OPEN ACCESS
Mathematical Models and Methods in Applied Sciences
Vol. 33, No. 5 (2023) 971–1008

DOI: 10.1142/S0218202523500215

Convergence of a particle method for a regularized spatially homogeneous Landau equation

José A. Carrillo*

Mathematical Institute, University of Oxford,
Andrew Wiles Building, Woodstock Road, Oxford OX2 6GG, UK
carrillo@maths.ox.ac.uk

Matias G. Delgadino

Department of Mathematics,
The University of Texas at Austin,
2515 Speedway, PMA 8.100 Austin, Texas 78712, USA
matias.delgadino@math.utexas.edu

Jeremy S. H. Wu

Mathematical Sciences Building,
University of California,
Los Angeles, 520 Portola Plaza
Los Angeles, California 90095, USA
jeremywu@math.ucla.edu

Received 13 November 2022
Accepted 25 January 2023
Published 28 March 2023
Communicated by N. Bellomo and F. Brezzi

We study a regularized version of the Landau equation, which was recently introduced in [J. A. Carrillo, J. Hu, L. Wang and J. Wu, A particle method for the homogeneous Landau equation, *J. Comput. Phys. X* **7** (2020) 100066, 24] to numerically approximate the Landau equation with good accuracy at reasonable computational cost. We develop the existence and uniqueness theory for weak solutions, and we reinforce the numerical findings in the above-mentioned paper by rigorously proving the validity of particle approximations to the regularized Landau equation.

Keywords: Kinetic equations; particle approximation; mean field limit.

AMS Subject Classification: 35Q92, 35Q49, 82C40

*Corresponding author.

1. Introduction

The Landau equation,[32] originally derived to approximate the Boltzmann operator when collisions between charged particles in a plasma are grazing, is one of the fundamental kinetic equations in plasma physics. Efficient computational methods for the full Vlasov–Maxwell–Landau system are of tremendous importance for modelling future fusion reactors and they represent a central conundrum in computational plasma physics. An important foundation to achieve such an ambitious goal is to provide accurate numerical methods with low computational cost to solve the collisional step in these computations, that is, to solve the spatially homogeneous Landau equation given by

$$\partial_t f = Q(f,f) := \nabla_v \cdot \left\{ \int_{\mathbb{R}^d} A(v-v_*)(f(v_*)\nabla_v f(v) - f(v)\nabla_{v_*} f(v_*))\mathrm{d}v_* \right\}, \tag{1.1}$$

with the collision kernel given by $A(z) = |z|^\gamma(|z|^2 I_d - z\otimes z) = |z|^{\gamma+2}\Pi(z)$ with I_d being the identity matrix, $\Pi(z)$ the projection matrix onto $\{z\}^\perp$, $-d-1 \le \gamma \le 1$, and $d \ge 2$. The most important case corresponds to $d = 3$ with $\gamma = -3$ associated with the physical interaction in plasmas. This case is usually called the Coulomb case because it can be derived from the Boltzmann equation in the grazing collision limit when particles interact via Coulomb forces.[11, 16, 39] The main formal properties of Q rely on the following reformulation:

$$Q(f,f) = \nabla_v \cdot \left\{ \int_{\mathbb{R}^d} A(v-v_*) f f_* (\nabla_v \log f - \nabla_{v_*}\log f_*)\mathrm{d}v_* \right\},$$

where $f = f(v)$, $f_* = f(v_*)$ are used; and its weak form acting on appropriate test functions $\phi = \phi(v)$

$$\begin{aligned}\int_{\mathbb{R}^d} Q(f,f)\phi\,\mathrm{d}v = -\frac{1}{2}\iint_{\mathbb{R}^{2d}} &(\nabla_v\phi - \nabla_{v_*}\phi_*)\cdot A(v-v_*) \\ &\times(\nabla_v \log f - \nabla_{v_*}\log f_*) f f_*\,\mathrm{d}v\,\mathrm{d}v_*. \end{aligned}\tag{1.2}$$

Then choosing $\phi(v) = 1, v, |v|^2$, one achieves conservation of mass, momentum and energy. Inserting $\phi(v) = \log f(v)$, one obtains the formal entropy decay with dissipation given by

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^d} f\log f\,\mathrm{d}v = -D(f(t,\cdot)) := -\frac{1}{2}\iint_{\mathbb{R}^{2d}} B_{v,v_*}\cdot A(v-v_*)B_{v,v_*} f f_*\,\mathrm{d}v\mathrm{d}v_* \le 0,$$

since A is symmetric and semipositive definite, with $B_{v,v_*} := \nabla_v \log f - \nabla_{v_*}\log f_*$. The equilibrium distributions are given by the Maxwellian

$$\mathscr{M}_{\rho,u,T} = \frac{\rho}{(2\pi T)^{d/2}}\exp\left(-\frac{|v-u|^2}{2T}\right),$$

for some constants ρ, T determining the density and the temperature of the particle ensemble, and mean velocity vector u, see Refs. 39 and 21.

Deterministic numerical methods based on particle approximations to (1.1) have been recently proposed in Ref. 13 keeping all the structural properties of the Landau equation described above: nonnegativity, conservation of mass, momentum and energy, and entropy dissipation at a semidiscrete level. This paper gives a theoretical underpinning to the numerical scheme introduced in Ref. 13. The main strategy is to delocalize the gradient operators in the weak form (1.2) while keeping intact the variational structure behind the equation rigorously developed in Ref. 10. This is reminiscent of similar approaches to approximate nonlinear diffusion models by nonlocal equations[9] while keeping their variational structure. More precisely, we analyse the Landau gradient flow of the regularized entropy,[10, 13] given by

$$\partial_t f = \nabla_v \cdot \left\{ f(v) \int_{\mathbb{R}^d} f(w) A(v-w)(\nabla G^\varepsilon * \log[f * G^\varepsilon](v) \right.$$
$$\left. - \nabla G^\varepsilon * \log[f * G^\varepsilon](w)) \mathrm{d}w \right\}, \tag{1.3}$$

where $G^\varepsilon \in C^\infty(\mathbb{R}^d)$ is a mollifier for fixed $\varepsilon > 0$. More specifically,

$$G^\varepsilon(v) = \frac{1}{\varepsilon^d} G\left(\frac{v}{\varepsilon}\right), \quad \int_{\mathbb{R}^d} G(v) dv = 1,$$

with $0 \le G \in C^\infty(\mathbb{R}^d)$, so that G^ε approximates the Dirac at the origin, δ_0, as $\varepsilon \downarrow 0$. Therefore, as $\varepsilon \downarrow 0$ (1.3) formally converges to the Landau equation. For technical reasons (cf. Lemma 2.4), we choose $G(v) = Ce^{-(1+|v|^2)^{1/2}}$ with $C > 0$ a normalization constant as in Ref. 10. However, we note that from the numerical point of view,[13] Gaussian mollifiers are simpler to deal with.

Our approach is to provide an existence theory for (1.3) as well as a particle approximation to the solution by interpreting (1.3) as a continuity equation with solution-dependent velocity fields. In particular, to introduce notation, we define a generalised interaction kernel for probability measures $g \in \mathscr{P}(\mathbb{R}^d)$ and $v, w \in \mathbb{R}^d$

$$K_g(v,w) := -|v-w|^{2+\gamma} \Pi[v-w](\nabla G^\varepsilon * \log[g * G^\varepsilon](v) - \nabla G^\varepsilon * \log[g * G^\varepsilon](w)).$$

Additionally, for $f \in \mathscr{P}(\mathbb{R}^d)$, we define the measure-dependent velocity

$$U^\varepsilon[g,f](v) := \int_{\mathbb{R}^d} K_g(v,w) \mathrm{d}f(w), \quad U^\varepsilon[f] := U^\varepsilon[f,f].$$

In this way, (1.3) can be written as

$$\partial_t f + \nabla \cdot (U^\varepsilon[f] f) = 0. \tag{1.4}$$

Formally speaking, by approximating an initial data by a finite number of atomic measures, we expect the solution of (1.4) to be approximated by a finite number of Dirac masses following the local velocity of particles. More precisely, suppose we are given initial data $f^0 \in \mathscr{P}(\mathbb{R}^d)$ for (1.3) and we can approximate f^0 by a sequence of empirical measures $\mu_0^N = \frac{1}{N}\sum_{i=1}^N \delta_{v_0^i}$, with equal weights for simplicity, where $v_0^i \in \mathbb{R}^d$ for $i = 1, \ldots, N$. We expect the solution to (1.4) to be given by the empirical measure with equal weights

$$\mu^N(t) = \frac{1}{N}\sum_{i=1}^{N} \delta_{v^i(t)}\,,$$

where $v^i(t)$ is the solution of the ODE system

$$\dot{v}^i(t) = U[\mu^N(t)](v^i(t)).$$

In fact, $\mu^N(t)$ is a distributional solution to (1.4) with μ_0^N as initial data. The results in Ref. 10 do not provide a well-posedness of measure solutions to (1.3) with measure initial data, ensuring only the existence by compactness. Due to the lack of continuous dependence with respect to initial data to (1.4) in the general probability measure setting, showing the convergence of the mean field limit is important from the numerical viewpoint.[13] More specifically, we show that μ^N converges towards the unique weak solution f of (1.4) in the limit $N \to \infty$.

In the simplified setting of equally weighted particles, the main result of this paper can be summarized as follows. Suppose that the initial data are well approximated in the sense that

$$W_\infty(\mu_0^N, f^0) \to 0 \quad \text{as } N \to \infty \text{ fast enough,}$$

where W_∞ denotes the ∞-Wasserstein metric,[35] see Hypothesis (**B1**) and (**B2**). As mentioned above, the evolution of μ_0^N through (1.3) is characterised entirely by the evolution of the "particles" starting at v_0^i according to the ODE system for $i = 1, \ldots, N$

$$\begin{aligned}\dot{v}^i &= -\frac{1}{N}\sum_{j=1}^{N} |v^i - v^j|^{2+\gamma}\Pi[v^i - v^j](\nabla G^\varepsilon * \log[\mu^N * G^\varepsilon](v^i) \\ &\quad - \nabla G^\varepsilon * \log[\mu^N * G^\varepsilon](v^j)).\end{aligned} \tag{1.5}$$

We will prove, at least for short times depending on the value of $\gamma \in (-3, 0]$, that $f = f(t) = f_t$ and $\mu^N = \mu^N(t) = \mu_t^N$ exist (cf. Theorem 1.1 and Lemma 1.1, respectively) and solve (1.3) according to the initial conditions f^0 and μ_0^N, respectively. Given the existence of such curves f, μ^N and the fact that $\mu_0^N \to f^0$ as $N \to \infty$, we seek to prove the mean field limit (cf. Theorem 1.2)

$$W_\infty(\mu^N(t), f(t)) \to 0, \quad \text{for } t \in (0, T_m) \quad \text{as } N \to \infty,$$

where $T_m > 0$ is the maximal existence time of f (cf. Theorem 1.1).

The mean-field limit has attracted lots of attention in the last years in different settings for aggregation-diffusion and Vlasov type kinetic equations. Different approaches have been taken leading to a very lively interaction between different communities of researchers in analysis and probability. We refer to Refs. 4, 17, 19, 33, 38 for the classical approaches in the field. Recent advances in non-Lipschitz settings and with applications to models with alignment have been done in Refs. 3, 8, 14, 15, for the aggregation-diffusion and Vlasov-type equations in Refs. 5–7, 12, 18, 20, 24–31, 34, 37, and for incompressible fluid problems in Refs. 23 and 22.

We prove the mean field limit to the regularized Landau equation (1.3) following the strategy and ideas from Refs. 23 and 7. The main difference with these references is the fact that Eq. (1.3) is more nonlocal, and it can be interpreted as a transport equation with a highly nonlocal nonlinear mobility depending quadratically on the density f. Let us finally mention that our result does not give quantitative bounds on the mean-field limit depending on N and ε compared to recent works.[6, 28–30] This is certainly an important open question of great importance from the numerical viewpoint.

1.1. *Main results*

The proof of the mean field limit convergence of (1.3) for fixed $\varepsilon > 0$ is achieved with the following strategy borrowed from Ref. 7: we first show the existence and uniqueness of the continuity equation (1.4) for some maximal time horizon $T_m > 0$ in Secs. 2 and 3, then we show that the particle system does also exist in Appendix C. We finally conclude by estimating the distance between the two systems in the W_∞ metric when $N \to \infty$ as well as establishing a lower bound on the existence of the particle system, Sec. 4. Let us point out that since the kernel A is singular or grows at infinity, these properties of the continuity equation and the associated particle system are not obvious.

Continuity equations of the form (1.4) have been extensively studied.[2, 19] We will see how the following assumptions on f^0 yield good properties for $U^\varepsilon[f]$.

(**A1**) The initial condition f^0 belongs to $\mathscr{P}_c(\mathbb{R}^d)$, the space of compactly supported probability measures on $\mathbb{R}^d$.

(**A2**) For $\gamma \in (-3,-2)$, there exists $p > 1$ such that $\frac{p}{p-1}(2+\gamma) > -d$ and f^0 belongs to $L^p(\mathbb{R}^d)$.

Theorem 1.1. (Existence of mean field limit) *Fix* $\varepsilon > 0$, $\gamma \in (-3, 0]$, *and initial data* $f^0 \in \mathscr{P}(\mathbb{R}^d)$ *satisfying* (**A1**) *and* (**A2**). *Then, there is a time horizon* $T = T(\gamma, \varepsilon, f^0) > 0$ *such that there is a unique weak solution* f *to* (1.3) *given in*

$$f \in \begin{cases} C([0,T]; \mathscr{P}_c(\mathbb{R}^d)), & \gamma \in [-2, 0], \\ C([0,T]; \mathscr{P}_c(\mathbb{R}^d)) \cap L^\infty(0,T; L^p(\mathbb{R}^d)), & \gamma \in (-3,-2), \end{cases}$$

where $f|_{t=0} = f^0$, *and the exponent* $p > 1$ *is the same as in* (**A2**).

In the case $\gamma \in [-2, 0]$, *the maximal time of existence* $T_M = +\infty$ *is infinite. While for the case* $\gamma \in (-3,-2)$, *either the maximal time of existence is infinite* $T_M = +\infty$, *or the* L^p *norm of the solution blows up*

$$\mathrm{esssup}_{s\in[0,t)} \|f(s)\|_{L^p} \uparrow +\infty \quad \text{as } t \uparrow T_M.$$

The notion of weak solution f to (1.4) (equivalently (1.3)) means that, for any $\phi \in C_c^\infty(\mathbb{R}^d)$, the following equality holds

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_{\mathbb{R}^d} \phi(v)\mathrm{d}f_t(v) = \int_{\mathbb{R}^d} \nabla\phi(v) \cdot U^\varepsilon[f(t)](v)\mathrm{d}f_t(v).$$

The particle solution (1.5) is in fact a solution to the previous equation when the initial condition is a convex combinations of delta measures (1.4). More precisely, for every $N \in \mathbb{N}$, take initial points $\{v_0^{i,N}\}_{i=1}^N \subset \mathbb{R}^d$ and positive weights $\{m_{i,N}\}_{i=1}^N$ satisfying

$$\sum_{i=1}^N m_{i,N} = 1, \quad m_{i,N} \geq 0, \quad \forall\, i = 1, \ldots, N.$$

The N-particle ODE system we consider is

$$\begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}t} v^i(t) = U^\varepsilon[\mu^N(t)](v^i) = \sum_{j=1}^N m_j K_{\mu^N(t)}(v^i, v^j), \\ &\mu^N(t) = \sum_{i=1}^N m_i \delta_{v^i(t)}, \\ &v^i\big|_{t=0} = v_0^i. \end{aligned} \tag{1.6}$$

Lemma 1.1. (Existence of particle solutions) *For any* $\varepsilon > 0$, *and* $\gamma \in (-3, 0]$, *there exists a time horizon* $T = T\big(\varepsilon, \gamma, \{v_0^{i,N}\}_{i=1}^N\big) > 0$ *and a curve* $v^{i,N} \in C^1([0,T]; \mathbb{R}^d)$ *which satisfies* (1.6). *For* $\gamma \in [-2, 0]$, *the solution to* (1.6) *is unique, and the time horizon* T *can be arbitrarily large.*

The well-posedness of the inter-particle system (1.6) is proven in Appendix C. The case $\gamma \in [-2, 0]$ is an application of Theorem 1.1, while for $\gamma \in (-3, -2)$ a standard Peano existence argument is used.

