Analysis and mean-field derivation of a porous-medium equation with fractional diffusion

Li Chen\textsuperscript{a}, Alexandra Holzinger\textsuperscript{b}, Ansgar Jüngel\textsuperscript{b}, and Nicola Zamponi\textsuperscript{a}

\textsuperscript{a}School of Business Informatics and Mathematics, University of Mannheim, Mannheim, Germany; \textsuperscript{b}Institute for Analysis and Scientific Computing, Vienna University of Technology, Wiedner Hauptstraße, Wien, Austria

\section*{ABSTRACT}
A mean-field-type limit from stochastic moderately interacting many-particle systems with singular Riesz potential is performed, leading to nonlocal porous-medium equations in the whole space. The nonlocality is given by the inverse of a fractional Laplacian, and the limit equation can be interpreted as a transport equation with a fractional pressure. The proof is based on Oelschläger’s approach and \textit{a priori} estimates for the associated diffusion equations, coming from energy-type and entropy inequalities as well as parabolic regularity. An existence analysis of the fractional porous-medium equation is also provided, based on a careful regularization procedure, new variants of fractional Gagliardo–Nirenberg inequalities, and the div-curl lemma. A consequence of the mean-field limit estimates is the propagation of chaos property.

\section*{ARTICLE HISTORY}
Received 17 September 2021
Accepted 25 August 2022

\section*{KEYWORDS}
Existence analysis; fractional diffusion; interacting particle systems; mean-field limit; nonlocal porous-medium equation; propagation of chaos

\section*{2000 MATHEMATICS SUBJECT CLASSIFICATION}
60H10; 60H30; 35K65; 35R11

\section*{1. Introduction}
The aim of this article is to derive and analyze the following nonlocal porous-medium equation:

$$\partial_t \rho = \text{div}(\rho \nabla P), \quad P = (-\Delta)^{-s}f(\rho), \quad \rho(0) = \rho^0 \quad \text{in} \ \mathbb{R}^d,$$

where $0 < s < 1$, $d \geq 2$, and $f \in C^1([0, \infty))$ is a nondecreasing function satisfying $f(0) = 0$. This model describes a particle system that evolves according to a continuity equation for the density $\rho(x, t)$ with velocity $v = -\nabla P$. The velocity is assumed to be the gradient of a potential, which expresses Darcy’s law. The pressure $P$ is related to the density in a nonlinear and nonlocal way through $P = (-\Delta)^{-s}f(\rho)$. The nonlocal operator $(-\Delta)^{-s}$ can be written as a convolution operator with a singular kernel,

$$(-\Delta)^{-s}u = \mathcal{K} * u, \quad \mathcal{K}(x) = c d_{d-s}|x|^{2s-d}, \quad x \in \mathbb{R}^d,$$

where $c d_{d-s} = \Gamma(d/2-s)/(4^s \pi^{d/2}\Gamma(s))$ and $\Gamma$ denotes the Gamma function [1, theorem 5].

\section*{CONTACT}
Ansgar Jüngel (juengel@tuwien.ac.at) Institute for Analysis and Scientific Computing, Vienna University of Technology, Wiedner Hauptstraße, Wien, Austria.

\section*{COPYRIGHT AND REPRINTS}
© 2022 The Author(s). Published with license by Taylor & Francis Group, LLC This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
If \( s = 0 \), we recover the porous-medium equation (for nonnegative solutions), while the case \( s = 1 \) was investigated in [2, 3] with \( f(u) = u \) for the evolution of the vortex density in a superconductor. Related models (with \( f(u) = u \)) appear in the dynamics of dislocations (line defects) in crystals [4, (1.5)]. Other applications include particle systems with long-range interactions [5, subsection 6.2]. The case \( 0 < s < 1 \) corresponds to long-range repulsive interactions. This model, still with \( f(u) = u \), was investigated in [4], but a mathematical justification is missing. In this article, we provide a rigorous derivation from an interacting particle system for general functions \( f(u) \). In this way, we aim to contribute to the understanding of mean-field limits involving nonquadratic nonlinearities.

Equation (1) was first analyzed in [6] with \( f(u) = u \) for nonnegative solutions and in [7] with \( f(u) = |u|^{m-2}u \ (m > 1) \) for sign-changing solutions. The nonnegative solutions have the interesting property that they propagate with finite speed, which is not common in other fractional diffusion models [6, 8]. Equation (1) was probabilistically interpreted in [9], and it was shown that the probability density of a so-called random flight process is given by a Barenblatt-type profile. Previous mean-field limits leading to (1) were concerned with the linear case \( f(u) = u \) only; see [10] (using the technique of [11]) and [12] (including additional diffusion as in (7) below). In [13], Eq. (1) (with \( f(u) = u \)) was derived in the high-force regime from the Euler–Riesz equations, which can be derived in the mean-field limit from interacting particle systems [14]. A direct derivation from particle systems with Lévy noise was proved in [15] for cross-diffusion systems, but still with \( f(u) = u \). Up to our knowledge, a rigorous derivation of (1) from stochastic interacting particle systems for general nonlinearities \( f(u) \) like power functions is missing in the literature. In this article, we fill this gap.

### 1.1. Problem setting

Equation (1) is derived from an interacting particle system with \( N \) particles, moving in the whole space \( \mathbb{R}^d \). Because of the singularity of the integral kernel and the degeneracy of the nonlinearity, we approximate (1) using three levels. First, we introduce a parabolic regularization adding a Brownian motion to the particle system with diffusivity \( \sigma \in (0, 1) \) and replacing \( f \) by a smooth approximation \( f_\sigma \). Second, we replace the interaction kernel \( K \) by a smooth kernel \( K_\zeta \) with compact support, where \( \zeta > 0 \). Third, we consider interaction functions \( W_\beta \) with \( \beta \in (0, 1) \), which approximate the delta distribution. (We refer to Subsection 1.3 for the precise definitions.)

The particle positions are represented on the microscopic level by the stochastic processes \( X^N_i(t) \) evolving according to

\[
\begin{align*}
    dX^N_i(t) &= -\nabla K_\zeta * f_\sigma \left( \frac{1}{N} \sum_{j=1, j\neq i}^N W_\beta \left( X^N_j(t) - X^N_i(t) \right) \right) dt + \sqrt{2\sigma} dB^N_i(t), \\
    X^N_i(0) &= \xi_i, \quad i = 1, \ldots, N,
\end{align*}
\]

where the convolution has to be understood with respect to \( x_\sigma \) \( (B^N_i(t))_{t \geq 0} \) are independent \( d \)-dimensional Brownian motions defined on a filtered probability space \((\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})\), and
\( \xi_i \) are independent identically distributed random variables in \( \mathbb{R}^d \) with the same probability density function \( \rho_\sigma^0 \) (defined in (12) below).

The mean-field-type limit is performed in three steps. First, for fixed \((\sigma, \beta, \zeta)\), system (3) is approximated for \( N \to \infty \) on the intermediate level by

\[
\begin{aligned}
d\hat{X}_i^N(t) &= -\nabla K_\zeta * f_\sigma\left(W_\beta * \rho_{\sigma,\beta,\zeta}\left(\hat{X}_i^N(t), \cdot \right)\right) dt + \sqrt{2\sigma}d\tilde{B}_i^N(t), \\
\hat{X}_i^N(0) &= \xi_i, \quad i = 1, \ldots, N,
\end{aligned}
\]

where \( \rho_{\sigma,\beta,\zeta} \) is the probability density function of \( \hat{X}_i^N \) and a strong solution to

\[
\partial_t \rho_{\sigma,\beta,\zeta} - \sigma \Delta \rho_{\sigma,\beta,\zeta} = \text{div}\left(\rho_{\sigma,\beta,\zeta} \nabla K_\zeta * f_\sigma\left(W_\beta * \rho_{\sigma,\beta,\zeta}\right)\right), \quad \rho_{\sigma,\beta,\zeta}(0) = \rho_{\sigma}^0 \quad \text{in} \ \mathbb{R}^d.
\]

System (4) is uncoupled, since \( \hat{X}_i^N \) depends on \( N \) only through the initial datum.

Second, passing to the limit \((\beta, \zeta) \to 0\) in the intermediate system leads on the macroscopic level to

\[
\begin{aligned}
d\hat{X}_i^N(t) &= -\nabla K \rho_{\sigma}(\hat{X}_i^N(t), \cdot) dt + \sqrt{2\sigma}d\tilde{B}_i^N(t), \\
\hat{X}_i^N(0) &= \xi_i, \quad i = 1, \ldots, N,
\end{aligned}
\]

where \( \rho_{\sigma} \) is the density function of \( \hat{X}_i^N \) and a weak solution to

\[
\partial_t \rho_{\sigma} = \sigma \Delta \rho_{\sigma} + \text{div}\left(\rho_{\sigma} \nabla (-\Delta)^{-\frac{1}{2}} f_\sigma(\rho_{\sigma})\right), \quad \rho_{\sigma}(0) = \rho_{\sigma}^0 \quad \text{in} \ \mathbb{R}^d.
\]

We perform the limits \( N \to \infty \) and \((\beta, \zeta) \to 0\) simultaneously. In this limit, we use the logarithmic scaling \( \beta \sim (\log N)^{-\mu} \) for some \( \mu > 0 \) between the strength of interaction \( \beta \) and the number of particles \( N \). This can be viewed as a moderately interacting particle system. For the smoothing parameter \( \zeta \) of the singularity from \( K \), we can even allow an algebraic dependence on \( N \), i.e., \( \zeta \sim N^{-\nu} \) for some \( \nu > 0 \); see Theorem 2 for details. Our approach also implies the two-step limit but leading to weak convergence only, compared to the convergence in expectation obtained in Theorem 3.

Third, the limit \( \sigma \to 0 \) is performed on the level of the diffusion equation, based on a priori estimates uniform in \( \sigma \) and the div-curl lemma.

The main result of this article is that the particles of system (3) become independent in the limit with a common density function that is a weak solution to (1)–(2).

1.2. State of the art

We already mentioned that the existence of weak solutions to (1) with \( f(u) = u \) was proved first in [6]. The convergence of the weak solution to a self-similar profile was shown by the same authors in [16]. The convergence becomes exponential, at least in one space dimension, when adding a confinement potential [17]. Equation (1) with \( f(u) = u \) was identified as the Wasserstein gradient flow of a square fractional Sobolev norm [18], implying time decay as well as energy and entropy estimates. The Hölder regularity of solutions to (1) was proved in [19] for \( f(u) = u \) and in [20] for \( f(u) = u^{m-1} \) and \( m \geq 2 \).

In the literature, related equations have been analyzed too. Equation (1) for \( f(u) = u \) and the limit case \( s = 1 \) was shown in [21] to be the Wasserstein gradient flow on the
space of probability measures, leading to the well-posedness of the equation and energy-dissipation inequalities. The existence of local smooth solutions to the regularized Eq. (7) are proved in [22]. The solutions \( \partial_t \rho = \text{div}(\rho^{m-1} \nabla P) \) with \( P = (-\Delta)^{-s} \rho \) in \( \mathbb{R}^d \) propagate with finite speed if and only if \( m \geq 2 \) [8]. The existence of weak solutions to this equation with \( P = (-\Delta)^{-s} (\rho^n) \) and \( n > 0 \) is proved in [23] (in bounded domains). While (1) has a parabolic-elliptic structure, parabolic-parabolic systems have been also investigated. For instance, the global existence of weak solutions to \( \partial_t \rho + (-\Delta)^s \rho = 0 \) in \( \mathbb{R}^d \) have been studied in the literature; see, e.g., [26]. Compared to (1), this problem has infinite speed of propagation. For a review and comparison of this model and (1), we refer to [27].

There is a huge literature concerning mean-field limits leading to diffusion equations, and the research started already in the 1980s; we refer to the reviews [28, 29] and the classical works of Sznitman [30, 31]. Oelschl"ager proved the mean-field limit in weakly interacting particle systems [32], leading to deterministic nonlinear processes, and moderately interacting particle systems [33], giving porous-medium-type equations with quadratic diffusion. First, investigations of moderate interactions in stochastic particle systems with nonlinear diffusion coefficients were performed in [34]. The approach of moderate interactions was extended in [35, 36] to multispecies systems, deriving population cross-diffusion systems. Reaction-diffusion equations with nonlocal terms were derived in the mean-field limit in [37]. The large population limit of point measure-valued Markov processes leads to nonlocal Lotka–Volterra systems with cross diffusion [38]. Further references can be found in [12, subsection 1.3].

Compared to previous works, we consider a singular kernel \( \mathcal{K} \) and derive a partial differential equation without Laplace diffusion by taking the limit \( \sigma \to 0 \). The authors of [39] derived the viscous porous-medium equation by starting from a stochastic particle system with a double convolution structure in the drift term, similar to (4). The main difference to our work is that (besides different techniques for the existence and regularity of solutions to the parabolic problems) we consider a singular kernel in one part of the convolution and a different scaling for the approximating regularized kernel \( \mathcal{K}_\xi = \mathcal{K} \omega_\xi * W_\xi \), where \( \omega_\xi \) is a \( W^{1,\infty}(\mathbb{R}^d) \) cutoff function (see Subsection 1.3 for the exact definition), in comparison to the interaction scaling \( W_\beta * \rho_\sigma, \beta, \xi \). The two different scalings \( \beta \) and \( \xi \) allow us to establish a result, for which the kernel regularization on the particle level does not need to be of logarithmic type but of power-law type only.

### 1.3. Main results and key ideas

We impose the following hypotheses:

1. **Data**:
   - \( 0 < s < 1, d \geq 2 \).

2. **Nonlinearity**:
   - \( \rho^0 \in L^\infty(\mathbb{R}^d) \cap L^1(\mathbb{R}^d) \) satisfies \( \rho^0 \geq 0 \) in \( \mathbb{R}^d \) and \( \int_{\mathbb{R}^d} \rho^0(x) |x|^{2d/(d-2s)} \, dx < \infty \).

3. **Nonlinearity**:
   - \( f \in C^1([0, \infty)) \) is nondecreasing, \( f(0) = 0 \), and \( u \mapsto uf(u) \) for \( u > 0 \) is strictly convex.
Let us discuss these assumptions. We assume that \( d \geq 2 \); the case \( d = 1 \) can be treated if \( s < 1/2 \); see [6]. Extending the range of \( s \) to \( s < 0 \) leads to the fractional (higher-order) thin-film equation, which is studied in [40]. The case \( 1 < s < d/2 \) may be considered too, since it yields better regularity results; we leave the details to the reader. Conversely, the case \( s \geq d/2 \) is more delicate since the multiplier in the definition of \((-\Delta)^{-s}\) using Fourier transforms does not define a tempered distribution. The case \( s = d/2 \) for \( d \leq 2 \) (with a logarithmic Riesz kernel) was analyzed in [10]. We need the moment bound for the initial datum \( \rho^0 \) to prove the same moment bound for \( \rho_\sigma \), which in turn is used several times, for instance to show the entropy balance and the convergence \( \rho_\sigma \to \rho \) as \( \sigma \to 0 \) in the sense of \( C^0_{\text{weak}}([0,T];L^1(\mathbb{R}^d)) \). The monotonicity of \( f \) and the strict convexity of \( u \mapsto uf(u) \) are needed to prove the strong convergence of \( (\rho_\sigma) \), which then allows us to identify the limit of \( (f_\sigma(\rho_\sigma)) \). An example of a function satisfying Hypothesis (H3) is \( f(u) = u^\beta \) with \( \beta \geq 1 \).

Our first result is concerned with the existence analysis of (1). This result is needed to prove the main theorem below. We write \( \| \cdot \|_\rho \) for the \( L^p(\mathbb{R}^d) \) norm and define the so-called entropy density \( h : [0,\infty) \to \mathbb{R} \) by

\[
h(u) = \int_0^u \int_1^v \frac{f'(w)}{w} dw dv \quad \text{for } u \geq 0.
\]

**Theorem 1** (Existence of weak solutions to (1)). Let Hypotheses (H1)–(H3) hold. Then, there exists a weak solution \( \rho \geq 0 \) to (1) satisfying (i) the regularity

\[
\rho \in L^\infty(0,\infty;L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)), \quad \nabla(-\Delta)^{-s/2}f(\rho) \in L^2(0,\infty;L^2(\mathbb{R}^d)),
\]

\[
\partial_t \rho \in L^2(0,\infty;H^{-1}(\mathbb{R}^d)),
\]

(ii) the weak formulation

\[
\int_0^T \langle \partial_t \rho, \phi \rangle dt + \int_0^T \int_{\mathbb{R}^d} \rho \nabla(-\Delta)^{-s}f(\rho) \cdot \nabla \phi dx dt = 0 \tag{8}
\]

for all \( \phi \in L^2(0,T;H^1(\mathbb{R}^d)) \) and \( T > 0 \), (iii) the initial datum \( \rho(0) = \rho^0 \) in the sense of \( H^{-1}(\mathbb{R}^d) \), and (iv) the following properties for \( t > 0 \):

- **Mass conservation:** \( \| \rho(t) \|_1 = \| \rho^0 \|_1 \),

- **Dissipation of the \( L^\infty \) norm:** \( \| \rho(t) \|_\infty \leq \| \rho^0 \|_\infty \),

- **Moment estimate:** \( \sup_{0 < t < T} \int_{\mathbb{R}^d} |\rho(x,t)| x^{2d/(d-2s)} dx \leq C(T) \),

- **Entropy inequality:**

\[
\int_{\mathbb{R}^d} h(\rho(t)) dx + \int_0^t \int_{\mathbb{R}^d} \nabla(-\Delta)^{-s/2}f(\rho)^2 dx dt \leq \int_{\mathbb{R}^d} h(\rho^0) dx.
\]

Note that the Hardy–Littlewood–Sobolev-type inequality (B3) (see Appendix B) implies that

\[
\| \rho \nabla(-\Delta)^{-s}f(\rho) \|_2 = \| \rho^{-s/2} \left[ \nabla(-\Delta)^{-s/2}f(\rho) \right] \|_2 \leq C \| \rho \|_{d/(2s)} \| \nabla(-\Delta)^{-s/2}f(\rho) \|_2,
\]

such that \( \rho \nabla(-\Delta)^{-s/2}f(\rho) \in L^2(\mathbb{R}^d) \), and the weak formulation (8) is well defined.
The main ideas of the proof of Theorem 1 are as follows. A priori estimates for strong solutions $\rho_{\sigma}$ to the regularized Eq. (7) are derived from mass conservation, the entropy inequality, and energy-type bounds. The energy-type bound allows us to show, for sufficiently small $\sigma > 0$, that the $L^\infty$ norm of $\rho_{\sigma}$ is bounded by the $L^\infty$ norm of $\rho^0$, up to some factor depending on the moment bound for $\rho^0$. The existence of a strong solution $\rho_{\sigma}$ is proved by regularizing (7) in a careful way to deal with the singular kernel. The regularized equation is solved locally in time by Banach’s fixed-point theorem. Entropy estimates allow us to extend this solution globally in time and to pass to the de-regularization limit. The second step is the limit $\sigma \to 0$ in (7). Since the bounds only provide weak convergence of (a subsequence of) $\rho_{\sigma}$, the main difficulty is the identification of the nonlinear limit $f_\sigma(\rho_{\sigma})$. This is done by applying the div-curl lemma and exploiting the monotonicity of $f$ and the strict convexity of $u \mapsto uf(u)$ [41].

We already mentioned that the existence of local smooth solutions $\rho_{\sigma}$ to (7) has been proven in [13]. However, we provide an independent proof that allows for global strong solutions and that yields a priori estimates needed in the mean-field limit.

Our second and main result is the propagation of chaos, which shows a mean-field-type convergence of the particle system (3) to a solution of (1). To define our particle system properly, we need some definitions. Introduce the smooth approximation

$$f_\sigma(u) = \int_0^u (\Gamma_{\sigma} \ast (f'1_{[0,\infty]}))(w)\Xi(\sigma w)dw \quad u \in \mathbb{R},$$

where the mollifier $\Gamma_{\sigma}$ for $\sigma > 0$ is given by $\Gamma_{\sigma}(x) = \sigma^{-1}\Gamma_1(x/\sigma)$, and $\Gamma_1 \in C_0^\infty(\mathbb{R})$ satisfies $\Gamma_1 \geq 0$, $\|\Gamma_1\|_1 = 1$, while the cutoff function $\tilde{\Xi} \in C_0^\infty(\mathbb{R})$ satisfies $0 \leq \tilde{\Xi} \leq 1$ in $\mathbb{R}$ and $\tilde{\Xi}(x) = 1$ for $|x| \leq 1$. Then, thanks to $\Gamma_{\sigma}$, we have $f_\sigma \in C^\infty(\mathbb{R})$. The cutoff function guarantees that the derivatives $D^k f_\sigma$ are bounded and compactly supported for all $k \geq 1$. Furthermore, it holds that $f'_\sigma \geq 0$, $f_\sigma(0) = 0$.

In a similar way, we introduce the mollifier function $W_{\beta}$ for $\beta > 0$ and $x \in \mathbb{R}^d$ by

$$W_{\beta}(x) = \beta^{-d}W_1(x/\beta), \quad W_1 \in C_0^\infty(\mathbb{R}^d) \text{ is symmetric, } W_1 \geq 0, \|W_1\|_1 = 1.$$

Let us define the cutoff version of the singular kernel $\mathcal{K}$ by

$$\tilde{\mathcal{K}}_\zeta := \mathcal{K}\omega_\zeta, \text{ where the cut-off function } \omega_\zeta \in W^{1,\infty}(\mathbb{R}^d) \text{ is such that }$$

$$0 \leq \omega_\zeta(x) \leq 1 \text{ for } x \in \mathbb{R}^d, \quad \|\nabla \omega_\zeta\|_\infty \leq 2\zeta,$$

$$\omega_\zeta(x) = 1 \text{ for all } |x| \leq \zeta^{-1}, \quad \omega_\zeta(x) = 0 \text{ for all } |x| \geq 2\zeta^{-1}.$$  

Then, the regularized kernel $\mathcal{K}_\zeta$ is given by

$$\mathcal{K}_\zeta(x) := \tilde{\mathcal{K}}_\zeta \ast W_\zeta(x) \quad \text{for all } x \in \mathbb{R}^d,$$

where $\zeta > 0$. Let the cutoff function $\Xi \in C_0^\infty(\mathbb{R}^d)$ satisfy $0 \leq \Xi \leq 1$ in $\mathbb{R}^d$ and $\Xi(x) = 1$ for $|x| \leq 1$. Then, we define the regularized initial datum for $x \in \mathbb{R}^d$ by

$$\rho^0_\sigma(x) = \kappa_\sigma(W_{\sigma} \ast \rho^0)(x)\Xi(\sigma x), \quad \text{where } \kappa_\sigma = \frac{\int_{\mathbb{R}^d} \rho^0(y)dy}{\int_{\mathbb{R}^d}(W_{\sigma} \ast \rho^0)(y)\Xi(\sigma y)dy}.$$  

This definition guarantees the mass conservation since $\|\rho^0_\sigma\|_1 = \|\rho^0\|_1$; see Subsection 2.1.
Note that our particle system (3) depends on 4 parameters: $N \in \mathbb{N}$ denotes the number of particles, $\beta > 0$ models the strength of interaction between the particles, $\zeta > 0$ describes the regularization of the singular kernel $\mathcal{K}$, and $\sigma > 0$ is a measure of the additional diffusion. The quantities $(\beta, \zeta, \sigma)$ are regularization parameters needed to overcome the singularity of the kernel $\mathcal{K}$ and the (possible) degeneracy of the nonlinearity $f$.

In the limit $N \to \infty$, $(\beta, \zeta, \sigma) \to 0$, we prove the following propagation-of-chaos result.

**Theorem 2** (Propagation of chaos). Let $\zeta^{-2s-1} \leq C_1 N^{1/4}$ and $\beta^{-3d-7} \leq \varepsilon \log N$ for some constants $C_1, \varepsilon > 0$, and let $P^k_{N, \sigma, \beta, \zeta}(t)$ be the joint distribution of $(X^N_i(t), \ldots, X^N_k(t))$ for $k \geq 1$ and $t \in (0, T)$. Then, there exists a subsequence in $\sigma$ such that

\[
\lim_{\sigma \to 0} \lim_{N \to \infty} P^k_{N, \sigma, \beta, \zeta}(t) = P^{\otimes k}(t) \quad \text{in the sense of distributions},
\]

where the limit is locally uniform in $t$, and the measure $P(t)$ is absolutely continuous with respect to the Lebesgue measure with the probability density function $\rho(t)$ that is a weak solution to (1).

It is well known (see, e.g., [31]) that the result of Theorem 2 implies the weak convergence of the empirical measure associated to the particle system (3) toward the deterministic measure $\rho(t)$, i.e.,

\[
\mu_{N, \sigma, \beta, \zeta}(t) = \sum_{i=1}^{N} \delta_{X_i^N(t)} \to \rho(t),
\]

for a subsequence in $\sigma$. Furthermore, Theorem 2 shows that at any time $t > 0$, in the limit $N \to \infty$, $(\beta, \zeta, \sigma) \to 0$, any finite selection of $k$ particles in (3) becomes independent with limiting distribution $P^{\otimes k}(t)$. If Eq. (1) was uniquely solvable, we would obtain the convergence of the whole sequence in $\sigma$. Unfortunately, the regularity of the solution $\rho$ to (1) is too weak to conclude the uniqueness of weak solutions. Up to our knowledge, none of the known methods, such as [42, 43], seem to be applicable to Eq. (1).

**Theorem 2** is proved in two steps: First, we show (strong) error estimates between particles of systems (3) and (6), respectively; see Proposition 3 below. Second, we show the weak convergence of (a subsequence of) $\rho_\sigma$ to a solution $\rho$ to (1); see Corollary 13.

**Proposition 3** (Error estimate for the stochastic system). Let $X_i^N$ and $\hat{X}_i^N$ be the solutions to (3) and (6), respectively. We assume that $\zeta^{-2s-1} \leq C_1 N^{1/4}$ for some constant $C_1 > 0$. Let $\delta \in (0, 1/4)$ and $a := \min\{1, d - 2s\} > 0$. Then, there exist constants $\varepsilon > 0$, depending on $\sigma$ and $\delta$, and $C_2 > 0$, depending on $\sigma$ and $T$, such that if $\beta^{-3d-7} \leq \varepsilon \log N$ then,

\[
\mathbb{E} \left( \sup_{0 \leq s \leq T} \max_{i=1, \ldots, N} \left( |X_i^N - \hat{X}_i^N(s)| \right) \right) \leq C_2 (\beta + \zeta^a) \to 0 \quad \text{as} \quad (N, \zeta, \beta) \to (\infty, 0, 0).
\]

The proposition is proved by estimating the differences

\[
E_1(t) := \mathbb{E} \left( \sup_{0 \leq s \leq T} \max_{i=1, \ldots, N} \left( |X_i^N - \hat{X}_i^N(s)| \right) \right),
\]

\[
E_2(t) := \mathbb{E} \left( \sup_{0 \leq s \leq T} \max_{i=1, \ldots, N} \left( |\hat{X}_i^N - \hat{X}_i^N(s)| \right) \right).
\]
and applying the triangle inequality. For the first difference, we estimate expressions like $k D^k K u$ for appropriate functions $u$ and $D^k W_{\beta}$ for $k \in \mathbb{N}$ in terms of negative powers of $\beta$ (here, $D^k$ denotes the $k$th-order partial derivatives). Using properties of Riesz potentials, in particular Hardy–Littlewood–Sobolev-type inequalities (see Lemmas 22 and 23), we show that for some $\mu_1 > 0$ ($i = 1, 2, 3$),

$$E_1(t) \leq C(\sigma) \beta^{-\mu_1} \int_0^t E_1(s) ds + C(\sigma) \beta^{-\mu_2} \zeta^{-\mu_3} N^{-1/2}.$$ 

By applying the Gronwall lemma and choosing a logarithmic scaling for $\beta$ and an algebraic scaling for $f$ with respect to $N$, we infer that $E_1(t) \leq C(\sigma) \beta_0 N^{l_4}$ for some $l_4 > 0$ ($i = 1, 2, 3$).

