge University Press, Oct 2011.

13 BORODIN, A. N. and SALMINEN, P. *Handbook of Brownian Motion - Facts and Formulae*. Probability and Its Applications. Birkhäuser Basel, Basel, 2002.

14 CALVEZ, V. and CARRILLO, J. A. Volume effects in the Keller–Segel model: energy estimates preventing blow-up. *Journal de mathématiques pures et appliquées 86*, 2 (2006), 155–175.

15 CATTIAUX, P. and PÉDÈCHES, L. The 2-D stochastic Keller-Segel particle model: existence and uniqueness. *ALEA 13*, 1 (2016), 447–463.

16 CAÑIZARES GARCÍA, A. and PICKL, P. Microscopic derivation of the Keller-Segel equation in the sub-critical regime. *ArXiv preprint arXiv:1703.04376* (Mar 2017).

17 CHAKRABORTY, A. Stochastic models for micromechanics and multiscale fracture analysis of functionally graded materials. *Ph.D. Thesis* (2008).

18 CHAVANIS, P.-H. Generalized thermodynamics and Fokker-Planck equations: Applications to stellar dynamics and two-dimensional turbulence. *Phys. Rev. E 68*, 3 (Sep 2003), 36108.

19 CHAYES, L. and PANFEROV, V. The McKean–Vlasov Equation in Finite Volume. *J Stat Phys 138*, 1-3 (Feb 2010), 351–380.

20 CHILDRESS, S. and PERCUS, J. K. Nonlinear aspects of chemotaxis. *Mathematical Biosciences 56*, 3 (Oct 1981), 217–237.

21 CIEŚLAK, T. and STINNER, C. Finite-time blowup and global-in-time unbounded solutions to a parabolic-parabolic quasilinear Keller-Segel system in higher dimensions. *Journal of Differential Equations 252*, 10 (May 2012), 5832–5851.

22 COHEN, M. H. and ROBERTSON, A. Wave propagation in the early stages of aggregation of cellular slime molds. *Journal of Theoretical Biology 31*, 1 (Apr 1971), 101–118.

23 CORRIAS, L., PERTHAME, B., and ZAAG, H. A chemotaxis model motivated by angiogenesis. *Comptes Rendus Mathematique 336*, 2 (Jan 2003), 141–146.

24 CORRIAS, L., PERTHAME, B., and ZAAG, H. Global Solutions of Some Chemotaxis and Angiogenesis Systems in High Space Dimensions. *Milan j. math. 72*, 1 (Oct 2004), 1–28.

25 CASTANEDA, P. P., TELEGA, J. J., and GAMBIN, B. *Nonlinear Homogenization and its Applications to Composites, Polycrystals and Smart Materials: Proceedings of the NATO Advanced Research Workshop, held in Warsaw, Poland, 23-26 June 2003*. Springer Science & Business Media, Feb 2006.

26 DOLBEAULT, J. and PERTHAME, B. Optimal critical mass in the two dimensional Keller–Segel model in R2. *Comptes Rendus Mathematique 339*, 9 (2004), 611–616.

27 DEMENGEL, F. and DEMENGEL, G. *Functional spaces for the theory of elliptic partial differential equations*. Springer, 2012.

28 EGAÑA FERNÁNDEZ, G. and MISCHLER, S. Uniqueness and Long Time Asymptotic for the Keller–Segel Equation: The Parabolic–Elliptic Case. *Arch Rational Mech Anal 220*, 3 (Jun 2016), 1159–1194.

29 EISENBACH, M. and LENGELER, J. W. *Chemotaxis*. Imperial College Press, London : River Edge, NJ, 2004.

30 FATKULLIN, I. A study of blow-ups in the Keller–Segel model of chemotaxis. *Nonlinearity 26* (2013), 81–94.

31 FORBES, N. S. Engineering the perfect (bacterial) cancer therapy. *Nat Rev Cancer 10*, 11 (Nov 2010), 785–794.

32 FOURNIER, N., HAURAY, M., and MISCHLER, S. Propagation of chaos for the 2d viscous vortex model. *Journal of the European Mathematical Society 016*, 7 (2014), 1423–1466.