For $N \in \mathbb{N}$ and trajectories $\{v^{i,N}(t)\}_{i=1}^N$ such as those constructed in Lemma 1.1, we define the minimum inter-particle distance for times t in the domain of existence

$$\eta_m^N(t) := \min_{i \neq j} |v^{i,N}(t) - v^{j,N}(t)|.$$

Taking the continuum and particle solutions f and μ^N from Theorem 1.1 and Lemma 1.1 respectively, we define

$$\eta^N(t) := W_\infty(\mu^N(t), f(t)).$$

The following assumptions are well-preparedness conditions on the initial data of the particle solution, see Ref. 7.

(**B1**) The initial particles $\{v_0^{i,N}\}_{i=1}^N \subset \mathbb{R}^d$ and weights $\{m_{i,N}\}_{i=1}^N \subset (0,1)$ satisfy $W_\infty(\mu_0^N, f^0) \to 0$ as $N \to \infty$.

(**B2**) For $\gamma \in (-3, -2)$, the initial particles moreover satisfy

$$\lim_{N\to\infty} \eta^N(0)^{\frac{d}{p'}} \eta_m^N(0)^{1+\gamma} = 0, \tag{1.7}$$

where the conjugate exponent p' satisfies $\frac{1}{p} + \frac{1}{p'} = 1$ with p from (**A2**).

The main result concerning the mean field limit can now be stated as:

Theorem 1.2. *Fix $\varepsilon > 0$, $\gamma \in (-3, 0]$, and initial data f^0 satisfying* (**A1**) *and* (**A2**). *We consider f the solution to* (1.3) *on the maximal time interval $[0, T_m]$ provided by Theorem* 1.1. *Given initial particle configurations $\{\mu_0^N\}_{n\in\mathbb{N}}$ satisfying* (**B1**) *and* (**B2**), *we consider $\{\mu^N\}_{N\in\mathbb{N}}$ particle solutions of* (1.3) *with maximal time of existence $T^N > 0$ provided by Lemma* 1.1. *Then $\liminf_{N\to\infty} T^N \geq T_m$, and the mean field limit holds*

$$\lim_{N\to\infty} \sup_{t\in[0,T]} W_\infty(\mu^N(t), f(t)) \to 0, \quad \forall T \in [0, T_m). \tag{1.8}$$

2. Estimates on the Velocity

This section collects the necessary estimates on the measure-dependent kernel and velocity, K and U^ε. To fix notation, we define the Lebesgue bracket

$$\langle v\rangle^2 = 1 + |v|^2, \quad v \in \mathbb{R}^d,$$

and the mollifying sequence by

$$G(v) = Ce^{-\langle v\rangle}, \quad \int_{\mathbb{R}^d} G(v)\mathrm{d}v = 1, \quad G^\varepsilon(v) = \frac{1}{\varepsilon^d} G(v/\varepsilon).$$

Moreover, we define the pth-order moment of a measure f by

$$M_p(f) = \int_{\mathbb{R}^d} \langle v\rangle^p \,\mathrm{d}f(v).$$

We will use the notation $a \lesssim_{\alpha,\beta,\ldots} b$ to represent the statement that there is a constant $C = C(\alpha, \beta, \ldots) > 0$ such that $a \leq Cb$.

Proposition 2.1. *Fix $\varepsilon > 0$ and $\gamma = -2$. Then, for every $f, g \in \mathscr{P}(\mathbb{R}^d)$, the functions $K_g(v, w)$ and $U^\varepsilon[g, f](v)$ are C^1, skew-symmetric and satisfy the estimates*

$$|K_g(v, w)| \leq \frac{2}{\varepsilon}, \quad |\nabla_v K_g(v, w)| \leq \frac{28}{\varepsilon^2},$$
$$|U^\varepsilon[g, f](v)| \leq \frac{2}{\varepsilon}, \quad |\nabla U^\varepsilon[g, f](v)| \leq \frac{28}{\varepsilon^2}.$$

Proposition 2.1 highlights the C^1-boundedness of the velocity field U^ε in the special case $\gamma = -2$. For this value of γ, the well-posedness of (1.4) follows by standard techniques.[19] Proposition 2.1 follows from the more general results Proposition 2.2, Lemma 2.3, and Proposition 2.3 where $\gamma \in [-3, 0]$. There, we shall see the precise dependence on γ. First, we recall a standard inequality for the Lebesgue bracket.

Lemma 2.1. (Peetre) *For any $p \in \mathbb{R}$ and $x, y \in \mathbb{R}^d$, we have*

$$\frac{\langle x\rangle^p}{\langle y\rangle^p} \leq 2^{|p|/2}\langle x - y\rangle^{|p|}.$$

Proof. A proof of this can be found in Ref. 10 or Ref. 1. □

Lemma 2.2. (log-derivative estimates) *For fixed $\varepsilon > 0$ we have the formula*

$$\nabla G^\varepsilon(v) = -\frac{1}{\varepsilon}\left\langle\frac{v}{\varepsilon}\right\rangle^{-1} G^\varepsilon(v)\frac{v}{\varepsilon}. \tag{2.1}$$

For $\mu \in \mathscr{P}(\mathbb{R}^d)$, denoting $\partial^i = \frac{\partial}{\partial v^i}$ and $\partial^{ij} = \frac{\partial^2}{\partial v^i \partial v^j}$, we obtain

$$|\nabla \log(\mu * G^\varepsilon)(v)| \le \frac{1}{\varepsilon}, \quad |\partial^{ij}\log(\mu * G^\varepsilon)(v)| \le \frac{4}{\varepsilon^2}. \tag{2.2}$$

Proof. This is proven in Ref. 10. □

Proposition 2.2. *Fix $\varepsilon > 0$, $g \in \mathscr{P}(\mathbb{R}^d)$, and $\gamma \in [-3, 0]$. We have the following estimate*:

$$|K_g(v,w)| \le \min\left(\frac{4}{\varepsilon^2}|v-w|^{3+\gamma}, \frac{2}{\varepsilon}|v-w|^{2+\gamma}\right).$$

Moreover, for fixed $f \in \mathscr{P}(\mathbb{R}^d)$, we have

$$|U^\varepsilon[g,f](v)| \lesssim_\varepsilon \begin{cases} M_{2+\gamma}(f)\langle v\rangle^{2+\gamma}, & \gamma \in (-2,0], \\ 1, & \gamma \in [-3,-2]. \end{cases}$$

Proof. We recall the expression

$$K_g(v,w) = -|v-w|^{2+\gamma}\Pi[v-w](\nabla G^\varepsilon * \log[g * G^\varepsilon](v) - \nabla G^\varepsilon * \log[g * G^\varepsilon](w)).$$

$|v-w| < 1$: Using the second-order estimate in (2.2), the difference of logarithms can be estimated by

$$|\nabla G^\varepsilon * \log[g * G^\varepsilon](v) - \nabla G^\varepsilon * \log[g * G^\varepsilon](w)| \le \frac{4}{\varepsilon^2}|v-w|,$$

giving the first estimate in the minimum.

$|v-w| \ge 1$: Bluntly apply the first-order estimate in (2.2) onto each of the logarithms

$$\|\nabla G^\varepsilon * \log[g * G^\varepsilon]\|_{L^\infty} \le \frac{1}{\varepsilon}.$$

The estimate for U^ε follows by recalling $U^\varepsilon[g,f](v) = \int_{\mathbb{R}^d} K_g(v,w)\mathrm{d}f(w)$ and Peetre's inequality from Lemma 2.1 for the case $\gamma \in (-2, 0]$. □

The following estimate is adapted from Ref. 23.

Lemma 2.3. (Pointwise difference in K) *Fix $\varepsilon > 0$, $g \in \mathscr{P}(\mathbb{R}^d)$, and $\gamma \in [-3, 0]$. We have*

$$|K_g(v_1,w) - K_g(v_2,w)| \lesssim_{\varepsilon,\gamma} |v_1 - v_2|\max(|v_1-w|^{2+\gamma}, |v_2-w|^{2+\gamma}).$$

For completeness, we refer to Appendix A for the proof of Lemma 2.3.

Proposition 2.3. (Hölder continuity of U) *Fix $\gamma \in [-3,-2]$ and $g \in \mathscr{P}(\mathbb{R}^d)$. Then we have*

$$|U^\varepsilon[g](v_1) - U^\varepsilon[g](v_2)| \lesssim_\varepsilon |v_1 - v_2|^{3+\gamma}.$$

Proof. We split the integration region into two cases

$$\begin{aligned} U^\varepsilon[g](v_1) - U^\varepsilon[g](v_2) &= \int_{\mathbb{R}^d} (K_g(v_1,w) - K_g(v_2,w))\mathrm{d}g(w) \\ &= \left(\int_{|v_1-v_2|\le \min(|v_1-w|,\,|v_2-w|)} + \int_{|v_1-v_2|>\min(|v_1-w|,\,|v_2-w|)} \right) \\ &\quad \times (K_g(v_1,w) - K_g(v_2,w))\mathrm{d}g(w) \\ &=: I_1 + I_2. \end{aligned}$$

We claim both $|I_1|$, $|I_2| \lesssim_\varepsilon |v_1-v_2|^{3+\gamma}$. Starting with I_1 where $|v_1-v_2| \le \min(|v_1-w|,\, |v_2-w|)$, we use Lemma 2.3 and the fact that $2+\gamma \le 0$ to deduce

$$\begin{aligned} |I_1| &\lesssim_\varepsilon |v_1-v_2| \int_{|v_1-v_2|\le\min(|v_1-w|,\,|v_2-w|)} \max(|v_1-w|^{2+\gamma},\, |v_2-w|^{2+\gamma})\mathrm{d}g(w) \\ &\le |v_1-v_2|^{3+\gamma}. \end{aligned}$$

Turning to I_2 corresponding to the other integration region, assume without loss of generality that $|v_1-v_2| > |v_1-w|$. By the triangle inequality we also have

$$|v_2 - w| \le 2|v_1-v_2|.$$

Putting these two estimates together and using the bound $|K_g(v,w)| \lesssim_\varepsilon \min(|v-w|^{3+\gamma},\, |v-w|^{2+\gamma})$ from Proposition 2.2, we have

$$\begin{aligned} |I_2| &\lesssim_\varepsilon \int_{|v_1-v_2|>\min(|v_1-w|,\,|v_2-w|)} \min(|v_1-w|^{3+\gamma}, |v_1-w|^{2+\gamma})\mathrm{d}g(w) + \cdots \\ &\quad + \int_{|v_1-v_2|>\min(|v_1-w|,\,|v_2-w|)} \min(|v_2-w|^{3+\gamma}, |v_2-w|^{2+\gamma})\mathrm{d}g(w) \\ &\lesssim |v_1-v_2|^{3+\gamma}. \end{aligned}$$

□

We can improve Proposition 2.3 to Lipschitz continuity by taking advantage of extra regularity properties of g.

Proposition 2.4. (Lipschitz continuity of U) *Fix $g \in \mathscr{P}(\mathbb{R}^d)$, $\varepsilon > 0$, and $\gamma \in (-3, 0]$. In the case $\gamma \in (-3,-2)$, assume further that g satisfies* (**A2**). *Then we have*

$$|U^\varepsilon[g](v^1) - U^\varepsilon[g](v^2)| \le \Lambda_\gamma(g, v^1, v^2)|v^1 - v^2|,$$

where

$$\Lambda_\gamma = \begin{cases} C_\varepsilon M_{2+\gamma}(g)(\langle v^1\rangle^{2+\gamma} + \langle v^2\rangle^{2+\gamma}), & \gamma \in [-2, 0], \\ C_{\varepsilon,\gamma,p',d}(1 + \|g\|_{L^p}), & \gamma \in (-3, -2) \end{cases}$$

and the constants $C > 0$ only depend on the quantities in the subscript.

Proof. The starting point is the application of Lemma 2.3 to first write

$$\begin{aligned} |U^\varepsilon[g](v^1) - U^\varepsilon[g](v^2)| &= \left| \int_{\mathbb{R}^d} (K_g(v^1, w) - K_g(v^2, w)) \mathrm{d}g(w) \right| \\ &\lesssim_\varepsilon |v^1 - v^2| \int_{\mathbb{R}^d} \max(|v^1 - w|^{2+\gamma}, |v^2 - w|^{2+\gamma}) \mathrm{d}g(w) \\ &\le |v^1 - v^2| \underbrace{\int_{\mathbb{R}^d} (|v^1 - w|^{2+\gamma} + |v^2 - w|^{2+\gamma}) \mathrm{d}g(w)}_{=:I}. \end{aligned}$$

We estimate I depending on the value of γ.

The case $\gamma \in (-3, -2)$: By splitting the integration region and using the fact that $g \in L^p$, we obtain

$$\begin{aligned} I &\lesssim 1 + \sup_{v \in \mathbb{R}^d} \int_{|v-w|\le 1} |v - w|^{2+\gamma} \, \mathrm{d}g(w) \\ &\le 1 + \sup_{v \in \mathbb{R}^d} \left(\int_{|v-w|\le 1} |v - w|^{(2+\gamma)p'} \, \mathrm{d}w \right)^{\frac{1}{p'}} \|g\|_{L^p} \\ &\lesssim_{\gamma,p,d} 1 + \|g\|_{L^p}. \end{aligned}$$

The case $\gamma \in [-2, 0]$: The integrand is no longer singular so we use Peetre's inequality from Lemma 2.1

$$|v - w|^{2+\gamma} \le \langle v - w\rangle^{2+\gamma} \lesssim_\gamma \langle v\rangle^{2+\gamma} \langle w\rangle^{2+\gamma}$$

for $v = v^1$, v^2 into I to get

$$I \lesssim \int_{\mathbb{R}^d} (\langle v^1\rangle^{2+\gamma} + \langle v^2\rangle^{2+\gamma}) \langle w\rangle^{2+\gamma} \, \mathrm{d}g(w). \qquad \square$$

2.1. *The velocity field as a function of measures*

The previous results established estimates for the pointwise variation of K_g and $U^\varepsilon[g]$ given a fixed measure g. We now investigate the measure-wise variation of K and U^ε given a fixed point.

Lemma 2.4. *Fix $\varepsilon > 0$ and let τ be an optimal transport map in W_∞ between $f, g \in \mathscr{P}_c(\mathbb{R}^d)$ so that $g = \tau\#f$. Then, we have*

$$|\log[g * G^\varepsilon](v) - \log[f * G^\varepsilon](v)| \le \frac{1}{\varepsilon} W_\infty(g, f), \quad \forall\, v \in \mathbb{R}^d.$$

Proof. We first note that

$$g * G^\varepsilon(v) = \int_{\mathbb{R}^d} G^\varepsilon(v-w)\mathrm{d}g(w) = \int_{\mathbb{R}^d} G^\varepsilon(v-\tau(w))\mathrm{d}f(w).$$

Using the fundamental theorem of calculus, we express the difference of the logarithms as

$$\begin{aligned}
&\log[g * G^\varepsilon](v) - \log[f * G^\varepsilon](v)\\
&\quad = \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t}\log\left[\int_{\mathbb{R}^d} G^\varepsilon(v-(t\tau(w)+(1-t)w))\mathrm{d}f(w)\right]\mathrm{d}t\\
&\quad = -\int_0^1 \frac{\int_{\mathbb{R}^d}(\tau(w)-w)\cdot\nabla G^\varepsilon(v-(t\tau(w)+(1-t)w))\mathrm{d}f(w)}{\int_{\mathbb{R}^d} G^\varepsilon(v-(t\tau(w)+(1-t)w)\mathrm{d}f(w)}\mathrm{d}t.
\end{aligned}$$

By definition, we have $|\tau(w)-w| \le W_\infty(g,f)$ and moreover recalling (2.1), we see that

$$|\nabla G^\varepsilon| \le \frac{1}{\varepsilon}G^\varepsilon.$$

Applying these two estimates into the previous computations, we have

$$|\log[g * G^\varepsilon](v) - \log[f * G^\varepsilon](v)| \le \frac{W_\infty(g,f)}{\varepsilon}. \qquad \square$$

The following technical estimates hinge on Lemma 2.4.