For the second difference $E_2$, we need the estimates $k W_{\beta} u$ for appropriate functions $u$ and $k K f$ for $k \geq N$ in terms of negative powers of $\beta$. Using properties of Riesz potentials, in particular Hardy–Littlewood–Sobolev-type inequalities (see Lemmas 22 and 23), we show that for some $l_4 > 0$ ($i = 1, 2, 3$),

$$E_2(t) \leq C(\sigma) (\beta + \zeta^a) + C(\sigma) \int_0^t E_2(s) ds,$$

and we conclude with Gronwall’s lemma that $E_2(t) \leq C(\sigma) (\beta + \zeta^a)$.

The article is organized as follows. The existence of global nonnegative weak solutions to (1) is proved in Section 2 by establishing an existence analysis for (7) and performing the limit $\sigma \to 0$. Some uniform estimates for the solution $\rho_{\sigma, \beta, \zeta}$ to (5) and for the difference $\rho_{\sigma, \beta, \zeta} - \rho_{\sigma}$ are shown in Section 3. Section 4 is devoted to the proof of the error estimate in Theorem 3 and the propagation of chaos in Theorem 2. In Appendices A–C we recall some auxiliary results and Hardy–Littlewood–Sobolev inequalities that are recalled in Lemmas 22 and 23. Then, after suitable computations, $E_2(t) \leq C(\sigma) (\beta + \zeta^a)$.

1.3.1. Notation

We write $\| \cdot \|_p$ for the $L^p(\mathbb{R}^d)$ or $L^p(\mathbb{R})$ norm with $1 \leq p \leq \infty$. The ball around the origin with radius $R > 0$ is denoted by $B_R$. The partial derivative $\partial / \partial x_i$ is abbreviated as $\partial_i$ for $i = 1, \ldots, d$, and $D^x$ denotes a partial derivative of order $|x|$, where $x \in \mathbb{N}_0^d$ is a multi-index. The notation $D^k$ refers to the $k$th-order tensor of partial derivatives of order $k \in \mathbb{N}$. In this situation, the norm $\|D^k u\|_p$ is the sum of all $L^p$ norms of partial derivatives of $u$ of order $k$. Finally, $C_0 > 0$, $C_1 > 0$, etc. denote generic constants with values changing from line to line.

2. Analysis of Eq. (1)

In this section, we prove the existence of global nonnegative weak solutions to (1). We first prove the existence of a solution $\rho_\sigma$ to (7) by a fixed-point argument, and then, perform the limit $\sigma \to 0$. In Subsection 2.1, we prove some basic estimates for a strong
solution \( \rho_\sigma \) to (7). Entropy and moment estimates as well as higher-order estimates are derived in Subsections 2.2 and 2.3, respectively. The existence of a unique strong solution to (7) is proved in Subsection 2.4 using a regularized version of (7) and Banach’s fixed-point theorem. The strong \( L^1(\mathbb{R}^d) \) limit \( \sigma \to 0 \) is performed in Subsection 2.5 using the div-curl lemma. Finally, Subsection 2.6 is concerned with the proof of a time-uniform weak \( L^1(\mathbb{R}^d) \) limit of \( (\rho_\sigma) \), which is needed in the proof of Proposition 3.

Recall definition (12) of the number \( \kappa_\sigma \), which is stated in (iv) below.

**Proposition 4.** Let Hypotheses (H1)-(H3) hold. Then, for all \( \sigma > 0 \), there exists a unique weak solution \( \rho_\sigma \geq 0 \) to (7) satisfying (i) the regularity
\[
\rho_\sigma \in L^\infty(0, \infty; L^1(\mathbb{R}^d)) \cap L^\infty(\mathbb{R}^d) \cap C^0 \left( [0, \infty); L^2(\mathbb{R}^d) \right),
\]
\[
\nabla \rho_\sigma \in L^2(0, \infty; L^2(\mathbb{R}^d)), \quad \partial_t \rho_\sigma \in L^2(0, \infty; H^{-1}(\mathbb{R}^d)),
\]

(ii) the weak formulation of (7) with test functions \( \phi \in L^2(0, T; H^1(\mathbb{R}^d)) \), (iii) the initial datum \( \rho_\sigma(0) = \rho_\sigma^0 \) in \( L^2(\mathbb{R}^d) \), and (vi) the following properties for \( t > 0 \), which are uniform in \( \sigma \) for sufficiently small \( \sigma > 0 \):

- **Mass conservation:** \( \| \rho_\sigma(t) \|_1 = \| \rho^0 \|_1 \).
- **Dissipation of the \( L^\infty \) norm:** \( \| \rho_\sigma \|_{L^\infty(0, \infty; L^\infty(\mathbb{R}^d))} \leq \kappa_\sigma \| \rho^0 \|_{L^\infty(\mathbb{R}^d)} \leq C \| \rho^0 \|_{L^\infty(\mathbb{R}^d)} \).
- **Moment estimate:** \( \sup_{t \in [0, \infty)} \int_{\mathbb{R}^d} \rho_\sigma(x,t) |x|^q dx \leq C_T \).
- **Entropy inequality:**
\[
\int_{\mathbb{R}^d} h(\rho_\sigma(T)) dx + 4\sigma \int_0^T \int_{\mathbb{R}^d} f_\sigma(\rho_\sigma) \sqrt{\nabla \rho_\sigma}^2 dx dt
\]
\[
+ \int_0^T \int_{\mathbb{R}^d} \nabla (-\Delta)^{\gamma/2} f_\sigma(\rho_\sigma) \sqrt{\nabla \rho_\sigma}^2 dx dt \leq \int_{\mathbb{R}^d} h(\rho_\sigma^0) dx \quad \text{for all } T > 0.
\]

Additionally, for any \( T > 0, 1 < p < \infty, \text{ and } 2 \leq q < \infty \), there exists \( C > 0 \), depending on \( T, \sigma, p, \text{ and } q \), such that
\[
\| \rho_\sigma \|_{L^p(0, T; W^{1,p}(\mathbb{R}^d))} + \| \partial_t \rho_\sigma \|_{L^p(0, T; W^{1,p}(\mathbb{R}^d))} + \| \rho_\sigma \|_{C^0([0, T]; W^{2,1}(\mathbb{R}^d) \cap W^{3,q}(\mathbb{R}^d))} \leq C,
\]
i.e., \( \rho_\sigma \) is even a strong solution to (7) and \( \rho_\sigma \in C^0([0, T]; W^{2,1}(\mathbb{R}^d) \cap W^{3,q}(\mathbb{R}^d)) \) for \( q \geq 2 \).

**2.1. Basic estimates for \( \rho_\sigma \)**

We prove a priori estimates in \( L^p \) spaces and an energy-type estimate. Let \( \sigma \in (0, 1) \) and let \( \rho_\sigma \) be a nonnegative strong solution to (7). Integration of (7) in \( \mathbb{R}^d \) and the definition of \( \rho_\sigma^0 \) immediately yield the mass conservation
\[
\| \rho_\sigma(t) \|_1 = \| \rho_\sigma^0 \|_1 = \| \rho^0 \|_1 \quad \text{for } t > 0.
\]

**Lemma 5** (Energy-type estimate). Let \( F \in C^2([0, \infty)) \) be convex and let \( F(\rho_\sigma^0) \in L^1(\mathbb{R}^d) \). Then,
\[
\frac{d}{dt} \int_{\mathbb{R}^d} F(\rho_\sigma) \, dx = -\sigma \int_{\mathbb{R}^d} F''(\rho_\sigma) |\nabla \rho_\sigma|^2 \, dx \\
- \frac{c_{d,1-s}}{2} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \frac{(G(\rho_\sigma(x)) - G(\rho_\sigma(y))) (f_\sigma(\rho_\sigma(x)) - f_\sigma(\rho_\sigma(y)))}{|x-y|^{d+2(1-s)}} \, dx \, dy \leq 0,
\]

(14)

where \( G(u) := \int_0^u vF''(v) \, dv \) for \( u \geq 0 \) and \( c_{d,1-s} \) is defined after (2).

**Proof.** First, we assume that \( F'' \) is additionally bounded. Then, \( F'(\rho_\sigma) - F'(0) \) is an admissible test function in the weak formulation of (7), since \( |F'(\rho_\sigma) - F'(0)| \leq \|F''\|_\infty \|\rho_\sigma\| \). It follows from definition (B1) of the fractional Laplacian and integration by parts that

\[
\frac{d}{dt} \int_{\mathbb{R}^d} F(\rho_\sigma) \, dx + \sigma \int_{\mathbb{R}^d} F''(\rho_\sigma) |\nabla \rho_\sigma|^2 \, dx = -\int_{\mathbb{R}^d} F''(\rho_\sigma) \rho_\sigma \nabla \rho_\sigma \cdot \nabla (-\Delta)^{-s} f_\sigma(\rho_\sigma) \, dx \\
- \int_{\mathbb{R}^d} G(\rho_\sigma) \nabla (-\Delta)^{-s} f_\sigma(\rho_\sigma) \, dx = -c_{d,1-s} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} G(\rho_\sigma(x)) \frac{f_\sigma(\rho_\sigma(x)) - f_\sigma(\rho_\sigma(y))}{|x-y|^{d+2(1-s)}} \, dx \, dy.
\]

A symmetrization of the last integral yields (14).

In the general case, we introduce \( F_k(u) = F(0) + F'(0)u + \int_0^u \int_0^v \min\{F''(w), k\} \, dw \, dv \) for \( k > 0 \). Then, \( F_k(u) \) is bounded and (14) follows for \( F \) replaced by \( F_k \). The result follows after taking the limit \( k \to \infty \) using monotone convergence. \( \square \)

We need a bound on \( \kappa_\sigma \), defined in (12), to derive uniform \( L^\infty(\mathbb{R}^d) \) bounds for \( \rho_\sigma \).

**Lemma 6 (Bound for \( \kappa_\sigma \)).** There exists \( C > 0 \) such that, for sufficiently small \( \sigma > 0 \),

\[
1 \leq \kappa_\sigma \leq \frac{1}{1 - C\sigma E}, \quad \text{where } E := \frac{1}{\|\rho^0\|_1} \int_{\mathbb{R}^d} \left(1 + |x|^{2d/(d-2s)}\right) \rho^0(x) \, dx.
\]

**Proof.** By Young’s convolution inequality (Lemma 19), we have

\[
\int_{\mathbb{R}^d} (W_\sigma * \rho^0)(x) \Xi(\sigma x) \, dx \leq \| W_\sigma * \rho^0 \|_1 \leq \| W_\sigma \|_1 \| \rho^0 \|_1 = \| \rho^0 \|_1,
\]

which shows that \( \kappa_\sigma \geq 1 \). To prove the upper bound, we use the triangle inequality \( |x| \leq |x-y| + |y| \):

\[
\int_{\mathbb{R}^d} (W_\sigma * \rho^0)(x) \Xi(\sigma x) \, dx \geq \int_{\{|x| \leq 1/\sigma\}} \int_{\mathbb{R}^d} W_\sigma(x-y) \rho^0(y) \, dy \, dx \\
= \int_{\mathbb{R}^d} \left( \int_{\mathbb{R}^d} W_\sigma(x-y) \, dx \right) \rho^0(y) \, dy - \int_{\{|x| > 1/\sigma\}} \int_{\mathbb{R}^d} W_\sigma(x-y) \rho^0(y) \, dy \, dx \\
\geq \int_{\mathbb{R}^d} \rho^0(y) \, dy - \sigma^{2d/(d-2s)} \int_{\{|x| > 1/\sigma\}} \int_{\mathbb{R}^d} |x|^{2d/(d-2s)} W_\sigma(x-y) \rho^0(y) \, dy \, dx \\
\geq \int_{\mathbb{R}^d} \rho^0(y) \, dy - \sigma^{2d/(d-2s)} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |x-y|^{2d/(d-2s)} W_\sigma(x-y) \rho^0(y) \, dy \, dx \\
- \sigma^{2d/(d-2s)} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |y|^{2d/(d-2s)} W_\sigma(x-y) \rho^0(y) \, dy \, dx.
\]
Using the property \( \int_{\mathbb{R}^d} |z|^{2d/(d-2s)} W_\sigma(z) \, dz \leq C \sigma^{2d/(d-2s)} \) for the second term on the right-hand side and \( \|W_p\|_{L^1(\mathbb{R}^d)} = 1 \) for the third term, we find that
\[
\int_{\mathbb{R}^d} (W_\sigma * \rho^0)(x) \Xi(\sigma x) \, dx \geq \int_{\mathbb{R}^d} \rho^0(y) \, dy - C \sigma^{2d/(d-2s)} \int_{\mathbb{R}^d} \rho^0(y) \, dy
\]
\[
- \sigma^{2d/(d-2s)} \int_{\mathbb{R}^d} |y|^{2d/(d-2s)} \rho^0(y) \, dy.
\]
Because of \( \sigma^{2d/(d-2s)} \leq \sigma \leq 1 \), we obtain
\[
\frac{\|\rho^0\|_1}{\kappa_\sigma} = \int_{\mathbb{R}^d} (W_\sigma * \rho^0)(x) \Xi(\sigma x) \, dx \geq \int_{\mathbb{R}^d} \rho^0(y) \, dy - C \sigma \int_{\mathbb{R}^d} \left( 1 + |y|^{2d/(d-2s)} \right) \rho^0(y) \, dy
\]
\[
\geq \int_{\mathbb{R}^d} \rho^0(y) \, dy - C \sigma \int_{\mathbb{R}^d} \rho^0(y) \, dy \cdot E = \| \rho^0 \|_1 (1 - C \sigma E),
\]
which proves the lemma.

\[
\square
\]

**Lemma 7** (Bounds for \( \rho_\sigma \)). The following bounds hold:
\[
\| \rho_\sigma(t) \|_\infty \leq \kappa_\sigma \| \rho^0 \|_\infty \leq C \| \rho^0 \|_\infty, \quad t > 0,
\]
\[
\sqrt{\sigma} \| \rho_\sigma \|_{L^2(0,T;H^1(\mathbb{R}^d))} \leq \| \rho^0 \|_2,
\]
where (15) holds for sufficiently small \( \sigma > 0 \).

**Lemma 7** and mass conservation imply that \( \| \rho_\sigma(t) \|_p \) is bounded for all \( t > 0 \) and \( 1 \leq p \leq \infty \). Observe that \( \kappa_\sigma \to 1 \) as \( \sigma \to 0 \). So, if \( \rho_\sigma(t) \to \rho(t) \) a.e., the dissipation of the \( L^\infty \) norm follows, as stated in **Theorem 1** (iv).

**Proof.** The convexity of \( F \) shows that \( G \), defined in **Lemma 5**, is nondecreasing. Therefore, \( (d/dt) \int_{\mathbb{R}^d} F(\rho_\sigma) \, dx \leq 0 \) and
\[
\sup_{t>0} \int_{\mathbb{R}^d} F(\rho_\sigma(t)) \, dx \leq \int_{\mathbb{R}^d} F(\rho^0) \, dx.
\]
We choose a convex function \( F \in C^2([0,\infty)) \) such that \( F(u) = 0 \) for \( u \leq \| \rho^0_\sigma \|_\infty \), \( F(u) > 0 \) for \( u > \| \rho^0_\sigma \|_\infty \) and satisfying \( F(u) \leq Cu \) for \( u \to \infty \). Then,
\[
0 \leq \int_{\mathbb{R}^d} F(\rho_\sigma(t)) \, dx \leq \int_{\mathbb{R}^d} F(\rho^0_\sigma) \, dx = 0 \quad \text{for} \ t > 0.
\]
Consequently, \( \rho_\sigma(x,t) \leq \| \rho^0_\sigma \|_\infty \leq \kappa_\sigma \| \rho^0 \|_\infty \) for \( t > 0 \), showing the \( L^\infty(\mathbb{R}^d) \) bound. Finally, choosing \( F(u) = u^2 \) in **Lemma 5**, the \( L^2(0,T;H^1(\mathbb{R}^d)) \) estimate follows.

\[
\square
\]

### 2.2. Entropy and moment estimates

We need a fractional derivative estimate for \( f_\sigma(\rho_\sigma) \), which is not an immediate consequence of **Lemma 5**. To this end, we define the entropy density
\[
h_\sigma(u) = \int_0^u \int_1^w f_\sigma'(w) \, dw \, dv, \quad u \geq 0.
\]
The last inequality follows after integration of $f$.

In particular, for all $T > 0$, there exists $C > 0$ such that

$$
\|f_\sigma(\rho_\sigma)\|_{L^2(0,T; H^{-1}(\mathbb{R}^d))} \leq C.
$$

**Proof.** The idea is to apply Lemma 5. Since $h_\sigma \not\in C^2([0,\infty))$, we cannot use the lemma directly. Instead, we apply it to the regularized function

$$
h_\sigma^\delta(u) = \int_0^u \int_1^w f'_\sigma(w) \frac{dwdv}{w + \delta}, \quad u \geq 0,
$$

where $\delta > 0$. Choosing $F = h_\sigma^\delta$ in Lemma 5 gives

$$
\frac{d}{dt} \int_{\mathbb{R}^d} h_\sigma^\delta(\rho_\sigma) \, d\mathbf{x} + 4\sigma \int_{\mathbb{R}^d} f'_\sigma(\rho_\sigma) \left| \nabla \rho_\sigma^{1/2} \right|^2 \, d\mathbf{x} + \int_{\mathbb{R}^d} \left| \nabla(-\Delta)^{-s/2} f_\sigma(\rho_\sigma) \right|^2 \, d\mathbf{x} = 0.
$$

**Step 1:** Estimate of $h_\sigma^\delta$. The pointwise limit $h_\sigma^\delta(\rho_\sigma) \rightarrow h_\sigma(\rho_\sigma)$ holds a.e. in $\mathbb{R}^d \times (0,T)$ as $\delta \rightarrow 0$. We observe that for all $0 < u \leq 1$,

$$
|h_\sigma^\delta(u)| \leq \sup_{0 < v < 1} f'(v) \int_0^u \int_v^1 \frac{dw}{w} \, dv \leq C(u \log u + 1),
$$

while for all $u > 1$, since $f'_\sigma \geq 0$ in $[0,\infty)$,

$$
|h_\sigma^\delta(u)| \leq \int_0^1 \int_v^1 f'_\sigma(w) \frac{dwdv}{w + \delta} + \int_1^u \int_1^w f'_\sigma(w) \frac{dwdv}{w + \delta} \leq C + \int_1^u f_\sigma(v) \, dv \leq C + uf_\sigma(u).
$$

The last inequality follows after integration of $f_\sigma(v) \leq f_\sigma(v) + vf_\sigma'(v) = (vf_\sigma(v))'$ in $(0,u)$. Therefore, since $\rho_\sigma \leq \|\rho_\sigma^\theta\|_\infty$ a.e. in $\mathbb{R}^d \times (0,\infty)$, we find that

$$
|h_\sigma^\delta(\rho_\sigma)| \leq C\rho_\sigma\left(\|\log \rho_\sigma\|_1 + 1\right) \mathbf{1}_{\{\rho_\sigma \leq 1\}} + C\mathbf{1}_{\{\rho_\sigma > 1\}} \leq C \left(\rho_\sigma^\theta + \rho_\sigma\right),
$$

where $\theta \in (0,1)$ is arbitrary, and consequently, because of mass conservation,

$$
\int_{\mathbb{R}^d} |h_\sigma^\delta(\rho_\sigma)| \, d\mathbf{x} \leq C + \int_{\mathbb{R}^d} \rho_\sigma^\theta \, d\mathbf{x}.
$$

**Step 2:** Estimate of $\int_{\mathbb{R}^d} \rho_\sigma^\theta \, d\mathbf{x}$. Let $0 < \alpha < 1$ and $d/(d + \alpha) < \theta < 1$. Then, by Young’s inequality,
\[
\int_{\mathbb{R}^d} \rho_\sigma^0 dx = \int_{\mathbb{R}^d} (1 + |x|^2)^{x/2} \rho_\sigma^0 (1 + |x|^2)^{-x/2} dx
\]
\[
\leq \int_{\mathbb{R}^d} (1 + |x|^2)^{x/2} \rho_\sigma dx + C \int_{\mathbb{R}^d} (1 + |x|^2)^{-x/(2(1-\theta))} dx
\]
\[
\leq \int_{\mathbb{R}^d} (1 + |x|^2)^{x/2} \rho_\sigma dx + C,
\]

since the choice of \( \theta \) guarantees that 
\(-x/(2(1-\theta)) < -d/2 \), so \( \int_{\mathbb{R}^d} (1 + |x|^2)^{-x/(2(1-\theta))} dx < \infty \). To control the right-hand side, we need to bound a suitable moment of \( \rho_\sigma \).

For this, we use the test function \((1 + |x|^2)^{x/2} \xi_k\) in the weak formulation of (7), where \( \xi_k \in C_0^2(\mathbb{R}^d) \) is a cutoff function with the properties

\[
\xi_k(x) = 1 \quad \text{for} \quad |x| \leq k, \quad \xi_k(x) = 0 \quad \text{for} \quad |x| \geq 2k, \quad k|\nabla \xi_k(x)| + k^2|\Delta \xi_k(x)| \leq C, \quad 0 \leq \xi_k(x) \leq 1 \quad \text{for} \quad x \in \mathbb{R}^d,
\]

and \( k > 1 \) is arbitrary. We find that

\[
\int_{\mathbb{R}^d} (1 + |x|^2)^{x/2} \xi_k \rho_\sigma(t) dx = \int_{\mathbb{R}^d} (1 + |x|^2)^{x/2} \xi_k \rho_\sigma^0 dx + \int_0^t \int_{\mathbb{R}^d} \rho_\sigma \xi_k \Delta (1 + |x|^2)^{x/2} dx ds
\]
\[
+ \sigma \int_0^t \int_{\mathbb{R}^d} \rho_\sigma (2 \nabla \left[ (1 + |x|^2)^{x/2} \right] \cdot \nabla \xi_k + (1 + |x|^2)^{x/2} \Delta \xi_k) dx ds
\]
\[- \alpha \int_0^t \int_{\mathbb{R}^d} \rho_\sigma \xi_k (1 + |x|^2)^{x/2-1} x \cdot \nabla (-\Delta)^{-s/2} f_\sigma(\rho_\sigma) dx ds
\]
\[- \int_0^t \int_{\mathbb{R}^d} \rho_\sigma (1 + |x|^2)^{x/2} \nabla \xi_k \cdot \nabla (-\Delta)^{-s/2} f_\sigma(\rho_\sigma) dx ds.
\]

Since \( \alpha < 1 \) and \( 0 \leq \xi_k \leq 1 \) in \( \mathbb{R}^d \), the terms involving \( \Delta (1 + |x|^2)^{x/2} \) and \( x(1 + |x|^2)^{x/2-1} \) are bounded in \( \mathbb{R}^d \). It follows from the choice of \( \xi_k \) that

\[
\left| \nabla \left[ (1 + |x|^2)^{x/2} \right] \cdot \nabla \xi_k \right| + \left| (1 + |x|^2)^{x/2} \Delta \xi_k \right| \leq C k^{x-2}, \quad (1 + |x|^2)^{x/2} |\nabla \xi_k| \leq C k^{x-1}.
\]

Thus, taking into account the assumption on \( \rho^0 \) and mass conservation,

\[
\sup_{0 < t < T} \int_{\mathbb{R}^d} (1 + |x|^2)^{x/2} \xi_k \rho_\sigma(t) dx \leq C + C \int_0^T \int_{\mathbb{R}^d} \rho_\sigma (-\Delta)^{-s/2} |\nabla (-\Delta)^{-s/2} f_\sigma(\rho_\sigma)| dx dt.
\]

Next, we apply the Hardy–Littlewood–Sobolev inequality (see Lemma 22) and the Hölder inequality and use the fact that \( \rho_\sigma(t) \) is bounded in any \( L^p(\mathbb{R}^d) \):

\[
\sup_{0 < t < T} \int_{\mathbb{R}^d} (1 + |x|^2)^{x/2} \rho_\sigma(x, t) \xi_k(x) dx
\]
\[
\leq C + C \int_0^T \| \rho_\sigma \|_{2d/(d+2s)} \left\| (-\Delta)^{-s/2} \left[ \nabla (-\Delta)^{-s/2} f_\sigma(\rho_\sigma) \right] \right\|_{2d/(d-2s)} dt
\]
\[
\leq C + \int_0^T \| \rho_\sigma \|_{2d/(d+2s)} \| \nabla (-\Delta)^{-s/2} f_\sigma(\rho_\sigma) \|_2 dt
\]
\[
\leq C(\eta) + \eta \int_0^T \| \nabla (-\Delta)^{-s/2} f_\sigma(\rho_\sigma) \|_2^2 dt
\]
for all \( \eta > 0 \). Since \( \xi_k(x) \leq \xi_{k+1}(x) \) for \( x \in \mathbb{R}^d \), \( k > 1 \), and \( \xi_k \to 1 \) a.e. in \( \mathbb{R}^d \) as \( k \to \infty \), we deduce from monotone convergence that in the limit \( k \to \infty \),

\[
\sup_{0 < t < T} \int_{\mathbb{R}^d} (1 + |x|^2)^{s/2} \rho_\sigma(x, t) dx \leq C(\eta) + \eta \int_0^T \| \nabla (-\Delta)^{-s/2} f_\sigma(\rho_\sigma) \|^2 dt
\]

for all \( \eta > 0 \). This proves that

\[
\int_{\mathbb{R}^d} \rho_\sigma^0 dx \leq C(\eta) + \eta \int_0^T \| \nabla (-\Delta)^{-s/2} f_\sigma(\rho_\sigma) \|^2 dt.
\]

**Step 3: A priori estimate.** Inserting the previous estimate into (19) leads to

\[
\sup_{0 < t < T} \int_{\mathbb{R}^d} |h_\sigma^0(\rho_\sigma(x, t))| dx \leq C(\eta) + \eta \int_0^T \| \nabla (-\Delta)^{-s/2} f_\sigma(\rho_\sigma) \|^2 dt.
\]

We integrate (18) in time and use the previous estimate:

\[
4\sigma \int_0^T \int_{\mathbb{R}^d} f_\sigma^t(\rho_\sigma) \frac{\rho_\sigma}{\rho_\sigma + \delta} |\nabla \rho_\sigma^{1/2}|^2 dx dt
\]

\[
+ \frac{c_{d,1-s}}{2} \int_0^T \int_{\mathbb{R}^d} \left( f_\sigma(\rho_\sigma(x)) - f_\sigma(\rho_\sigma(y)) \right) \left( \rho_\sigma(\rho_\sigma(x)) - \rho_\sigma(\rho_\sigma(y)) \right) \frac{1}{|x - y|^{d+2(1-s)}} dx dy dt
\]

\[
\leq \int_{\mathbb{R}^d} |h_\sigma^0(\rho_\sigma(T))| dx + \int_{\mathbb{R}^d} |h_\sigma^0(\rho_\sigma)| dx \leq C(\eta) + \eta \int_0^T \| \nabla (-\Delta)^{-s/2} f_\sigma(\rho_\sigma) \|^2 dt.
\]

We wish to pass to the limit \( \delta \to 0 \) in the previous inequality. We deduce from dominated convergence that \( f_\sigma^0(\rho_\sigma) \to f_\sigma(\rho_\sigma) \) a.e. in \( \mathbb{R}^d \times [0, \infty) \). The integrand of the second term on the left-hand side is nonnegative, and we obtain from Fatou’s lemma that

\[
4\sigma \int_0^T \int_{\mathbb{R}^d} f_\sigma^t(\rho_\sigma) |\nabla \rho_\sigma^{1/2}|^2 dx dt + \frac{c_{d,1-s}}{2} \int_0^T \int_{\mathbb{R}^d} \frac{\left( f_\sigma(\rho_\sigma(x)) - f_\sigma(\rho_\sigma(y)) \right)^2}{|x - y|^{d+2(1-s)}} dx dy dt
\]

\[
\leq C(\eta) + \eta \int_0^T \| \nabla (-\Delta)^{-s/2} f_\sigma(\rho_\sigma) \|^2 dt.
\]

By the integral representation of the fractional Laplacian,

\[
\frac{c_{d,1-s}}{2} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \frac{(f_\sigma(\rho_\sigma(x)) - f_\sigma(\rho_\sigma(y)))^2}{|x - y|^{d+2(1-s)}} dx dy = \| \nabla (-\Delta)^{-s/2} f_\sigma(\rho_\sigma) \|^2_{L^2},
\]

the last term in (20) can be absorbed for sufficiently small \( \eta > 0 \) by the second term on the left-hand side. This leads to the estimate

\[
4\sigma \int_0^T \int_{\mathbb{R}^d} f_\sigma^t(\rho_\sigma) |\nabla \rho_\sigma^{1/2}|^2 dx dt + \int_0^T \int_{\mathbb{R}^d} |\nabla (-\Delta)^{-s/2} f_\sigma(\rho_\sigma)|^2 dx dt \leq C.
\]

Thus, we can pass to the limit \( \delta \to 0 \) in (18) giving the desired entropy balance. Finally, bound (17) follows from the definition of the \( H^{1-s}(\mathbb{R}^d) \) norm and the facts that
Lemma 9 (Moment estimate). It holds that
\[
\sup_{0 < t < T} \int_{\mathbb{R}^d} \rho_\sigma(x, t) |x|^{2d/(d-2s)} \, dx \leq C,
\]
where \( C > 0 \) depends on \( T \) and the \( L^1(\mathbb{R}^d) \) norms of \( \rho^0 \) and \( | \cdot |^{2d/(d-2s)} \rho^0 \).