33 FOURNIER, N. and JOURDAIN, B. Stochastic particle approximation of the Keller-Segel equation and two-dimensional generalization of Bessel processes. *ArXiv preprint arXiv:1507.01087* (Jul 2015).

34 GODINHO, D. and QUIÑINAO, C. Propagation of chaos for a subcritical Keller–Segel model. *Ann. Inst. H. Poincaré Probab. Statist. 51*, 3 (Aug 2015), 965–992.

35 HAGOOD, J. W. and THOMSON, B. S. Recovering a function from a Dini derivative. *The American Mathematical Monthly 113*, 1 (2006), 34–46.

36 HAURAY, M. and MISCHLER, S. On Kac's Chaos And Related Problems. *Journal of Functional Analysis 266*, 10 (May 2014), 6055–6157.

37 HAŠKOVEC, J. and SCHMEISER, C. Stochastic Particle Approximation for Measure Valued Solutions of the 2d Keller-Segel System. *J Stat Phys 135*, 1 (Apr 2009), 133–151.

38 HAŠKOVEC, J. and SCHMEISER, C. Convergence analysis of a stochastic particle approximation for measure valued solutions of the 2d Keller-Segel system. *Comm. PDE 36* (2011), 940–960.

39 HERRERO, M. A. and VELÁZQUEZ, J. J. L. Singularity patterns in a chemotaxis model. *Math. Ann. 306*, 1 (Sep 1996), 583–623.

40 HILLEN, T. and PAINTER, K. J. A user's guide to PDE models for chemotaxis. *Journal of Mathematical Biology 58*, 1-2 (Jan 2009), 183–217.

41 HORSTMANN, D. From 1970 until present: the Keller-Segel model in chemotaxis and its consequences. (Jan 2003).

42 HORSTMANN, D. Generalizing the Keller–Segel Model: Lyapunov Functionals, Steady State Analysis, and Blow-Up Results for Multi-species Chemotaxis Models in the Presence of Attraction and Repulsion Between Competitive Interacting Species. *J Nonlinear Sci 21*, 2 (Apr 2011), 231–270.

43 HUANG, H. and LIU, J.-G. Error estimate of a random particle blob method for the Keller-Segel equation. *Mathematics of Computation* (May 2016).

44 IBRAGIMOV, A. I., MCNEAL, C. J., RITTER, L. R., and WALTON, J. R. A mathematical model of atherogenesis as an inflammatory response. *Math Med Biol 22*, 4 (Dec 2005), 305–333.

45 IRIMIA, D. Microfluidic Technologies for Temporal Perturbations of Chemotaxis. *Annu Rev Biomed Eng 12* (Aug 2010), 259–284.

46 JAGER, W. and LUCKHAUS, S. On Explosions of Solutions to a System of Partial Differential Equations Modelling Chemotaxis. *Transactions of the American Mathematical Society 329*, 2 (Feb 1992), 819.

47 K, ASHOK, R. *Toxicity and Waste Management Using Bioremediation*. IGI Global, Dec 2015.

48 KAC, M. Foundations of Kinetic Theory. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability* (1956), vol. 3, Univ of California Press, 171–197.

49 KELLER, E. F. and SEGEL, L. A. Initiation of slime mold aggregation viewed as an instability. *Journal of Theoretical Biology 26*, 3 (Mar 1970), 399–415.

50 KOWALCZYK, R. Preventing blow-up in a chemotaxis model. *Journal of Mathematical Analysis and Applications 305*, 2 (May 2005), 566–588.

51 LAZAROVICI, D. and PICKL, P. A mean-field limit for the Vlasov-Poisson system. *Arxiv preprint arXiv:1502.04608 [math-ph]* (Feb 2015).