Lemma 2.5. (Measure-wise difference in K) *Fix $g^i \in \mathscr{P}_c(\mathbb{R}^d)$ for $i = 1, 2$ and $\gamma \in [-3,0]$. Then for every v, $w \in \mathbb{R}^d$, we have the estimate*

$$|K_{g^1}(v,w) - K_{g^2}(v,w)| \lesssim \min\left(\frac{1}{\varepsilon^3}|v-w|^{3+\gamma},\ \frac{2}{\varepsilon^2}|v-w|^{2+\gamma}\right)W_\infty(g^1,\, g^2).$$

Proof. Here, we need to estimate

$$\begin{aligned}
&K_{g^1}(v,w) - K_{g^2}(v,w)\\
&\quad = -|v-w|^{2+\gamma}\Pi[v-w](\nabla G^\varepsilon * \{\log[g^1 * G^\varepsilon] - \log[g^2 * G^\varepsilon]\}(v)\\
&\qquad - \nabla G^\varepsilon * \{\log[g^1 * G^\varepsilon] - \log[g^2 * G^\varepsilon]\}(w)).
\end{aligned}$$

By the fundamental theorem of calculus, we have an estimate for this difference

$$\begin{aligned}
&|v-w|^{2+\gamma}\left|\int_0^1 \frac{\mathrm{d}}{\mathrm{d}t}\nabla G^\varepsilon * \{\log[g^1 * G^\varepsilon] - \log[g^2 * G^\varepsilon]\}(tv+(1-t)w)\mathrm{d}t\right|\\
&\quad \le |v-w|^{3+\gamma}\left\|\nabla^2 G^\varepsilon * \{\log[g^1 * G^\varepsilon] - \log[g^2 * G^\varepsilon]\}\right\|_{L^\infty}.
\end{aligned}$$

Using Lemma 2.4 and the comparison $|\nabla^2 G^\varepsilon| \lesssim \frac{1}{\varepsilon^2}G^\varepsilon$ (an extension of (2.1)), we apply Young's convolution inequality to deduce

$$|K_{g^1}(v,w) - K_{g^2}(v,w)| \lesssim \frac{1}{\varepsilon^3}|v-w|^{3+\gamma}W_\infty(g^1,g^2).$$

On the other hand, without estimating second-order derivatives, we can bluntly prove

$$\begin{aligned}|K_{g^1}(v,w) - K_{g^2}(v,w)| &\le |v-w|^{2+\gamma}(|\nabla G^\varepsilon * \{\log[g^1 * G^\varepsilon] - \log[g^2 * G^\varepsilon]\}(v)| \\ &\quad + |\nabla G^\varepsilon * \{\log[g^1 * G^\varepsilon] - \log[g^2 * G^\varepsilon]\}(w)|) \\ &\le 2|v-w|^{2+\gamma}\|\nabla G^\varepsilon * \{\log[g^1 * G^\varepsilon] - \log[g^2 * G^\varepsilon]\}\|_{L^\infty}.\end{aligned}$$

Recalling (2.1) and Lemma 2.4 which say

$$|\nabla G^\varepsilon| \le \frac{1}{\varepsilon} G^\varepsilon, \quad |\log[g^1 * G^\varepsilon] - \log[g^2 * G^\varepsilon]| \le \frac{1}{\varepsilon} W_\infty(g^1, g^2),$$

we use Young's convolution inequality again to get

$$|K_{g^1}(v,w) - K_{g^2}(v,w)| \le \frac{2}{\varepsilon^2}|v-w|^{2+\gamma} W_\infty(g^1, g^2). \qquad \square$$

Under minimal assumptions on the probability measures, we can obtain a Hölder estimate with respect to the W_∞ metric.

Lemma 2.6. *Fix g^i, $f^i \in \mathscr{P}(\mathbb{R}^d)$ for $i = 1, 2$ and $\gamma \in (-3, -2]$. Then, we have the estimate*

$$|U^\varepsilon[g^1, f^1](v) - U^\varepsilon[g^2, f^2](v)| \lesssim W_\infty(g^1, g^2) + W_\infty(f^1, f^2)^{3+\gamma}.$$

Proof. Starting from the definition, we have

$$\begin{aligned}U^\varepsilon[g^1,\, f^1](v) - U^\varepsilon[g^2,\, f^2](v) &= \int_{\mathbb{R}^d} K_{g^1}(v,w)\mathrm{d}f^1(w) - \int_{\mathbb{R}^d} K_{g^2}(v,w)\mathrm{d}f^2(w) \\ &= \int_{\mathbb{R}^d} (K_{g^1}(v,w) - K_{g^2}(v,w))\mathrm{d}f^1(w) \\ &\quad + \int_{\mathbb{R}^d} K_{g^2}(v,w)\mathrm{d}(f^1 - f^2)(w) \\ &=: I_1 + I_2.\end{aligned}$$

We claim the following estimates

$$|I_1| \lesssim_\varepsilon W_\infty(g^1, g^2), \quad |I_2| \lesssim_\varepsilon W_\infty(f^1, f^2)^{3+\gamma}.$$

The term I_1 is almost completely treated by Lemma 2.5. We can further estimate the minimum by

$$\begin{aligned}|I_1| &\lesssim_\varepsilon W_\infty(g^1, g^2)\int_{\mathbb{R}^d} \min(|v-w|^{3+\gamma}, |v-w|^{2+\gamma})\mathrm{d}f^1(w) \\ &\lesssim W_\infty(g^1, g^2)\int_{\mathbb{R}^d} \langle v-w\rangle^{2+\gamma}\, \mathrm{d}f^1(w).\end{aligned}$$

When $\gamma \in (-3, -2]$, simply estimate $\langle v-w\rangle^{2+\gamma} \le 1$. This takes care of I_1 so we focus on I_2 for the rest of this proof.

Firstly, take τ an optimal transport map in W_∞ between f^1 and f^2 i.e.

$$W_\infty(f^1, f^2) = \mathrm{esssup}_{w\in\mathbb{R}^d}|\tau(w) - w|.$$

Moreover, the following identity holds in $\mathscr{P}(\mathbb{R}^d)$; $f^2 = \tau\# f^1$. This allows us to rewrite the difference I_2 as

$$I_2 = \int_{\mathbb{R}^d}[K_{g^2}(v,w) - K_{g^2}(v,\tau(w))]\mathrm{d}f^1(w). \tag{2.3}$$

We split the integration region in (2.3) into $\mathscr{A} = \{w \in \mathbb{R}^d \,|\, |v-w| < 2W_\infty(f^1, f^2)\}$ and its complement $\mathbb{R}^d\backslash\mathscr{A}$. In the set $\mathscr{A}$, we begin with the blunt L^∞ bound on K from Proposition 2.2 which implies

$$\begin{aligned}|K_{g^2}(v,w) - K_{g^2}(v,\tau(w))| &\lesssim_\varepsilon \min(|v-w|^{3+\gamma}, |v-w|^{2+\gamma}) \\ &\quad + \min(|v-\tau(w)|^{3+\gamma}, |v-\tau(w)|^{2+\gamma}) \\ &\lesssim W_\infty(f^1,f^2)^{3+\gamma} + |v-\tau(w)|^{3+\gamma}.\end{aligned}$$

For the second term, we simply use the triangle inequality

$$|v-\tau(w)| \le |v-w| + |\tau(w) - w| \le 3W_\infty(f^1, f^2).$$

This gives

$$|K_{g^2}(v,w) - K_{g^2}(v,\tau(w))| \lesssim_\varepsilon W_\infty(f^1,f^2)^{3+\gamma} \tag{2.4}$$

which is independent of w so we have

$$\int_{\mathscr{A}} |K_{g^2}(v,w) - K_{g^2}(v,\tau(w))|\,\mathrm{d}f^1(w) \lesssim_\varepsilon W_\infty(f^1,f^2)^{3+\gamma}.$$

Turning to the complement region $\mathbb{R}^d\backslash\mathscr{A}$ given as $\{w \in \mathbb{R}^d \,|\, |v-w| \ge 2W_\infty(f^1, f^2)\}$, we use Lemma 2.3 to obtain

$$|K_{g^2}(v,w) - K_{g^2}(v,\tau(w))| \lesssim_\varepsilon |\tau(w) - w| \max(|v-w|^{2+\gamma},\ |v-\tau(w)|^{2+\gamma}).$$

Recalling that $2+\gamma \le 0$, the reverse triangle inequality yields

$$|v-\tau(w)| \ge |v-w| - |\tau(w)-w| \ge W_\infty(f^1, f^2),$$

because $|v-w| \ge 2W_\infty(f^1, f^2)$.

Therefore, from the previous estimate, we obtain

$$|K_{g^2}(v,w) - K_{g^2}(v,\tau(w))| \lesssim_\varepsilon |\tau(w) - w|W_\infty(f^1,f^2)^{2+\gamma} \le W_\infty(f^1,f^2)^{3+\gamma}.$$

This is exactly the same as (2.4) for $\mathscr{A}$. Integrating both inequalities against f^1 yields

$$\begin{aligned}|I_2| \ &\le \ \left(\int_{\mathscr{A}} + \int_{\mathbb{R}^d\backslash\mathscr{A}}\right)|K_{g^2}(v,w) - K_{g^2}(v,\tau(w))|\mathrm{d}f^1(w) \\ &\lesssim_\varepsilon W_\infty(f^1,f^2)^{3+\gamma}.\end{aligned}$$

□

If we impose more assumptions on the probability measures, in particular (**A2**), we can derive linear stability with respect to the W_∞ metric.

Proposition 2.5. (Linear stability) *Fix $g^i, f^i \in \mathscr{P}_c(\mathbb{R}^d)$ for $i = 1, 2$ and $\gamma \in (-3, 0]$. For $\gamma \in (-3, -2)$, assume further that f^i satisfies* (**A2**). *Then we have the estimate*

$$
\begin{aligned}
&|U^\varepsilon[g^1, f^1](v) - U^\varepsilon[g^2, f^2](v)| \\
&\quad \lesssim_\varepsilon \begin{cases} \langle v\rangle^{2+\gamma} [M_{2+\gamma}(f^1) W_\infty(g^1, g^2) \\ \qquad + \{M_{2+\gamma}(f^1) + M_{2+\gamma}(f^2)\} W_\infty(f^1, f^2)], & \gamma \in [-2, 0], \\ W_\infty(g^1, g^2) + (1 + \|f^1\|_{L^p} + \|f^2\|_{L^p}) W_\infty(f^1, f^2), & \gamma \in (-3, -2). \end{cases}
\end{aligned}
$$

Proof. Our starting point repeats the proof of Lemma 2.6 above. Using the same notation from there, we split

$$
\begin{aligned}
U^\varepsilon[g^1, f^1](v) - U^\varepsilon[g^2, f^2](v) &= \int_{\mathbb{R}^d} (K_{g^1}(v, w) - K_{g^2}(v, w)) \mathrm{d}f^1(w) \\
&\quad + \int_{\mathbb{R}^d} K_{g^2}(v, w) \mathrm{d}(f^1 - f^2)(w) \\
&=: I_1 + I_2.
\end{aligned}
$$

We inherit the estimate for I_1 from the proof of Lemma 2.6 which reads, using Peetre's inequality in Lemma 2.1 for $\gamma \in [-2, 0]$,

$$
|I_1| \lesssim_\varepsilon W_\infty(g^1, g^2) \times \begin{cases} 1, & \gamma \in (-3, -2), \\ M_{2+\gamma}(f^1) \langle v\rangle^{2+\gamma}, & \gamma \in [-2, 0]. \end{cases}
$$

We focus entirely on I_2; in the case $\gamma \in (-3, -2)$, we claim that

$$
|I_2| \lesssim_\varepsilon (1 + \|f^1\|_{L^p} + \|f^2\|_{L^p}) W_\infty(f^1, f^2).
$$

In the case $\gamma \in [-2, 0]$, we claim that

$$
|I_2| \lesssim_\varepsilon \langle v\rangle^{2+\gamma} (M_{2+\gamma}(f^1) + M_{2+\gamma}(f^2)) W_\infty(f^1, f^2).
$$

In both cases, we rewrite I_2 in the following way; take τ an optimal transport map between f^1 and f^2 in W_∞ so that we have

$$
I_2 = \int_{\mathbb{R}^d} (K_{g^2}(v, w) - K_{g^2}(v, \tau(w))) \mathrm{d}f^1(w), \quad f^2 = \tau \# f^1.
$$

Applying Lemma 2.3 and recalling the (anti-)symmetry of $K_{g^2}(v, w) = -K_{g^2}(w, v)$, we have

$$
\begin{aligned}
|I_2| &\lesssim_\varepsilon \int_{\mathbb{R}^d} |w - \tau(w)| \max(|v - w|^{2+\gamma}, |v - \tau(w)|^{2+\gamma}) \mathrm{d}f^1(w) \\
&\leq W_\infty(f^1, f^2) \left(\int_{\mathbb{R}^d} (|v - w|^{2+\gamma} + |v - \tau(w)|^{2+\gamma}) \mathrm{d}f^1(w) \right).
\end{aligned}
$$

We split the sum and reformulate the second term in terms of f^2 using τ to obtain

$$|I_2| \lesssim_\varepsilon W_\infty(f^1,\, f^2)\left(\int_{\mathbb{R}^d} |v-w|^{2+\gamma}\,\mathrm{d}f^1(w) + \int_{\mathbb{R}^d} |v-w|^{2+\gamma}\,\mathrm{d}f^2(w)\right). \tag{2.5}$$

The case $\gamma \in (-3,-2)$: By partitioning $\mathbb{R}^d$ into $\{w \in \mathbb{R}^d \,|\, |v-w| \le 1\}$ and its complement, a standard application of Hölder's inequality gives

$$\int_{\mathbb{R}^d} |v-w|^{2+\gamma}\,\mathrm{d}f^1(w) \le 1 + \left(\int_{|v-w|\le 1} |v-w|^{(2+\gamma)p'}\right)^{\frac{1}{p'}} \|f^1\|_{L^p}$$

and similarly for f^2. Using the assumption $(2+\gamma)p' > -d$, inequality (2.5) is further refined to

$$|I_2| \lesssim_{\varepsilon,\, d,\, \gamma,\, p} (1 + \|f^1\|_{L^p} + \|f^2\|_{L^p}) W_\infty(f^1,\, f^2).$$

The case $\gamma \in [-2, 0]$: Since $2+\gamma \ge 0$, we use Peetre's inequality Lemma 2.1 to estimate

$$|v-w|^{2+\gamma} \le \langle v-w\rangle^{2+\gamma} \lesssim \langle v\rangle^{2+\gamma}\langle w\rangle^{2+\gamma}.$$

Inserting this into (2.5), we get

$$|I_2| \lesssim_\varepsilon \langle v\rangle^{2+\gamma} W_\infty(f^1,\, f^2)\left(\int_{\mathbb{R}^d} \langle w\rangle^{2+\gamma}\,\mathrm{d}(f^1+f^2)(w)\right). \qquad \square$$

3. The Continuum Model

This section is devoted to the proof of Theorem 1.1; the well-posedness of (1.3). To fix notation, we seek solutions in the following spaces:

$$X_\gamma = X_\gamma(T) := \begin{cases} C([0,T]; \mathscr{P}_c(\mathbb{R}^d)), & \gamma \in [-2,0], \\ C([0,T]; \mathscr{P}_c(\mathbb{R}^d)) \cap L^\infty(0,T; L^p(\mathbb{R}^d)), & \gamma \in (-3,-2). \end{cases}$$

For $\gamma \in (-3,-2)$, the exponent p corresponds to that of (**A2**). In particular, we endow X_γ with the metric

$$d_\infty(f^1,\, f^2) := \sup_{t\in[0,T]} W_\infty(f^1(t), f^2(t)), \quad \forall\, f^1,\, f^2 \in X_\gamma.$$

Given $g \in X_\gamma$, we first want to find $f \in X_\gamma$ solving

$$\partial_t f + \nabla\cdot(fU^\varepsilon[g]) = 0, \quad f(t=0) = f^0. \tag{P}$$

Well-posedness of (1.3) then comes from ensuring the map $g \mapsto f$ just described has a unique fixed point in a closed subspace of X_γ.