Proof. For the following computations, we would need to use cutoff functions to make the calculations rigorous. We leave the details to the reader, as we wish to simplify the presentation. Let \( m = 2d/(d-2s) \). Since \( | \cdot |^m \rho^0 \in L^1(\mathbb{R}^d) \) by assumption, we can compute
\[
\frac{d}{dt} \int_{\mathbb{R}^d} \rho_\sigma(t) \frac{|x|^m}{m} \, dx = \sigma(m - 2 + d) \int_{\mathbb{R}^d} |x|^{m-2} \rho_\sigma dx - \int_{\mathbb{R}^d} \rho_\sigma |x|^{m-2} x \cdot \nabla (-\Delta)^{-s} f_\sigma(\rho_\sigma) \, dx \\
\leq C \| | \rho_\sigma - 1 | + \| | |^{-1} \rho_\sigma \|_{2d/(d+2s)} \| \nabla (-\Delta)^{-s} f_\sigma(\rho_\sigma) \|_{2d/(d-2s)}.
\]

By Young’s inequality and mass conservation, we have
\[
\| | \rho_\sigma \|_1 \leq C \int_{\mathbb{R}^d} (1 + |x|^m) \rho_\sigma dx \leq C \int_{\mathbb{R}^d} |x|^m \rho_\sigma dx.
\]
It follows from (17) that \( \nabla (-\Delta)^{-s} f_\sigma(\rho_\sigma) \) is bounded in \( L^2(0, T; H^s(\mathbb{R}^d)) \). In particular, because of the Sobolev embedding \( H^s(\mathbb{R}^d) \hookrightarrow L^m(\mathbb{R}^d) \),
\[
\| \nabla (-\Delta)^{-s} f_\sigma(\rho_\sigma) \|_{L^2(0, T; L^m(\mathbb{R}^d))} \leq C.
\]
Furthermore, using \( \rho_\sigma \in L^\infty(0, \infty; L^\infty(\mathbb{R}^d)) \), Young’s inequality, and the property \( 2d/(d+2s) \geq 1 \) (recall that \( d \geq 2 \))
\[
\| | \rho_\sigma \|_{2d/(d+2s)} \leq \int_{\mathbb{R}^d} \rho_\sigma^{2d/(d+2s)} |x|^{2d(m-1)/(d+2s)} \, dx \\
\leq C \int_{\mathbb{R}^d} \rho_\sigma |x|^{2d(m-1)/(d+2s)} \, dx.
\]
Thus, we infer from (21) and the identity \( 2d(m-1)/(d+2s) = m \) that
\[
\frac{d}{dt} \int_{\mathbb{R}^d} \rho_\sigma(t) \frac{|x|^m}{m} \, dx \leq C \int_{\mathbb{R}^d} \rho_\sigma(t) |x|^m \, dx,
\]
and Gronwall’s lemma concludes the proof.

2.3. Higher-order estimate

We need some estimates in higher-order Sobolev spaces.

Proposition 10 (Higher-order regularity). Let \( T > 0, 1 < p < \infty, \) and \( 2 \leq q < \infty. \) Then, there exists \( C > 0, \) depending on \( T, \sigma, p, \) and \( q, \) such that
\[
f_\sigma(\rho_\sigma) \in L^2(\mathbb{R}^d) \quad \text{since} \quad f_\sigma \text{ is locally Lipschitz continuous,} \quad f_\sigma(0) = 0, \quad \text{and} \quad \rho_\sigma \text{ is bounded both in } L^\infty(\mathbb{R}^d) \quad \text{and} \quad L^2(\mathbb{R}^d) \end{align*}

independently of \( \sigma. \) \( \square \)
\[
\|\rho_\sigma\|_{L^p(0,T;W^{3,p}(\mathbb{R}^d))} + \|\partial_t \rho_\sigma\|_{L^p(0,T;W^{1,p}(\mathbb{R}^d))} + \|\rho_\sigma\|_{C^0([0,T];W^{2,q}(\mathbb{R}^d))} \leq C.
\]

**Proof.**  

**Step 1:** Case \( s > 1/2. \) If \( s > 1/2, \) then, \( w := \rho_\sigma \nabla(-\Delta)^{-s} f_\sigma(\rho_\sigma) \) does not involve any derivative of \( \rho_\sigma. \) Thus, \( w \in L^p(0,T;L^p(\mathbb{R}^d)) \) for \( p < \infty \) and Lemma 26 in Appendix C implies that \( \rho_\sigma \in L^p(0,T;W^{1,p}(\mathbb{R}^d)). \) Iterating the argument leads to the conclusion. Thus, in the following, we can assume that \( 0 < s \leq 1/2. \)

**Step 2:** Estimate of \( \text{div} w \) in \( L^p(0,T;W^{-1,p}(\mathbb{R}^d)) \). We claim that \( w \) can be estimated in \( L^p(0,T;L^p(\mathbb{R}^d)) \) for any \( p < \infty. \) Then, by Lemma 26, \( \nabla \rho_\sigma \in L^p(0,T;L^p(\mathbb{R}^d)). \) We use the \( L^\infty \) bound for \( \rho_\sigma, \) the fractional Gagliardo–Nirenberg inequality (Lemma 24), and Young’s inequality to find that

\[
\|w\|_p \leq C\|\nabla(-\Delta)^{-s} f_\sigma(\rho_\sigma)\|_p \leq C\|f_\sigma(\rho_\sigma)\|_2^2 \|\nabla f_\sigma(\rho_\sigma)\|_p^{1-2s} \leq C(\eta) + \eta\|\nabla \rho_\sigma\|_p,
\]

where \( \eta > 0 \) is arbitrary. By estimate (C2) in Lemma 26,

\[
\|\rho_\sigma \nabla(-\Delta)^{-s} f_\sigma(\rho_\sigma)\|_p = \|w\|_p \leq C(\eta) + \eta\left(\|\rho_\sigma \nabla(-\Delta)^{-s} f_\sigma(\rho_\sigma)\|_p + T^{1/p}\|\nabla \rho^0\|_p\right).
\]

Choosing \( \eta > 0 \) sufficiently small shows the claim.

**Step 3:** Estimate of \( \text{div} w \) in \( L^p(0,T;L^p(\mathbb{R}^d)) \). We use Hölder’s inequality with \( 1/p = 2s/(d+p) + 1/q \) to obtain

\[
\|\text{div} w\|_p \leq \|\nabla \rho_\sigma \cdot \nabla(-\Delta)^{-s} f_\sigma(\rho_\sigma)\|_p + \|\rho_\sigma(-\Delta)^{-s} f_\sigma(\rho_\sigma)\|_p \\
\leq \|\nabla \rho_\sigma\|_{(d+p)/(2s)} \|\nabla(-\Delta)^{-s} f_\sigma(\rho_\sigma)\|_q + C\|(-\Delta)^{-s} f_\sigma(\rho_\sigma)\|_p.
\]

By the fractional Gagliardo–Nirenberg inequality (Lemma 25 with \( \theta = 1 + d/p - d/q - 2s \) and Lemma 24 with \( s \) replaced by \( 1 - s \)) and Young’s inequality, it follows that

\[
\|\text{div} w\|_p \leq C\|\nabla \rho_\sigma\|_{(d+p)/(2s)} \|f_\sigma(\rho_\sigma)\|_{p}^{1-\theta} \|\nabla f_\sigma(\rho_\sigma)\|_p^\theta + C\|f_\sigma(\rho_\sigma)\|_p^\theta \|D^2 f_\sigma(\rho_\sigma)\|_p^{1-s} \\
\leq C\|\nabla \rho_\sigma\|_{(d+p)/(2s)} \|\nabla \rho_\sigma\|_{p}^\theta + C\|f_\sigma(\rho_\sigma)\|_p^\theta D^2 \rho_\sigma + f_\sigma''(\rho_\sigma)\nabla \rho_\sigma \otimes \nabla \rho_\sigma\|_p^{1-s} \\
\leq C(\eta) + C\|\nabla \rho_\sigma\|_{(d+p)/(2s)} \|\nabla \rho_\sigma\|_p + C\|\nabla \rho_\sigma\|_{2p}^2 + \eta\|D^2 \rho_\sigma\|_p^	heta,
\]

where \( \eta > 0 \) is arbitrary. Taking the \( L^p(0,T) \) norm of the previous inequality and observing that \( p/(1-\theta) = (d+p)/(2s) \) (because of \( \theta = d(1/p - 1/q) + 1 - 2s \)), it follows that

\[
\|\text{div} w\|_{L^p(0,T;L^p(\mathbb{R}^d))} \leq C + C\|\nabla \rho_\sigma\|_{L^p(0,T;L^p(\mathbb{R}^d))}^{1/(1-\theta)} \|\nabla \rho_\sigma\|_{L^p(0,T;L^p(\mathbb{R}^d))} + C\|\nabla \rho_\sigma\|_{L^p(0,T;L^p(\mathbb{R}^d))} \\
+ C\|\nabla \rho_\sigma\|_{L^p(0,T;L^p(\mathbb{R}^d))}^{2} + \eta\|D^2 \rho_\sigma\|_{L^p(0,T;L^p(\mathbb{R}^d))}.
\]

Lemma 26 and Step 2 (\( \nabla \rho_\sigma \in L^p(0,T;L^p(\mathbb{R}^d)) \)) show that

\[
\|\partial_t \rho_\sigma\|_{L^p(0,T;L^p(\mathbb{R}^d))} + (1 - C\eta)\|D^2 \rho_\sigma\|_{L^p(0,T;L^p(\mathbb{R}^d))} \leq C.
\]

Choosing \( \eta > 0 \) sufficiently small, this yields \( \partial_t \rho_\sigma \in L^p(0,T;L^p(\mathbb{R}^d)) \) and \( \rho_\sigma \in L^p(0,T;W^{2,p}(\mathbb{R}^d)). \) We deduce from Lemma 20, applied to \( \nabla \rho_\sigma, \) that \( \nabla \rho_\sigma \in L^\infty(0,T;L^q(\mathbb{R}^d)) \) for any \( 2 \leq q < \infty. \) (At this point, we need the restriction \( q \geq 2. \))
Step 4: Higher-order regularity. To improve the regularity of $\rho_\sigma$, we differentiate (7) in space. Recall that $\partial_i = \partial/\partial x_i$, $i = 1, \ldots, d$. Then,

$$
\begin{aligned}
\partial_t \partial_i \rho_\sigma - \sigma \Delta \partial_i \rho_\sigma &= \sum_{j=1}^d \partial_i \partial_j \left( \rho_\sigma \partial_j (\Delta)^{-1} f_\sigma (\rho_\sigma) \right) = \sum_{j=1}^d \left( \partial_{ij}^2 \rho_\sigma \partial_j (\Delta)^{-1} f_\sigma (\rho_\sigma) + \partial_t \rho_\sigma \partial_{ij}^2 (\Delta)^{-1} f_\sigma (\rho_\sigma) + \partial_{ij} \rho_\sigma \partial_j (\Delta)^{-1} f_\sigma (\rho_\sigma) \right).
\end{aligned}
$$

We estimate the right-hand side term by term. Let $0 < s \leq 1/2$. First, by Hölder’s inequality with $1/p = 1/q + 1/r$, $1 < p < q < \infty$, $\max\{2,p\} < r < \infty$ and the fractional Gagliardo–Nirenberg inequality (Lemma 24),

$$
\begin{aligned}
\|\partial_{ij}^2 \rho_\sigma \partial_j (\Delta)^{-1} f_\sigma (\rho_\sigma) \|^p_{L^p(0,T;L^q(\mathbb{R}^d))} &\leq \int_0^T \|\partial_{ij}^2 \rho_\sigma \|_q^p \|\partial_j (\Delta)^{-1} f_\sigma (\rho_\sigma) \|^p \| \, dt \\
&\leq C \int_0^T \|\partial_{ij}^2 \rho_\sigma \|_q^p \|f_\sigma (\rho_\sigma) \|(1-2s/p) \|\nabla f_\sigma (\rho_\sigma) \|^{2p} \| \, dt \\
&\leq C \|f_\sigma (\rho_\sigma) \|_{L^\infty(0,T;L^q(\mathbb{R}^d))} \|\nabla f_\sigma (\rho_\sigma) \|^{2p} \int_0^T \|\partial_{ij}^2 \rho_\sigma \|_q^p \, dt \leq C.
\end{aligned}
$$

The second and third term on the right-hand side of (22) can be treated in a similar way, observing that $\partial_{ij}^2 (\Delta)^{-1} = \partial_j (\Delta)^{-1} \partial_i$. The last term is estimated according to

$$
\begin{aligned}
\|\partial_{ij} \rho_\sigma \partial_j (\Delta)^{-1} f_\sigma (\rho_\sigma) \|_p &\leq C \|\partial_{ij} \rho_\sigma \|_q \|\partial_j (\Delta)^{-1} f_\sigma (\rho_\sigma) \|_p \\
&\leq C \|\partial_{ij}^2 f_\sigma (\rho_\sigma) \|_p \|\nabla \partial_{ij}^2 f_\sigma (\rho_\sigma) \|^{1-2s} \| \|_p \\
&\leq C(\eta) \|\partial_{ij}^2 f_\sigma (\rho_\sigma) \|_p + \eta \|\nabla \partial_{ij}^2 f_\sigma (\rho_\sigma) \|^{1-2s} \| \|_p,
\end{aligned}
$$

and the last expression can be absorbed by the corresponding estimate of $\Delta \partial_i \rho_\sigma$ from the left-hand side of (22). Then, we deduce from Lemma 26 that $\partial_t \partial_i \rho_\sigma$, $\partial_{ij} \rho_\sigma \in L^p(0,T;L^p(\mathbb{R}^d))$ for all $p > 1$ and Lemma 20, applied to $\partial_{ij}^2 \rho_\sigma$, yields $\partial_{ij} \rho_\sigma \in C^0([0,T];L^q(\mathbb{R}^d))$ for all $q \geq 2$.

Next, if $1/2 < s < 1$, we use the second inequality in Lemma 24 and argue similarly as before. This finishes the proof.

Lemma 11. Under the assumptions of Proposition 10, for every $q \geq 2$, there exists a constant $C = C(q) > 0$, depending on $\sigma$, such that

$$
\|\rho_\sigma \|_{C^0([0,T];W^{3,q}(\mathbb{R}^d)) \cap W^{3,q}(\mathbb{R}^d)} \leq C.
$$

The embedding $W^{3,q}(\mathbb{R}^d) \hookrightarrow W^{2,\infty}(\mathbb{R}^d)$ for $q > d$ yields a bound for $\rho_\sigma$ in $C^0([0,T];W^{2,\infty}(\mathbb{R}^d))$.

Proof. We first prove the bound in $C^0([0,T];W^{3,q}(\mathbb{R}^d))$. By differentiating (7) twice in space, estimating similarly as in Step 4 of the previous proof, and using the regularity results of Proposition 10, we can show that $\rho_\sigma$ is bounded in $L^\infty(0,T;W^{3,q}(\mathbb{R}^d))$ for any $q \geq 2$.

It remains to show the $C^0([0,T];W^{2,1}(\mathbb{R}^d))$ bound for $\rho_\sigma$. In view of mass conservation and Gagliardo–Nirenberg–Sobolev’s inequality, it suffices to show a bound for $D^2 \rho_\sigma$ in $L^\infty(0,T;L^1(\mathbb{R}^d))$. To this end, we define the weights $\gamma_n = (1 + |x|^2)^{n/2}$ for $n \geq 0$ and test Eq. (7) for $\rho_\sigma$ with $v_n := \gamma_n \rho_\sigma$. Then,
\[ \partial_t v_n - \sigma \Delta v_n = \text{div} \left( v_n \nabla K \ast f_\sigma (\rho_n) \right) + I_n, \quad v_n(0) = \gamma_n^0 \rho_\sigma^0 \quad \text{in } \mathbb{R}^d, \]

where \( I_n = -2 \sigma \nabla \gamma_n \cdot \nabla \rho_\sigma - \sigma \rho_\sigma \Delta \gamma_n - \rho_\sigma \nabla \gamma_n \cdot \nabla K \ast f_\sigma (\rho_n). \)

Arguing as in Step 4 of the previous proof, we can find a bound in \( L^\infty \left( 0, T; W^{2, p}(\mathbb{R}^d) \right) \) for \( v_n \). Indeed, we can proceed by induction over \( n \), since the additional terms in \( I_n \) can be controlled by Sobolev norms of \( v_0, \ldots, v_{n-1} \). The definition of \( \rho_\sigma^0 \) implies that \( \gamma_n^0, \gamma_n^0 \nabla \rho_\sigma^0 \in L^\infty(\mathbb{R}^d) \cap L^1(\mathbb{R}^d) \) for every \( n \geq 0 \). Then, choosing \( n > d \) yields, for \( 0 \leq t \leq T \), that
\[
\| \gamma_n \partial_t^2 \rho_\sigma \|_p \leq \| \partial_t^2 (\gamma_n \rho_\sigma) \|_p + 2 \| \nabla \gamma_n \cdot \nabla \rho_\sigma \|_p + \| \rho_\sigma \partial_t \nabla \gamma_n \|_p \leq C(T).
\]

We conclude from \( \gamma_n^{-1} \in L^\infty(\mathbb{R}^d) \cap L^1(\mathbb{R}^d) \) that
\[
\| \partial_t \rho_\sigma \|_1 \leq \| \gamma_n^{-1} \|_{\mathbb{P}/(p-1)} \| \partial_t \rho_\sigma \|_p \leq C(T).
\]

This proves the desired bound. \( \square \)

### 2.4. Existence of solutions to (7)

We show that the regularized Eq. (7) possesses a unique strong solution \( \rho_\sigma \).

**Step 1: Existence for an approximated system.** Let \( T > 0 \) arbitrary, define the spaces

\[
X_T := L^2 \left( 0, T; H^1(\mathbb{R}^d) \right) \cap H^1 \left( 0, T; H^{-1}(\mathbb{R}^d) \right) \hookrightarrow Y_T := C^0(\{0, T\}; L^2(\mathbb{R}^d)),
\]

\[
Y_{T, R} := \{ u \in Y_T : \| u \|_T \leq R \},
\]

and consider the mapping \( S : v \in Y_T \mapsto u \in Y_T, \)

\[
\partial_t u - \sigma \Delta u = \text{div} \left( u \nabla \mathcal{K}_s(\delta) \ast f_\sigma(\eta) (v) \right) \quad \text{in } \mathbb{R}^d \times (0, T),
\]

\[
u(0) = \rho_\sigma^0 \quad \text{in } \mathbb{R}^d,
\]

where \( \mathcal{K}_s(\delta) : \mathbb{R}^d \rightarrow \mathbb{R}_+ \) is a regularized version of \( \mathcal{K}_s \), defined by

\[
\mathcal{K}_s(\delta) = \mathcal{K}_s(\delta) \ast \mathcal{K}_s(\delta/2) \mathcal{K}_s(\delta/2) (x) = c_{d, s/2} \left\{ \begin{array}{ll}
\delta^{d-s} + (s-d) \delta^{d-1} \left| x \right| & \text{for } |x| < \delta, \\
|x|^{d-s} & \text{for } \delta \leq |x| \leq \delta^{-1}, \\
\left[ \delta^{d-s} + (s-d) \delta^{d+1-s} \left| x \right| \right]_+ & \text{for } |x| > \delta^{-1},
\end{array} \right.
\]

and \( f_\sigma(\eta) \) is given by

\[
f_\sigma(\eta)(\rho) = \int_0^{|\rho|} f_\sigma'(u) \min(1, u \eta^{-1}) du + \frac{\eta^2}{2} \rho^2, \quad \rho \in \mathbb{R}.
\]

The regularization with parameter \( \eta \) is needed for the entropy estimates.

**Lemma 12.** For any \( 0 < s < 1 \) and a.e. \( x \in \mathbb{R}^d \), the function \( \delta \mapsto \mathcal{K}_s(\delta)(x) \) is nonincreasing for \( \delta \in (0, 1) \).
Proof. Let $r^*_\delta = (d - s + 1)/(d - s)\delta$. We can write $\Phi_{\delta}(x) = c_{d-s/2}\Phi_{\delta}(|x|)$ with

$$
\Phi_{\delta}(r) = \left\{
\begin{array}{ll}
\delta^{d-1} + (s - d)\delta^{d-2}(r - \delta) & \text{for } r < \delta, \\
r^{d-1} & \text{for } \delta \leq r \leq \delta^{-1}, \\
\delta^{d-1} + (s - d)\delta^{d-1}(r - \delta^{-1}) & \text{for } \delta^{-1} < r < r^*_\delta, \\
0 & \text{for } r \geq r^*_\delta.
\end{array}
\right.
$$

Then, $\Phi_{\delta} \in C^0([0,\infty)) \cap C^1(0,r^*_\delta)$, and its derivative equals

$$
\Phi'_{\delta}(r) = \left\{
\begin{array}{ll}
-(d-s)\max\{r,\delta\}^{d-1} & 0 \leq r \leq 1, \\
-(d-s)\min\{r,\delta^{-1}\}^{d-1} & 1 \leq r < r^*_\delta.
\end{array}
\right.
$$

We show that $\Phi_{\delta}(r)$ is nonincreasing in $\delta \in (0,1)$ for $r \geq 1$. We have for $1 \leq r < r^*_\delta$,

$$
\Phi_{\delta}(r) = \Phi_{\delta}(1) + \int_1^r \Phi'_{\delta}(u)du = 1 - (d-s)\int_1^r \min\{u,\delta^{-1}\}^{d-1}du.
$$

Furthermore, we have $\Phi_{\delta}(r^*_\delta) = 0$, while $\min\{u,\delta^{-1}\} = \delta^{-1} > 0$ for $u > r^*_\delta$, so it holds that

$$
\Phi_{\delta}(r) = \left(1 - (d-s)\int_1^r \min\{u,\delta^{-1}\}^{d-1}du\right)_+ \text{ for } r \geq 1.
$$

At this point, the above representation formula together with elementary monotonicity considerations show that $\Phi_{\delta}(r)$ is nonincreasing in $\delta \in (0,1)$ for $r \geq 1$. It remains to show that $\Phi_{\delta}(r)$ is nonincreasing in $\delta \in (0,1)$ for $0 \leq r < 1$. It holds that

$$
\Phi_{\delta}(r) = \Phi_{\delta}(1) - \int_r^1 \Phi'_{\delta}(u)du = 1 + (d-s)\int_r^1 \max\{u,\delta^{-1}\}^{d-1}du \text{ for } 0 \leq r < 1.
$$

Once again, we conclude from the above representation formula together with elementary monotonicity considerations that $\Phi_{\delta}(r)$ is nonincreasing in $\delta \in (0,1)$ for $0 \leq r < 1$. This finishes the proof. \qed

We derive some estimates for $f^{(\eta)}_{\sigma}$. First, we have $0 \leq f^{(\eta)}_{\sigma}(\rho) \leq C_\eta \rho^2$ for $\rho \in \mathbb{R}$, since

$$
f^{(\eta)}_{\sigma}(\rho) \leq \left(\eta + \eta^{-1}\max\{0,\eta\}\frac{\rho^2}{2}\right) \text{ for } |\rho| \leq \eta,
$$

$$
f^{(\eta)}_{\sigma}(\rho) \leq f_{\sigma}(|\rho|) + \frac{\eta}{2}\rho^2 \leq \left(\|f_\sigma\|_\infty \eta^{-2} + \frac{\eta}{2}\right)\rho^2 \text{ for } |\rho| > \eta.
$$

Furthermore,

$$
|Df^{(\eta)}_{\sigma}(\rho)| = \frac{\rho}{|\rho|}f'_{\sigma}(|\rho|)\min(1,|\rho|\eta^{-1}) + \eta|\rho| \leq (\eta + \eta^{-1}\|f'_\sigma\|_\infty)|\rho|,
$$

which implies that $|Df^{(\eta)}_{\sigma}(\rho)| \leq C_\eta |\rho|$ for $\rho \in \mathbb{R}$. This shows that there exists $C(\eta) > 0$ such that for any $\rho_1, \rho_2 \in \mathbb{R}$,

$$
|f^{(\eta)}_{\sigma}(\rho_1) - f^{(\eta)}_{\sigma}(\rho_2)| \leq C(\eta)(|\rho_1| + |\rho_2|)|\rho_1 - \rho_2|.
$$

It follows that $f^{(\eta)}_{\sigma}(v) \in L^\infty(0,T;L^1(\mathbb{R}^d))$ for $v \in Y_T$. 
Since $\nabla K_s^{(\delta)} \in L^\infty(\mathbb{R}^d)$, a standard argument shows that (23) has a unique solution $u \in X_T \hookrightarrow Y_T$. Therefore, the mapping $S$ is well-defined. Additionally, the nonnegativity of $u$ follows immediately after by testing (23) with $\min(0, u)$.