52 LIU, J.-G. and YANG, R. A random particle blob method for the Keller-Segel equation and convergence analysis. *Mathematics of Computation 86*, 304 (2017), 725–745.

53 LIU, J.-G. and ZHANG, Y. Convergence of stochastic interacting particle systems in probability under a Sobolev norm. *Annals of Mathematical Sciences and Applications 1*, 2 (2016), 251–299.

54 LUCA, M., CHAVEZ-ROSS, A., EDELSTEIN-KESHET, L., and MOGILNER, A. Chemotactic signaling, microglia, and Alzheimer's disease senile plaques: Is there a connection? *Bull. Math. Biol. 65*, 4 (Jul 2003), 693–730.

55 MARX, R. B. and AITKEN, M. D. Bacterial Chemotaxis Enhances Naphthalene Degradation in a Heterogeneous Aqueous System. *Environ. Sci. Technol. 34*, 16 (Aug 2000), 3379–3383.

56 MCSHANE, E. J. *Integration*. Princeton University Press, 1944.

57 NANJUNDIAH, V. Chemotaxis, signal relaying and aggregation morphology. *Journal of Theoretical Biology 42*, 1 (Nov 1973), 63–105.

58 OSADA, H. Propagation of chaos for the two dimensional Navier-Stokes equation. *Probabilistic methods in mathematical physics : proceedings of the Taniguchi International Symposium, Katata and Kyoto, 1985* (1987), 303–334.

59 OKSENDAL, B. *Stochastic differential equations*. Springer, 2003.

60 PANDEY, G. and JAIN, R. K. Bacterial Chemotaxis toward Environmental Pollutants: Role in Bioremediation. *Appl. Environ. Microbiol. 68*, 12 (Dec 2002), 5789–5795.

61 PATLAK, C. S. Random walk with persistence and external bias. *Bulletin of Mathematical Biophysics 15*, 3 (Sep 1953), 311–338.

62 PERTHAME, B. *Transport Equations in Biology*. Frontiers in Mathematics. Birkhäuser Basel, 2007.

63 PINSKY, R. G. *Positive harmonic functions and diffusion*. Cambridge University Press, 1995.

64 RITTER, L. R. A short course in the modeling of chemotaxis. 2004.

65 REVUZ, D. and YOR, M. *Continuous Martingales and Brownian Motion*. Springer Berlin Heidelberg, Berlin, Heidelberg, 1999.

66 STEVENS, A. The Derivation of Chemotaxis Equations as Limit Dynamics of Moderately Interacting Stochastic Many-Particle Systems. *SIAM J. Appl. Math. 61*, 1 (Jan 2000), 183–212.

67 SZNITMAN, A.-S. Topics in propagation of chaos. *Springer* (1991), 165–251.

68 TKACHENKO, N., WEISSMANN, J. D., PETERSEN, W. P., LAKE, G., ZOLLIKOFER, C. P. E., and CALLEGARI, S. Individual-based modelling of population growth and diffusion in discrete time. *PLoS One 12*, 4 (Apr 2017).

69 VELÁZQUEZ, J. Point Dynamics in a Singular Limit of the Keller-Segel Model 1: Motion of the Concentration Regions. *SIAM J. Appl. Math. 64*, 4 (Jan 2004), 1198–1223.

70 VELÁZQUEZ, J. J. Well-posedness of a model of point dynamics for a limit of the Keller–Segel system. *Journal of Differential Equations 206*, 2 (2004), 315–352.

71 VILLANI, C. *Topics in Optimal Transportation*. American Mathematical Soc., 2003.

72 WANG, G. and WEI, J. Steady State Solutions of a Reaction-Diffusion System Modeling Chemotaxis. *Math. Nachr. 233-234*, 1 (Jan 2002), 221–236.

73 WOLANSKY, G. On the evolution of self-interacting clusters and applications to semilinear equations with exponential nonlinearity. *J. Anal. Math. 59*, 1 (Dec 1992), 251–272.