Given a curve $g \in X_\gamma$, we denote by Φ_g the characteristic flow corresponding to (P) satisfying

$$\frac{\mathrm{d}}{\mathrm{d}t}\Phi_g(t,v) = U^\varepsilon[g](\Phi_g(t,v)), \quad \Phi_g(0,v) = v \in \mathbb{R}^d. \tag{3.1}$$

Proposition 3.1. *Fix $\varepsilon > 0, \gamma \in (-3,0]$, $g \in X_\gamma$, and initial condition f^0 satisfying* (**A1**) *and* (**A2**)*. Then, $f(t) = \Phi_g(t,\cdot)\#f^0$ is the unique weak solution in $C([0,T];\mathscr{P}(\mathbb{R}^d))$ to* (P)*.*

Proof. The (local) Lipschitz continuity of $U^\varepsilon[g]$ is provided by Proposition 2.4 so the characteristic system (3.1) has a unique solution, Φ_g, up to the flow map's maximal time of existence $T^* > 0$. By Ref. 19 $f(t) = \Phi_g(t,\cdot)\#f^0$ is the unique weak solution in $C([0,T^*];\mathscr{P}(\mathbb{R}^d))$. For $\gamma \in (-3,-2)$, Proposition 2.4, implies global Lipschitz regularity of U^ε hence Φ_g is globally defined and we can directly take $T^* = T$. For $\gamma \in [-2,0]$, Lemma 3.1 excludes blow up of Φ_g so we can take $T^* = T$ here too. □

Lemma 3.1. *For $\varepsilon > 0$ and $g \in X_\gamma$, let Φ_g be the flow map of* (3.1) *with maximal time of existence $T^* > 0$. Then, we have the estimates*

$$\langle \Phi_g \rangle \leq \begin{cases} \langle v \rangle \exp\left\{C_\varepsilon\left[\sup_{s\in[0,T]} M_2(g(s))\right]t\right\}, & \gamma \in (-2,0], \\ \langle v \rangle + C_\varepsilon t, & \gamma \in (-3,-2], \end{cases} \qquad \forall\, t \in [0,T^*].$$

Here, $C_\varepsilon > 0$ is a constant depending only on $\varepsilon > 0$. In particular, Φ_g extends to a global solution of (3.1) *on $[0,T]$.*

Proof. We begin, for $\gamma \in (-2,0]$, by differentiating $\frac{1}{2}|\Phi_g(t,v)|^2$ with respect to $t \in (0,T^*)$. Expanding the definition of U^ε, we obtain

$$\begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{2}|\Phi_g(t,v)|^2 \\ &\quad = \Phi_g(t,v)\cdot U^\varepsilon[g](\Phi_g(t,v)) \\ &\quad = \Phi_g(t,v)\cdot \int_{\mathbb{R}^d} \underbrace{|\Phi_g(t,v)-w|^{2+\gamma}\Pi[\Phi_g(t,v)-w]B^\varepsilon(\Phi_g(t,v),w)}_{=:I}\,\mathrm{d}g(w), \end{aligned} \tag{3.2}$$

where we have abbreviated

$$B^\varepsilon(v,w) := \nabla G^\varepsilon * \log[g*G^\varepsilon](v) - \nabla G^\varepsilon * \log[g*G^\varepsilon](w).$$

Notice that $|B^\varepsilon| \lesssim_\varepsilon 1$ by (2.2). Our goal is to show $\int I\,\mathrm{d}g(w) \lesssim |\Phi_g|$ and then apply Grönwall's inequality. First, we split the integral into regions where $|\Phi_g - w| \leq 1$ and $|\Phi_g - w| > 1$. As $2+\gamma > 0$, the former piece can be easily estimated as follows:

$$\begin{aligned} \Phi_g \cdot \int_{\mathbb{R}^d} I\,\mathrm{d}g(w) &= \Phi_g \cdot \left(\int_{|\Phi_g - w|\leq 1} I\,\mathrm{d}g(w) + \int_{|\Phi_g-w|>1} I\,\mathrm{d}g(w)\right) \\ &\lesssim_\varepsilon |\Phi_g| + \underbrace{\int_{|\Phi_g-w|>1} |\Phi_g - w|^{2+\gamma}|\Pi[\Phi_g - w]\Phi_g|\,\mathrm{d}g(w)}_{=:II}. \end{aligned} \tag{3.3}$$

Turning to II, we use the fact that $\Pi[v-w]v = \Pi[v-w]w$ and $|\Phi_g - w|^2 \leq 2|\Phi_g|^2 + 2|w|^2$ to estimate

$$II \lesssim |\Phi_g|^2 \int_{|\Phi_g - w|>1} |\Phi_g - w|^\gamma |\Pi[\Phi_g - w]w| \,\mathrm{d}g(w)$$
$$+ \int_{|\Phi_g - w|>1} |w|^2 |\Phi_g - w|^\gamma |\Pi[\Phi_g - w]\Phi_g| \,\mathrm{d}g(w).$$

Finally, since $|\Phi_g - w| > 1$ and $\gamma \leq 0$, we can bluntly estimate the remaining contributions by

$$II \leq M_1(g)|\Phi_g|^2 + M_2(g)|\Phi_g|.$$

Putting this together with (3.3) and (3.2), we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{2}\langle \Phi_g(t,v)\rangle^2 \lesssim_\varepsilon M_2(g)\langle \Phi_g(t,v)\rangle^2 .$$

Grönwall's inequality gives the inequality for $\gamma \in (-2, 0]$.

Turning to the case $\gamma \in (-3, -2]$, we write the integral form of (3.1)

$$|\Phi_g(t,v) - v| = \left| \int_0^t U^\varepsilon[g](\Phi_g(s,v))\mathrm{d}s \right| \lesssim_\varepsilon t.$$

The final estimate comes from applying Proposition 2.2. □

Proposition 3.2. *The space (X_γ, d_∞) for $\gamma \in (-3, 0]$ is a complete metric space.*

Proof. This can be proven from the fact that $(\mathscr{P}_c(\mathbb{R}^d), W_\infty)$ is complete and metrises weak convergence.[35] Moreover, the L^p norm is lower semi-continuous with respect to this topology. □

3.1. *Moment and L^p propagation*

In this section, we derive the moment and L^p propagation estimates we will need for the fixed point argument to prove Theorem 1.1.

Proposition 3.3. *Fix $\varepsilon > 0, \gamma \in (-3, 0]$, and initial condition f^0 satisfying* (**A1**) *and* (**A2**) *and $g \in X_\gamma$. The unique weak solution $f(t) = \Phi_g(t, \cdot)\#f^0$ to* (P) *belongs to $C([0,T]; \mathscr{P}_c(\mathbb{R}^d))$ with second moment growth estimate*

$$M_2(f(t)) \leq M_2(f^0)\exp\left\{C_\varepsilon \left(\sup_{s\in[0,T]} M_2(g(s))\right) t\right\}, \quad \forall\, t \in [0,T].$$

If, moreover, $f = g$ (i.e. f solves (1.3)), then $M_2(f(t)) = M_2(f^0)$ for all $t \in [0,T]$.

Remark 3.1. The same propagation result applies for higher order moments. The constant C_ε grows linearly with the order of the moment.

The following L^p estimate can be derived directly from standard facts about solutions to the continuity equation which can be found, for example, in Refs. 19, 2. For completeness, we prove Lemma 3.2 in Appendix B.

Lemma 3.2. *For fixed $\varepsilon > 0$, $\gamma \in (-3,-2)$ and $g \in X_\gamma$, we have the following L^p estimate for $f(t) = \Phi_g(t,\cdot)\#f^0$ where $f^0 \in L^p$ and Φ_g are as in (P) and (3.1), respectively.*

$$\|f(t)\|_{L^p} \le \|f^0\|_{L^p} \exp\left\{C_{\varepsilon,\gamma,d}\big(1 + \operatorname{esssup}_{s\in[0,T]}\|g(s)\|_{L^p}\big)t\right\}, \quad \forall\, t \in [0,T].$$

In particular, $f \in X_\gamma$.

Proof of Proposition 3.3. We begin by writing the weak formulation of (P) against test functions $\phi \in C_c^\infty(\mathbb{R}^d)$,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^d}\phi(v)\mathrm{d}f_t(v) &= \int_{\mathbb{R}^d}\nabla\phi(v)\cdot U^\varepsilon[g(t)](v)\mathrm{d}f_t(v)\\ &= \iint_{\mathbb{R}^{2d}}|v-w|^{2+\gamma}\nabla\phi(v)\cdot\Pi[v-w]B^\varepsilon\,\mathrm{d}f(v)\mathrm{d}g(w).\end{aligned}\tag{3.4}$$

We will derive the desired estimate by making use of $\langle v\rangle^2\phi_R$ as a test function in (3.4), where

$$\phi_R(v) = \phi\left(\frac{v}{R}\right), \quad \phi(v) = \Phi(|v|) = \begin{cases}1 & |v|\le 1\\ 0 & |v| > 2\end{cases}, \quad \Phi \ge 0, \quad \Phi \in C_c^\infty(\mathbb{R}).$$

Firstly, notice that $\nabla(\langle v\rangle^2\phi_R(v))$ is supported in $|v| \le 2R$ and takes the form

$$\nabla(\langle v\rangle^2\phi_R(v)) = \underbrace{\left(2\phi_R(v) + \frac{1}{R}\langle v\rangle^2\frac{\Phi'(|v|/R)}{|v|}\right)}_{=:P(v)}v. \tag{3.5}$$

In particular, since $\frac{1}{R} \lesssim \frac{1}{|v|} \lesssim \frac{1}{\langle v\rangle}$ for large $R \gg 1$, the bound for Φ' gives

$$\nabla(\langle v\rangle^2\phi_R(v)) = P(v)v, \quad \text{with } |P(v)| \lesssim 1. \tag{3.6}$$

We start with the easier case of $\gamma \in (-3,-2)$. By interpolating the estimates in (2.2), the function $\nabla\log[g * G^\varepsilon]$ is Hölder continuous so we can deduce

$$|v-w|^{2+\gamma}|B^\varepsilon| \lesssim_\varepsilon 1, \quad B^\varepsilon = \nabla G^\varepsilon * \log[g * G^\varepsilon](v) - \nabla G^\varepsilon * \log[g * G^\varepsilon](w).$$

This greatly simplifies the double integral to the following:

$$\begin{aligned}&\left|\iint_{\mathbb{R}^{2d}}|v-w|^{2+\gamma}\nabla(\langle v\rangle^2\phi_R(v))\cdot\Pi[v-w]B^\varepsilon\,\mathrm{d}f(v)\mathrm{d}g(w)\right|\\ &\qquad \le C_\varepsilon\iint_{\mathbb{R}^{2d}}\langle v\rangle\,\mathrm{d}f(v)\mathrm{d}g(w) \le C_\varepsilon M_1(f).\end{aligned}$$

This establishes (a stronger version of) the result for $\gamma \in (-3,-2)$. Turning to the case $\gamma \in [-2,0]$, we split the inner integral in v of (3.4) into regions where $|v-w| \le 1$ and $|v-w| \ge 1$ obtaining a first reduction using (2.2) and (3.6)

$$\begin{aligned}
&\int_{w\in\mathbb{R}^d}\left(\int_{\{v:|v-w|\le 1\}}+\int_{\{v:|v-w|\ge 1\}}\right)|v-w|^{2+\gamma}P(v)v\cdot\Pi[v-w]B^{\varepsilon}\,\mathrm{d}f(v)\mathrm{d}g(w)\\
&\qquad\lesssim_{\varepsilon} M_0(g)M_1(f)\\
&\qquad\quad+\underbrace{\int_{w\in\mathbb{R}^d}\int_{\{v:|v-w|\ge 1\}}|v-w|^{2+\gamma}P(v)v\cdot\Pi[v-w]B^{\varepsilon}\,\mathrm{d}f(v)\mathrm{d}g(w)}_{=:I}.
\end{aligned} \tag{3.7}$$

For the term I, we use the identity $\Pi[v-w]v = \Pi[v-w]w$, $|B^{\varepsilon}| \lesssim_{\varepsilon} 1$, (3.6), and Young's inequality (cf. the proof of Lemma 3.1) which give

$$\begin{aligned}
|I| &\lesssim_{\varepsilon} \int_{w\in\mathbb{R}^d}\int_{\{v:|v-w|\ge 1\}}(|v|^2+|w|^2)|v-w|^{\gamma}|\Pi[v-w]v|\,\mathrm{d}f(v)\mathrm{d}g(w)\\
&\le \int_{w\in\mathbb{R}^d}\int_{\{v:|v-w|\ge 1\}}|v|^2|\Pi[v-w]w|\,\mathrm{d}f(v)\mathrm{d}g(w)\\
&\quad+\int_{w\in\mathbb{R}^d}|w|^2\int_{\{v:|v-w|\ge 1\}}|\Pi[v-w]v|\,\mathrm{d}f(v)\mathrm{d}g(w)\\
&\le M_1(g)M_2(f)+M_2(g)M_1(f).
\end{aligned}$$

Collecting this estimate with (3.7), we have

$$\begin{aligned}
&\left|\iint_{\mathbb{R}^{2d}}|v-w|^{2+\gamma}\nabla(\langle v\rangle^2\,\phi_R(v))\cdot\Pi[v-w]B^{\varepsilon}\,\mathrm{d}f(v)\mathrm{d}g(w)\right|\\
&\qquad\lesssim_{\varepsilon}(M_0(g)M_1(f)+M_1(g)M_2(f)+M_2(g)M_1(f))\le M_2(g)M_2(f).
\end{aligned}$$

Inserting this estimate into (3.4) and integrating in time, we get

$$\int_{\mathbb{R}^d}\langle v\rangle^2\,\phi_R(v)\mathrm{d}f_t(v)\le\int_{\mathbb{R}^d}\langle v\rangle^2\,\phi_R(v)\mathrm{d}f^0(v)+C_{\varepsilon}\int_0^t M_2(g(s))M_2(f(s))\mathrm{d}s.$$

By Monotone Convergence, passing to the limit $R\to\infty$ with Grönwall's inequality gives the stated *a priori* estimate on the growth of the second moment of f.

Concerning the statement that $f\in C([0,T];\mathscr{P}_c(\mathbb{R}^d))$, notice that Φ_g is bounded in $[0,T]$ when $g\in X_{\gamma}$ according to Lemma 3.1. Moreover, since f^0 has compact support, f has compact support as a push-forward of f^0 through a bounded flow map.

Finally, in the case $f = g$, take $\phi(v) = \langle v\rangle^2 = 1 + |v|^2$ as a test function (justified by the previous estimates) so that the right-hand side of (3.4) reads

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}M_2(f(t)) &= \frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^d}(1+|v|^2)\mathrm{d}f_t(v)\\ &= 2\iint_{\mathbb{R}^d}|v-w|^{2+\gamma}v\cdot\Pi[v-w]B^\varepsilon\,\mathrm{d}f_t(v)\mathrm{d}f_t(w)\\ &= \iint_{\mathbb{R}^d}|v-w|^{2+\gamma}(v-w)\cdot\Pi[v-w]B^\varepsilon\,\mathrm{d}f_t(v)\mathrm{d}f_t(w) = 0.\end{aligned}$$

In plain words, $M_2(f(t))$ is constant. □

3.2. *Well-posedness of* (1.3)

In this section, we prove Theorem 1.1 by applying the contraction mapping theorem to the solution map of (P) denoted by $S_\gamma : g \in X_\gamma \mapsto f(t) = \Phi_g(t,\cdot)\#f^0 \in X_\gamma$ (cf. Proposition 3.3 and Lemma 3.2). We will leverage the propagation estimates from Sec. 3.1 to the following closed subspace of X_γ. For $T > 0$, define

$$M_\gamma = M_\gamma(T) := \begin{cases}\{f\in X_\gamma \,|\, \sup_{t\in[0,T]} M_2(f(t)) \le 2M_2(f^0)\}, & \gamma\in[-2,0],\\ \{f\in X_\gamma \,|\, \operatorname{esssup}_{t\in[0,T]}\|f(t)\|_{L^p} \le 3\|f^0\|_{L^p}\}, & \gamma\in(-3,-2).\end{cases}$$

Remark 3.2. Fix $b > 0$ and $k > a > 0$. For $T > 0$, define the function

$$F_T : [0,k] \to [0,+\infty), \quad F_T(x) = ae^{bTx}.$$

Then, for every $T \le T_C := \frac{1}{bk}\log\frac{k}{a}$, it holds $F_T \le k$. In other words, $F_T : [0,k] \to [0,k]$.