We show now that $S$ is a contraction on $Y_{T, R}$ for sufficiently small $T > 0$. We start with a preparation. By testing (23) with $u$ and taking into account the $L^\infty$ bound for $\nabla K_s^{(\delta)}$, we deduce from Young’s inequality for products and convolutions that
\[
\int_{\mathbb{R}^d} u(t)^2 \, dx + \frac{\sigma}{2} \int_0^t \int_{\mathbb{R}^d} |\nabla u(t)|^2 \, dx \, dt \leq C(\delta, \eta, \sigma) \int_0^t \|u(\tau)^2\|_2 \, d\tau,
\]
since $\|f_\sigma^{(\eta)}(v)\|_1 \leq C_\eta \|v\|_2^3$ for $v \in Y_T$. Then, if $v \in Y_{T, R}$, we infer from Gronwall’s lemma that
\[
\int_{\mathbb{R}^d} u(t)^2 \, dx + \int_0^t \int_{\mathbb{R}^d} |\nabla u(t)|^2 \, dx \, dt \leq e^{C(\sigma, \delta, \eta)R^t t} \int_{\mathbb{R}^d} |\rho_\sigma^{(\eta)}(x)| \, dx \quad \text{for } 0 \leq t \leq T. \tag{24}
\]

Let $v_i \in Y_{T, R}$ and set $u_i = S(v_i)$, $i = 1, 2$. We compute
\[
\|u_1 \nabla K_s^{(\delta)} \ast f_\sigma^{(\eta)}(v_1) - u_2 \nabla K_s^{(\delta)} \ast f_\sigma^{(\eta)}(v_2)\|_2
\]
\[
\leq \|u_1 - u_2\|_2 \|\nabla K_s^{(\delta)} \ast f_\sigma^{(\eta)}(v_1)\|_2 + \|u_2 \nabla K_s^{(\delta)} \ast f_\sigma^{(\eta)}(v_1) - f_\sigma^{(\eta)}(v_1) - f_\sigma^{(\eta)}(v_2)\|_2
\]
\[
\leq \|u_1 - u_2\|_2 \|\nabla K_s^{(\delta)} \ast f_\sigma^{(\eta)}(v_1)\|_\infty + \|u_2\|_2 \|\nabla K_s^{(\delta)} \ast f_\sigma^{(\eta)}(v_1) - f_\sigma^{(\eta)}(v_1) - f_\sigma^{(\eta)}(v_2)\|_1
\]
\[
\leq C(\delta, \eta) \left(\|u_1 - u_2\|_2 \|v_1\|_2 + \|u_2\|_2 \|v_1 - v_2\|_1 \right).
\]

Therefore, using (24), for $v_1, v_2 \in Y_{T, R}$,
\[
\|u_1 \nabla K_s^{(\delta)} \ast f_\sigma^{(\eta)}(v_1) - u_2 \nabla K_s^{(\delta)} \ast f_\sigma^{(\eta)}(v_2)\|_2 \leq C(\delta, \eta, R, T) \left(\|u_1 - u_2\|_2 + \|v_1 - v_2\|_1 \right). \tag{25}
\]

Next, we write (23) for $(u_i, v_i)$ in place of $(u, v)$, $i = 1, 2$, take the difference between the two equations, and test the resulting equation with $u_1 - u_2$:
\[
\frac{1}{2} \|(u_1 - u_2)(t)\|_2^2 + \sigma \int_0^t \int_{\mathbb{R}^d} |\nabla (u_1 - u_2)|^2 \, dx \, dt
\]
\[
= - \int_0^t \int_{\mathbb{R}^d} (u_1 \nabla K_s^{(\delta)} \ast f_\sigma^{(\eta)}(v_1) - u_2 \nabla K_s^{(\delta)} \ast f_\sigma^{(\eta)}(v_2)) \, dx \, dt
\]
\[
\leq \frac{\sigma}{2} \int_0^t \int_{\mathbb{R}^d} |\nabla (u_1 - u_2)|^2 \, dx \, dt + \frac{1}{2\sigma} \int_0^t \|u_1 \nabla K_s^{(\delta)} \ast f_\sigma^{(\eta)}(v_1) - u_2 \nabla K_s^{(\delta)} \ast f_\sigma^{(\eta)}(v_2)\|_2^2 \, dt.
\]

It follows from (25) that
\[
\|(u_1 - u_2)(t)\|_2^2 + \sigma \int_0^t \int_{\mathbb{R}^d} |\nabla (u_1 - u_2)|^2 \, dx \, dt
\]
\[
\leq C(\delta, \eta, R, T, \sigma) \int_0^t \left(\|(u_1 - u_2\|_2^2 + \|v_1 - v_2\|_1^2\right) \, dt,
\]
and we conclude from Gronwall’s lemma that
This inequality implies that $S$ is a contraction in $Y_{T,R}$, provided that $T$ is sufficiently small. Therefore, by Banach’s theorem, $S$ admits a unique fixed point $u \in Y_{T,R} \subset Y_T$ for $T > 0$ sufficiently small.

It remains to show that the local solution can be extended to a global one. To this end, we note that the function $u \in X_T$ satisfies (23) with $v = u$:

$$
\partial_t u - \sigma \Delta u = \text{div} \left( u \nabla \mathcal{K}_s(\delta) \ast f_\sigma^{(\eta)}(u) \right) \quad \text{in } \mathbb{R}^d \times (0, T),
$$

$$
u(\cdot, 0) = \rho_0' \quad \text{in } \mathbb{R}^d.
$$

Then, defining the truncated entropy density

$$
h^{(\eta)}(\rho) = \int_0^\rho \int_0^\rho D_f^{(\eta)}(v) v^{-1} dvdu, \quad \rho \geq 0,
$$

and testing (26) with $Dh^{(\eta)}(u)$ yields, in view of the definition of $\mathcal{K}_s^{(\delta)}$, that

$$
\int_{\mathbb{R}^d} h^{(\eta)}(u(t)) dx + \sigma \int_0^t \int_{\mathbb{R}^d} D_f^{(\eta)}(u) u^{-1} |\nabla u|^2 dxdt + (27)
\int_0^t \int_{\mathbb{R}^d} |\nabla \mathcal{K}_{s/2}^{(\delta)} \ast f_\sigma^{(\eta)}(u)|^2 dxdt = \int_{\mathbb{R}^d} h^{(\eta)}(\rho_0') dx
$$

for $0 \leq t \leq T$. This inequality and the definitions of $f_\sigma^{(\eta)}$ and $h^{(\eta)}$ yield a $(\delta, T)$-uniform bound for $u$ in $L^2(0, T; H^1(\mathbb{R}^d))$, which in turn (together with (26)) implies a $(\delta, T)$-uniform bound for $u$ in $X_T$, and a fortiori in $Y_T$. This means that the solution $u$ can be prolonged to the whole time interval $[0, \infty)$ and exists for all times.

Finally, we prove out that, since $\nabla \mathcal{K}_s^{(\delta)} \in L^2(\mathbb{R}^d)$, then, $\nabla \mathcal{K}_s^{(\delta)} \ast f_\sigma^{(\eta)}(u) \in L^\infty(0, T; L^2(\mathbb{R}^d))$ and so $u\nabla \mathcal{K}_s^{(\delta)} \ast f_\sigma^{(\eta)}(u) \in L^\infty(0, T; L^1(\mathbb{R}^d))$. This fact yields the conservation of mass for $u$, i.e., $\int_{\mathbb{R}^d} u(t) dx = \int_{\mathbb{R}^d} \rho_0' dx$ for $t > 0$. Indeed, it is sufficient to test (26) with a cutoff $\psi_R \in C_0^1(\mathbb{R}^d)$ satisfying $\psi_R(x) = 1$ for $|x| < R$, $\psi_R(x) = 0$ for $|x| > 2R$, $\nabla \psi_R(x) \leq CR^{-1}$ for $x \in \mathbb{R}^d$, and then, to take the limit $R \to \infty$.

Step 2: Limit $\delta \to 0$. Let $u^{(\delta)}$ be the solution to (26). An adaption of the proof of [24, lemma 1] shows that the embedding $H^1(\mathbb{R}^d) \cap L^1(\mathbb{R}^d, (1 + |x|^2)^{\kappa/2}) \hookrightarrow L^2(\mathbb{R}^d)$ is compact. Thus, because of the $\delta$-uniform bounds for $u^{(\delta)}$, the Aubin–Lions Lemma implies that (up to a subsequence) $u^{(\delta)} \to u$ strongly in $L^2(0, T; L^2(\mathbb{R}^d))$ for every $T > 0$. We wish now to study the convergence of the nonlinear and nonlocal terms in (26)–(27) as $\delta \to 0$.

It follows from (27) that (up to a subsequence)

$$
\nabla \mathcal{K}_{s/2}^{(\delta)} \ast f_\sigma^{(\eta)}(u^{(\delta)}) \to U \quad \text{weakly in } L^2(\mathbb{R}^d \times (0, T)) \quad \text{as } \delta \to 0.
$$

To identify the limit $U$, we first notice that, by construction, $0 \leq \mathcal{K}_{s/2}^{(\delta)} / \mathcal{K}_{s/2}$ a.e. in $\mathbb{R}^d$. Furthermore, the Hardy–Littlewood–Sobolev inequality, the bound for $f_\sigma^{(\eta)}$, and then, the Gagliardo–Nirenberg–Sobolev inequality yield that
\[
\|\mathcal{K}_{s/2} \ast f^{(n)}_{\sigma}(u)\|_{(d+2)/(d-s)} \leq C\|f^{(n)}_{\sigma}(u)\|_{(d+2)/(d+2s/d)} \leq C(\eta)\|u\|_{(2d+4)/(d+2s/d)}^2 \leq C(\eta)\|u\|_{(2d+2)/(d+2)}^2.
\]

Therefore, since \(u \in L^\infty(0, T; L^2(\mathbb{R}^d)) \cap L^2(0, T; H^1(\mathbb{R}^d))\),
\[
\int_0^T \|\mathcal{K}_{s/2} \ast f^{(n)}_{\sigma}(u)\|_{(d+2)/(d-s)} dt \leq C(\eta)\|u\|_{L^\infty(0, T; L^2(\mathbb{R}^d))}^2 \int_0^T \|\nabla u\|_{(d-s)/(d+2)}^2 dt \leq C(\eta, T),
\]
meaning that \(\mathcal{K}_{s/2} \ast f^{(n)}_{\sigma}(u) \in L^{(d+2)/(d-s)}(\mathbb{R}^d \times (0, T))\). Taking into account that \(f^{(n)}_{\sigma}(u) \geq 0\) and that \(\delta \mapsto \mathcal{K}_{s/2}(\delta) \in \mathbb{R}\) is nonincreasing (see Lemma 12), we deduce from monotone convergence that
\[
\mathcal{K}_{s/2} \ast f^{(n)}_{\sigma}(u) \rightarrow \mathcal{K}_{s/2} \ast f^{(n)}_{\sigma}(u) \quad \text{strongly in } L^{(d+2)/(d-s)}(\mathbb{R}^d \times (0, T)).
\] (29)

Furthermore, arguing as before and using the estimates for \(Df^{(n)}_{\sigma}\) leads to
\[
\left\|\mathcal{K}_{s/2} \ast \left(f^{(n)}_{\sigma}(u^{(\delta)}) - f^{(n)}_{\sigma}(u)\right)\right\|_{(d+2)/(d-s)} \leq \left\|\mathcal{K}_{s/2} \ast \left(f^{(n)}_{\sigma}(u^{(\delta)}) - f^{(n)}_{\sigma}(u)\right)\right\|_{(d+2)/(d-s)} \leq C\|f^{(n)}_{\sigma}(u^{(\delta)}) - f^{(n)}_{\sigma}(u)\|_{(d+2)/(d+2s/d)} \leq C(\eta)\|u^{(\delta)} - u\|_{(d+2)/(d+2s/d)} \leq C(\eta)\left(\|u\|_{L^2(\mathbb{R}^d)}^{(s+2)/(d+2)} \|\nabla u\|_{L^2(\mathbb{R}^d)}^{(d-s)/(d+2)} + \|u^{(\delta)}\|_{L^2(\mathbb{R}^d)}^{(s+2)/(d+2)} \|\nabla u^{(\delta)}\|_{L^2(\mathbb{R}^d)}^{(d-s)/(d+2)} \right). \]

Since \(u^{(\delta)}\) is bounded in \(L^\infty(0, T; L^2(\mathbb{R}^d)) \cap L^2(0, T; H^1(\mathbb{R}^d))\), it follows that (up to a subsequence) \(\mathcal{K}_{s/2} \ast \left(f^{(n)}_{\sigma}(u^{(\delta)}) - f^{(n)}_{\sigma}(u)\right)\) converges weakly to some limit in \(L^{(d+2)/(d-s)}(\mathbb{R}^d)\) \((\mathbb{R}^d \times (0, T))\). However, Hölder’s inequality and the fact that \(u^{(\delta)} \rightarrow u\) strongly in \(L^p(0, T; L^2(\mathbb{R}^d))\) for every \(2 \leq p < \infty\), which follows from
\[
\int_0^T \|u^{(\delta)} - u\|_{L^2(\mathbb{R}^d)}^2 dt \leq \sup_{0 < t < T} \|u^{(\delta)} - u\|_{L^2(\mathbb{R}^d)}^2 \int_0^T \|u^{(\delta)} - u\|_{L^2(\mathbb{R}^d)}^2 dt \rightarrow 0 \quad \text{as } \delta \rightarrow 0,
\]
imply that
\[
\mathcal{K}_{s/2} \ast \left(f^{(n)}_{\sigma}(u^{(\delta)}) - f^{(n)}_{\sigma}(u)\right) \rightarrow 0 \quad \text{strongly in } L^p(0, T; L^{(d+2)/(d-s)}(\mathbb{R}^d)), \quad p < \frac{d+2}{d-s}.
\]

We conclude that
\[
\mathcal{K}_{s/2} \ast \left(f^{(n)}_{\sigma}(u^{(\delta)}) - f^{(n)}_{\sigma}(u)\right) \rightarrow 0 \quad \text{weakly in } L^{(d+2)/(d-s)}(\mathbb{R}^d \times (0, T)).
\] (30)

We deduce from (29) to (30) that
\[
\mathcal{K}_{s/2} \ast f^{(n)}_{\sigma}(u^{(\delta)}) - \mathcal{K}_{s/2} \ast f^{(n)}_{\sigma}(u) = \left(\mathcal{K}_{s/2} \ast f^{(n)}_{\sigma}(u) - \mathcal{K}_{s/2} \ast f^{(n)}_{\sigma}(u)\right) + \mathcal{K}_{s/2} \ast \left(f^{(n)}_{\sigma}(u^{(\delta)}) - f^{(n)}_{\sigma}(u)\right) \rightarrow 0 \quad \text{weakly in } L^{(d+2)/(d-s)}(\mathbb{R}^d \times (0, T)),
\]
which, together with (28), implies that \( U = \nabla K_s/2 * f_\sigma^{(\eta)}(u) \), that is,
\[
\nabla \tilde{K}^{(\delta)}_{s/2} * f_\sigma^{(\eta)}(u^{(\delta)}) \to \nabla K_{s/2} * f_\sigma^{(\eta)}(u) \quad \text{weakly in } L^2(\mathbb{R}^d \times (0, T)).
\]
(31)

Let \( \psi \in C_0^\infty(\mathbb{R}^d \times (0, T)) \). Because of
\[
\nabla K^{(\delta)}_s * f_\sigma^{(\eta)}(u^{(\delta)}) = \tilde{K}^{(\delta)}_{s/2} * \left( \nabla K^{(\delta)}_{s/2} * f_\sigma^{(\eta)}(u^{(\delta)}) \right),
\]
we find that
\[
\int_0^T \int_{\mathbb{R}^d} \psi \cdot \nabla K^{(\delta)}_s * f_\sigma^{(\eta)}(u^{(\delta)}) \, dx \, dt = \int_0^T \int_{\mathbb{R}^d} \left( \nabla K^{(\delta)}_{s/2} * f_\sigma^{(\eta)}(u^{(\delta)}) \right) \cdot \left( \nabla K^{(\delta)}_{s/2} * \psi \right) \, dx \, dt.
\]

Our goal is to show that \( \tilde{K}^{(\delta)}_{s/2} * \psi \to K_{s/2} * \psi \) strongly in \( L^2(\mathbb{R}^d \times (0, T)) \) as \( \delta \to 0 \). We can assume without loss of generality that \( \psi \geq 0 \) a.e. in \( \mathbb{R}^d \times (0, T) \). Indeed, for general functions \( \psi \), we may write \( \psi = \psi_+ + \psi_- \), where \( \psi_+ = \max\{0, \psi\} \) and \( \psi_- = \min\{0, \psi\} \), and we have \( \tilde{K}^{(\delta)}_{s/2} * \psi = \tilde{K}^{(\delta)}_{s/2} * \psi_+ - \tilde{K}^{(\delta)}_{s/2} * (-\psi_-) \). Once again, since \( \tilde{K}^{(\delta)}_{s/2} \to \tilde{K}_{s/2} \) a.e. in \( \mathbb{R}^d \), it is sufficient to show that \( K_{s/2} * \psi \in L^2(\mathbb{R}^d \times (0, T)) \). The Hardy–Littlewood–Sobolev inequality (see Appendix B) yields
\[
\int_0^T \| K_{s/2} * \psi \|_2^2 \, dt \leq C \int_0^T \| \psi \|_{2^{d/(d+2)}}^2 \, dt.
\]

It follows from (31), the previous argument, and the fact that \( K_s * u = (-\Delta)^{-1} u = K_{s/2} * K_{s/2} * u \) that
\[
\int_0^T \int_{\mathbb{R}^d} \psi \cdot \nabla K^{(\delta)}_s * f_\sigma^{(\eta)}(u^{(\delta)}) \, dx \, dt \to \int_0^T \int_{\mathbb{R}^d} \left( \nabla K^{(\delta)}_{s/2} * f_\sigma^{(\eta)}(u) \right) \cdot \left( K_{s/2} * \psi \right) \, dx \, dt
\]
\[
= \int_0^T \int_{\mathbb{R}^d} \psi \cdot \nabla K_s * f_\sigma^{(\eta)}(u) \, dx \, dt
\]
for every \( \psi \in L^2(0, T; L^{2d/(d+2)}(\mathbb{R}^d)) \), which means that
\[
\nabla K^{(\delta)}_s * f_\sigma^{(\eta)}(u^{(\delta)}) \to \nabla K_s * f_\sigma^{(\eta)}(u) \quad \text{weakly in } L^2(0, T; L^{2d/(d-2s)}(\mathbb{R}^d)).
\]
(32)

Since \( u^{(\delta)} \to u \) strongly in \( L^2(0, T; L^2(\mathbb{R}^d)) \) and \( (u^{(\delta)}) \) is bounded in \( L^\infty(0, T; L^1(\mathbb{R}^d)) \) (via mass conservation), it also holds that \( u^{(\delta)} \to u \) strongly in \( L^2(0, T; L^{2d/(d+2s)}(\mathbb{R}^d)) \). Therefore, the convergence (32) is sufficient to pass to the limit \( \delta \to 0 \) in (26).

**Step 3: Limit \( \eta \to 0 \) and conclusion.** The limit \( \delta \to 0 \) in (26) shows that the limit \( u \) solves
\[
\partial_t u - \sigma \Delta u = \text{div} \left( u \nabla K_s * f_\sigma^{(\eta)}(u) \right) \quad \text{in } \mathbb{R}^d \times (0, T),
\]
\[
u(t, 0) = \rho_0^0 \quad \text{in } \mathbb{R}^d.
\]
(33)

Fatou’s Lemma and the weakly lower semicontinuity of the \( L^2 \) norm allow us to infer from (27) that for \( t > 0 \),
\begin{equation}
\int_{\mathbb{R}^d} h(\eta(x,t)) dx + \sigma \int_0^t \int_{\mathbb{R}^d} Df_{\sigma}(u) u^{-1} |\nabla u|^2 dx dt
+ \int_0^t \int_{\mathbb{R}^d} |\nabla K_{s/2} * f_{\sigma}(u(t))|^2 dx dt \leq \int_{\mathbb{R}^d} h(\eta(0)) dx.
\end{equation}

At this point, all the bounds for \( u \), derived in the previous subsections, and the moment estimate, contained in Lemma 9, can be proved like in Subsections 2.1–2.2. All these estimates are uniform in \( \eta \). It is rather straightforward to perform the limit \( \eta \to 0 \) in (33)–(34) to obtain a weak solution to (7). However, the higher regularity bounds obtained in Subsection 2.3 imply that \( u \) is actually a strong solution to (7), which in turn yields the uniqueness of \( u \) as a weak solution to (7). This finishes the proof of Theorem 4.

2.5. Limit \( \sigma \to 0 \)

We prove that there exists a subsequence of \( \rho(\sigma) \) that converges strongly in \( L^1(\mathbb{R}^d \times (0,T)) \) to a weak solution \( \rho \) to (1).

The uniform \( L^\infty(\mathbb{R}^d \times (0,T)) \) bound for \( \rho(\sigma) \) in Lemma 7 implies that, up to a subsequence, \( \rho(\sigma) \to \rho \) weakly* in \( L^\infty(\mathbb{R}^d \times (0,T)) \) as \( \sigma \to 0 \). We deduce from the uniform \( L^\infty(0,T;L^1(\mathbb{R}^d)) \) bound (13) and the moment bound for \( \rho(\sigma) \) in Lemma 9 that \( \rho(\sigma) \) is equi-integrable. Thus, by the Dunford–Pettis theorem, again up to a subsequence, \( \rho(\sigma) \to \rho \) weakly in \( L^1(\mathbb{R}^d \times (0,T)) \). It follows from the \( L^2(0,T;H^1(\mathbb{R}^d)) \) estimate (16) that \( \sigma \Delta \rho(\sigma) \to 0 \) strongly in \( L^2(0,T;H^{-1}(\mathbb{R}^d)) \). The estimates in (17) and Lemma 7 show that \( \partial_t \rho(\sigma) \) is bounded in \( L^2(0,T;H^{-1}(\mathbb{R}^d)) \) and consequently, up to a subsequence, \( \partial_t \rho(\sigma) \to \partial_t \rho \) weakly in \( L^2(0,T;H^{-1}(\mathbb{R}^d)) \). Therefore, the limit \( \sigma \to 0 \) in (7) leads to
\begin{equation}
\partial_t \rho = \text{div} \left( \rho \nabla (\nabla f(\rho)) \right) \quad \text{in } L^2(0,T;H^{-1}(\mathbb{R}^d)),
\end{equation}
where the overline denotes the weak limit of the corresponding sequence.

We need to identify the weak limit on the right-hand side. The idea is to use the div-curl lemma [41, theorem 10.21]. For this, we define the vector fields with \( d+1 \) components
\[ U_\sigma := \left( \rho \nabla (\nabla f(\rho)), V_\sigma := (f(\rho,0,\ldots,0). \right\]
Let \( R > 0 \) be arbitrary and write \( B_R \) for the ball around the origin with radius \( R \). The \( L^\infty(\mathbb{R}^d) \) bound (15) for \( \rho \) and the \( L^2(0,T;H^{-1}(\mathbb{R}^d)) \) bound (17) for \( f(\rho) \) show that \( U_\sigma \) is bounded in \( L^p(B_R \times (0,T)) \) for some \( p > 1 \), while \( V_\sigma \) is bounded in \( L^\infty(B_R \times (0,T)) \). Furthermore, by (17),
\[ \text{div}(U_\sigma) = \sigma \Delta \rho(\sigma) \to 0 \quad \text{strongly in } L^2(0,T;H^{-1}(B_R)) \hookrightarrow H^{-1}(B_R \times (0,T)), \]
\[ \| \nabla f(\rho) \|_{L^2(0,T;H^{-1}(B_R))} \leq C \| \nabla f(\rho) \|_{L^2(0,T;H^{-1}(B_R))} \leq C, \]
where \( \text{curl}(V_\sigma) \) is the antisymmetric part of the Jacobian matrix of \( V_\sigma \). Hence, by the compact embedding \( H^{-s}(B_R \times (0,T)) \hookrightarrow W^{-1,1}(B_R \times (0,T)) \) (since \( L^2(0,T;H^{-1}(B_R)) \subset \)
We use the properties that

\[ \text{setting} \]

Then, since

\[ C \]

Taking into account the fundamental theorem of calculus for the function

\[ f(u) = f'(w) dw \tilde{\Xi}(\sigma u) du \]

This means that

\[ \text{uniformly bounded, we conclude that} \]

\[ \rho_\sigma f_\sigma(\rho_\sigma) = \rho f(\rho) \quad \text{a.e.
}\]

By definition (9) of \( f_\sigma(\rho_\sigma) \), it follows for arbitrary \( \rho_\sigma \in [0, L] \) and sufficiently large \( L > 0 \), that

\[ f_\sigma(\rho_\sigma) = \int_0^{\rho_\sigma} \left( \Gamma(1, [0, \infty)) (u) \tilde{\Xi}(u) du = \int_0^{\rho_\sigma} \Gamma_\sigma(u-w) f'(w) dw \tilde{\Xi}(\sigma u) du \]

\[ = \int_0^{\rho_\sigma} \Gamma_\sigma(u-w) f'(w) dw \tilde{\Xi}(\sigma u) du = \int_0^{\rho_\sigma} \left( \int_0^{\rho_\sigma} \Gamma_\sigma(u-w) \tilde{\Xi}(\sigma u) du \right) f(w) dw. \]

We use the properties that \( (\rho_\sigma) \) is uniformly bounded and \( \tilde{\Xi} = 1 \) in \([-1, 1]\). Then, choosing \( \sigma > 0 \) sufficiently small,

\[ f_\sigma(\rho_\sigma) = \int_0^{\rho_\sigma} \left( \int_0^{\rho_\sigma} \Gamma_\sigma(u-w) du \right) f(w) dw \]

\[ = \int_0^{\rho_\sigma} \Gamma_\sigma(u-w) f(w) dw - \int_0^{\rho_\sigma} \Gamma_\sigma(w) f(w) dw \]

\[ = \int_0^{\rho_\sigma} \Gamma_\sigma(u-w) f(w) dw, \]

setting \( \tilde{f} := f(1, [0, \infty)) \). Hence, using \( f(0) = 0 \), we find that

\[ f_\sigma(\rho_\sigma) - f(\rho) = \int_\mathbb{R} \Gamma_\sigma(u) \left( \tilde{f}(u + \rho_\sigma) - f(\rho_\sigma) \right) du - \int_\mathbb{R} \Gamma_\sigma(w) \left( \tilde{f}(w) - \tilde{f}(0) \right) dw. \]

Taking into account the fundamental theorem of calculus for the function \( \tilde{f} \in C^0 \cap W^{1,1}(\mathbb{R}) \), we can estimate as follows:

\[ |f_\sigma(\rho_\sigma) - f(\rho_\sigma)| \leq \text{ess sup}_{u \in \text{supp}(\Gamma_\sigma) \cap [0, \infty)} \left( \frac{\int_0^{\rho_\sigma} \Gamma_\sigma(u) \left( \tilde{f}(u + \rho_\sigma) - f(\rho_\sigma) \right) du}{u} \right) \int_\mathbb{R} \Gamma_\sigma(w)|w|dw \]

\[ \leq \left( \text{max}_{\xi \in \text{supp}(\Gamma_\sigma) \cap [0, \infty)} \left( f'(\xi + \rho_\sigma) + f'(\xi) \right) \right) \int_\mathbb{R} \Gamma_\sigma(w)|w|dw. \]

Then, since \( \Gamma_\sigma(u) = \sigma^{-1} \Gamma_1(\sigma^{-1} u) \), \( \text{supp}(\Gamma_\sigma) \subset B_\sigma(0) \) is compact, \( f \in C^1((0, \infty)) \), and \( (\rho_\sigma) \) is uniformly bounded, we conclude that

\[ |f_\sigma(\rho_\sigma) - f(\rho_\sigma)| \leq C\sigma. \]

This means that \( f_\sigma(\rho_\sigma) - f(\rho_\sigma) \to 0 \) strongly in \( L^\infty(B_R \times (0, T)) \), and it shows that \( \rho_\sigma f(\rho_\sigma) = \rho f(\rho) \) a.e. in \( B_R \times (0, T) \). As \( f \) is nondecreasing, we can apply [41, theorem 10.19] to infer that \( f(\rho_\sigma) = f(\rho) \) a.e. in \( B_R \times (0, T) \). Consequently, \( \rho_\sigma f(\rho_\sigma) = \rho f(\rho) \). As \( u \mapsto uf(u) \) is assumed to be strictly convex, we conclude from [41, theorem 10.20] that \( (\rho_\sigma) \) converges a.e. in \( B_R \times (0, T) \). Since \( (\rho_\sigma) \) is bounded in \( L^\infty(\mathbb{R}^d \times (0, T)) \), it follows that \( \rho_\sigma \to \rho \) strongly in \( L^p(B_R \times (0, T)) \) for all \( p < \infty \). Using the moment estimate from Lemma 9, we infer from

\[ \text{Lemma 9, we infer from} \]

\[ \text{Lemma 9, we infer from} \]
\[
\limsup_{\sigma \to 0} \int_0^T \int_{\mathbb{R}^d} |\rho_\sigma - \rho| dx dt = \limsup_{\sigma \to 0} \int_0^T \int_{\mathbb{R}^d \setminus B_R} |\rho_\sigma - \rho| dx dt \\
\leq R^{-2d/(d-2s)} \limsup_{\sigma \to 0} \int_0^T \int_{\mathbb{R}^d \setminus B_R} \rho_\sigma(t,x)|x|^{2d/(d-2s)} dx \\
\leq R^{-2d/(d-2s)} C \to 0 \quad \text{as} \quad R \to \infty
\]

that \(\rho_\sigma \to \rho\) strongly in \(L^p\left(\mathbb{R}^d \times (0, T)\right)\) for all \(p < \infty\). The strong convergences of \(\rho_\sigma\) and \(f_\sigma(\rho_\sigma)\) in \(L^p(\mathbb{R}^d \times (0, T))\) for all \(p < \infty\) allow us to identify the weak limit in (35), proving the weak formulation (8).

Finally, we deduce from the uniform \(L^2\left(0, T; H^{-1}(\mathbb{R}^d)\right)\) bound for \(\partial_t \rho_\sigma\) and the fact that \(\rho_\sigma \to \rho\) strongly in \(L^p(\mathbb{R}^d)\) for any \(p < \infty\) that \(\rho(0) = \rho^0\) in the sense of \(H^{-1}(\mathbb{R}^d)\). Properties (iv) of Theorem 1 follow from the corresponding expressions satisfied by \(\rho_\sigma\) in the limit \(\sigma \to 0\).

### 2.6. Time-uniform convergence of \((\rho_\sigma)\)

The following lemma is needed in the proof of Proposition 3. It is essentially a consequence of the \(L^2\left(0, T; H^{-1}(\mathbb{R}^d)\right)\) bound of \(\partial_t \rho_\sigma\) and the Ascoli–Arzelà theorem.

**Corollary 13.** Under the assumptions of Theorem 1, it holds for all \(\phi \in L^\infty(\mathbb{R}^d)\) that, possibly for a subsequence,

\[
\int_{\mathbb{R}^d} \rho_\sigma \phi dx \to \int_{\mathbb{R}^d} \rho \phi dx \quad \text{uniformly in} \quad [0, T].
\]

**Proof.** Let \(\phi \in C_0^1(\mathbb{R}^d)\) and \(0 \leq t_1 < t_2 \leq T\). The uniform \(L^2\left(0, T; H^{-1}(\mathbb{R}^d)\right)\) bound of \(\partial_t \rho_\sigma\) implies that

\[
\left| \int_{\mathbb{R}^d} \rho_\sigma(t_2) \phi dx - \int_{\mathbb{R}^d} \rho_\sigma(t_1) \phi dx \right| = \left| \int_{t_1}^{t_2} \langle \partial_t \rho_\sigma, \phi \rangle dt \right| \\
\leq |t_2 - t_1|^{1/2} \| \partial_t \rho_\sigma \|_{L^2(0, T; H^{-1}(\mathbb{R}^d))} \| \phi \|_{H^1(\mathbb{R}^d)} \leq C |t_2 - t_1|^{1/2} \| \phi \|_{H^1(\mathbb{R}^d)}.
\]

Hence, the sequence of functions \(t \mapsto \int_{\mathbb{R}^d} \rho_\sigma(t) \phi ds\) is bounded and equicontinuous in \([0, T]\). By the Ascoli–Arzelà theorem, up to a \(\phi\)-depending subsequence, \(\int_{\mathbb{R}^d} \rho_\sigma \phi dx \to \xi_\phi\), strongly in \(C^0([0, T])\) as \(\sigma \to 0\). Since \(\rho_\sigma \rightharpoonup^* \rho\) weakly* in \(L^\infty(0, T; L^\infty(\mathbb{R}^d))\), we can identify the limit, \(\xi_\phi = \int_{\mathbb{R}^d} \rho \phi dx\). Since \(H^1(\mathbb{R}^d)\) is separable, a Cantor diagonal argument together with a density argument allows us to find a subsequence (which is not relabeled) such that for all \(\phi \in H^1(\mathbb{R}^d)\),

\[
\int_{\mathbb{R}^d} \rho_\sigma \phi dx \to \int_{\mathbb{R}^d} \rho \phi dx \quad \text{strongly in} \quad C^0([0, T]).
\]

Since \((\rho_\sigma)\) is bounded in \(L^\infty(0, T; L^2(\mathbb{R}^d))\), another density argument shows that this limit also holds for all \(\phi \in L^2(\mathbb{R}^d)\).
Now, let $\phi \in L^\infty(\mathbb{R}^d)$. Using $\phi 1_{\{ |x| < R \}} \in L^2(\mathbb{R}^d)$, it follows from (36) and the moment estimate for $\rho_\sigma$ that
\[
\limsup_{\sigma \to 0} \sup_{0 < t < T} \left| \int_{\mathbb{R}^d} \rho_\sigma(t) \phi dx - \int_{\mathbb{R}^d} \rho(t) \phi dx \right| 
\leq \limsup_{\sigma \to 0} \sup_{0 < t < T} \left| \int_{\mathbb{R}^d} \rho_\sigma(t) \phi 1_{\{ |x| > R \}} dx - \int_{\mathbb{R}^d} \rho(t) \phi 1_{\{ |x| > R \}} dx \right|
\leq R^{-2d/(d-2s)} \| \phi \|_\infty \limsup_{\sigma \to 0} \sup_{0 < t < T} \int_{\mathbb{R}^d} (\rho_\sigma(x,t) + \rho(x,t)) |x|^{2d/(d-2s)} dx
\leq C(T) R^{-2d/(d-2s)} \| \phi \|_\infty \to 0 \quad \text{as } R \to \infty.
\]
This shows that
\[
\limsup_{\sigma \to 0} \sup_{0 < t < T} \left| \int_{\mathbb{R}^d} \rho_\sigma(t) \phi dx - \int_{\mathbb{R}^d} \rho(t) \phi dx \right| = 0,
\]
concluding the proof.

\section{Analysis of Eq. (5)}

This section is devoted to the analysis of Eq. (5),
\[
\partial_t \rho_{\sigma, \beta, \zeta} - \sigma \Delta \rho_{\sigma, \beta, \zeta} = \text{div} \left( \rho_{\sigma, \beta, \zeta} \nabla K_\zeta * f_\sigma(W_\beta * \rho_{\sigma, \beta, \zeta}) \right), \quad t > 0,
\]
\[
\rho_{\sigma, \beta, \zeta}(0) = \rho_\sigma^0 \quad \text{in } \mathbb{R}^d,
\]
where $K_\zeta = \tilde{K}_\zeta * W_\zeta$ and $W_\beta$ is defined in (10), as well as to an estimate for the difference $\rho_{\sigma, \beta, \zeta} - \rho_\sigma$, which is needed in the mean-field analysis. The existence and uniqueness of a strong solution to (37) follows from standard parabolic theory, since we regularized the singular kernel and smoothed the nonlinearity.

\begin{proposition} \textbf{(Uniform estimates).}
Let Hypotheses (H1)-(H3) hold and let $T > 0$, $p > d$. Set $a := \min\{1, d - 2s\}$, let $\rho_\sigma$ be the strong solution to (7), and let $\rho_{\sigma, \beta, \zeta}$ be the strong solution to (5). Then, there exist constants $C_1 > 0$, and $\varepsilon_0 > 0$, both depending on $\sigma$, $p$, and $T$, such that if $\beta + \zeta^a < \varepsilon_0$ then,
\[
\|\rho_{\sigma, \beta, \zeta} - \rho_\sigma\|_{L^\infty(0,T;W^{2,p}(\mathbb{R}^d))} \leq C_1(\beta + \zeta^a),
\]
\[
\|\rho_{\sigma, \beta, \zeta}\|_{L^\infty(0,T;W^{2,p}(\mathbb{R}^d))} \leq C_1.
\]
Furthermore, for every $q \geq 2$, there exists $C_2 = C_2(q) > 0$, depending on $\sigma$ and $T$, such that
\[
\| (K_\zeta - K) * \rho_\sigma \|_{L^\infty(0,T;L^\infty(\mathbb{R}^d))} \leq C_2 \zeta^a,
\]
\[
\|\rho_{\sigma, \beta, \zeta}\|_{L^\infty(0,T;W^{2,1}(\mathbb{R}^d) \cap W^{3,q}(\mathbb{R}^d))} \leq C_2.
\]
\end{proposition}

The proof is presented in the following subsections. The most difficult part is the proof of (38) in Subsection 3.1. We first prove an estimate for $D^2(\rho_{\sigma, \beta, \zeta} - \rho_\sigma)$ that depends on a lower-order estimate of this difference. Second, this lower-order estimate is shown
by testing the equation satisfied by the difference \( \rho_{\sigma, \beta, \zeta} - \rho_{\sigma} \) with a suitable nonlinear test function. Based on the arguments of this section, estimates (39)–(41) are, then, shown in Subsections 3.2–3.4, respectively.

### 3.1. Proof of (38)

We introduce the difference \( u := \rho_{\sigma, \beta, \zeta} - \rho_{\sigma} \), which satisfies

\[
\partial_t u - \sigma \Delta u = \text{div} \left[ (u + \rho_{\sigma}) \nabla K * f_{\sigma}(W_\beta * (u + \rho_{\sigma})) - \rho_{\sigma} \nabla K * f_{\sigma}(\rho_{\sigma}) \right] = D[u] + R[\rho_{\sigma}, u] + S[\rho_{\sigma}, u] \quad \text{in} \ \mathbb{R}^d, \ t > 0,
\]

and the initial datum \( u(0) = 0 \) in \( \mathbb{R}^d \), where

\[
D[\rho_{\sigma} u] = \text{div} \left[ u \nabla K * f_{\sigma}(W_\beta * u) \right], \quad R[\rho_{\sigma}, u] = \text{div} \left[ u \nabla K * (f_{\sigma}(W_\beta * (u + \rho_{\sigma})) - f_{\sigma}(W_\beta * u)) \right. \\
+ \rho_{\sigma} \nabla K * (f_{\sigma}(W_\beta * (u + \rho_{\sigma})) - f_{\sigma}(W_\beta * \rho_{\sigma})) + \rho_{\sigma} \nabla K * (f_{\sigma}(W_\beta * \rho_{\sigma}) - f_{\sigma}(\rho_{\sigma})) \left. \right], \quad S[\rho_{\sigma}, u] = \text{div} \left[ (u + \rho_{\sigma}) \nabla (K_{\zeta} - K) * f_{\sigma}(W_\beta * (u + \rho_{\sigma})) \right].
\]

We show first an estimate for \( D^2 u \) that depends on a lower-order estimate for \( u \).

**Lemma 15** (Conditional estimate for \( u \)). For any \( p > d \), there exists a number \( \Gamma_p \in (0, 1) \) such that, if \( \sup_{0 < t < T} \| u(t) \|_{W^{1,p}(\mathbb{R}^d)} \leq \Gamma_p \), then,

\[
\| D^2 u \|_{L^p(0, T; L^p(\mathbb{R}^d))} \leq C \left( \| u \|_{L^p(0, T; W^{1,p}(\mathbb{R}^d))} + \beta + \zeta^a \right),
\]

recalling that \( a = \min\{1, d - 2s\} \), and where \( C > 0 \) is independent of \( u, \beta, \) and \( \zeta \), but may depend on \( \sigma \).

**Proof.** Let \( \Gamma_p \in (0, 1) \) be such that \( \sup_{0 < t < T} \| u(t) \|_{W^{1,p}(\mathbb{R}^d)} \leq \Gamma_p \). We will find a constraint for \( \Gamma_p \) at the end of the proof. The aim is to derive an estimate for the right-hand side of (42) in \( L^p(0, T; L^p(\mathbb{R}^d)) \). We observe that \( \| u(t) \|_1 \leq \| \rho_{\sigma, \beta, \zeta} \|_1 + \| \rho_{\sigma} \|_1 \leq 2\| \rho^0 \|_1 \) for \( t \in [0, T] \). In the following, we denote by \( C > 0 \) a generic constant that may depend on \( \sigma \), without making this explicit. Furthermore, we denote by \( \mu \) a generic exponent in \((0, 1), \) whose value may vary from line to line.

**Step 1: Estimate of \( D[u] \).** Let \( 1/2 < s < 1 \). Then, by the Hardy–Littlewood–Sobolev-type inequality (B3),

\[
\| D[u] \|_p \leq \| \nabla u \cdot \nabla K * f_{\sigma}(W_\beta * u) \|_p + \| u \nabla K * [f'_{\sigma}(W_\beta * u)W_\beta * \nabla u] \|_p \\
\leq C\| \nabla u \|_p\| f_{\sigma}(W_\beta * u) \|_{d/(2s-1)} + C\| u \|_{d/(2s-1)}\| f'_{\sigma}(W_\beta * u) \|_{\infty}\| \nabla u \|_p.
\]

We use the Young convolution inequality, the Gagliardo–Nirenberg inequality, the smoothness of \( f_{\sigma} \), the property \( f_{\sigma}(0) = 0 \), and the fact \( \| W_{\beta} \|_{L^1(\mathbb{R}^d)} = 1 \) to estimate the terms on the right-hand side:
\[ \|W_\beta \ast u\|_\infty \leq \|u\|_\infty \leq \|u\|_{1-\frac{2}{d}} \|\nabla u\|_{\frac{d}{p}}^\lambda \leq C \Gamma_p^\lambda \leq C, \]

\[ \|f_\sigma(W_\beta \ast u)\|_\infty \leq \max_U |f'_\sigma| \|W_\beta \ast u\|_\infty \leq C, \]

\[ \|f_\sigma(W_\beta \ast u)\|_\infty \leq |f'_\sigma(0)| + \max_U |f''_\sigma| \|W_\beta \ast u\|_\infty \leq C, \]

where \( U := \left[ -||W_\beta \ast u\|_\infty, ||W_\beta \ast u\|_\infty \right] \) and \( \lambda > 0, \mu > 0. \) Therefore, \( \|D[u]\|_p \leq C\|\nabla u\|_p \)

Next, let \( 0 < s \leq 1/2. \) Then, we write

\[ D[u] = \nabla u \cdot \nabla + u^\lambda \nabla^2 u = D_1 + D_2 + D_3. \]

By the Hardy–Littlewood–Sobolev-type inequality (Lemma 22),

\[ \|D_1\|_p \leq C\|\nabla u\|_{d/(2s)} \|f'_\sigma(W_\beta \ast u)W_\beta \ast \nabla u\|_p \leq C\|\nabla u\|_{d/(2s)} \|\nabla u\|_p. \]

Next, we apply the Gagliardo–Nirenberg inequality with \( \lambda = (1 + 1/d - 2s/d)/(1 + 2/d - 1/p); \)

\[ \|\nabla u\|_{d/(2s)} \leq C\|u\|_{1-\frac{2}{d}} \|D^2 u\|_{\frac{d}{p}}^\lambda \leq C\|D^2 u\|_{\frac{d}{p}}^\lambda, \]

which is possible as long as \( \lambda \geq 1/2 \) or equivalently \( d \geq 2s, \) which is true. Consequently, using \( \Gamma_p \leq 1, \)

\[ \|D_1\|_p \leq C\|\nabla u\|_p \|D^2 u\|_{\frac{d}{p}}^\lambda \leq C\Gamma_p^\lambda \|\nabla u\|_{1-\frac{2}{d}} \|D^2 u\|_{\frac{d}{p}}^\lambda \leq C(\delta) \|\nabla u\|_p + \delta \|D^2 u\|_p, \]

where \( \delta > 0 \) is arbitrary. It follows from the Hardy–Littlewood–Sobolev-type inequality and the Gagliardo–Nirenberg inequality

\[ \|\nabla u\|_{2p}^2 \leq C\|D^2 u\|_{d/p} \|\nabla u\|_{p}^{2-2/d} \leq C\Gamma_p \|D^2 u\|_{p}^{d/p} \|\nabla u\|_{p}^{1-d/p} \]

that

\[ \|D_2\|_p \leq C\|u\|_{d/2s} \|\nabla u\|_{d/p} \|\nabla u\|_{p}^{1-d/p} \leq C(\delta) \|\nabla u\|_p + \delta \|D^2 u\|_p. \]

Finally, using similar ideas, we obtain

\[ \|D_3\|_p \leq C\|u\|_{d/(2s)} \|\Delta u\|_p \leq C\Gamma_p^\mu \|D^2 u\|_p. \]

Summarizing the estimates for \( D_1, D_2, \) and \( D_3 \) and integrating in time leads to

\[ \|D[u]\|_{L^p(0,T;L^p(\mathbb{R}^d))} \leq C\|u\|_{L^p(0,T;W^{1,p}(\mathbb{R}^d))} + C\Gamma_p^\mu \|D^2 u\|_{L^p(0,T;L^p(\mathbb{R}^d))}. \]

**Step 2: Estimate of \( R[p_\sigma, u] \).** We write \( R[p_\sigma, u] = R_1 + R_2 + R_3 \) for the three terms in the definition of \( R[p_\sigma, u] \) below (42).

**Step 2a: Estimate of \( R_1 \).** If \( s > 1/2 \), we can argue similarly as in the derivation of (43), which gives
We infer that (possibly with a different $k_R$)
\[
\left| f_\sigma''(W_\beta * (u + \rho_\sigma)) - f_\sigma''(W_\beta * u) \right| W_\beta * \Delta \rho_\sigma \leq C \left| f_\sigma''(W_\beta * (u + \rho_\sigma)) - f_\sigma''(W_\beta * u) \right| W_\beta * \Delta \rho_\sigma \leq C \Gamma_R^\mu \|D^2 u\|_p.
\]

All terms except the last one can be treated by the Hardy–Littlewood–Sobolev and Gagliardo–Nirenberg inequalities as before. For the last term, we use these inequalities and the $L^\infty(\mathbb{R}^d)$ bound for $\rho_\sigma$:
\[
\|R_{16}\|_p \leq C \|u\|_{d/2(3)} \left| f_\sigma''(W_\beta * (u + \rho_\sigma)) - f_\sigma''(W_\beta * u) \right| W_\beta * \Delta \rho_\sigma \leq C \|u\|_{d/2(3)} \left| f_\sigma''(W_\beta * (u + \rho_\sigma)) - f_\sigma''(W_\beta * u) \right| W_\beta * \Delta \rho_\sigma \leq C \|u\|_{d/2(3)} \|\Delta u\|_p \leq C \Gamma_R^\mu \|D^2 u\|_p.
\]

We infer that (possibly with a different $\mu > 0$ than before)
\[
\|R_1\|_{L^p(0,T;L^p(\mathbb{R}^d))} \leq C \|u\|_{L^p(0,T;L^p(\mathbb{R}^{2d}))} + C \Gamma_R^\mu \|D^2 u\|_{L^p(0,T;L^p(\mathbb{R}^d))}.
\]

**Step 2b: Estimate of $R_2$.** Since $|f_\sigma''|$ is bounded on the interval $[-\|\rho_\sigma\|_\infty, \|\rho_\sigma\|_\infty]$, we obtain for $s > 1/2$,
\[
\|R_2\|_{L^p(0,T;L^p(\mathbb{R}^d))} \leq C \|u\|_{L^p(0,T;L^p(\mathbb{R}^d))}.
\]

For $0 < s \leq 1/2$, we write $R_2 = R_{21} + \cdots + R_{27}$, where
\[
R_{21} = \nabla \rho_\sigma \cdot \mathcal{K} \left[ f_\sigma''(W_\beta * (u + \rho_\sigma)) W_\beta \cdot \nabla u \right],
\]
\[
R_{22} = \rho_\sigma \mathcal{K} \left[ f_\sigma''(W_\beta * (u + \rho_\sigma)) W_\beta \cdot \nabla (u + \rho_\sigma) \cdot (W_\beta \cdot \nabla u) \right],
\]
\[
R_{23} = \rho_\sigma \mathcal{K} \left[ f_\sigma''(W_\beta * (u + \rho_\sigma)) W_\beta \cdot \Delta u \right],
\]
\[
R_{24} = \nabla \rho_\sigma \cdot \mathcal{K} \left[ (f_\sigma''(W_\beta * (u + \rho_\sigma)) - f_\sigma''(W_\beta * \rho_\sigma)) W_\beta \cdot \nabla \rho_\sigma \right],
\]
\[
R_{25} = \rho_\sigma \mathcal{K} \left[ f_\sigma''(W_\beta * (u + \rho_\sigma)) (W_\beta \cdot \nabla u) \cdot (W_\beta \cdot \nabla \rho_\sigma) \right],
\]
\[
R_{26} = \rho_\sigma \mathcal{K} \left[ (f_\sigma''(W_\beta * (u + \rho_\sigma)) - f_\sigma''(W_\beta * \rho_\sigma)) W_\beta \cdot \Delta \rho_\sigma \right],
\]
\[
R_{27} = \rho_\sigma \mathcal{K} \left[ (f_\sigma''(W_\beta * (u + \rho_\sigma)) - f_\sigma''(W_\beta * \rho_\sigma)) W_\beta \cdot \Delta \rho_\sigma \right].
\]

Similar estimations as before allow us to treat all terms except the third one:
\[
\|R_{23}\|_p \leq \rho_\sigma \mathcal{K} \left[ (f_\sigma''(W_\beta * (u + \rho_\sigma)) - f_\sigma''(W_\beta * \rho_\sigma)) W_\beta \cdot \Delta u \right] \|_p
+ \rho_\sigma \mathcal{K} \left[ f_\sigma''(W_\beta * \rho_\sigma) W_\beta \cdot \Delta u \right] \|_p \leq: Q_{231} + Q_{232}.
\]

The first term can be estimated similarly as above by $Q_{231} \leq C \Gamma_R^\mu \|D^2 u\|_p$, while
\[ Q_{232} \leq \| \rho_\sigma \Delta K \ast [f'_\sigma (W_\beta \ast \rho_\sigma) W_\beta \ast u] \|_p + \| \rho_\sigma K \ast [\Delta f'_\sigma (W_\beta \ast \rho_\sigma) W_\beta \ast u] \|_p + 2 \| \rho_\sigma K \ast [\nabla f'_\sigma (W_\beta \ast \rho_\sigma) \cdot (W_\beta \ast \nabla u)] \|_p. \]

It follows from \(-\Delta K \ast v = (-\Delta)^{1-s} v\) and the fractional Gagliardo–Nirenberg inequality (Lemma 24) that
\[
Q_{232} \leq C \| u \|_{W^{1,r} (\mathbb{R}^d)} + \| \rho_\sigma (-\Delta)^{1-s} [f''_\sigma (W_\beta \ast \rho_\sigma) W_\beta \ast u] \|_p \\
\leq C \| u \|_{W^{1,r} (\mathbb{R}^d)} + C \| \rho_\sigma \|_\infty \| f''_\sigma (W_\beta \ast \rho_\sigma) W_\beta \ast u \|_p^2 \| D^2 [f''_\sigma (W_\beta \ast \rho_\sigma) W_\beta \ast u] \|_{p}^{1-s} \\
\leq C \| u \|_{W^{1,r} (\mathbb{R}^d)} + C \| u \|_p \left( \| u \|_p^{1-s} + \| \nabla u \|_p^{1-s} + \| D^2 u \|_p^{1-s} \right) \\
\leq C \| u \|_{W^{1,r} (\mathbb{R}^d)} + C \Gamma_p \| D^2 u \|_p.
\]

This shows that \(\| R_{23} \|_p \leq C \| u \|_{W^{1,r} (\mathbb{R}^d)} + C \Gamma_p \| D^2 u \|_p\), and we conclude that
\[
\| R_2 \|_{L^p (0, T; L^p (\mathbb{R}^d))} \leq C \| u \|_{L^p (0, T; W^{1,r} (\mathbb{R}^d))} + C \Gamma_p \| D^2 u \|_{L^p (0, T; L^p (\mathbb{R}^d))}.
\]

**Step 2c: Estimate of \(R_3\).** We write \(R_3 = R_{31} + \cdots + R_{37}\), where
\[
R_{31} = \nabla \rho_\sigma \cdot K \ast \left( (f'_\sigma (W_\beta \ast \rho_\sigma) - f'_\sigma (\rho_\sigma)) W_\beta \ast \nabla \rho_\sigma \right), \\
R_{32} = \rho_\sigma K \ast \left( f''_\sigma (W_\beta \ast \rho_\sigma) - f''_\sigma (\rho_\sigma) \right) | W_\beta \ast \nabla \rho_\sigma |^2, \\
R_{33} = \rho_\sigma K \ast \left( f''_\sigma (\rho_\sigma) (W_\beta \ast \nabla \rho_\sigma - \nabla \rho_\sigma) \cdot (W_\beta \ast \nabla \rho_\sigma) \right), \\
R_{34} = \nabla \rho_\sigma \cdot K \ast \left( f'_\sigma \rho_\sigma (W_\beta \ast \nabla \rho_\sigma - \nabla \rho_\sigma) \right), \\
R_{35} = \rho_\sigma K \ast \left( f''_\sigma \rho_\sigma \nabla \rho_\sigma \cdot (W_\beta \ast \nabla \rho_\sigma - \nabla \rho_\sigma) \right), \\
R_{36} = \rho_\sigma K \ast \left( f'_\sigma \rho_\sigma (W_\beta \ast \Delta \rho_\sigma - \Delta \rho_\sigma) \right), \\
R_{37} = \rho_\sigma K \ast \left( (f'_\sigma (W_\beta \ast \rho_\sigma) - f'_\sigma (\rho_\sigma)) W_\beta \ast \Delta \rho_\sigma \right).
\]

We start with the estimate of \(R_{31}\). We use the Hardy–Littlewood–Sobolev inequality (Lemma 22) and Lemma 21 to estimate \(W_\beta \ast \rho_\sigma - \rho_\sigma\):
\[
R_{31} \leq C \| \nabla \rho_\sigma \|_{d/s} \| f'_\sigma (W_\beta \ast \rho_\sigma) - f'_\sigma (\rho_\sigma) \|_p \| W_\beta \ast \nabla \rho_\sigma \|_{d/s} \\
\leq C \| \nabla \rho_\sigma \|_{d/s} \max_{\| \rho_\sigma \|_\infty} \| f''_\sigma \|_p \| W_\beta \ast \rho_\sigma - \rho_\sigma \|_p \leq C(\sigma) \beta,
\]
also taking into account the \(L^\infty (0, T; L^q (\mathbb{R}^d))\) bound for \(\nabla \rho_\sigma\); see Proposition 10. With this regularity, we can estimate all other terms except \(R_{34}\) and \(R_{36}\). Since they have similar structures, we only treat \(R_{34}\). This term is delicate since the factor \(f'_\sigma (\rho_\sigma)\) cannot be bounded in \(L^q (\mathbb{R}^d)\) for any \(q < \infty\). Therefore, one might obtain via Hardy–Littlewood–Sobolev’s inequality factors like \(\| \nabla \rho_\sigma \|_{q_1}\) and \(\| D^2 \rho_\sigma \|_{q_2}\) with either \(q_1 < 2\) or \(q_2 < 2\). However, for such factors, an \(L^\infty\) bound in time is currently lacking (Proposition 10 provides such a bound only for \(q \geq 2\)). Our idea is to add and subtract the term \(f'_\sigma (0)\) since
\[
| f'_\sigma (\rho_\sigma) - f'_\sigma (0) | \leq \rho_\sigma \max_{\| \rho_\sigma \|_\infty} | f''_\sigma | \leq C \rho_\sigma
\]
can be controlled. This leads to
\[
\|R_{34}\|_p \leq \|\nabla \rho_\sigma \cdot \mathcal{K} \ast \left( \left(f'_\sigma(\rho_\sigma) - f'_\sigma(0)\right)(W_\beta \ast \nabla \rho_\beta - \nabla \rho_\beta) \right)\|_p
\]
\[
+ \|f'_\sigma(0) \nabla \rho_\sigma \cdot \mathcal{K} \ast (W_\beta \ast \nabla \rho_\sigma - \nabla \rho_\sigma)\|_p
\]
\[
\leq C\beta + |f'_\rho(0)||\nabla \rho_\sigma \cdot \mathcal{K} \ast (W_\beta \ast \nabla \rho_\sigma - \nabla \rho_\sigma)\|_p =: C\beta + Q_{341},
\]
as the first term can be estimated in a standard way. For the estimate of $Q_{341}$, we need
to distinguish two cases.