Motivated by Remark 3.2, we define the time horizon

$$T_C := \begin{cases}\dfrac{\log 2}{2C_\varepsilon M_2(f^0)}, & \gamma\in[-2,0],\\[2ex] \min\left(\dfrac{\log 2}{C_{\varepsilon,\gamma,d}}, \dfrac{\log\frac{3}{2}}{6C_{\varepsilon,\gamma,d}\|f^0\|_{L^p}}\right), & \gamma\in(-3,-2),\end{cases}$$

where C_ε, $C_{\varepsilon,\gamma,d} > 0$ are the constants appearing in the exponential in Proposition 3.3 and Lemma 3.2, respectively. The plan of the following proof is to show that S_γ is a contraction from M_γ to itself.

Proof of Theorem 1.1. For fixed $g \in M_\gamma$, we denote $f = S_\gamma g$ (i.e. $f(t) = \Phi_g(t,\cdot)\#f^0$). Let us first consider $\gamma \in [-2,0]$ and show $f \in M_\gamma$. The estimate of Proposition 3.3 reads

$$\sup_{s\in[0,T]} M_2(f(s)) \le F_T\left(\sup_{s\in[0,T]} M_2(g(s))\right),$$

where F_T is the function from Remark 3.2 and the constants are $a = M_2(f^0)$, $b = C_\varepsilon$. Moreover, we take the constant $k = 2M_2(f^0) > a = M_2(f^0)$. This fulfils the criteria of Remark 3.2 so that for $T \leq T_C = \frac{\log 2}{2C_\varepsilon M_2(f^0)}$, we have

$$\sup_{s\in[0,T]} M_2(f(s)) \leq 2M_2(f^0).$$

This proves $f \in M_\gamma$ in the case $\gamma \in [-2, 0]$. The case $\gamma \in (-3, -2)$ follows similarly from Lemma 3.2 (which replaces Proposition 3.3) and Remark 3.2.

Existence and uniqueness: We now prove that S_γ is a contraction on M_γ with respect to the metric d_∞. We need to show that there is some universal constant $\kappa \in (0, 1)$ such that for every $g^1, g^2 \in M_\gamma$, we have

$$d_\infty(S_\gamma g^1,\, S_\gamma g^2) = \sup_{t\in[0,T]} W_\infty(S_\gamma g^1(t),\, S_\gamma g^2(t)) \leq \kappa d_\infty(g^1,\, g^2).$$

Let us denote $f^i = S_\gamma g^i$ for $i = 1,\, 2$ so that $f^i \in M_\gamma$ solves (P) induced by g^i. Let us fix $t \in [0, T]$ and suppress the time dependence. We can use $(\Phi_{g^1} \times \Phi_{g^2})\#f^0$ as an admissible transport plan between f^1 and f^2 and the following estimate[36]:

$$\begin{aligned} W_\infty(f^1, f^2) &= \lim_{p\to\infty} W_p(\Phi_{g^1}\#f^0, \Phi_{g^2}\#f^0) \\ &\leq \lim_{p\to\infty} \left(\int_{\mathbb{R}^d} |\Phi_{g^1}(v) - \Phi_{g^2}(v)|^p \,\mathrm{d}f^0(v) \right)^{\frac{1}{p}}. \end{aligned} \tag{3.8}$$

The crucial quantity to estimate is the difference between the flow maps. Notice that we are only concerned with the difference for $v \in \mathrm{supp} f^0 \subset B_R$ from (**A1**). In particular, for $g^1, g^2 \in M_\gamma$ and $\gamma \in (-2, 0]$, we will use the following growth estimate from Lemma 3.1

$$\left\langle \Phi_{g^i} \right\rangle \leq \langle R\rangle \exp\left\{C_\varepsilon M_2(f^0)T\right\}, \quad i = 1, 2. \tag{3.9}$$

Using the fact that the flow maps satisfy (3.1), we write again for fixed $t \in [0, T]$

$$\begin{aligned} \Phi_{g^1}(v) - \Phi_{g^2}(v) &= \int_0^t U^\varepsilon[g^1](\Phi_{g^1}(s, v)) - U^\varepsilon[g^2](\Phi_{g^2}(s, v))\mathrm{d}s \\ &= \int_0^t \underbrace{U^\varepsilon[g^1](\Phi_{g^1}(s, v)) - U^\varepsilon[g^1](\Phi_{g^2}(s, v))}_{=:I_1} \mathrm{d}s \\ &\quad + \int_0^t \underbrace{U^\varepsilon[g^1](\Phi_{g^2}(s, v)) - U^\varepsilon[g^2](\Phi_{g^2}(s, v))}_{=:I_2} \mathrm{d}s. \end{aligned} \tag{3.10}$$

Starting with the difference I_1, we use the Lipschitz regularity of U^ε from Proposition 2.4 to get

$$|I_1| \leq \Lambda_\gamma(g^1,\, \Phi_{g^1}(s, v),\, \Phi_{g^2}(s, v))|\Phi_{g^1}(s, v) - \Phi_{g^2}(s, v)|.$$

In particular, absorbing more constants into C_ε, (3.9) gives the bound

$$|\Lambda_\gamma| \leq \begin{cases} C_\varepsilon M_2(f^0)\, \langle R\rangle^{2+\gamma} \exp(C_\varepsilon(2+\gamma)M_2(f^0)T), & \gamma \in [-2,0], \\ C_{\varepsilon,\gamma,p',d}(1+\|f^0\|_{L^p}), & \gamma \in (-3,-2). \end{cases}$$

Turning to I_2, we need to estimate the measure-wise difference of the velocity fields U^ε. An application of Proposition 2.5 and (3.9) gives

$$|I_2| \lesssim_\varepsilon W_\infty(g^1(s),\, g^2(s)) \begin{cases} \langle R\rangle^{2+\gamma} M_2(f^0) \\ \quad \times \exp\left(C_\varepsilon(2+\gamma)M_2(f^0)T\right), & \gamma \in [-2,0], \\ (1+\|f^0\|_{L^p}), & \gamma \in (-3,-2). \end{cases}$$

Inserting these estimates for I_1 and I_2 back into (3.10), we have

$$|\Phi_{g^1}(v) - \Phi_{g^2}(v)| \lesssim_{\varepsilon,\, f^0,\, \gamma,\, T} \int_0^t |\Phi_{g^1}(s,v) - \Phi_{g^2}(s,v)| + W_\infty(g^1(s), g^2(s))\mathrm{d}s.$$

Denoting the constant on the right-hand side by $c = c(\varepsilon, f^0)$ (the dependence on γ and T is bounded), an application of Grönwall's inequality gives

$$|\Phi_{g^1}(v) - \Phi_{g^2}(v)| \leq c\int_0^t e^{c(t-s)} W_\infty(g^1(s), g^2(s))\mathrm{d}s.$$

Notice that the previous estimate is independent of v and $p > 1$, hence when it is substituted into (3.8), we obtain

$$W_\infty(f^1(t), f^2(t)) \leq c\int_0^t e^{c(t-s)} W_\infty(g^1(s), g^2(s))\mathrm{d}s \leq (e^{cT}-1)d_\infty(g^1, g^2).$$

By reducing $T > 0$ even further (depending only on ε and f^0), we can ensure that the Lipschitz constant $e^{cT} - 1 =: \kappa < 1$.

Maximal time of existence: Having finished with the short time existence and uniqueness of solutions to (1.3), we turn to the statement concerning the maximal time of existence, T_M. The dichotomy for $\gamma \in (-3,-2)$ comes from the standard Cauchy–Lipschitz theory. For the case $\gamma \in [-2,0]$, we were able to construct solutions provided we could ensure

$$\sup_{t\in[0,T_M]} M_2(f(t)) \leq 2M_2(f^0).$$

On the other hand, the last statement of Proposition 3.3 shows that second moments are conserved by solutions of (1.3); $M_2(f(t)) = M_2(f^0)$, for every $t \in [0, T_M]$. Thus, we can indefinitely repeat the contraction mapping argument and extend the solution globally to any finite time horizon. □

4. The Mean Field Limit

This section is dedicated to the proof of Theorem 1.2. The initial computations for both cases $\gamma \in [-2,0]$ and $\gamma \in (-3,-2)$ are the same which we present now until they diverge. To fix notation, let f and $\mu^N = \sum_{i=1}^N m_i \delta_{v^i(t)}$ denote the continuum and (any) empirical solution constructed from Sec. 3 and Appendix C, respectively. f^0 denotes the initial data to f satisfying (**A1**) and (**A2**) while $\mu_0^N = \sum_{i=1}^N m_i \delta_{v_0^i}$ denotes the initial data of μ^N satisfying (**B1**) and (**B2**). We define the following "discrete" flow

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t} F^N(t,s;v) = U^\varepsilon[\mu^N](F^N(t,s;v)), \\ F^N(s,s;v) = v \in \mathbb{R}^d, \end{cases} \qquad t,s \in [0,T^N], \tag{4.1}$$

such that $\eta_m(t) = \min_{i\neq j} |v^i(t) - v^j(t)| > 0$ for $t \in [0,T^N]$ together with the "continuous" flow

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t} F(t,s;v) = U^\varepsilon[f](F(t,s;v)), \\ F(s,s;v) = v \in \mathbb{R}^d, \end{cases} \qquad t,s \in [0,T_m]. \tag{4.2}$$

Notice that $T^N > 0$ may be taken as *any* arbitrary time horizon for $\gamma \in [-2,0]$ since μ^N is defined for all times and Proposition C.1 asserts that $\eta_m^N(t) > 0$ for all times. The dependence on $N \in \mathbb{N}$ for the time horizon T^N in (4.1) is only relevant for $\gamma \in (-3,-2)$ and this is investigated in the sequel. Here $T_m > 0$ is the maximal time of existence of the continuum limit solving (1.3) as indicated in Theorem 1.1. We may choose $T_m = +\infty$ for $\gamma \in [-2,0]$ while *a priori* it may be finite for $\gamma \in (-3,-2)$. From the discussion in Sec. 3 and Appendix C, the flows in (4.1) and (4.2) are well defined. Fix $0 < t_0 < \min(T_m, T^N)$. Take τ^0 an optimal transport map in W_∞ between $f(t_0)$ and $\mu^N(t_0)$ i.e. $\mu^N(t_0) = \tau^0 \# f(t_0)$. From the construction of f in Sec. 3, we have that $f(t) = F(t,\, t_0;\, \cdot)\# f(t_0)$. Moreover, we also have $\mu^N(t) = F^N(t,\, t_0;\, \cdot)\#\mu^N(t_0)$. Using a composition of all these maps, we can define a candidate transport map to estimate the W_∞ distance between $f(t)$ and $\mu^N(t)$ by

$$\tau^t \# f(t) = \mu^N(t), \quad \text{where } \tau^t = F^N(t,\, t_0;\, \cdot) \circ \tau^0 \circ F(t_0,\, t;\, \cdot).$$

For any $1 \le p < \infty$, the W_p Wasserstein distance can be estimated by

$$W_p^p(\mu^N(t), f(t)) \le \int_{\mathbb{R}^d} |F(t,\, t_0;\, v) - F^N(t,\, t_0;\, \tau^0(v))|^p \,\mathrm{d}f_{t_0}(v).$$

The limit $p = \infty$ is then given by

$$\eta^N(t) = W_\infty(\mu^N(t), f(t)) \le \|F(t,\, t_0;\, \cdot) - F^N(t,\, t_0;\, \tau^0(\cdot))\|_{L^\infty(f_{t_0})}.$$

By definition of the flows defined in (4.1) and (4.2), we have

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=t_0^+} F^N(t,\, t_0;\, \tau^0(v)) - F(t,\, t_0;\, v) = U^\varepsilon[\mu^N(t_0)](\tau^0(v)) - U^\varepsilon[f(t_0)](v).$$

Therefore, we can estimate

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=t_0^+} \eta^N(t) \le \|U[\mu^N(t_0)](\tau^0(v)) - U[f(t_0)](v)\|_{L_v^\infty(f_{t_0})}. \tag{4.3}$$

We expand the velocity difference in (4.3) using the fact that τ^0 transports f_{t_0} to $\mu^N_{t_0}$

$$\begin{aligned}
&U^\varepsilon[\mu^N(t_0)](\tau^0(v)) - U^\varepsilon[f(t_0)](v)\\
&\quad= \int_{\mathbb{R}^d} K_{\mu^N(t_0)}(\tau^0(v), w)d\mu^N_{t_0}(w) - \int_{\mathbb{R}^d} K_{f(t_0)}(v,w)\mathrm{d}f_{t_0}(w)\\
&\quad= \int_{\mathbb{R}^d} (K_{\mu^N(t_0)}(\tau^0(v),\tau^0(w)) - K_{f(t_0)}(v,w))\mathrm{d}f_{t_0}(w)\\
&\quad= \int_{\mathbb{R}^d} \underbrace{\{K_{\mu^N_{t_0}}(\tau^0(v),\tau^0(w)) - K_{\mu^N_{t_0}}(\tau^0(v),w) + K_{\mu^N_{t_0}}(\tau^0(v),w) - K_{\mu^N_{t_0}}(v,w)\}}_{=:D_1}\\
&\qquad\times \mathrm{d}f_{t_0}(w)\\
&\qquad+ \int_{\mathbb{R}^d} \underbrace{\{K_{\mu^N_{t_0}}(v,w) - K_{f_{t_0}}(v,w)\}}_{=:D_2}\,\mathrm{d}f_{t_0}(w).
\end{aligned} \tag{4.4}$$

Starting with the D_2 term, we need only concern ourselves with bounded v and w in the integrations owing to Lemma 3.1. In particular, the regions of integration can be restricted to

$$|v|,\,|w| \le \begin{cases} R\exp(C_\varepsilon t_0), & \gamma\in(-2,0],\\ R + C_\varepsilon t_0, & \gamma\in(-3,-2). \end{cases} \tag{4.5}$$

By Lemmas 2.1 and 2.5, we can estimate the D_2 term by

$$\begin{aligned}
\int_{\mathbb{R}^d} |D_2|\,\mathrm{d}f_{t_0}(w) &\lesssim_\varepsilon W_\infty(\mu^N_{t_0}, f_{t_0}) \int \min(|v-w|^{3+\gamma},\, |v-w|^{2+\gamma})\mathrm{d}f_{t_0}(w)\\
&\le W_\infty(\mu^N_{t_0}, f_{t_0})\\
&\quad\times \begin{cases} R^{2+\gamma}\exp(C_\varepsilon(2+\gamma)t_0)M_{2+\gamma}(f_{t_0}), & \gamma\in(-2,0],\\ 1, & \gamma\in(-3,-2). \end{cases}
\end{aligned} \tag{4.6}$$

As for the D_1 term, we can complete the proof of Theorem 1.2 for $\gamma\in(-2,0]$.