If $1/2 < s \leq 1$, we infer from the Hardy–Littlewood–Sobolev-type inequality (B3) that
\[
Q_{341} \leq C\|\nabla \rho_\sigma\|_{d/(2s-1)}\|W_\beta \ast \rho_\sigma - \rho_\sigma\|_p \leq C\|\nabla \rho_\sigma\|_{d/(2s-1)}\|\nabla \rho_\sigma\|_p \leq C\beta.
\]
Next, let $0 < s \leq 1/2$. Then, we apply the Hardy–Littlewood–Sobolev-type inequality
(B2), the standard Gagliardo–Nirenberg inequality for some $\lambda > 0$, and Lemma 21:
\[
Q_{341} \leq C\|\nabla \rho_\sigma\|_{d/(2s)}\|W_\beta \ast \nabla \rho_\sigma - \nabla \rho_\sigma\|_p \leq C\|\rho_\sigma\|^{1-\lambda}_1\|D^2 \rho_\sigma\|_p (\lambda\|D^2 \rho_\sigma\|_p) \leq C\beta.
\]
We conclude that $\|R_{34}\|_p \leq C\beta$ and eventually
\[
\|R_3\|_{L^p(0,T;L^p(\mathbb{R}^d))} \leq C\beta.
\]
Summarizing the estimates for $R_1$, $R_2$, and $R_3$ finishes this step:
\[
\|R[\rho_\sigma,u]\|_{L^p(0,T;L^p(\mathbb{R}^d))} \leq C\|u\|_{L^p(0,T;W^{1,p}(\mathbb{R}^d))} + C\beta + C\Gamma_q^\mu\|D^2 u\|_{L^p(0,T;W^{1,p}(\mathbb{R}^d))}.
\]  

**Step 3: Estimate of $S[\rho_\sigma,u]$.** We formulate this term as $S[\rho_\sigma,u] = S_1 + \cdots + S_4$, where
\[
S_1 = \text{div} \left[ u \nabla (\mathcal{K}_\xi - \mathcal{K}) \ast \left( f_\sigma(W_\beta \ast (u + \rho_\sigma)) - f_\sigma(W_\beta \ast \rho_\sigma) \right) \right],
\]
\[
S_2 = \text{div} \left[ u \nabla (\mathcal{K}_\xi - \mathcal{K}) \ast f_\sigma(W_\beta \ast \rho_\sigma) \right],
\]
\[
S_3 = \text{div} \left[ \rho_\sigma \nabla (\mathcal{K}_\xi - \mathcal{K}) \ast \left( f_\sigma(W_\beta \ast (u + \rho_\sigma)) - f_\sigma(W_\beta \ast \rho_\sigma) \right) \right],
\]
\[
S_4 = \text{div} \left[ \rho_\sigma \nabla (\mathcal{K}_\xi - \mathcal{K}) \ast f_\sigma(W_\beta \ast \rho_\sigma) \right].
\]
The terms $S_1$, $S_2$, and $S_3$ can be treated as the terms in $R[\rho_\sigma,u]$, since they have the
same structure and the techniques used to estimate integrals involving $\mathcal{K}$ can be applied
to those involving $\mathcal{K}_\xi$. This leads to (for some $\mu > 0$)
\[
\|S_1 + S_2 + S_3\|_{L^p(0,T;L^p(\mathbb{R}^d))} \leq C\|u\|_{L^p(0,T;W^{1,p}(\mathbb{R}^d))} + C\Gamma_q^\mu\|D^2 u\|_{L^p(0,T;L^p(\mathbb{R}^d))}.
\]  

It remains to estimate $S_4$. We write $S_4 = S_{41} + S_{42} + S_{43}$, where
\[
S_{41} = \nabla \rho_\sigma \cdot (\mathcal{K}_\xi - \mathcal{K}) \ast \left[ f''_\sigma(W_\beta \ast \rho_\sigma)W_\beta \ast \nabla \rho_\sigma \right],
\]
\[
S_{42} = \rho_\sigma(\mathcal{K}_\xi - \mathcal{K}) \ast \left[ f''_\sigma(W_\beta \ast \rho_\sigma)W_\beta \ast \nabla \rho_\sigma \right]^2,
\]
\[
S_{43} = \rho_\sigma(\mathcal{K}_\xi - \mathcal{K}) \ast \left[ f''_\sigma(W_\beta \ast \rho_\sigma)W_\beta \ast \Delta \rho_\sigma \right].
\]
Observe that, because of the definition of $\mathcal{K}_\xi = \tilde{\mathcal{K}}_\xi \ast W_\xi$ with $\tilde{\mathcal{K}}_\xi = \mathcal{K}_\omega_\xi$ (defined in (11)),
we have $(\mathcal{K}_\xi - \mathcal{K}) \ast \nu = \mathcal{K} \ast (W_\xi \ast \nu - \nu) - (\mathcal{K}(1 - \omega_\xi)) \ast W_\xi \ast \nu$ for every function $\nu$
for which the convolution is defined, and therefore, by the Hardy–Littlewood–Sobolev-type
inequality (B2), Young’s convolution inequality, and Lemma 21,
\[
\|\rho_s(K\zeta - K) * v\|_p \leq C\|\rho_s\|_{d/(2s)} \|W_s * v - v\|_p + C\|\rho_s\|_p (K(1 - \omega_s) * v)\|_\infty
\]
\[
\leq C\|\rho_s\|_{d/(2s)} \|\nabla v\|_p \zeta + C\|\rho_s\|_p \|K1_{R^d \setminus B(0, \zeta^{-1})}\|_\infty \|v\|_1
\]
\[
\leq C\|\rho_s\|_{d/(2s)} \|\nabla v\|_p \zeta + C\zeta^{d-2s}\|\rho_s\|_p \|v\|_1,
\]

Given the regularity properties of \(\rho_s\) (see Lemma 11) and the assumptions on \(f_\sigma\), it follows that
\[
\|S_s\|_{L^p(0,T;L^p(\mathbb{R}^d))} \leq C_s^{\min\{1,d-2s\}}. \tag{47}
\]

We conclude from (46) and (47) that
\[
\|S[\rho_s, u]\|_{L^p(0,T;L^p(\mathbb{R}^d))} \leq C\|u\|_{L^p(0,T;W^{1,p}(\mathbb{R}^d))} + C_s^a + C\Gamma^a_p \|D^2 u\|_{L^p(0,T;L^p(\mathbb{R}^d))}, \tag{48}
\]
where \(a := \min\{1,d-2s\}\).

Step 4: End of the proof. Summarizing (44, 45), and (48), we infer that the right-hand side of (42) can be bounded (for some \(\mu > 0\)) by
\[
\|D[u] + R[\rho_s, u] + S[\rho_s, u]\|_{L^p(0,T;L^p(\mathbb{R}^d))} \leq C\|u\|_{L^p(0,T;W^{1,p}(\mathbb{R}^d))} + C(\beta + \zeta^a) + C\Gamma^a_p \|D^2 u\|_{L^p(0,T;L^p(\mathbb{R}^d))}.
\]

By parabolic regularity (C1),
\[
\|D^2 u\|_{L^p(0,T;L^p(\mathbb{R}^d))} \leq C\|u\|_{L^p(0,T;W^{1,p}(\mathbb{R}^d))} + C(\beta + \zeta^a) + C\Gamma^a_p \|D^2 u\|_{L^p(0,T;L^p(\mathbb{R}^d))}.
\]

Choosing \(\Gamma_p > 0\) sufficiently small finishes the proof. \(\square\)

It remains to estimate the \(L^p(0,T;W^{1,p}(\mathbb{R}^d))\) norm of \(u\). This is done in the following lemma.

Lemma 16 (Unconditional estimate for \(u\)). For any \(p > d\), there exist constants \(C > 0\), and \(\varepsilon_0 > 0\), both depending on \(\sigma, p,\) and \(T\), such that for \(\beta + \zeta^a < \varepsilon_0\),
\[
\|u\|_{L^\infty(0,T;W^{1,p}(\mathbb{R}^d))} \leq C(\beta + \zeta^a).
\]

recalling that \(a := \min\{1,d-2s\}\).

Proof. The idea is to test (42) with \(p|u|^{p-2} u - p \text{div} (|\nabla u|^{p-2} \nabla u)\). Integration by parts and some elementary computations lead to
\[
\int_{\mathbb{R}^d} p \text{div} (|\nabla u|^{p-2} \nabla u) \Delta u dx = -p \sum_{i,j} \int_{\mathbb{R}^d} |\nabla u|^{p-2} \partial_i u \partial_j \partial_j^2 u dx
\]
\[
= p \sum_{i,j} \int_{\mathbb{R}^d} \partial_j (|\nabla u|^{p-2} \partial_i u) \partial_j^2 u dx
\]
\[
= p \int_{\mathbb{R}^d} |\nabla u|^{p-2} |D^2 u|^2 dx + p \sum_j \int_{\mathbb{R}^d} \partial_j (|\nabla u|^{p-2}) \partial_j (|\nabla u|^2) dx
\]
\[
= p \int_{\mathbb{R}^d} |\nabla u|^{p-2} |D^2 u|^2 dx + \sum_j \int_{\mathbb{R}^d} \frac{4}{p} (p - 2) \left( \partial_j (|\nabla u|^{p/2}) \right)^2 dx.
\]

Consequently, we have
Choosing Lemmas 15 and 16 leads to

\[ p \|u(t)\|_{W^{1,p}(\mathbb{R}^d)}^p + \sigma p(p - 1) \int_0^t \int_{\mathbb{R}^d} |u|^{p-2} \left| \nabla u \right|^2 dx ds \]

\[ + \sigma \int_0^t \int_{\mathbb{R}^d} \left( p |\nabla u|^{p-2} |D^2 u|^2 + 4(p - 2) p^{-1} \left| \nabla \left( \nabla u^{p/2} \right) \right|^2 \right) dx ds \]

\[ = p \int_0^t \int_{\mathbb{R}^d} \left( |u|^{p-2} u - \text{div} \left( |\nabla u|^{p-2} \nabla u \right) \right) (Du) + R[\rho_\alpha, u] + S[\rho_\alpha, u] dx ds \]

\[ =: Q[u]. \]

We infer from Lemmas 20 and 26 that \( u \in C^0([0, T]; W^{1,p}(\mathbb{R}^d)) \). Therefore, since \( u(0) = 0 \), it holds that \( \|u(t)\|_{W^{1,p}(\mathbb{R}^d)} \leq \Gamma_p \) for all \( t \in [0, T^*) \) and \( T^* := \sup\{t_0 \in (0, T) : \|u(t)\|_{W^{1,p}(\mathbb{R}^d)} \leq \Gamma_p \} \) for \( 0 \leq t \leq t_0 \). Let \( t \in [0, T^*) \). We have shown in the proof of the previous lemma that

\[ \|D[u] + R[\rho_\alpha, u] + S[\rho_\alpha, u]\|_{L^p(0, t; L^p(\mathbb{R}^d))} \leq C\|u\|_{L^p(0, t; W^{1,p}(\mathbb{R}^d))} + C(\beta + \zeta^a). \]

Hence, we can estimate the right-hand side \( Q[u] \) of (49) as follows:

\[ Q[u] \leq C \int_0^t \int_{\mathbb{R}^d} \left( |u|^{p-1} + \left| \nabla u \right|^{p-2} |D^2 u| \right) |D[u] + R[\rho_\alpha, u] + S[\rho_\alpha, u]| dx \]

\[ \leq C \left( \|u\|^{p-1}_{L^p(0, t; L^p(\mathbb{R}^d))} + \|\nabla u\|^{p-2}_{L^p(0, t; L^p(\mathbb{R}^d))} \right) \|\nabla u\|^{p/2-1}_{L^p(0, t; L^p(\mathbb{R}^d))} \left( \|D^2 u\|_{L^2(0, t; L^2(\mathbb{R}^d))} \right) \]

\[ \times \left( \|u\|_{L^p(0, t; W^{1,p}(\mathbb{R}^d))} + \beta + \zeta^a \right) \]

\[ \leq C(\delta, p, t) \left( \|u\|^{p}_{L^p(0, t; W^{1,p}(\mathbb{R}^d))} + (\beta + \zeta^a)^p \right) + \delta \|\nabla u\|^{p/2-1}_{L^p(0, t; L^p(\mathbb{R}^d))} \left( \|D^2 u\|_{L^2(0, t; L^2(\mathbb{R}^d))} \right)^2, \]

where \( \delta > 0 \). Choosing \( \delta \) sufficiently small, the last term is absorbed by the corresponding expression on the left-hand side of (49), and we infer from (49) that for \( 0 \leq t \leq T^* \),

\[ \|u(t)\|^p_{W^{1,p}(\mathbb{R}^d)} \leq C(p, t) \int_0^t \|u\|^p_{W^{1,p}(\mathbb{R}^d)} ds + C(p, t)(\beta + \zeta^a)^p. \]

We assume without loss of generality that \( C(p, t) \) is nondecreasing in \( t \). Then, Gronwall’s lemma implies that for \( 0 \leq t \leq T^* \),

\[ \|u(t)\|^p_{W^{1,p}(\mathbb{R}^d)} \leq C(p, T)(\beta + \zeta^a)^p \int_0^t e^{C(p, T)(t-s)} ds \leq (\beta + \zeta^a)e^{C(p, T)t}. \]

Choosing \( \varepsilon_0 = \frac{1}{2} \Gamma_p \exp \left( -C(p, T)T/p \right) < 1 \), we find that \( \|u(t)\|_{W^{1,p}(\mathbb{R}^d)} \leq \Gamma_p/2 \) for \( \beta + \zeta^a < \varepsilon_0 \) and \( 0 \leq t \leq T^* \). By definition of \( T^* \), it follows that \( T^* = T \). In particular, \( \|u(t)\|_{W^{1,p}(\mathbb{R}^d)} \leq C(\beta + \zeta^a) \) for \( 0 < t < T \), which finishes the proof.\( \square \)

### 3.2. Proof of (38) and (39)

Combining Lemmas 15 and 16 leads to

\[ \|u\|_{L^p(0, T; W^{2,p}(\mathbb{R}^d))} \leq C(\sigma, p, T)(\beta + \zeta^a), \quad \text{where } a = \min\{1, d - 2s\}, \]

(50)
as long as $\beta + \zeta^a < \varepsilon_0$ and $p > d$. Next, we differentiate (42) with respect to $x_i$ (writing $\partial_i$ for $\partial/\partial x_i$):

$$\partial_i (\partial_i u) - \sigma \Delta (\partial_i u) = \partial_i (D[u] + R[\rho_\sigma, u] + S[\rho_\sigma, u]), \quad \partial_i u(0) = 0 \quad \text{in} \quad \mathbb{R}^d.$$ 

Taking into account estimate (50) and arguing as in the proof of Lemma 15, we can show that for $\delta > 0$,

$$\|\partial_i (D[u] + R[\rho_\sigma, u] + S[\rho_\sigma, u])\|_{L^p(0, T; L^p(\mathbb{R}^d))} \leq C(p, \sigma, \delta)(\beta + \zeta^a) + \delta \|D^3 u\|_{L^p(0, T; L^p(\mathbb{R}^d))}.$$ 

We infer from parabolic regularity (Lemma 26) for sufficiently small $\delta > 0$ that

$$\|\partial_i Du\|_{L^p(0, T; L^p(\mathbb{R}^d))} + \|D^3 u\|_{L^p(0, T; L^p(\mathbb{R}^d))} \leq C(p, \sigma)(\beta + \zeta^a).$$

Then, Lemma 20, applied to $Du$, leads to (38), which with Proposition 10 implies (39).

### 3.3. Proof of (40)

Let $x \in \mathbb{R}^d$. We use the definitions of $K_\zeta$ and $W_\zeta$ to find that

$$|(K_\zeta - K) \ast \rho_\sigma(x)| = \left| \int_{\mathbb{R}^d} W_\zeta(x - y) ((K \ast \rho_\sigma)(x) - ((K_\zeta \ast \rho_\sigma)(y)) \, dy \right|$$

$$\leq \int_{\mathbb{R}^d} W_\zeta(x - y) |x - y| \left| \frac{(K \ast \rho_\sigma)(x) - (K_\zeta \ast \rho_\sigma)(y)}{|x - y|} \right| \, dy + \| (K(1 - \omega_\zeta)) \ast \rho_\sigma \|_{\infty}$$

$$\leq \| \nabla K \ast \rho_\sigma \|_{\infty} \int_{\mathbb{R}^d} W_\zeta(z) |z| \, dz + \| K_1 \chi_{1 - B(0, \zeta^{-1})} \|_{\infty} \| \rho_\sigma \|_1$$

$$\leq \zeta \| \nabla K \ast \rho_\sigma \|_{\infty} \int_{\mathbb{R}^d} W_1(y) |y| \, dy + \zeta^{d-2s} \| \rho_\sigma \|_1.$$ 

Let $\phi \in C_0^\infty(\mathbb{R}^d)$ be such that $\text{supp}(\phi) \subset B_2$ and $\phi = 1$ in $B_1$. Then, (since we can assume without loss of generality that $\zeta < 1$), by arguing like in the derivation of (47), we obtain

$$|(K_\zeta - K) \ast \rho_\sigma(x)| \leq C \zeta^{\min\{1, d-2s\}} \left( \| \nabla (K \phi) \ast \rho_\sigma \|_{\infty} + \| \nabla (K(1 - \phi)) \ast \rho_\sigma \|_{\infty} + \| \rho_\sigma \|_1 \right),$$

A computation shows that for $p > \max\{d/(2s), 2\}$,

$$\| \nabla (K \phi) \ast \rho_\sigma \|_{\infty} = \| (K \phi) \ast \nabla \rho_\sigma \|_{\infty} \leq \| K \phi \|_{p/(p-1)} \| \nabla \rho_\sigma \|_p \leq C \| \nabla \rho_\sigma \|_p,$$

$$\| \nabla (K(1 - \phi)) \ast \rho_\sigma \|_{\infty} \leq \| \nabla (K(1 - \phi)) \|_{\infty} \| \rho_\sigma \|_1 \leq C \| \rho_\sigma \|_1,$$

where we note that $K_1 \chi_{B_2} \in L^{p/(p-1)}$ if $p > d/(2s)$. Then, in view of the regularity of $\rho_\sigma$ in Lemma 11, we find that

$$\| (K_\zeta - K) \ast \rho_\sigma \|_{L^\infty(0, T; L^\infty(\mathbb{R}^d))} \leq C \zeta^{a}.$$ 

### 3.4. Proof of (41)

The $L^\infty(0, T; W^{2,1}(\mathbb{R}^d) \cap W^{3, q}(\mathbb{R}^d))$ bound for $\rho_{\sigma, \beta, \zeta}$ is shown in a similar way as the corresponding bound for $\rho_\sigma$ in Lemma 11.
4. Mean-field analysis

This section is devoted to the proof of Proposition 3 and Theorem 2. The existence of solutions to (4) and (6) as well as the existence of density functions is shown in Subsection 4.1. In Subsection 4.2, we estimate the difference $X_N^i - \bar{X}_i^N$ of the processes of the original system (3) and the intermediate system (4), while the difference $\bar{X}_i^N - \hat{X}_i^N$ of the processes of the intermediate system (4) and the macroscopic system (6) is estimated in Subsection 4.3. These estimates are combined in Subsection 4.4 to conclude with the proof of Proposition 3 and Theorem 2.

4.1. Existence of density functions for (4) and (6)

First, we show that the coefficients of the stochastic differential Eq. (6), satisfied by $\bar{X}_i^N$, are globally Lipschitz continuous and of at most linear growth. The latter condition follows from

$$\begin{align*}
|\nabla K * f_\sigma(\rho_\sigma(x,t))| &\leq \|K * \nabla f_\sigma(\rho_\sigma)\|_{L^\infty\left(0,T;L^1(\mathbb{R}^d)\right)} \\
&\leq C\|K * \nabla f_\sigma(\rho_\sigma)\|_{L^\infty\left(0,T;W^{1,p}(\mathbb{R}^d)\right)} \leq C\|\nabla f_\sigma(\rho_\sigma)\|_{L^\infty\left(0,T;W^{1,p}(\mathbb{R}^d)\right)} \leq C(\sigma),
\end{align*}$$

where $p > d$ and $r = dp/(d + 2s)$ according to the Hardy–Littlewood–Sobolev inequality, and we used the regularity bounds for $\rho_\sigma$ from Lemma 26. The global Lipschitz continuity is a consequence of the mean-value theorem, the Hardy–Littlewood–Sobolev inequality, and the $W^{2,\infty}(\mathbb{R}^d)$ regularity of $\rho_\sigma$ from Lemma 11:

$$\begin{align*}
\sup_{0 < t < T} |\nabla K * f_\sigma(\rho_\sigma(x,t)) - \nabla K * f_\sigma(\rho_\sigma(y,t))| &\leq \sup_{0 < t < T} \|\nabla^2 K * f_\sigma(\rho_\sigma(\cdot,t))\|_{L^\infty}|x - y| \\
&= \sup_{0 < t < T} \|K * \left(f''_\sigma(\rho_\sigma)\nabla_\rho_\sigma \otimes \nabla_\rho_\sigma + f'_\sigma(\rho_\sigma)\nabla^2_\rho_\sigma\right)(\cdot,t)\|_{L^\infty}|x - y| \leq C(\sigma)|x - y|.
\end{align*}$$

These two conditions yield the existence and uniqueness of solutions to the associated particle systems [44, theorems 2.5 and 2.9]. Moreover, by [45, theorem 2.3.1], the law of the process $\bar{X}_i^N$ is absolutely continuous with respect to the Lebesgue measure. By Radon-Nikodym’s theorem, there exists a density function $\hat{u}(t)$ for all $t > 0$ on $\mathbb{R}^d$, which is measurable and integrable with respect to the Lebesgue measure. (Since all $\bar{X}_i^N$ are copies of the same process, their density functions are the same almost everywhere.)

The processes $\bar{X}_i^N(t)$ have continuous paths, which implies the continuity of the distribution function of $\bar{X}_i^N(t)$ with respect to time, and this implies in turn the Bochner measurability of $\hat{u}(t)$. Clearly, we have $\sup_{0 < t < T} \|\hat{u}(t)\|_{L^1(\mathbb{R}^d)} = 1$, which shows that $\hat{u} \in L^\infty(0,T;L^1(\mathbb{R}^d))$.

Similar arguments show that $\bar{X}_i^N(t)$ has a density function $\bar{u} \in L^\infty(0,T;L^1(\mathbb{R}^d))$.

Next, we show that $\hat{u}$ and $\bar{u}$ can be identified with the weak solutions $\rho_\sigma$ and $\rho_{\sigma,\beta,\gamma}$, respectively, using Itô’s lemma. Indeed, let $\phi \in C^\infty_c(\mathbb{R}^d \times [0,T])$. We infer from Itô’s formula that
\[
\begin{align*}
\phi(\tilde{X}_i^N(t), t) &= \phi(\tilde{X}_i^N(0), 0) + \sum_{l=1}^t \partial_t \phi(\tilde{X}_i^N(s), s) ds + \sigma \int_0^t \Delta \phi(\tilde{X}_i^N(s), s) ds \\
&\quad - \int_0^t \nabla K \ast f_\sigma(\rho_\sigma(\tilde{X}_i^N(s), s)) \cdot \nabla \phi(\tilde{X}_i^N(s), s) ds + \sqrt{2\sigma} \sum_{l=1}^t \nabla \phi(\tilde{X}_i^N(s), s) \cdot dB_i^N(s).
\end{align*}
\]

Taking the expectation, the Itô integral vanishes, and we end up with

\[
\int_{\mathbb{R}^d} \phi(x, t) \tilde{u}(x, t) dx = \int_{\mathbb{R}^d} \phi(x, 0) \rho_\sigma(x) dx + \int_0^t \int_{\mathbb{R}^d} \partial_t \phi(x, s) \tilde{u}(x, s) dx ds \\
+ \sigma \int_0^t \int_{\mathbb{R}^d} \Delta \phi(x, s) \tilde{u}(x, s) dx ds - \int_0^t \int_{\mathbb{R}^d} \nabla K \ast f_\sigma(\rho_\sigma(x, s)) \cdot \nabla \phi(x, s) \tilde{u}(x, s) dx ds.
\]

Hence, \(\tilde{u}\) is a weak solution in the space \(L^\infty(0, T; L^1(\mathbb{R}^d))\) to the linear equation

\[
\partial_t \tilde{u} = \sigma \Delta \tilde{u} + \text{div} (\tilde{u} \nabla K \ast f_\sigma(\rho_\sigma)), \quad \tilde{u}(0) = \rho_\sigma^0 \quad \text{in} \ \mathbb{R}^d,
\]

where \(\rho_\sigma\) is the unique solution to (7).

It can be shown that (52) is uniquely solvable in the class of functions in \(L^\infty(0, T; L^1(\mathbb{R}^d))\). This implies that \(\tilde{u} = \rho_\sigma\) in \(\mathbb{R}^d \times (0, T)\) (and similarly \(\tilde{u} = \rho_\sigma, 0, \tilde{u}\)). The proof is technical but standard; see, e.g., [35, theorem 7] for a sketch of a proof.

Another approach is as follows. Because of the linearity of (51), it is sufficient to prove that \(\tilde{u} \equiv 0\) in \(\mathbb{R}^d \times (0, T)\) if \(\rho_\sigma^0 = 0\). First, we verify that \(v := \nabla K \ast f_\sigma(\rho_\sigma) \in L^\infty(0, T; W^{1,\infty}(\mathbb{R}^d))\) and \(\tilde{u} \in L^p(0, T; L^q(\mathbb{R}^d))\) for \(p < d/(d-1)\). Then, by density, (51) holds for all \(\phi \in W^{1,q}(0, T; L^q(\mathbb{R}^d)) \cap L^2(0, T; W^{2,q}(\mathbb{R}^d))\) with \(q > d\) and \(\phi(T) = 0\). Choosing \(\psi\) to be the unique strong solution to the dual problem

\[
\partial_t \psi + \sigma \Delta \psi = v \cdot \nabla \psi + g, \quad \psi(T) = 0 \quad \text{in} \ \mathbb{R}^d
\]

in the very weak formulation of (51), we find that \(\int_0^T \int_{\mathbb{R}^d} \tilde{u} g dx dt = 0\) for all \(g \in C^\infty_0(\mathbb{R}^d \times (0, T))\), which implies that \(\tilde{u} = 0\).

### 4.2. Estimate of \(X_i^N - \tilde{X}_i^N\)

We derive an estimate for the expectation of the difference \(X_i^N - \tilde{X}_i^N\). To this end, we need to estimate the difference of the microscopic average \(N^{-1} \sum_{j=1, j \neq i}^{N} W_\beta(\tilde{X}_j^N - X_i^N)\) and the macroscopic average \(W_\beta \ast \rho_{\beta, \zeta, \rho}(\tilde{X}_i^N)\). By a careful choice of \(\beta\) and \(\zeta\), we show that this estimate is of the order \(N^{-1/4+\delta}\) for \(\delta > 0\).