Proof. (Proof of Theorem 1.2 for moderately soft potentials) Building on the previous discussion, we only need to estimate the D_1 term. Using Lemma 2.3 twice, we get

$$\begin{aligned}
&\int_{\mathbb{R}^d} |D_1|\,\mathrm{d}f_{t_0}(w)\\
&\quad\lesssim_\varepsilon \int_{B_{R\exp(C_\varepsilon t_0)}} |\tau^0(w)-w| \min(|\tau^0(v)-\tau^0(w)|^{2+\gamma}, |\tau^0(v)-w|^{2+\gamma})\mathrm{d}f_{t_0}(w)\\
&\qquad+ |\tau^0(v)-v| \int_{B_{R\exp(C_\varepsilon t_0)}} \min(|\tau^0(v)-w|^{2+\gamma}, |v-w|^{2+\gamma})\mathrm{d}f_{t_0}(w).
\end{aligned}$$

The assumptions (**B1**) and (**A1**) say $W_\infty(\mu_0^N, f_0) \to 0$ and $\mathrm{supp} f_0 \subset B_R$, thus for sufficiently large $N \gg 1$, we must have $\mathrm{supp}\mu_0^N \subset B_{R+1}$. Moreover, τ^0 pushes forward f_{t_0} to $\mu_{t_0}^N$, so we obtain $\mathrm{Im}\tau^0 \subset B_{(R+1)\exp(C_\varepsilon t_0)}$ from (4.5). Hence, we can bluntly estimate the minimum terms using Lemma 2.1 to obtain

$$\int_{\mathbb{R}^d} |D_1| \,\mathrm{d}f_{t_0}(w) \le R^{2+\gamma}\exp(C_\varepsilon(2+\gamma)t_0)\left(\int |\tau^0(w) - w|\,\mathrm{d}f_{t_0}(w) + |\tau^0(v) - v|\right)$$
$$\le 2R^{2+\gamma}\exp(C_\varepsilon(2+\gamma)t_0)W_\infty(\mu_{t_0}^N, f_{t_0}).$$

Collecting this and (4.6), plugging them into (4.4) and then (4.3), we arrive at

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=t_0^+} \eta^N(t) \lesssim_\varepsilon R^{2+\gamma} e^{C_\varepsilon(2+\gamma)t_0}\eta^N(t_0).$$

As $t_0 \in (0, T)$ was chosen arbitrarily, a direct application of Grönwall's inequality gives

$$W_\infty(\mu^N(t), f(t)) \le W_\infty(\mu_0^N, f^0)\exp\left\{C_\varepsilon^1 R^{2+\gamma} e^{C_\varepsilon(2+\gamma)T} t\right\}, \quad \forall\, t \in [0, T].$$

This implies the mean field limit for $\gamma \in (-2, 0]$. □

Proof. (Proof of Theorem 1.2 for very soft potentials) For $\gamma \in (-3, -2)$, the same method to estimate the D_1 term in (4.4) does not work. Moreover, the construction of the particle solutions μ^N is only local in time up to some time horizon $T^N > 0$ (cf. Appendix C) which may be strictly less than T_m and may also degenerate to 0 as $N \to +\infty$. We overcome these issues to show the mean field limit by repeating the inter-particle distance analysis from.[7] The general steps from[7] to show $\eta^N(t) \to 0$ are to couple the evolutions of η^N and η_m^N together in the following way where we recall η_m^N denotes the inter-minimum particle distance for the particles in μ^N

$$\eta_m^N(t) := \min_{i \ne j} |v^i(t) - v^j(t)|.$$

(1) We show first in Sec. 4.1 the growth estimate of η coupled with η_m

$$\frac{\mathrm{d}}{\mathrm{d}t}\eta^N \lesssim_{\varepsilon,\gamma} \eta^N(1 + \|f\|_{L^p})\big(1 + (\eta^N)^{\frac{d}{p'}}(\eta_m^N)^{1+\gamma}\big). \tag{4.7}$$

(2) Then in Sec. 4.2 we obtain the decay estimate of η_m coupled with η

$$\frac{\mathrm{d}}{\mathrm{d}t}\eta_m^N \gtrsim_{\varepsilon,\gamma} -\eta_m^N(1 + \|f\|_{L^p})\big(1 + (\eta^N)^{\frac{d}{p'}}(\eta_m^N)^{1+\gamma}\big). \tag{4.8}$$

(3) The coupled system (4.7) and (4.8) together with (**B2**) allow us to deduce both $\liminf_{N\to\infty} T^N \ge T_m$ and $\eta^N(t) \to 0$ for all times $t \in [0, T_m)$. This is performed in Appendix D based on the argument in Ref. 7. □

Since the velocity vector field U^ε depends on the solution, we cannot directly repeat the arguments from Ref. 7 to establish (4.7) and (4.8). Nevertheless, once these estimates are proven, step (3) follows exactly as in Ref. 7 which we leave to Appendix D for completeness.

4.1. *Step* 1

We focus on the D_1 term from (4.4) recalling that (4.6) implies

$$\int_{\mathbb{R}^d} |D_2| \, \mathrm{d}f_{t_0}(w) \lesssim_\varepsilon \eta^N(t_0).$$

Proposition 4.1. *For fixed $\varepsilon > 0$ and $\gamma \in (-3,-2)$, we have the estimate*

$$\int_{\mathbb{R}^d} |D_1| \, \mathrm{d}f_{t_0}(w) \lesssim_{\varepsilon,\gamma,p'} \eta^N(1 + \|f_{t_0}\|_{L^p})(1 + (\eta^N)^{\frac{d}{p'}}(\eta_m^N)^{1+\gamma}).$$

Substituting these estimates for D_1 and D_2 into (4.4) and then (4.3) gives (4.7) completing the first step.

Proof. Using (2.3), we obtain

$$\begin{aligned}|D_1| \lesssim_{\varepsilon,\gamma} &|\tau^0(w) - w| \max(|\tau^0(v) - w|^{2+\gamma}, |\tau^0(v) - \tau^0(w)|^{2+\gamma}) \\ &+ |\tau^0(v) - v| \max(|\tau^0(v) - w|^{2+\gamma}, |v - w|^{2+\gamma}).\end{aligned}$$

Integration region $|v - w| \geq 4\eta^N$: We first deduce

$$|\tau^0(v) - \tau^0(w)| \geq |v - w| - |\tau^0(v) - v| - |\tau^0(w) - w| \geq |v - w| - 2\eta^N \geq \frac{|v - w|}{2}.$$

The second inequality is obtained by remembering τ^0 is an optimal transport map in W_∞ between $f(t_0)$ and $\mu^N(t_0)$. Similarly, we have the estimate

$$|\tau^0(v) - w| \geq |v - w| - |\tau^0(v) - v| \geq |v - w| - \eta^N \geq \frac{3|v - w|}{4}.$$

Overall, these estimates lead to

$$|D_1| \lesssim_{\varepsilon,\gamma} \eta^N |v - w|^{2+\gamma}$$

and integrating over $\{w \in \mathbb{R}^d \,|\, |v - w| \geq 4\eta^N\}$ yields

$$\begin{aligned}\int_{|v-w|\geq 4\eta^N} |D_1| \, \mathrm{d}f_{t_0}(w) \; &\lesssim_{\varepsilon,\gamma} \; \eta^N \left(\int_{4\eta^N \leq |v-w| \leq 1} + \int_{|v-w|>1} \right) |v - w|^{2+\gamma} \, \mathrm{d}f_{t_0}(w) \\ &\lesssim_{p',\gamma} \eta^N(\|f_{t_0}\|_{L^p} + 1).\end{aligned}$$

Integration region $|v - w| < 4\eta^N$: Here, we do not use the cancellations in

$$D_1 = K_{\mu^N}(\tau^0(v), \tau^0(w)) - K_{\mu^N}(v, w),$$

instead, we estimate each term using Proposition 2.2. Since $\mathrm{Im}\tau^0 \subset \{v^i\}_{i=1}^N$, if $\tau^0(v) = \tau^0(w)$, then the Hölder regularity of $\log[\mu^N * G^\varepsilon]$ (after interpolating the estimates in (2.2)) gives $K_{\mu^N}(\tau^0(v), \tau^0(w)) = 0$. Otherwise, we use $|\tau^0(v) - \tau^0(w)| \geq \eta_m$ and Proposition 2.2 to deduce

$$|D_1| \lesssim_\varepsilon \left(\eta_m^N\right)^{2+\gamma} + |v - w|^{2+\gamma}.$$

By Hölder's inequality, the integral can be estimated by

$$\begin{aligned}
\int_{|v-w|<4\eta^N} |D_1|\,\mathrm{d}f_{t_0}(w) &\lesssim_\varepsilon \int_{|v-w|<4\eta^N} |v-w|^{2+\gamma}\,\mathrm{d}f_{t_0}(w) + \left(\eta_m^N\right)^{2+\gamma} \\
&\quad \times \int_{|v-w|<4\eta^N} \mathrm{d}f_{t_0}(w) \\
&\leq \|f_{t_0}\|_{L^p} \left(\left(\int_{|v-w|<4\eta^N} |v-w|^{(2+\gamma)p'}\,\mathrm{d}w\right)^{\frac{1}{p'}}\right. \\
&\quad \left. + \left(\int_{|v-w|<4\eta^N} 1\,\mathrm{d}w\right)^{\frac{1}{p'}} (\eta_m^N)^{2+\gamma}\right) \\
&\lesssim_{\gamma,p'} \|f_{t_0}\|_{L^p}\left((\eta^N)^{\frac{d}{p'}+2+\gamma} + (\eta_m^N)^{2+\gamma}(\eta^N)^{\frac{d}{p'}}\right).
\end{aligned}$$

Finally, choose two indices i, j such that $\eta_m^N = |v^i - v^j|$. We seek to estimate η_m against η by looking at where τ^0 sends the midpoint $\frac{v^i+v^j}{2}$. In the case that $\frac{v^i+v^j}{2} \notin \mathrm{supp} f$, then we can define τ^0 to be whatever we want as it does not affect the W_∞ distance. In particular, we can assign $\tau^0(v) \in \{v^i\}_{i=1}^N$ for every $v \notin \mathrm{supp} f^0$ without changing the transport cost. Suppose

$$\tau^0\left(\frac{v^i+v^j}{2}\right) = v^k \in \{v^i\}_{i=1}^N.$$

Without loss of generality $v_k \neq v_i$, and we have the following lower bound

$$\begin{aligned}
\eta^N &\geq \left|\tau^0\left(\frac{v^i+v^j}{2}\right) - \frac{v^i+v^j}{2}\right| = \left|v^k - \frac{v^i+v^j}{2}\right| = \left|v^k - v^i + \frac{v^i-v^j}{2}\right| \\
&\geq |v^k - v^i| - \frac{1}{2}|v^i - v^j| \geq \frac{1}{2}\eta_m^N.
\end{aligned}$$

This implies $\eta_m^N \leq 2\eta^N$ which simplifies

$$\int_{|v-w|<4\eta^N} |D_1|\,\mathrm{d}f_{t_0}(w) \lesssim_{\varepsilon,\gamma,p'} (\eta^N)^{\frac{d}{p'}+1}\left(\eta_m^N\right)^{1+\gamma}\|f_{t_0}\|_{L^p}.$$

□

4.2. *Step 2*

Having derived an upper bound for the growth of η^N coupled with η_m^N, we need to find a corresponding lower bound for the decrease of η_m^N coupled with η^N to close the system.

Proposition 4.2. *The minimum inter-particle distance satisfies the lower bound for its decay*

$$\frac{\mathrm{d}}{\mathrm{d}t}\eta_m^N \gtrsim_{\varepsilon,\gamma} -\eta_m^N(1+\|f\|_{L^p})(1+(\eta^N)^{\frac{d}{p'}}\left(\eta_m^N\right)^{1+\gamma}).$$

Proof. Choose two indices i, $j = 1, \dots, N$ such that $|v^i - v^j| = \eta_m^N$ where we will suppress time dependence for simplicity. We have

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}|v^i - v^j| &\geq -|U^\varepsilon[\mu^N](v^i) - U^\varepsilon[\mu^N](v^j)| \\
&\geq -\int_{\mathbb{R}^3} |K_{\mu^N}(v^i, w) - K_{\mu^N}(v^j, w)|\,\mathrm{d}\mu^N(w) \\
&= -\int_{\mathbb{R}^3} |K_{\mu^N}(v^i, \tau(w)) - K_{\mu^N}(v^j, \tau(w))|\,\mathrm{d}f(w).
\end{aligned}$$

Here, we have set τ as an optimal transfer map in W_∞ such that $\mu^N(t) = \tau\# f(t)$ for $t \in [0, \min(T, T^N))$. We split the integration into the following domains:

$$\mathscr{A} = \{w : \min(|v^i - w|, |v^j - w|) \geq 2\eta^N\}, \quad \mathscr{B} = \mathbb{R}^d \backslash \mathscr{A}.$$

Starting with $\mathscr{A}$, we use the inequality

$$|v^i - \tau(w)| \geq |v^i - w| - |w - \tau(w)| \geq |v^i - w| - \eta^N \geq \frac{|v^i - w|}{2}$$

and Lemma 2.3 to deduce

$$\begin{aligned}
&\int_{\mathscr{A}} |K_{\mu^N}(v^i, \tau(w)) - K_{\mu^N}(v^j, \tau(w))|\,\mathrm{d}f(w) \\
&\quad \lesssim_{\varepsilon,\gamma} |v^i - v^j| \int_{\mathscr{A}} \max(|v^i - \tau(w)|^{2+\gamma}, |v^j - \tau(w)|^{2+\gamma})\mathrm{d}f(w) \\
&\quad \leq 2^{-(2+\gamma)}|v^i - v^j| \int_{\mathscr{A}} (|v^i - w|^{2+\gamma} + |v^j - w|^{2+\gamma})\mathrm{d}f(w) \\
&\quad \lesssim_\gamma \eta_m^N(1 + \|f\|_{L^p}).
\end{aligned}$$

In the last line, we have bluntly estimated

$$\begin{aligned}
\int_{\mathscr{A}} |v^i - w|^{2+\gamma}\,\mathrm{d}f(w) &\leq \int_{\mathbb{R}^d} |v^i - w|^{2+\gamma}\,\mathrm{d}f(w) \\
&\leq \int_{|v^i - w| \geq 1} \mathrm{d}f(w) + \int_{|v^i - w| < 1} |v^i - w|^{2+\gamma}\,\mathrm{d}f(w)
\end{aligned}$$

with the usual Hölder's inequality for the second term and similarly for v^j.

Turning to the region $\mathscr{B}$, since $\mathrm{Im}\tau \subset \{v^i\}_{i=1}^N$, as soon as $v^i \neq \tau(w)$, we must have

$$|v^i - \tau(w)| \geq \eta_m^N,$$

with a similar estimate for v^j. By further blunting the L^∞ estimate in Proposition 2.2, we obtain

$$|K_{\mu^N}(v^i, \tau(w))| \lesssim_\varepsilon |v^i - \tau(w)|^{2+\gamma} \leq \left(\eta_m^N\right)^{2+\gamma}.$$

If $v^i = \tau(w)$, then the Hölder regularity of $\nabla \log[\mu^N * G^\varepsilon]$ from (2.2) gives $K_{\mu^N}(v^i, \tau(w)) = 0$. The familiar method using Hölder's inequality gives

$$\int_{\mathscr{B}} \mathrm{d}f(w) \leq \left(\int_{\mathscr{B}} \mathrm{d}w\right)^{\frac{1}{p'}} \|f\|_{L^p} \lesssim (\eta^N)^{\frac{d}{p'}} \|f\|_{L^p}.$$

Putting these two estimates together, we treat the full integral over $\mathscr{B}$ by

$$\begin{aligned}\int_{\mathscr{B}} |K_{\mu^N}(v^i, \tau(w)) - K_{\mu^N}(v^j, \tau(w))| \,\mathrm{d}f(w) &\lesssim \left(\eta^N_m\right)^{2+\gamma} \int_{\mathscr{B}} \mathrm{d}f(w) \\ &\lesssim (\eta^N)^{\frac{d}{p'}} \left(\eta^N_m\right)^{2+\gamma} \|f\|_{L^p}.\end{aligned}$$

Finally, we add up the integrals over $\mathscr{A}$ and $\mathscr{B}$ to get

$$\begin{aligned}&\int_{\mathbb{R}^d} |K_{\mu^N}(v^i, \tau(w)) - K_{\mu^N}(v^j, \tau(w))| \,\mathrm{d}f(w) \\ &\quad\lesssim_{\varepsilon,\gamma} \eta^N_m (1 + \|f\|_{L^p})(1 + (\eta^N)^{\frac{d}{p'}} \left(\eta^N_m\right)^{1+\gamma}).\end{aligned}$$

□

Appendix A. Proof of Lemma 2.3

The structure of our kernel is more general than those considered in Ref. 7, but the idea is the same and we provide the details for completeness.