**Lemma 17.** Let \(X_i^N\) and \(\tilde{X}_i^N\) be the solutions to (3) and (4), respectively, and let \(\delta \in (0, 1/4)\). Under the assumptions of Theorem 3 on \(\beta\) and \(\zeta\), it holds that

\[
\mathbb{E} \left( \sup_{0 < s < T} \max_{t=1, \ldots, N} \left| (X_i^N - \tilde{X}_i^N)(s) \right| \right) \leq CN^{-1/4+\delta}.
\]
Proof. To simplify the presentation, we set

$$\Psi(x, t) := f_\sigma \left( \frac{1}{N} \sum_{j=1, j \neq i}^N W_\beta \left( X_j^N(t) - x \right) \right), \quad \bar{\Psi}(x, t) := f_\sigma \left( \frac{1}{N} \sum_{j=1, j \neq i}^N W_\beta \left( \bar{X}_j^N(t) - x \right) \right),$$

and we write $\rho := \rho_{\alpha, \beta, \varsigma}$. Taking the difference of Eqs. (3) and (4) in the integral formulation leads to

$$\sup_{0 < s < t} \left| \left( X_i^N - \bar{X}_i^N \right)(s) \right| \leq \int_0^t \left| \nabla K_\varsigma \ast \left( \Psi(X_i^N(s), s) - \Psi(\bar{X}_i^N(s), s) \right) \right| ds \leq \int_0^t \left| \nabla K_\varsigma \ast \left( \Psi(X_i^N(s), s) - \bar{\Psi}(\bar{X}_i^N(s), s) \right) \right| ds + \int_0^t \left| \nabla K_\varsigma \ast \left( \bar{\Psi}(\bar{X}_i^N(s), s) - \bar{\Psi}(\bar{X}_i^N(s), s) \right) \right| ds \leq C_1 + C_2,$$

(53)

Step 1: Estimate of $I_1$. To estimate $I_1$, we formulate $I_1 = I_{11} + I_{12} + I_{13}$, where

$$I_{11} = \int_0^t \left| \nabla K_\varsigma \ast \left( \Psi(X_i^N(s), s) - \Psi(\bar{X}_i^N(s), s) \right) \right| ds,$$

$$I_{12} = \int_0^t \left| \nabla K_\varsigma \ast \left( \Psi(\bar{X}_i^N(s), s) - \bar{\Psi}(\bar{X}_i^N(s), s) \right) \right| ds,$$

$$I_{13} = \int_0^t \left| \nabla K_\varsigma \ast \left( \bar{\Psi}(\bar{X}_i^N(s), s) - \bar{\Psi}(\bar{X}_i^N(s), s) \right) \right| ds.$$

We start with the first integral:

$$I_{11} \leq \int_0^t \left\| D^2 K_\varsigma \ast \Psi(\cdot, s) \right\|_\infty \sup_{0 < r < s} \max_{i=1, \ldots, N} \left| \left( X_i^N - \bar{X}_i^N \right)(r) \right| ds.$$

We claim that

$$\left\| D^k K_\varsigma \ast \Psi(\cdot, s) \right\|_\infty \leq C(\sigma) \beta^{-(k+1)(d+k)-1}, \quad k \in \mathbb{N}. \quad (54)$$

For the proof, we introduce

$$\Phi(x, y) := f_\sigma \left( \frac{1}{N} \sum_{j=1}^{N-1} W_\beta (y_j - x) \right) \quad \text{for} \quad x \in \mathbb{R}^d, \quad y = (y_1, \ldots, y_{N-1}) \in \mathbb{R}^{(N-1)d}.$$ 

Then, by definition of $K_\varsigma$,

$$\left\| D^k K_\varsigma \ast \Psi(\cdot, t) \right\|_\infty \leq \sup_{y \in \mathbb{R}^{N-1}} \left\| W_\varsigma \ast \left( K \omega_\varsigma \ast D^k \Phi(\cdot, y) \right) \right\|_\infty.$$

We estimate the right-hand side:

$$\left\| W_\varsigma \ast \left( K \omega_\varsigma \ast D^k \Phi(\cdot, y) \right) \right\|_\infty \leq \left\| W_\varsigma \right\|_1 \left\| K \omega_\varsigma \ast D^k \Phi(\cdot, y) \right\|_\infty \leq C \left\| K \omega_\varsigma \ast D^k \Phi(\cdot, y) \right\|_{W^{1,p}(\mathbb{R}^d)} \leq C \left\| D^k \Phi(\cdot, y) \right\|_p + C \left\| D^{k+1} \Phi(\cdot, y) \right\|_p \leq C \left\| D^k \Phi(\cdot, y) \right\|_r + C \left\| D^{k+1} \Phi(\cdot, y) \right\|_r,$$
The estimate of the remaining term is symmetric, we find that

\[ k \text{ since } \]

\[ \text{that the last term is estimated according to } \]

\[ \text{last step. It follows from the Faàdi Bruno formula, after an elementary computation that the last term is estimated according to } \]

\[ \|D^{k+1}\Phi(\cdot,y)\|_{r}^{r} = \left| \int_{\mathbb{R}^{d}} D^{k+1} \left( f_{\sigma} \left( \frac{1}{N} \sum_{j=1}^{N-1} W_{\beta}(y_{j} - x) \right) \right) \right| \]

\[ \leq C(k, N) \max_{\ell=1, \ldots, k+1} \|f_{\sigma}^{(\ell)}\|_{\infty} \|D^{kr} W_{\beta}\|_{\infty} \max_{0 \leq j \leq k} \left| D^{j+1} W_{\beta}(x) \right|^{r} dx \]

\[ \leq C(k, N) \max_{\ell=1, \ldots, k+1} \|f_{\sigma}^{(\ell)}\|_{\infty} \beta^{-(d+k)kr} \beta^{-(d+k+1)r+d} \leq C(k, N, \sigma) \beta^{-(d+k)(k+1)r-r}, \]

since \( \|D^{k} W_{\beta}\|_{\infty} \leq C \beta^{-(d+k)} \) and \( \|D^{j+1} W_{\beta}\|_{r} \leq C \beta^{-(d+j+1)+d/r} \). This verifies (54). We infer from (54) with \( k = 2 \) that

\[ I_{11} \leq C \beta^{-3d-7} \int_{0}^{t} \sup_{0 < r < s} \max_{i=1, \ldots, N} \left| (X_{i}^{N} - \bar{X}_{i}^{N})(r) \right| ds. \]

The term \( I_{13} \) is estimated in a similar way, with \( \Psi \) replaced by \( \bar{\Psi} \):

\[ I_{13} \leq C \beta^{-3d-7} \int_{0}^{t} \sup_{0 < r < s} \max_{i=1, \ldots, N} \left| (X_{i}^{N} - \bar{X}_{i}^{N})(r) \right| ds. \]

The estimate of the remaining term \( I_{12} \) is more involved. Since \( W_{\beta} \) is assumed to be symmetric, we find that

\[ I_{12} = \left| \int_{0}^{t} \int_{\mathbb{R}^{d}} \mathcal{K}_{\zeta}(y) \nabla \left\{ f_{\sigma} \left( \frac{1}{N} \sum_{j=1, j \neq i}^{N} W_{\beta} \left( X_{j}^{N}(s) - \bar{X}_{i}^{N}(s) + y \right) \right) \right. \]

\[ - f_{\sigma} \left( \frac{1}{N} \sum_{j=1, j \neq i}^{N} W_{\beta} \left( \bar{X}_{j}^{N}(s) - X_{i}^{N}(s) + y \right) \right) \left\{ \right. \} dy ds \]

\[ \leq C \int_{0}^{t} \int_{\mathbb{R}^{d}} \mathcal{K}_{\zeta}(y) \left| f_{\sigma}^{(\ell)} \left( \frac{1}{N} \sum_{j=1, j \neq i}^{N} W_{\beta} \left( X_{j}^{N}(s) - \bar{X}_{i}^{N}(s) + y \right) \right) \right. \]

\[ \times \left. \frac{1}{N} \sum_{j \neq i} \nabla \left( W_{\beta} \left( X_{j}^{N}(s) - X_{i}^{N}(s) + y \right) - W_{\beta} \left( \bar{X}_{j}^{N}(s) - X_{i}^{N}(s) + y \right) \right) \right) \]

\[ + \left\{ f_{\sigma}^{(\ell)} \left( \frac{1}{N} \sum_{j \neq i} W_{\beta} \left( \bar{X}_{j}^{N}(s) - X_{i}^{N}(s) + y \right) \right) - f_{\sigma} \left( \frac{1}{N} \sum_{j \neq i} W_{\beta} \left( \bar{X}_{j}^{N}(s) - X_{i}^{N}(s) + y \right) \right) \right\} \]

\[ \times \left. \frac{1}{N} \sum_{j \neq i} \nabla W_{\beta} \left( \bar{X}_{j}^{N}(s) - X_{i}^{N}(s) + y \right) \right| dy ds \]

\[ \leq C \|f_{\sigma}^{(\ell)}\|_{\infty} \int_{0}^{t} \sup_{0 < r < s} \max_{i=1, \ldots, N} \left| (X_{i}^{N} - \bar{X}_{i}^{N})(s) \right| \frac{1}{N} \sum_{j \neq i} \int_{\mathbb{R}^{d}} \mathcal{K}_{\zeta}(y) |D^{2} W_{\beta}(y + \xi_{ij}(s))| dy ds \]

\[ + C \|f_{\sigma}^{(\ell)}\|_{\infty} \int_{0}^{t} \sup_{0 < r < s} \max_{i=1, \ldots, N} \left| (X_{i}^{N} - \bar{X}_{i}^{N})(s) \right| \left| \mathcal{K}_{\zeta} * \nabla W_{\beta} \right|_{\infty} ds, \]
where \( \xi_i(s) \) is a random value. We write \( \mathcal{K}^1 = \mathcal{K}|_{B_1}, \mathcal{K}^2 = \mathcal{K}|_{\mathbb{R}^d \setminus B_1} \) and note that \( \mathcal{K}_\zeta \leq \mathcal{K} \) for all \( \zeta > 0 \). Then,

\[
\int_{\mathbb{R}^d} \mathcal{K}_\zeta(y)|D^2 W_\beta(y + \xi_i(s))|dy \leq \int_{B_1 \setminus \zeta} (\mathcal{K}^1 \ast W_\zeta)(y)|D^2 W_\beta(y + \xi_i(s))|dy \\
+ \int_{\mathbb{R}^d \setminus B_1 \setminus \zeta} (\mathcal{K}^2 \ast W_\zeta)(y)|D^2 W_\beta(y + \xi_i(s))|dy \\
\leq \|\mathcal{K}^1 \ast W_\zeta\|_{L^{\theta/(\theta-1)}(B_1 \setminus \zeta)}\|D^2 W_\beta(\cdot + \xi_i(s))\|_{L^\theta(B_1 \setminus \zeta)} \\
+ \|\mathcal{K}^2 \ast W_\zeta\|_{\infty}\|D^2 W_\beta(\cdot + \xi_i(s))\|_{L^1(\mathbb{R}^d \setminus B_1 \setminus \zeta)} \\
\leq \|\mathcal{K}^1\|_{L^{\theta/(\theta-1)}(B_1)}\|D^2 W_\beta(\cdot + \xi_i(s))\|_{L^\theta(B_1 \setminus \zeta)} + \|\mathcal{K}^2\|_{\infty}\|D^2 W_\beta(\cdot + \xi_i(s))\|_{L^1(\mathbb{R}^d)} \\
\leq C(\|D^2 W_\beta\|_{\infty} + \|D^2 W_\beta\|_1) \leq C\beta^{-d-2}.
\]

Observe that we did not use the compact support for \( \mathcal{K}_\zeta \) (which depends on \( \zeta \)), because a negative exponent of \( \zeta \) at this point would lead to a logarithmic connection between \( \zeta \) and \( N \) in the end, which we wish to avoid.

Furthermore, by the convolution, Sobolev, and Hardy–Littlewood–Sobolev inequalities as well as the fact that \( |\mathcal{K}_\zeta \ast \nabla W_\beta| = |(\mathcal{K}W_\zeta) \ast \nabla W_\beta| \leq \mathcal{K} \ast |W_\zeta| \ast |\nabla W_\beta|, \)

\[
\|\mathcal{K}_\zeta \ast \nabla W_\beta\|_{\infty} = \|W_\zeta \ast \mathcal{K}_\zeta \ast \nabla W_\beta\|_{\infty} \leq \|\mathcal{K}_\zeta \ast \nabla W_\beta\|_{\infty} \leq \|\mathcal{K}_\zeta \ast \nabla W_\beta\|_{\infty} \\
\leq C\|\mathcal{K}_\zeta \ast \nabla W_\beta\|_{W^{1,p}(\mathbb{R}^d)} \leq C\left(\|\mathcal{K} \ast |\nabla W_\beta|\|_p + \|\mathcal{K} \ast |D^2 W_\beta|\|_p\right)^{1/p} \\
\leq C\|\nabla W_\beta\|_{W^{1,p}(\mathbb{R}^d)} \leq C\beta^{-d-2+d/r},
\]

where we recall that \( r > d/(2s) \) and we choose \( p > d \) satisfying \( 1/p = 2s/d - 1/r \). The previous two estimates lead to

\[
I_{12} \leq C(\sigma)\beta^{-d-2} \int_0^t \sup_{0 < r < s} \max_{i=1,\ldots,N} \left| \left( X_i^N - \bar{X}_i^N \right)(r) \right| ds.
\]

We summarize:

\[
I_1 \leq C(\sigma)\beta^{-3d-7} \int_0^t \sup_{0 < r < s} \max_{i=1,\ldots,N} \left| \left( X_i^N - \bar{X}_i^N \right)(r) \right| ds. \tag{55}
\]

**Step 2: Estimate of \( I_2 \).** We take the expectation of \( I_2 \) and use the mean-value theorem:

\[
\mathbb{E}(I_2) = \int_0^t \mathbb{E} \mathbb{E} \left\{ f_\sigma \left( \frac{1}{N} \sum_{j \neq i} W_\beta \left( \bar{X}_j^N(s) - Y_i(s) \right) \right) - f_\sigma \left( W_\beta \ast \rho \left( \bar{X}_i^N(s) - Y_i(s) \right) \right) \right\} dy ds \\
\leq N^{-1}\|f_\sigma\|_{\infty}\|\mathcal{K}_\zeta \ast \nabla W_\zeta\|_1 \int_0^t \sup_{0 \leq y \leq \mathbb{R}^d} \left( \sum_{j \neq i} |b_{ij}(y,s)| \right) ds,
\]

where

\[
b_{ij}(y,s) = W_\beta \left( \bar{X}_j^N(s) - Y_i(s) \right) - \frac{N}{N-1} W_\beta \ast \rho \left( \bar{X}_i^N(s) - Y_i(s) \right).
\]
We deduce from \( \| \nabla W_\zeta \|_{L^1(\mathbb{R}^d)} \leq C_\zeta^{-1} \) that
\[
\| \tilde{K}_\zeta \ast \nabla W_\zeta \|_1 \leq C_\zeta^{-1} \| \tilde{K}_\zeta \|_1 \leq C_\zeta^{-2s-1},
\]
due to the compact support of \( \tilde{K}_\zeta(x) = |x|^{2s-d} \omega_\zeta(x) \leq C |x|^{2s-d} 1_{|x| \leq 2\zeta^{-1}} \) and
\[
\int_{\{|x| < 2/\zeta\}} |x|^{2s-d} \, dx = \int_{\{|y| < 2\}} \zeta^{-d} |y/\zeta|^{2s-d} \, dy = C\zeta^{-2s}.
\]
We claim that \( \mathbb{E} \left( \sum_{j \neq i} |b_{ji}(y, s)| \right) \leq C(\sigma) \beta^{-d/2} N^{1/2} \) for all \( y \in \mathbb{R}^d \). To show the claim, we compute the expectation \( \mathbb{E} \left[ \left( \sum_{j \neq i} b_{ij}(y, s) \right)^2 \right] \). We estimate first the terms with \( k \neq j \) (omitting the argument \( y, s \) to simplify the notation). Then, an elementary but tedious computation leads to
\[
\mathbb{E}(b_{ji} b_{ki}) = \int_{\mathbb{R}^d} \left( W_\beta(x_j - x_i + y) - \frac{N}{N-1} W_\beta \ast \rho(x_i - y) \right) \times \left( W_\beta(x_k - x_i + y) - \frac{N}{N-1} W_\beta \ast \rho(x_i - y) \right) \rho(x_i) \rho(x_j) \rho(x_k) \, dx_idx_j dx_k
\]
\[
= \int_{\mathbb{R}^d} \left( W_\beta \ast \rho(x_i - y) - \frac{N}{N-1} W_\beta \ast \rho(x_i - y) \right)^2 \rho(x_i) \, dx_i \leq N^{-2} \| \rho \|_{L^\infty(0,T;L^\infty(\mathbb{R}^d))} \| W_\beta \ast \rho \|_{L^2(0,T;L^2(\mathbb{R}^d))}^2 \leq C(\sigma) N^{-2} \| W_\beta \|_1^2 \leq C(\sigma) N^{-2}.
\]
The diagonal terms contribute in the following way:
\[
\mathbb{E}(b_{ji}^2) = \int_{\mathbb{R}^d} \left( W_\beta(x_j - x_i + y) - \frac{N}{N-1} W_\beta \ast \rho(x_i - y) \right)^2 \rho(x_i) \rho(x_j) \, dx_idx_j
\]
\[
= \int_{\mathbb{R}^d} \left( W_\beta \ast \rho(x_i - y) - \frac{2N}{N-1} (W_\beta \ast \rho)(x_i - y)^2 \right)^2 \rho(x_i) \, dx_i \leq C(\sigma) \left( \| W_\beta \ast \rho \|_{L^\infty(0,T;L^\infty(\mathbb{R}^d))} + \| W_\beta \ast \rho \|_{L^2(0,T;L^2(\mathbb{R}^d))} \right) \leq C(\sigma) \beta^{-d},
\]
since \( \| W_\beta \ast \rho \|_2 \leq \| W_\beta \|_1 \| \rho \|_2 \leq C \| W_\beta \|_2 \leq \beta^{-d} C \). This shows that
\[
\mathbb{E} \left( \sum_{j \neq i} |b_{ji}(y, s)| \right) \leq \left( \mathbb{E} \left[ \sum_{j \neq i} b_{ji}(y, s) \right]^2 \right)^{1/2} \leq C(\sigma) \beta^{-d/2} N^{1/2}.
\]
We infer that (56) becomes
\[
I_2 \leq C(\sigma) \zeta^{-2s-1} \beta^{-d/2} N^{-1/2}.
\]
Step 3: End of the proof. We insert (55) and (57) into (53) to infer that
\[
E_1(t) := \mathbb{E}\left( \sup_{0 < s < t} \max_{i = 1, \ldots, N} \left| (X_i^N - \tilde{X}_i^N)(s) \right| \right)
\leq C(\sigma) \beta^{-3d-7} \int_0^t E_1(s) ds + C(\sigma) \zeta^{-2s-1} \beta^{-d/2} N^{-1/2}.
\]

By Gronwall’s lemma,
\[
E_1(t) \leq C(\sigma) \zeta^{-2s-1} \beta^{-d/2} N^{-1/2} \exp \left( C(\sigma) \beta^{-3d-7} T \right), \quad 0 \leq t \leq T.
\]

We choose \( \varepsilon = \tilde{\delta}/(C(\sigma)T) \) for some arbitrary \( \tilde{\delta} \in (0, 1/4) \). Then, since by assumption, \( \beta^{-d/2} \leq \beta^{-3d-7} \leq \varepsilon \log N \) and \( \zeta^{-2s-1} \leq C_1 N^{1/4} \), we find that
\[
E_1(t) \leq C(\sigma) C_1 \varepsilon \log (N) N^{-1/4} \exp \left( C(\sigma) T \varepsilon \log N \right) = \frac{C_1 \tilde{\delta}}{T} \log (N) N^{-1/4+\tilde{\delta}},
\]
proving the result. \( \square \)

4.3. Estimate of \( X_i^N - \hat{X}_i^N \)

Next, we compute the expectation of \( \bar{X}_i^N - \hat{X}_i^N \) by estimating the difference between \( \nabla K_\zeta \) and \( \nabla K \) as well as the difference between \( W_\beta * \rho(\bar{X}_i^N) \) and \( \rho_\sigma(\hat{X}_i^N) \). The estimate depends on \( \beta \) and \( \zeta \).

Lemma 18. Let \( \bar{X}_i^N \) and \( \hat{X}_i^N \) be the solutions to (4) and (6), respectively. Then, there exists a constant \( C > 0 \), depending on \( \sigma \), such that
\[
\mathbb{E}\left( \sup_{0 < t < T} \max_{i = 1, \ldots, N} \left| (\bar{X}_i^N - \hat{X}_i^N)(t) \right| \right) \leq C(\beta + \zeta^a),
\]
where \( a := \min\{1, d - 2s\} \).

Proof. We compute the difference
\[
\left| (\bar{X}_i^N - \hat{X}_i^N)(t) \right| = \left| \int_0^t \left( \nabla K_\zeta * f_\sigma(W_\beta * \rho(\bar{X}_i^N)(s), s) \right) - \nabla K * f_\sigma(\rho_\sigma(\hat{X}_i^N)(s), s) \right) ds \right| \leq J_1 + J_2 + J_3,
\]
where \( \rho := \rho_\sigma(\beta, \zeta) \), the convolution is taken with respect to \( x_n \), and
\[
J_1 = \left| \int_0^t \nabla K_\zeta * \left( f_\sigma(W_\beta * \rho(\bar{X}_i^N)(s), s) \right) - f_\sigma(W_\beta * \rho(\hat{X}_i^N)(s), s) \right) ds \right|,
\]
\[
J_2 = \left| \int_0^t \nabla K_\zeta * \left( f_\sigma(W_\beta * \rho(\bar{X}_i^N)(s), s) \right) - f_\sigma(\rho_\sigma(\hat{X}_i^N)(s), s) \right) ds \right|,
\]
\[
J_3 = \left| \int_0^t \nabla(\zeta - K) * f_\sigma(\rho_\sigma(\hat{X}_i^N)(s), s) \right) ds \right|.
\]

Step 1: Estimate of \( J_1 \). We write \( \nabla K_\zeta * f_\sigma(\cdots) = K_\zeta * \nabla f_\sigma \) and add and subtract the expression \( f_\sigma'(W_\beta * \rho(\bar{X}_i^N) - y) \nabla W_\beta * \rho(\hat{X}_i^N - y) \):
\[ J_1 = \int_{0}^{t} \int_{\mathbb{R}^d} \mathcal{K}_{\zeta} (y) \left( f_{\sigma} \left( W_{\beta} \ast \rho \left( \hat{X}_{i}^{N} (s) - y \right) \right) \nabla W_{\beta} \ast \left[ \rho \left( \hat{X}_{i}^{N} (s) - y \right) \right] \right. \\
\left. \quad - f_{\sigma} \left( W_{\beta} \ast \rho \left( \hat{X}_{i}^{N} (s) - y \right) \right) - f_{\sigma} \left( W_{\beta} \ast \rho \left( \hat{X}_{i}^{N} (s) - y \right) \right) \nabla W_{\beta} \ast \left[ \rho \left( \hat{X}_{i}^{N} (s) - y \right) \right] \right) \, dy \, ds \]

By the mean-value theorem and using \( \| W_{\beta} \|_1 = 1 \), we obtain for some random variable \( \xi_{\theta} (s) \),

\[
J_1 \leq \| f_{\sigma} \|_{W^{2, \infty} (\mathbb{R})} \| \nabla \rho \|_{L^\infty (0, T; L^\infty (\mathbb{R}^d))} \int_{0}^{t} \sup_{0 < r < s \leq 1, \ldots, N} \left| \left( \hat{X}_{i}^{N} - \hat{X}_{i}^{N} \right) (r) \right| \\
\times \int_{\mathbb{R}^d} \sum_{k=1}^{2} |\mathcal{K}_{\zeta} (y) | D^k W_{\beta} \ast \rho (y + \xi_{ij} (s), s) | \, dy \, ds.
\]  

We need to estimate the last integral. For this, we write for \( k = 1, 2 \)

\[
\int_{\mathbb{R}^d} |\mathcal{K}_{\zeta} (y) | D^k W_{\beta} \ast \rho (y + \xi_{\theta} (s), s) | \, dy \leq K_1^k + K_2^k,
\]

where

\[
K_1^k := \int_{B_{i+\zeta}} |\mathcal{K}^1 | W_{\zeta} (y) | D^k W_{\beta} \ast \rho (y + \xi_{\theta} (s), s) | \, dy,
\]

\[
K_2^k := \int_{\mathbb{R}^d \setminus B_{i-\zeta}} |\mathcal{K}^2 | W_{\zeta} (y) | D^k W_{\beta} \ast \rho (y + \xi_{ij} (s), s) | \, dy,
\]

where \( \mathcal{K}^1 = \mathcal{K}^1_{B_{i+\zeta}} \) and \( \mathcal{K}^2 = \mathcal{K}^2_{\mathbb{R}^d \setminus B_{i}} \). Note that \( \mathcal{K}_{\zeta} \leq \mathcal{K} \). A similar argument as for the estimate of \( I_{12} \) in the proof of Lemma 17 shows that for \( \theta > \max \{ d / (2s), d \} \),

\[
K_1^k + K_2^k \leq C \left( \| D^k W_{\beta} \ast \rho \|_{L^\infty (0, T; L^\theta (\mathbb{R}^d))} + \| D^k W_{\beta} \ast \rho \|_{L^\infty (0, T; L^1 (\mathbb{R}^d))} \right) \\
\leq C \left( \| D^k \rho \|_{L^\infty (0, T; L^\theta (\mathbb{R}^d))} + \| D^k \rho \|_{L^\infty (0, T; L^1 (\mathbb{R}^d))} \right) \leq C (\sigma),
\]

where we used Proposition 14 ((39) and (41)) with \( p = \theta \) in the last inequality. We conclude from (58) that

\[
J_1 \leq C (\sigma) \int_{0}^{t} \sup_{0 < r < s \leq 1, \ldots, N} \left| \left( \hat{X}_{i}^{N} - \hat{X}_{i}^{N} \right) (r) \right| \, ds.
\]  

**Step 2: Estimate of \( J_2 \).** We treat the two cases \( s < 1/2 \) and \( s \geq 1/2 \) separately. Let first \( s \geq 1/2 \). Then,

\[
J_2 = \left| \int_{0}^{t} \nabla \mathcal{K}_{\zeta} \ast W_{\zeta} \ast \left( f_{\sigma} \left( W_{\beta} \ast \rho \left( \hat{X}_{i}^{N} (s) \right) \right) \right) - f_{\sigma} \left( \rho_{\sigma} \left( \hat{X}_{i}^{N} (s) \right) \right) \right| \, ds \\
\leq T \| \nabla \mathcal{K}_{\zeta} \ast \left( f_{\sigma} (W_{\beta} \ast \rho) - f_{\sigma} (\rho_{\sigma}) \right) \|_{L^\infty (0, T; L^\infty (\mathbb{R}^d))}.
\]
Recalling the definition of $\tilde{K}_\zeta = K \omega_\zeta$ in (11) and writing $\nabla \tilde{K}_\zeta * u = \nabla K * u - [(1 - \omega_\zeta) \nabla K] * u + [K \nabla \omega_\zeta] * u$ for $u = f_\sigma(W_\beta * \rho) - f_\sigma(\rho_\sigma)$, we find that

$$J_2 \leq C(T) \left( \|\nabla K * u\|_{L^\infty([0,T];L^\infty(\mathbb{R}^d))} + \|[(1 - \omega_\zeta) \nabla K] * u\|_{L^\infty([0,T];L^\infty(\mathbb{R}^d))} + \| [K \nabla \omega_\zeta] * u\|_{L^\infty([0,T];L^\infty(\mathbb{R}^d))} \right).$$

(60)

We estimate the right-hand side term by term. Because of

$$\nabla K * v \begin{cases} \nabla(-\Delta)^{-1/2} \nu & \text{for } s = 1/2 \\ (\nabla K) * \nu & \text{for } s > 1/2, \end{cases}$$

we use Sobolev’s embedding $W^{1,p}(\mathbb{R}^d) \hookrightarrow L^\infty(\mathbb{R}^d)$ for any $p > d$, and then, the boundedness of the Riesz operator $\nabla(-\Delta)^{-1/2} : L^p(\mathbb{R}^d) \to L^p(\mathbb{R}^d)$ [46, chapter IV, §3.1] in case $s = 1/2$ or the Hardy–Littlewood–Sobolev inequality for $\tilde{x} = x - 1/2 > 0$ (see Lemma 22) in case $s > 1/2$ to control the first norm in (60) by

$$\|\nabla K * u\|_{L^\infty(0,T;L^\infty(\mathbb{R}^d))} \leq C \left( \|\nabla K * u\|_{L^\infty(0,T;L^p(\mathbb{R}^d))} + \sum_{j=1}^{d} \|\nabla K * D^j u\|_{L^\infty(0,T;L^{p_j}(\mathbb{R}^d))} \right),$$

where $r = p$ in case $s = 1/2$ and $r = pd/(d + 2s - 1)$ in case $s > 1/2$. Choosing $p > d + (2s - 1)$ guarantees that $r > d$ always holds.