The case $\gamma \in [-2, 0]$: The fundamental theorem of calculus with Proposition 2.2 gives

$$\begin{aligned}|K_g(v_1, w) - K_g(v_2, w)| &= \left| \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t} K_g(tv_1 + (1-t)v_2, w) \mathrm{d}t \right| \\ &\lesssim_\varepsilon |v_1 - v_2| \int_0^1 |tv_1 + (1-t)v_2 - w|^{2+\gamma} \mathrm{d}t.\end{aligned}$$

Up to a constant depending on γ, the integrand can be estimated by

$$\begin{aligned}|tv_1 + (1-t)v_2 - w|^{2+\gamma} &\lesssim_\gamma |v_1 - w|^{2+\gamma} + |v_2 - w|^{2+\gamma} \\ &\lesssim \max(|v_1 - w|^{2+\gamma}, |v_2 - w|^{2+\gamma}).\end{aligned}$$

The case $\gamma \in [-3, -2)$: Set $\Gamma(t) = (1-t)v_1 + tv_2 - w$ and we separate into further cases.

Case 1 - For every $t \in [0, 1]$, we have $|\Gamma(t)| \geq \frac{1}{4} \min(|v_1 - w|, |v_2 - w|)$: We can repeat the previous computations almost exactly and recover the desired estimate.

Case 2 - There is a $t \in [0, 1]$ such that $|\Gamma(t)| < \frac{1}{4} \min(|v_1 - w|, |v_2 - w|)$: We need to perturb the original contour Γ to avoid the possible singularity. Notice that we can find $t \in (0, 1)$ such that $|\Gamma(t)| < \frac{1}{4} \min(|v_1 - w|, |v_2 - w|)$. We first take (the unique) $t_m \in (0, 1)$ such that

$$|\Gamma(t_m)| = \min_{t \in [0,1]} |\Gamma(t)|.$$

Next, define the other two time points where $|\Gamma(t)| = \frac{1}{4}\min(|v_1 - w|, |v_2 - w|)$,

$$t_i := \inf\left\{t \in [0,1] : |\Gamma(t)| = \frac{1}{4}\min(|v_1 - w|, |v_2 - w|)\right\},$$

$$t_s := \sup\left\{t \in [0,1] : |\Gamma(t)| = \frac{1}{4}\min(|v_1 - w|, |v_2 - w|)\right\}.$$

By continuity of $|\Gamma(t)|$, we have that all t_m, t_i, $t_s \in (0,1)$. The triangle formed by connecting the vectors $\Gamma(t_i)$, $\Gamma(t_s) - \Gamma(t_i)$, and $\Gamma(t_s)$ is *isosceles* so the following quantity is well defined (see Fig. A.1)

$$r := |\Gamma(t_i) - \Gamma(t_m)| = |\Gamma(t_m) - \Gamma(t_s)|.$$

We wish to apply the fundamental theorem by taking the contour connecting $v_1 - w$ to $v_2 - w$ that traces a semicircular arc from $\Gamma(t_i)$ to $\Gamma(t_s)$ in the direction furthest from the origin (the green arc in Fig. A.1). More precisely, the direction furthest away from the origin is defined as

$$e := \frac{\Gamma(t_m)}{|\Gamma(t_m)|},$$

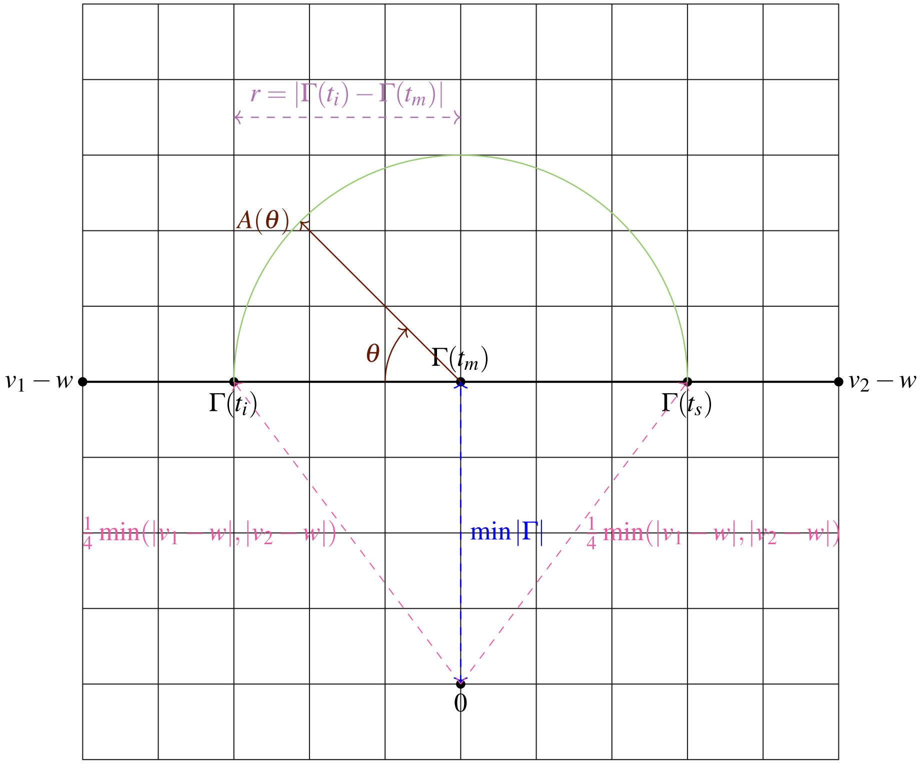

Fig. A.1. Simplistic visual perturbation of $\Gamma(t)$ to avoid the singularity.

or if $\Gamma(t_m) = 0$, take any $e \in \mathbb{S}^{d-1}$. For a given angle $\theta \in [0, \pi]$, the green arc can be parameterised by

$$A(\theta) := \Gamma(t_m) + r\left(\cos\theta \frac{\Gamma(t_i) - \Gamma(t_m)}{r} + \sin\theta\, e\right).$$

Observe that by the (reverse) triangle inequality and the fact that $e \perp \Gamma(t_i) - \Gamma(t_m)$, we have the lower bound for all $\theta \in [0, \pi]$

$$\begin{aligned} |A(\theta)| \geq |\Gamma(t_m)| - r = |\Gamma(t_m)| - |\Gamma(t_i) - \Gamma(t_m)| \geq |\Gamma(t_i)| \\ = \frac{1}{4}\min(|v_1 - w|, |v_2 - w|). \end{aligned} \tag{A.1}$$

Putting these pieces together, we define the perturbed contour $\tilde{\Gamma} : [0,\, t_i + \pi + 1 - t_s] \to \mathbb{R}^d$ by

$$\tilde{\Gamma}(t) := \begin{cases} \Gamma(t), & t \in [0,\, t_i], \\ A(t - t_i), & t \in [t_i,\, t_i + \pi], \\ \Gamma(t - t_i - \pi + t_s), & t \in [t_i + \pi,\, t_i + \pi + 1 - t_s]. \end{cases}$$

We will apply the fundamental theorem of calculus on each of the three pieces of $\tilde{\Gamma}$ to estimate the difference

$$\begin{aligned} &|K_g(v_1, w) - K_g(v_2, w)| \\ &\quad\leq |K_g(\tilde{\Gamma}(0) + w, w) - K_g(\tilde{\Gamma}(t_i) + w, w)| \\ &\qquad + |K_g(\tilde{\Gamma}(t_i) + w, w) - K_g(\tilde{\Gamma}(t_i + \pi) + w, w)| \\ &\qquad + |K_g(\tilde{\Gamma}(t_i + \pi) + w, w) - K_g(\tilde{\Gamma}(t_i + \pi + 1 - t_s) + w, w)| \\ &\quad\leq \int_0^{t_i} \left|\frac{\mathrm{d}}{\mathrm{d}t} K_g(\tilde{\Gamma}(t) + w, w)\right| \mathrm{d}w + \int_{t_i}^{t_i+\pi} \left|\frac{\mathrm{d}}{\mathrm{d}t} K_g(\tilde{\Gamma}(t) + w, w)\right| \mathrm{d}w \\ &\qquad + \int_{t_i+\pi}^{t_i+\pi+1-t_s} \left|\frac{\mathrm{d}}{\mathrm{d}t} K_g(\tilde{\Gamma}(t) + w, w)\right| \mathrm{d}w =: T_1 + T_2 + T_3. \end{aligned} \tag{A.2}$$

Starting with T_1, the chain rule gives

$$T_1 \leq |v_1 - v_2| \int_0^{t_i} |\nabla_v K_g(\tilde{\Gamma}(t) + w, w)|\, \mathrm{d}t.$$

Using the derivative estimate in Proposition 2.2 and the fact that $|\Gamma(t)| \geq \frac{1}{4}\min(|v_1 - w|, |v_2 - w|)$ for $t \in [0, t_i]$, we obtain

$$\begin{aligned} T_1 &\lesssim_\varepsilon |v_1 - v_2| \int_0^{t_i} |\tilde{\Gamma}(t)|^{2+\gamma} dt \\ &\leq |v_1 - v_2| \int_0^{t_i} \max(|v_1 - w|^{2+\gamma}, |v_2 - w|^{2+\gamma}) dt \\ &\leq t_i |v_1 - v_2| \max(|v_1 - w|^{2+\gamma}, |v_2 - w|^{2+\gamma}). \end{aligned} \tag{A.3}$$

Similarly for T_3, we have

$$T_3 \lesssim_\varepsilon (1-t_s)|v_1-v_2| \max(|v_1-w|^{2+\gamma}, |v_2-w|^{2+\gamma}). \tag{A.4}$$

We now turn to T_2, we substitute $A(t-t_i)$ into this piece and use the derivative estimate from Proposition 2.2 with the chain rule to get

$$T_2 \lesssim_\varepsilon \int_{t_i}^{t_i+\pi} \left|\frac{\mathrm{d}}{\mathrm{d}t}A(t-t_i)\right| |A(t-t_i)|^{2+\gamma}\mathrm{d}t.$$

Recalling the definitions of r (this is the length of a particular segment of $[v_1 - w, v_2 - w]$) and A together with the lower bound (A.1), we use

$$\left|\frac{\mathrm{d}}{\mathrm{d}t}A\right| = r \leq |v_1 - v_2| \quad \text{and} \quad |A| \geq \frac{1}{4}\min(|v_1-w|, |v_2-w|)$$

so that we have

$$T_2 \lesssim_\varepsilon \pi|v_1-v_2| \max(|v_1-w|^{2+\gamma}, |v_2-w|^{2+\gamma}).$$

Putting this inequality with (A.4) and (A.3) into (A.2), we achieve the desired result.

Appendix B. Proof of Lemma 3.2

By our abuse of notation from interchanging probability measures with their densities, we write down the explicit formula for $f(t,v)$ as a density

$$f(t,v) = \frac{f^0(\Phi_g^{-1}(t,v))}{|\det(\nabla\Phi_g(t, \Phi_g^{-1}(t,v)))|}. \tag{B.1}$$

Here, the inverse Φ_g^{-1} should be thought of as the "reverse" flow map to Φ_g where the direction of time has been reversed. Changing variables with (B.1), we have

$$\int |f(t,v)|^p \, \mathrm{d}v = \int \frac{|f^0(v)|^p}{|\det(\nabla\Phi_g(t,v))|^{p-1}} \, \mathrm{d}v. \tag{B.2}$$

We turn to estimating the denominator in the integrand of (B.2). Again, standard facts about the flow map Φ_g from Ref. 19 give the following formula:

$$\det\nabla\Phi_g(t,v) = \exp\left\{\int_0^t \nabla_v \cdot U^\varepsilon[g](\Phi_g(s,v))\mathrm{d}s\right\}.$$

From an application of the Dominated Convergence Theorem and Proposition 2.2, we have

$$\begin{aligned} |\nabla_v \cdot U^\varepsilon[g](v)| &\leq \int |\nabla_v \cdot K_g(v,w)| \, \mathrm{d}g(w) \\ &\lesssim_\varepsilon \int |v-w|^{2+\gamma} \, \mathrm{d}g(w) \leq 1 + C_{\gamma,d}\|g\|_{L^p}. \end{aligned}$$

The last computation is obtained by the usual method of splitting the integration region between $|v-w|<1$ and $|v-w|\geq 1$ recalling $2+\gamma<0$. Inserting this inequality into (B.2), we obtain the desired estimate

$$\int |f(t,v)|^p \mathrm{d}v \leq \left(\int |f^0(v)|^p \mathrm{d}v\right) \exp\{C_{\varepsilon,\gamma,d}(p-1)(1+\operatorname{esssup}_{s\in[0,T]}\|g(s)\|_{L^p})t\}.$$

Finally, Proposition 3.3 already proved $f\in C([0,T];\mathscr{P}_c(\mathbb{R}^d))$ and the $L^\infty_t L^p_v$ property is clear from the estimate we have just proved.

Appendix C. The Interacting Particle System

This section is concerned with proving Lemma 1.1; the well-posedness of the particle system described in (1.6). Throughout this section, the number $N\in\mathbb{N}$ of particles is fixed as well as the positive weights $\{m_i\}_{i=1}^N$ and initial points $\{v_0^i\}_{i=1}^N$. We denote the initial empirical data by $\mu_0^N=\sum_{i=1}^N m_{i,N}\delta_{v_0^i}$. We can apply the same arguments from Sec. 4 for $\gamma\in(-2,0]$.

Proof. (Proof of Lemma 1.1 for $\boldsymbol{-2<\gamma\leq 0}$) The initial empirical data μ_0^N satisfies (**A1**) with radius of support $R_N:=\max_{i=1,\ldots,N}|v_0^i|$. Applying Theorem 1.1 for any $T>0$, we have the unique solution $\mu^N(t)\in X_\gamma(T)$. Moreover, Proposition 3.1 says that $\mu^N(t)$ can be represented as

$$\mu^N(t)=\Phi_{\mu^N}(t,\cdot)\#\mu_0^N,$$

where Φ_{μ^N} is the (unique!) flow map in (3.1) induced by the curve μ^N. Since μ^N is the push-forward of μ_0^N, it is also an empirical measure with the form

$$\mu^N(t)=\sum_{i=1}^N m_{i,N}\delta_{\Phi_{\mu^N}(t,v_0^{i,N})}.$$

Moreover, for every $i=1,\ldots,N$, $\Phi_{\mu^N}(t,v_0^{i,N})$ solves precisely (1.6). □

The following proposition gives a lower bound on the minimum inter-particle distance

$$\eta_m^N(t)=\min_{i,j=1,\ldots,N}|v^i(t)-v^j(t)|.$$

Proposition C.1. (No collisions in finite time) *Fix* ε, $T>0$, $\gamma\in(-2,0]$, *and* $\eta_m^N(0)=\min_{i\neq j}|v_0^i-v_0^j|>0$. *Then, there is a constant* $C=C(\varepsilon,T,M_2(\mu_0^N))>0$ *such that the minimum inter-particle distance decays with exponential rate*

$$\eta_m^N(t)\gtrsim \eta_m^N(0)\exp\{-Ct\},\quad \forall\, t\in[0,T].$$

Proof. Choose two indices i, $j = 1, \dots, N$ such that $\eta_m^N = |v^i - v^j|$ where we will suppress the time dependence for simplicity. We have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}|v^i - v^j| &\geq -|U^\varepsilon[\mu^N](v^i) - U^\varepsilon[\mu^N](v^j)| \\ &\geq -\int_{\mathbb{R}^3} |K_{\mu^N}(v^i, w) - K_{\mu^N}(v^j, w)| d\mu^N(w).\end{aligned}$$

The goal is to estimate the integral. Firstly, we simplify the integration by recalling that $\mathrm{supp}\mu^N$ is bounded. Indeed, setting $R = \max_{i=1,\dots,N} |v_0^i|$, Lemma 3.1 implies

$$\langle v^i \rangle = \langle \Phi_{\mu^N}(v_0^i) \rangle \leq R \exp(C_\varepsilon M_2(\mu_0^N)t), \quad \forall\, t \in [0, T].$$

Applying Lemma 2.3 to the difference of the kernels, we have

$$\begin{aligned}&\int_{\mathrm{supp}\mu^N} |K_{\mu^N}(v^i, w) - K_{\mu^N}(v^j, w)| d\mu^N(w) \\ &\quad \lesssim_{\varepsilon,\gamma} |v^i - v^j| \int_{\mathrm{supp}\mu^N} \max(|v^i - w|^{2+\gamma}, |v^j - w|^{2+\gamma}) d\mu^N(w) \\ &\quad \lesssim_\gamma R \exp(C_\varepsilon M_2(\mu_0^N)T)|v^i - v^j|. \qquad \square\end{aligned}$$

In the case $\gamma \in (-3, -2)$, we can no longer apply Theorem 1.1 directly, since it requires an L^p assumption on the initial data μ_0^N which is not valid for empirical measures. In particular, the vector field is no longer Lipschitz regular (cf. Proposition 2.4) so we must make do with Hölder regularity (cf. Proposition 2.3).

Proof. (Proof of Lemma 1.1 for $-3 < \gamma < -2$) We revisit the proof of Peano's theorem using Schauder's fixed point theorem to construct solutions to (1.4). We set $X = C([0, T]; \mathbb{R}^d)$ and define the solution map $S : X \to X$ by

$$(Sv^i)(t) := v_0^i + \int_0^t U^\varepsilon[\mu^N(s)](v^i(s))\mathrm{d}s, \quad i = 1, \dots, N.$$

This is well defined and certainly $Sv^i \in X$ for each $v^i \in X$ owing to the uniform bound for U^ε in Proposition 2.2 when $\gamma \in (-3, -2)$. We seek to prove (1) S is continuous and (2) $S(X)$ is pre-compact.

S is continuous: For every $i = 1, \dots, N$ fix $v^{i,n}$, $v^i \in X$ such that $v^{i,n} \to v^i$ in X. We label their corresponding empirical measures

$$\mu^N(t) = \sum_{i=1}^N m_i \delta_{v^i(t)}, \quad \mu^{N,n}(t) = \sum_{i=1}^N m_i \delta_{v^{i,n}(t)}.$$

We have the estimate

$$\begin{aligned}|(Sv^{i,n})(t)-(Sv^i)(t)| &\le \int_0^t |U^\varepsilon[\mu^{N,n}(s)](v^{i,n}(s))-U^\varepsilon[\mu^N(s)](v^i(s))|\mathrm{d}s\\ &\le \int_0^t |U^\varepsilon[\mu^{N,n}(s)](v^{i,n}(s))-U^\varepsilon[\mu^{N,n}(s)](v^i(s))|\\ &\quad +|U^\varepsilon[\mu^{N,n}(s)](v^i(s))-U^\varepsilon[\mu^N(s)](v^i(s))|\mathrm{d}s.\end{aligned}$$

Applying Proposition 2.3 to the first difference and Lemma 2.6 to the second difference without being precise about the constants, we obtain

$$\begin{aligned}|(Sv^{i,n})(t)-(Sv^i)(t)| \lesssim_{\varepsilon,\gamma} &\int_0^t |v^{i,n}(s)-v^i(s)|^{3+\gamma}\mathrm{d}s\\ &+\int_0^t W_\infty(\mu^{N,n}(s),\mu^N(s))+W_\infty(\mu^{N,n}(s),\mu^N(s))^{3+\gamma}\mathrm{d}s.\end{aligned}$$

The first integral converges to 0 as $n\to\infty$. As well, the infinite Wasserstein distance is also continuous with respect to the particles; $v^{i,n}\to v^i$ in X for every $i=1,\dots,N$ as $n\to\infty$ implies $W_\infty(\mu^{N,n}(s),\mu^N(s))\to 0$ as $n\to\infty$.

$S(X)$ is pre-compact: We fix $v^i\in X$ for every $i=1,\dots,N$ in this step. Firstly, it is clear that $S(X)$ is bounded using Proposition 2.2

$$|(Sv)(t)|\le |v(0)|+C_\varepsilon t, \quad \forall\, v\in X.$$

Turning to equicontinuity, fix $t_1\le t_2$ both in $[0,T]$. Applying Proposition 2.2 again, we have

$$|(Sv^i)(t_1)-(Sv^i)(t_2)|\le \int_{t_1}^{t_2} |U^\varepsilon[\mu^N(s)](v^i(s))|\,\mathrm{d}s \lesssim_\varepsilon |t_1-t_2|. \qquad \square$$

Appendix D. Step 3

In this appendix, we prove step (3) from Sec. 4 which establishes Theorem 1.2. The results of Secs. 4.1 and 4.2 yield

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{d}t}\eta^N \lesssim_\varepsilon \eta^N(1+\|f\|_{L^p})(1+\left(\eta^N\right)^{\frac{d}{p'}}\left(\eta^N_m\right)^{1+\gamma}),\\ &\frac{\mathrm{d}}{\mathrm{d}t}\eta^N_m \gtrsim_\varepsilon -\eta^N_m(1+\|f\|_{L^p})(1+\left(\eta^N\right)^{\frac{d}{p'}}\left(\eta^N_m\right)^{1+\gamma}),\end{aligned} \tag{D.1}$$

when $t\in[0,\min(T_m,T^N))$. If $\left(\eta^N\right)^{\frac{d}{p'}}\left(\eta^N_m\right)^{1+\gamma}\le 1$, then we immediately obtain

$$\begin{aligned}&\eta^N(t)\le \eta^N(0)e^{C(1+\|f\|_{L^p})t},\\ &\eta^N_m(t)\ge \eta^N_m(0)e^{-C(1+\|f\|_{L^p})t}, \quad \forall\, t\in[0,\min(T_m,T^N)).\end{aligned} \tag{D.2}$$

We wish to show that (D.2) holds for all $t \in [0, T_m)$ as $N \to \infty$ which amounts to showing $T^N > T_m$ when N is sufficiently large. Define first

$$a(t) := \frac{\eta^N(t)}{\eta^N(0)}, \quad \eta_m(t) := \frac{\eta_m^N(t)}{\eta_m^N(0)}, \quad \xi_N := \eta^N(0)^{\frac{d}{p'}} \eta_m^N(0)^{1+\gamma}.$$

Thus, we rewrite (D.1) in terms of a, b, and ξ_N

$$\frac{\mathrm{d}}{\mathrm{d}t} a \lesssim_\varepsilon a(1 + \|f\|_{L^p})(1 + \xi_N a^{\frac{d}{p'}} b^{1+\gamma}),$$

$$\frac{\mathrm{d}}{\mathrm{d}t} b \gtrsim_\varepsilon -b(1 + \|f\|_{L^p})(1 + \xi_N a^{\frac{d}{p'}} b^{1+\gamma}).$$

Since $a(0) = b(0) = 1$ and we assume by (1.7) $\xi_N \to 0$ as $N \to \infty$, when N is sufficiently large, we can find $T_*^N (\le T^N)$ such that

$$\xi_N a^{\frac{d}{p'}} b^{1+\gamma} \le 1, \quad \forall t \in [0, T_*^N]. \tag{D.3}$$

Now by (D.2), we have similar estimates

$$a(t) \le e^{C(1+\|f\|_{L^p})t}, \quad b(t) \ge e^{-C(1+\|f\|_{L^p})t}, \quad \forall t \in [0, T_*^N].$$

Returning to (D.3), we obtain an estimate for T_*^N given by

$$\xi_N e^{C(1+\|f\|_{L^p})\left(\frac{d}{p'} - (1+\gamma)\right)t} \le 1 \Leftrightarrow t \le -\frac{\log \xi_N}{C(1 + \|f\|_{L^p})\left(\frac{d}{p'} - (1+\gamma)\right)}.$$

This means that T_*^N has the lower bound

$$-\frac{\log \xi_N}{C(1 + \|f\|_{L^p})\left(\frac{d}{p'} - (1+\gamma)\right)} \le T_*^N.$$

However, since (1.7) means $\xi_N \to 0$ as $N \to \infty$, this implies

$$\liminf_{N \to \infty} T_*^N = \infty.$$

Since $T_*^N < T^N$, we have that $T^N \ge T_m$ for $N \gg 1$ sufficiently large.

Acknowledgments

JAC was supported the Advanced Grant Nonlocal-CPD (Nonlocal PDEs for Complex Particle Dynamics: Phase Transitions, Patterns and Synchronization) of the European Research Council Executive Agency (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 883363). JAC was also partially supported by the EPSRC grant numbers EP/T022132/1 and EP/V051121/1. MGD was partially supported by NSF-DMS-2205937 and NSF-DMS RTG 1840314. JW was supported by the Mathematical Institute Award of the University of Oxford. The authors would like to thank the Isaac Newton Institute for Mathematical Sciences, Cambridge, for support and hospitality during the programme *Frontiers in Kinetic Theory* where work on this paper was undertaken. This work was supported by EPSRC grant no EP/R014604/1.

References

1. J. Barros-Neto, *An Introduction to the Theory of Distributions*, Pure and Applied Mathematics, Vol. 14 (Marcel Dekker, Inc., 1973).
2. A. L. Bertozzi, T. Laurent and J. Rosado, L^p theory for the multidimensional aggregation equation, *Comm. Pure Appl. Math.* **64** (2011) 45–83.
3. F. Bolley, J. A. Cañizo and J. A. Carrillo, Stochastic mean-field limit: Non-Lipschitz forces and swarming, *Math. Models Methods Appl. Sci.* **21** (2011) 2179–2210.
4. W. Braun and K. Hepp, The Vlasov dynamics and its fluctuations in the $1/N$ limit of interacting classical particles, *Comm. Math. Phys.* **56** (1977) 101–113.
5. D. Bresch, P.-E. Jabin and J. Soler, A new approach to the mean-field limit of Vlasov–Fokker–Planck equations, preprint (2022), arXiv:2203.15747.
6. D. Bresch, P.-E. Jabin and Z. Wang, On mean-field limits and quantitative estimates with a large class of singular kernels: Application to the Patlak–Keller–Segel model, *C. R. Math. Acad. Sci. Paris* **357** (2019) 708–720.
7. J. A. Carrillo, Y.-P. Choi and M. Hauray, The derivation of swarming models: Mean-field limit and Wasserstein distances, in *Collective Dynamics from Bacteria to Crowds*, CISM International Centre for Mathematical Sciences, Courses and Lectures, Vol. 553 (Springer, 2014), pp. 1–46.
8. J. A. Carrillo, Y.-P. Choi, M. Hauray and S. Salem, Mean-field limit for collective behavior models with sharp sensitivity regions, *J. Eur. Math. Soc.* **21** (2019) 121–161.
9. J. A. Carrillo, K. Craig and F. S. Patacchini, A blob method for diffusion, *Calc. Var. Partial Differential Equations* **58** (2019) Paper No. 53, 53.
10. J. A. Carrillo, M. G. Delgadino, L. Desvillettes and J. Wu, The landau equation as a gradient flow.
11. J. A. Carrillo, M. G. Delgadino and J. Wu, Boltzmann to Landau from the gradient flow perspective, *Nonlinear Anal.* **219** (2022) Paper No. 112824, 49.
12. J. A. Carrillo, M. DiFrancesco, A. Figalli, T. Laurent and D. Slepčev, Global-in-time weak measure solutions and finite-time aggregation for nonlocal interaction equations, *Duke Math. J.* **156** (2011) 229–271.
13. J. A. Carrillo, J. Hu, L. Wang and J. Wu, A particle method for the homogeneous Landau equation, *J. Comput. Phys. X* **7** (2020) 100066, 24.
14. Y.-P. Choi and S. Salem, Propagation of chaos for aggregation equations with no-flux boundary conditions and sharp sensing zones, *Math. Models Methods Appl. Sci.* **28** (2018) 223–258.
15. Y.-P. Choi and S. Salem, Collective behavior models with vision geometrical constraints: Truncated noises and propagation of chaos, *J. Differential Equations* **266** (2019) 6109–6148.
16. P. Degond and B. Lucquin-Desreux, The Fokker–Planck asymptotics of the Boltzmann collision operator in the Coulomb case, *Math. Models Methods Appl. Sci.* **2** (1992) 167–182.
17. R. L. Dobrušin, Vlasov equations, *Funktsional. Anal. i Prilozhen.* **13** (1979) 48–58, 96.
18. M. Duerinckx, Mean-field limits for some Riesz interaction gradient flows, *SIAM J. Math. Anal.* **48** (2016) 2269–2300.
19. F. Golse, The mean-field limit for the dynamics of large particle systems, in *Journées Équations aux Dérivées Partielles* (Univ. Nantes, Nantes, 2003), pp. Exp. No. IX, 47.
20. F. Golse, On the dynamics of large particle systems in the mean field limit, in *Macroscopic and Large Scale Phenomena: Coarse Graining, Mean Field Limits and Ergodicity*, Lecture Notes in Applied Mathematics and Mechanics, Vol. 3 (Springer, 2016), pp. 1–144.

21. M. P. Gualdani and N. Zamponi, Spectral gap and exponential convergence to equilibrium for a multi-species Landau system, *Bull. Sci. Math.* **141** (2017) 509–538.
22. D. Han-Kwan and M. Iacobelli, From Newton's second law to Euler's equations of perfect fluids, *Proc. Amer. Math. Soc.* **149** (2021) 3045–3061.
23. M. Hauray, Wasserstein distances for vortices approximation of Euler-type equations, *Math. Models Methods Appl. Sci.* **19** (2009) 1357–1384.
24. M. Hauray, Mean field limit for the one dimensional Vlasov-Poisson equation, in *Séminaire Laurent Schwartz==Équations aux dérivées partielles et applications.* Année 2012–2013 (École Polytech., Palaiseau, 2014), Sérmin. Équ. Dériv. Partielles, pp. Exp. No. XXI, 16.
25. M. Hauray and P.-E. Jabin, N-particles approximation of the Vlasov equations with singular potential, *Arch. Ration. Mech. Anal.* **183** (2007) 489–524.
26. M. Hauray and P.-E. Jabin, Particle approximation of Vlasov equations with singular forces: Propagation of chaos, *Ann. Sci. Éc. Norm. Supér.* (4) **48** (2015) 891–940.
27. P.-E. Jabin, A review of the mean field limits for Vlasov equations, *Kinet. Relat. Models* **7** (2014) 661–711.
28. P.-E. Jabin and Z. Wang, Mean field limit and propagation of chaos for Vlasov systems with bounded forces, *J. Funct. Anal.* **271** (2016) 3588–3627.
29. P.-E. Jabin and Z. Wang, Mean field limit for stochastic particle systems, in *Active Particles. Vol. 1. Advances in Theory, Models, and Applications*, Modeling and Simulation in Science, Engineering and Technology (Birkhäuser/Springer, 2017), pp. 379–402.
30. P.-E. Jabin and Z. Wang, Quantitative estimates of propagation of chaos for stochastic systems with $W^{-1,\infty}$ kernels, *Invent. Math.* **214** (2018) 523–591.
31. D. Lazarovici and P. Pickl, A mean field limit for the Vlasov–Poisson system, *Arch. Ration. Mech. Anal.* **225** (2017) 1201–1231.
32. E. M. Lifshitz, *Perspectives in Theoretical Physics* (Pergamon Press, 1992), the collected papers of E. M. Lifshitz [E. M. Lifshits], edited by L. P. Pitaevskiĭ, with an introduction by D. ter Haar, with a biography of Lifshitz by Ya. B. Zel'dovich and M. I. Kaganov, translated by J. B. Sykes.
33. H. Neunzert, An introduction to the nonlinear Boltzmann–Vlasov equation, in *Kinetic Theories and the Boltzmann Equation*, Lecture Notes in Mathematics, Vol. 1048 (Springer, 1984), pp. 60–110.
34. M. Petrache and S. Serfaty, Next order asymptotics and renormalized energy for Riesz interactions, *J. Inst. Math. Jussieu* **16** (2017) 501–569.
35. S. T. Rachev and L. Rüschendorf, *Mass Transportation Problems. Vol. I: Theory*, Probability and its Applications (Springer-Verlag, 1998).
36. F. Santambrogio, *Optimal Transport for Applied Mathematicians: Calculus of Variations, PDEs, and Modeling*, Progress in Nonlinear Differential Equations and their Applications, Vol. 87 (Birkhäuser/Springer, 2015).
37. S. Serfaty, Mean field limit for Coulomb-type flows, *Duke Math. J.* **169** (2020) 2887–2935, with an appendix by Mitia Duerinckx and Serfaty.
38. H. Spohn, *Large Scale Dynamics of Interacting Particles* (Springer Science & Business Media, 2012).
39. C. Villani, On a new class of weak solutions to the spatially homogeneous Boltzmann and Landau equations, *Arch. Ration. Mech. Anal.* **143** (1998) 273–307.