For the second norm in (60), Hölder’s inequality yields for $q > d$ and $1/q + 1/q' = 1$, for every $t > 0$,

$$\left\| [(1 - \omega_\zeta) \nabla K] * u(t) \right\|_{L^\infty(\mathbb{R}^d)} \leq \|1 - \omega_\zeta\|_{L^\infty(\mathbb{R}^d)} \|\nabla K\|_{L^{p_j}(\{|x| > 2\zeta^{-1}\})} \|u(t)\|_{L^q(\mathbb{R}^d)},$$

which can be bounded by $C\zeta^{1-2s+d/q} \|u(t)\|_{L^q(\mathbb{R}^d)}$, since

$$\|\nabla K\|_{L^{q_j}(\{|x| > 2\zeta^{-1}\})} \leq C \int_{\{|x| > 2\zeta^{-1}\}} |x|^{(2s-d-1)q_d} \, dx = C\zeta^{-d} \int_{\{|y| > 2\}} |y/\zeta|^{(2s-d-1)q_d} \, dy \leq C\zeta^{-d+d(1+d-2s)q_d}.$$
Using these estimates in (60), we arrive (for \( s \geq 1/2 \)) at
\[
J_2 \leq C(T) \left( \| f_\sigma(W_\beta \ast \rho) - f_\sigma(\rho_\sigma) \|_{L^\infty([0,T;W^{1,r}(\mathbb{R}^d)]^\star)} + \| f_\sigma(W_\beta \ast \rho) - f_\sigma(\rho_\sigma) \|_{L^\infty([0,T,L^d(\mathbb{R}^d)])} \right),
\]
where we recall that \( r, q > d \). These norms can be estimated by \( \| f_\sigma(W_\beta \ast \rho(t)) - f_\sigma(\rho_\sigma(t)) \|_{L^d(\mathbb{R}^d)} \leq \| f_\sigma \|_\infty \| W_\beta \ast \rho(t) - \rho_\sigma(t) \|_{L^d(\mathbb{R}^d)} \) and
\[
\| \nabla (f_\sigma(W_\beta \ast \rho) - f_\sigma(\rho_\sigma)) (t) \|_{L^d(\mathbb{R}^d)} \leq \| f_\sigma \|_\infty \| W_\beta \ast \nabla \rho - \nabla \rho_\sigma \|_{L^d(\mathbb{R}^d)}
+ \| f_\sigma \|_\infty \| (W_\beta \ast \rho - \rho_\sigma)(t) \|_{L^d(\mathbb{R}^d)} \| \nabla \rho_\sigma \|_{L^\infty(\mathbb{R}^d)}.
\]
The \( L^\infty(\mathbb{R}^d \times (0,T)) \) bound for \( \nabla \rho_\sigma \) from Lemma 11 and the definition of \( f_\sigma \) finally show for \( s \geq 1/2 \) and \( r, q > d \) that
\[
J_2 \leq C(\sigma, T) \left( \| W_\beta \ast \rho - \rho_\sigma \|_{L^\infty([0,T;W^{1,r}(\mathbb{R}^d)]^\star)} + \| W_\beta \ast \rho - \rho_\sigma \|_{L^\infty([0,T,L^d(\mathbb{R}^d)])} \right). \tag{61}
\]

Now, let \( s < 1/2 \). In this case, we cannot estimate \( \nabla K \) and put the gradient to the second factor of the convolution. Adding and subtracting an appropriate expression as in Step 1, using the embedding \( W^{1,p}(\mathbb{R}^d) \hookrightarrow L^\infty(\mathbb{R}^d) \) for \( p > d \), the estimate \( K_\xi \leq K \), and the Hardy–Littlewood–Sobolev inequality, we find that
\[
J_2 = \left| \int_0^t \int_{\mathbb{R}^d} K_\xi(y) \left( \left( f_\sigma(W_\beta \ast \rho(\tilde{X}_i^N(s) - y)) - f_\sigma(\rho_\sigma(\tilde{X}_i^N(s) - y)) \right) \nabla W_\beta \ast \rho(\tilde{X}_i^N(s) - y)
- f_\sigma(\rho_\sigma(\tilde{X}_i^N(s) - y)) \left( \nabla \rho_\sigma(\tilde{X}_i^N(s) - y) - \nabla W_\beta \ast \rho(\tilde{X}_i^N(s) - y) \right) \right) dyds \right|
\leq \| f_\sigma \|_\infty \| W_\beta \ast \nabla \rho \|_{L^\infty} \int_0^t \int_{\mathbb{R}^d} K_\xi(y) \left| \rho_\sigma(\tilde{X}_i^N(s) - y) - W_\beta \ast \rho(\tilde{X}_i^N(s) - y) \right| dyds
+ \| f_\sigma \|_\infty \int_0^t \int_{\mathbb{R}^d} K_\xi(y) \left| \nabla \rho_\sigma(\tilde{X}_i^N(s) - y) - W_\beta \ast \nabla \rho(\tilde{X}_i^N(s) - y) \right| dyds
\leq \max \left\{ \| \nabla \rho \|_{L^\infty([0,T,L^\infty(\mathbb{R}^d)]^\star)}, 1 \right\} \| f_\sigma \|_{W^{1,\infty}} T \left( \| K \ast (W_\beta \ast \rho - \rho_\sigma) \|_{L^\infty([0,T,L^\infty(\mathbb{R}^d)])}
+ \| K \ast (W_\beta \ast \nabla \rho - \nabla \rho_\sigma) \|_{L^\infty([0,T,L^\infty(\mathbb{R}^d)])} \right)
\leq C(\sigma, T) \left( \| \nabla \rho \|_{L^\infty([0,T,L^\infty(\mathbb{R}^d)])} + 1 \right) \sum_{|\xi| \leq 2} \| W_\beta \ast D^2 \rho - D^2 \rho_\sigma \|_{L^\infty([0,T,L^d(\mathbb{R}^d)])},
\]
where \( r > d \) is such that \( 1/r = 2s/d + 1/p \) (this is needed for the Hardy–Littlewood–Sobolev inequality) and \( p > d \) (because of Sobolev’s embedding). Note that \( r > d \) can be only guaranteed if \( s < 1/2 \). Together with the fact that \( \| \nabla \rho \|_{L^\infty([0,T,L^\infty(\mathbb{R}^d)])} \leq C(\sigma) \) (choose \( q > d \) in (41) and use Sobolev’s embedding), this shows that for \( s < 1/2 \),
\[
J_2 \leq C(\sigma, T) \sum_{|\xi| \leq 2} \| W_\beta \ast D^2 \rho - D^2 \rho_\sigma \|_{L^\infty([0,T,L^d(\mathbb{R}^d)])}. \tag{62}
\]
It follows from estimate (38) and Lemma 21 in Appendix A for \( p > d \) that
\[
\| (W_\beta \ast D^2 \rho - D^2 \rho_\sigma)(t) \|_{L^1(\mathbb{R}^d)} \leq C(\| D^2 \nabla \rho \|_{L^1(\mathbb{R}^d)} + \beta + \zeta^a) \leq C(\sigma, T)(\beta + \zeta^a),
\]
where we used the $L^\infty(0, T; W^{3, p}(\mathbb{R}^d))$ estimate for $\rho = \rho_{\sigma, \beta, \zeta}$ in (41). Then, we deduce from estimates (61) and (62) that for all $0 < s < 1$,

$$J_2 \leq C(\sigma, T)(\beta + \zeta^a),$$

where we recall that $a = \min\{1, d - 2s\}$.

**Step 3: Estimate of $J_3$ and end of the proof.** Arguing similarly as in Subsection 3.3, we have

$$\| (\mathcal{K}_\zeta - \mathcal{K}) \ast \nabla \rho_\sigma \|_{L^\infty(0, T; L^\infty(\mathbb{R}^d))} \leq C\zeta^a \left( \| D^2 \rho_\sigma \|_{L^\infty(0, T; L^1(\mathbb{R}^d))} + \| \nabla \rho_\sigma \|_{L^\infty(0, T; L^1(\mathbb{R}^d))} \right).$$

This implies that

$$J_3 \leq \| f_\sigma \|_{\infty} \| (\mathcal{K}_\zeta - \mathcal{K}) \ast \nabla \rho_\sigma \|_{L^\infty(0, T; L^\infty(\mathbb{R}^d))} \leq C(\sigma)\zeta^a. \quad (63)$$

Taking the expectation, we infer from (59) to (63) that

$$E_2(t) := \mathbb{E} \left( \sup_{0 < s < t} \max_{i = 1, \ldots, N} \left| X_i^N(t) - \hat{X}_i^N(t) \right| \right) \leq C(\sigma)(\beta + \zeta^a) + C(\sigma) \int_0^t E_2(s) ds,$$

An application of Gronwall’s lemma gives the result.

**4.4. Proof of Theorem 2 and Proposition 3**

Lemmas 17 and 18 show that

$$\mathbb{E} \left( \sup_{0 < s < t} \max_{i = 1, \ldots, N} \left| X_i^N(t) - \hat{X}_i^N(t) \right| \right) \leq C(N^{-1/4 + \delta} + \beta + \zeta^{\min\{1, d - 2s\}}),$$

and this expression converges to zero as $N \to \infty$ and $(\beta, \zeta) \to 0$ under the conditions stated in Theorem 3. This result implies the convergence in probability of the $k$-tuple $(X_i^N, \ldots, X_k^N)$ to $(\hat{X}_i^N, \ldots, \hat{X}_k^N)$. Since convergence in probability implies convergence in distribution, we obtain

$$\lim_{N \to \infty, (\beta, \zeta) \to 0} P_{N, \beta, \sigma}(t) = P_{\sigma}^{\otimes k}(t) \quad \text{locally uniform in time},$$

where $P_{N, \beta, \sigma}(t)$ and $P_{\sigma}^{\otimes k}(t)$ denote the joint distributions of $(X_i^N, \ldots, X_k^N)(t)$ and $(\hat{X}_i^N, \ldots, \hat{X}_k^N)(t)$, respectively. By Subsection 4.1, $P_{\sigma}(t)$ is absolutely continuous with the density function $\rho_\sigma(t)$. Using the test function $\phi = 1_{(-\infty, x]^d}$ in Corollary 13, we have, up to a subsequence,

$$P_{\sigma}(t, (-\infty, x]^d) = \int_{(-\infty, x]^d} \rho_\sigma(y, t) dy \to \int_{(-\infty, x]^d} \rho(y, t) dy =: P(t, (-\infty, x]^d)$$

locally uniformly for $t > 0$. Since the convergence also holds for the initial condition, the result is shown.
Funding

The second and third authors have been partially supported by the Austrian Science Fund (FWF), grants P30000, P33010, F65, and W1245. The fourth author acknowledges support from the Alexander von Humboldt Foundation. This work has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme, ERC Advanced Grant no. 101018153.

References

[1] Stinga, P. R. (2019). User’s guide to the fractional Laplacian and the method of semi-groups. In: Kochubei, A., Luchko, Y., eds. Handbook of Fractional Calculus with Applications, Vol. 2. Berlin: De Gruyter, pp. 235–266.

[2] Chapman, J., Rubinstein, J., Schatzman, M. (1996). A mean-field model for superconducting vortices. Eur. J. Appl. Math. 7(2):97–111. DOI: 10.1017/S0956792500002242.

[3] E, W. (1994). Dynamics of vortex liquids in Ginzburg–Landau theories with applications in superconductivity. Phys. Rev. B Condens. Matter. 50(2):1126–1135. DOI: 10.1103/physrevb.50.1126.

[4] Biler, P., Imbert, C., Karch, G. (2015). The nonlocal porous medium equation: Barenblatt profiles and other weak solutions. Arch. Rational Mech. Anal. 215(2):497–529. DOI: 10.1007/s00205-014-0786-1.

[5] Vázquez, J. L. (2017). The mathematical theories of diffusion. Nonlinear and fractional diffusion. In: Bonforte, M., Grillo, G., eds. Nonlocal and Nonlinear Diffusions and Interactions: New Methods and Directions. Lect. Notes Math. 2186. Cham: Springer, pp. 205–278.

[6] Caffarelli, L., Vázquez, J. L. (2011). Nonlinear porous medium flow with fractional potential pressure. Arch. Rational Mech. Anal. 202(2):537–565. DOI: 10.1007/s00205-011-0420-4.

[7] Biler, P., Imbert, C., Karch, G. (2011). Barenblatt profiles for a non local porous medium equation. C. R. Acad. Sci. Paris Ser. I 349(11–12):641–645. DOI: 10.1016/j.crma.2011.06.003.

[8] Stan, D., del Teso, F., Vázquez, J. L. (2014). Finite and infinite speed of propagation for porous medium equations with fractional pressure. C. R. Acad. Sci. Paris Ser. I 352(2):123–128. DOI: 10.1016/j.crma.2013.12.003.

[9] Plociniczak, L. (2019). Derivation of the nonlocal pressure form of the fractional porous medium equation in the hydrological setting. Commun. Nonlin. Sci. Numer. Simul. 76:66–70. DOI: 10.1016/j.cnsns.2019.04.014.

[10] Duerinckx, M. (2016). Mean-field limits for some Riesz interaction gradient flows. SIAM J. Math. Anal. 48(3):2269–2300. DOI: 10.1137/15M1042620.

[11] Serfaty, S. (2016). Mean-field limits of the Gross–Pitaevskii and parabolic Ginzburg–Landau equations. J. Amer. Math. Soc. 30(3):713–768. DOI: 10.1090/jams/872.

[12] Olivera, C., Richard, A., Tomašević, M. (2020). Quantitative particle approximation of nonlinear Fokker–Planck equations with singular kernel. Ann. Sc. Norm. Super. Pisa Cl. Sci.

[13] Choi, Y.-P., Jeong, I.-J. (2021). Relaxation to fractional porous medium equation from Euler–Riesz system. J. Nonlin. Sci. 31(95):28.

[14] Serfaty, S. (2020). Mean field limit for Coulomb-type flows. Duke Math. J. 169(15):2887–2935. DOI: 10.1215/00127094-2020-0019.

[15] Daus, E. S., Ptashnyk, M., Raithel, C. (2022). Derivation of a fractional cross-diffusion system as the limit of a stochastic many-particle system driven by Lévy noise. J. Diff. Eqs. 309:386–426.

[16] Caffarelli, L., Vázquez, J. L. (2011). Asymptotic behaviour of a porous medium equation with fractional diffusion. Discrete Contin. Dyn. Sys. 29(4):1393–1404. DOI: 10.3934/dcds.2011.29.1393.

[17] Carrillo, J. A., Huang, Y., Santos, M. C., Vazquez, J. L. (2015). Exponential convergence towards stationary states for the 1D porous medium equation with fractional pressure. J. Differ. Equ. 258(3):736–763. DOI: 10.1016/j.jde.2014.10.003.
Lisini, S., Mainini, E., Segatti, A. (2018). A gradient flow approach to the porous medium equation with fractional pressure. *Arch. Rational Mech. Anal.* 227(2):567–606. DOI: 10.1007/s00205-017-1168-2.

Caffarelli, L., Soria, F., Vázquez, J. L. (2013). Regularity of solutions of the fractional porous medium flow. *J. Eur. Math. Soc.* 15(5):1701–1746. DOI: 10.4171/JEMS/401.

Imbert, C., Tarhini, R., Vigneron, F. (2020). Regularity of solutions of a fractional porous medium equation. *Interfaces Free Bound* 22(4):401–442. DOI: 10.4171/IFB/445.

Ambrosio, L., Serfaty, S. (2008). A gradient flow approach to an evolution problem arising in superconductivity. *Commun. Pure Appl. Math.* 61(11):1495–1539. DOI: 10.1002/cpa.20223.

Caffarelli, L., Soria, F., Vázquez, J. L. (2013). Regularity of solutions of the fractional porous medium flow. *J. Eur. Math. Soc.* 15(5):1701–1746. DOI: 10.4171/JEMS/401.

Imbert, C., Tarhini, R., Vigneron, F. (2020). Regularity of solutions of a fractional porous medium equation. *Interfaces Free Bound* 22(4):401–442. DOI: 10.4171/IFB/445.

Ambrosio, L., Serfaty, S. (2008). A gradient flow approach to an evolution problem arising in superconductivity. *Commun. Pure Appl. Math.* 61(11):1495–1539. DOI: 10.1002/cpa.20223.

Choi, Y.-P., Jeong, I.-J. (2021). Classical solutions for fractional porous medium flow. *Nonlin. Anal.* 210 (112393):13.

Nguyen, Q.-H., Vazquez, J. L. (2018). Porous medium equation with nonlocal pressure in a bounded domain. *Commun. Partial Differ. Eqns.* 43(10):1502–1539. DOI: 10.1080/03605302.2018.1475492.

Imbert, C., Tarhini, R., Vigneron, F. (2020). Regularity of solutions of a fractional porous medium equation. *Interfaces Free Bound* 22(4):401–442. DOI: 10.4171/IFB/445.

Ambrosio, L., Serfaty, S. (2008). A gradient flow approach to an evolution problem arising in superconductivity. *Commun. Pure Appl. Math.* 61(11):1495–1539. DOI: 10.1002/cpa.20223.

Caffarelli, L., Gualdani, M., Zamponi, N. (2020). Existence of weak solutions to a continuity equation with space time nonlocal Darcy law. *Commun. Partial Differ. Equ.* 45(12):1799–1819. DOI: 10.1080/03605302.2020.1814325.

Daus, E. S., Gualdani, M., Zamponi, N. (2020). Long time behavior and weak-strong uniqueness for a nonlocal porous media equation. *J. Differ. Equ.* 268(4):1820–1839. DOI: 10.1016/j.jde.2019.09.029.

de Pablo, A., Quiros, F., Rodriguez, A., Vázquez, J. L. (2012). A general fractional porous medium equation. *Commun. Pure Appl. Math.* 65(9):1242–1284. DOI: 10.1002/cpa.21408.

Vázquez, J. L. (2012). Nonlinear diffusion with fractional Laplacian operators. In: Holden, H., Karlsen, K., eds. *Nonlinear Partial Differential Equations, Abel Symposia*, Vol. 7. Berlin: Springer, pp. 1–27.

Golse, F. (2003). The mean-field limit for the dynamics of large particle systems. *Journ. Equ. Dériv. Partielles* 1–47. DOI: 10.5802/jedp.623.

Jabin, P.-E., Wang, Z., Bellomo, N., Degond, P., Tadmor, E. (2017). Mean field limit for stochastic particle systems. In: *Active Particles*, Vol. 1. Boston: Springer, pp. 379–402.

Sznitman, A. S. (1984). Nonlinear reflecting diffusion process, and the propagation of chaos and fluctuations associated. *J. Funct. Anal.* 56(3):311–336. DOI: 10.1016/S0022-1236(84)90080-6.

Oelschläger, K. (1984). A martingale approach to the law of large numbers for weakly interacting stochastic processes. *Ann. Probab.* 12:458–479.

Oelschläger, K. (1990). Large systems of interacting particles and the porous medium equation. *J. Differ. Eqns.* 88(2):294–346. DOI: 10.1016/0022-0396(90)90101-T.

Jourdain, B., Méleard, S. (1998). Propagation of chaos and fluctuations for a moderate model with smooth initial data. *Ann. Inst. H. Poincaré Probab. Stat.* 34(6):727–766. DOI: 10.1016/S0246-0203(99)80002-8.

Chen, L., Daus, E. S., Holzinger, A., Jüngel, A. (2021). Rigorous derivation of population cross-diffusion systems from moderately interacting particle systems. *J. Nonlin. Sci.* 31(94):21. DOI: 10.1007/s00033-019-09101-T.

Chen, L., Daus, E. S., Jüngel, A. (2019). Rigorous mean-field limit and cross diffusion. *Z. Angew. Math. Phys.* 70(4):21. DOI: 10.1007/s00033-019-1170-7.

Ichikawa, K., Rouzimaimaiti, M., Suzuki, T. (2012). Reaction diffusion equation with nonlocal term arises as a mean field limit of the master equation. *Discrete Contin. Dyn. Syst. S* 5(1):115–126. DOI: 10.3934/dcdss.2012.5.115.

Fontbona, J., Méleard, S. (2015). Non local Lotka–Volterra system with cross-diffusion in an heterogeneous medium. *J. Math. Biol.* 70(4):829–854. DOI: 10.1007/s00285-014-0781-z.

Figalli, A., Philipowski, R. (2008). Convergence to the viscous porous medium equation and propagation of chaos. *Alea* 4:185–203.
Appendix A. Auxiliary results

We recall some known results. The following result is proved in [47, theorem 4.33].

Lemma 19 (Young’s convolution inequality). Let $1 \leq p, r \leq \infty$, $u \in L^p(\mathbb{R}^d)$, $v \in L^q(\mathbb{R}^d)$, and $1/p + 1/q = 1 + 1/r$. Then, $u * v \in L^r(\mathbb{R}^d)$ and
\[
\|u * v\|_r \leq \|u\|_p \|v\|_q.
\]

The following lemma slightly extends [48, lemma 7.3] from the $L^2$ to the $L^p$ setting.

Lemma 20. Let $p \geq 2$ and $T > 0$. Then, the following embedding is continuous:
\[
L^p(0, T; W^{1,p}(\mathbb{R}^d)) \cap W^{1,p}(0, T; W^{-1,p}(\mathbb{R}^d)) \hookrightarrow C^0([0, T]; L^p(\mathbb{R}^d)).
\]

Proof. Let $u \in L^p(0, T; W^{1,p}(\mathbb{R}^d)) \cap W^{1,p}(0, T; W^{-1,p}(\mathbb{R}^d))$ and $0 \leq t_1 \leq t_2 \leq T$. Then,
\[
\left| \int_{\mathbb{R}^d} |u(t_2)|^p \, dx - \int_{\mathbb{R}^d} |u(t_1)|^p \, dx \right| = \left| \int_{t_1}^{t_2} \langle \partial_t u, p |u|^{p-2} u \rangle \, dt \right|
\leq p \|\partial_t u\|_{L^p(t_1, t_2; W^{-1,p}(\mathbb{R}^d))} \|u|^{p-2} u\|_{L^p(t_1, t_2; W^{1,p}(\mathbb{R}^d))},
\]
where $p' = p/(p-1)$. Direct computations using Young’s inequality lead to
\[
\|u|^{p-2} u\|_{L^p(t_1, t_2; W^{1,p}(\mathbb{R}^d))} \geq C \int_{t_1}^{t_2} \int_{\mathbb{R}^d} \left| u|^p + |u|^{p(p-2)} |\nabla u|^p \right| \, dx \, dt
\leq C \int_{t_1}^{t_2} \|u(t)|^p_{W^{1,p}(\mathbb{R}^d)} \, dt.
\]

We infer from (A1) and the continuity of the integrals with respect to the time integration boundaries that $t \mapsto \|u(t)\|_p$ is continuous and
\[
\sup_{0 < t < T} \|u(t)\|_p \leq \|u(0)\|_p + C \|\partial_t u\|_{L^p(t_1, t_2; W^{-1,p}(\mathbb{R}^d))} + C \|u\|_{L^p(0, T; W^{1,p}(\mathbb{R}^d))}.
\]
Next, let \( t \in [0, T] \) be arbitrary and let \( \tau_n \to 0 \) as \( n \to \infty \) such that \( t + \tau_n \in [0, T] \). Estimate (A2) implies that \( (u(t + \tau_n))_{n \in \mathbb{N}} \) is bounded in \( L^p(\mathbb{R}^d) \). Thus, there exists a subsequence \( (\tau_{n'}^j) \) of \( (\tau_n) \) such that \( u(t + \tau_{n'}^j) \to v(t) \) weakly in \( L^p(\mathbb{R}^d) \) as \( n' \to \infty \) for some \( v(t) \in L^p(\mathbb{R}^d) \). We can show, using estimate (A2) and dominated convergence for the integral

\[
\int_0^T \int_{\mathbb{R}^d} (u(t + \tau_{n'}, x) - v(t, x)) \phi(t, x) dx \quad \text{for} \quad \phi \in C^\infty_0(\mathbb{R}^d \times (0, T))
\]

that in the limit \( n' \to \infty \)

\[
\int_0^T \int_{\mathbb{R}^d} (u(t, x) - v(t, x)) \phi(t, x) dx = 0,
\]

which yields \( v(t) = u(t) \).

Moreover, since \( t \mapsto \|u(t)\|_p \) is continuous, we have \( \|u(t + \tau_{n'})\|_p \to \|u(t)\|_p \). Since \( L^p(\mathbb{R}^d) \) is uniformly convex, we deduce from [47, prop. 3.32] that \( u(t + \tau_{n'}) \to u(t) \) strongly in \( L^p(\mathbb{R}^d) \). Since the limit is unique, the whole sequence converges. Together with (A2), this concludes the proof.

Let \( W_1 \in C^\infty_0(\mathbb{R}^d) \) be nonnegative with \( \int_{\mathbb{R}^d} W_1(x) dx = 1 \) and define \( W_\beta(x) = \beta^{-d} W_1(x/\beta) \) for \( x \in \mathbb{R}^d \) and \( \beta > 0 \).

**Lemma 21.** Let \( 1 \leq p < \infty \) and \( u \in W^{1,p}(\mathbb{R}^d) \). Then,

\[
\|W_\beta * u - u\|_p \leq C\|\nabla u\|_p.
\]

**Proof.** We use Hölder’s inequality and the fact that \( \|W_\beta\|_{L^1(\mathbb{R}^d)} = 1 \) to find that

\[
\|W_\beta * u - u\|_p^p = \int_{\mathbb{R}^d} \left| \int_{\mathbb{R}^d} W_\beta(x - y)(u(x) - u(y)) dy \right|^p dx
\]

\[
\leq \int_{\mathbb{R}^d} \left( \int_{\mathbb{R}^d} W_\beta(x - y) dy \right)^{p-1} \left( \int_{\mathbb{R}^d} W_\beta(x - y) |u(x) - u(y)|^p dy \right) dx
\]

\[
= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} W_\beta(x - y) |u(y) - u(y)|^p dy dx
\]

\[
\leq \|\nabla u\|_p^p \int_{\mathbb{R}^d} W_\beta(z) |z|^p dz \leq C\beta^p \|\nabla u\|_p^p,
\]

which shows the lemma. \( \square \)

### Appendix B. Fractional Laplacian

We recall that the fractional Laplacian \((-\Delta)^s \) for \( 0 < s < 1 \) can be written as the pointwise formula

\[
(-\Delta)^s u(x) = c_{d,s} \int_{\mathbb{R}^d} \frac{u(x) - u(y)}{|x - y|^{d+2s}} dy,
\]

where \( c_{d,s} = \frac{4^s \Gamma(d/2 + s)}{\pi^{d/2} \Gamma(-s)} \), \( u \in H^s(\mathbb{R}^d) \), and the integral is understood as principal value if \( 1/2 \leq s < 1 \) [1, theorem 2]. The inverse fractional Laplacian \((-\Delta)^{-s} \) is defined in (2). The following lemma can be found in [46, chapter V, subsection 1.2].

**Lemma 22** (Hardy–Littlewood–Sobolev inequality). Let \( 0 < s < 1 \) and \( 1 < p < \infty \). Then, there exists a constant \( C > 0 \) such that for all \( u \in L^p(\mathbb{R}^d) \),
\[ \|(-\Delta)^{-\gamma} u\|_q \leq C\|u\|_p, \quad \text{where} \quad \frac{1}{p} = \frac{1}{q} + \frac{2s}{d}. \]

Applying Hölder’s, and then, Hardy–Littlewood–Sobolev’s inequality gives the following result.

**Lemma 23.** Let \(0 < s < 1\) and \(1 \leq p < q < \infty\). Then, there exists \(C > 0\) such that for all \(u \in L^s(\mathbb{R}^d), v \in L^r(\mathbb{R}^d),\)

\[ \|u(-\Delta)^{-s}v\|_p \leq C\|u\|_q\|v\|_r, \quad \frac{1}{q} + \frac{1}{r} = \frac{1}{p} + \frac{2s}{d}, \quad (p, q, r) \in \mathbb{R}^8 \]

**Lemma 25.** Let \(d \geq 2\) and \(1 < p < \infty\). Then, there exists \(C > 0\) such that for all \(u \in W^{1,p}(\mathbb{R}^d)\) or \(u \in W^{2,p}(\mathbb{R}^d),\)

\[ \|(-\Delta)^{s}u\|_p \leq C\|u\|_p\|\nabla u\|^{2s}_p \quad \text{if} \quad 0 < s \leq \frac{1}{2}, \]

\[ \|(-\Delta)^{s}u\|_p \leq C\|u\|_p\|D^2 u\|_p \quad \text{if} \quad 1/2 < s \leq 1. \]

**Proof.** It follows from the properties of the Riesz and Bessel potentials [46, theorem 3, p. 96] that the operator \((-\Delta)^{s} : W^{1,p}(\mathbb{R}^d) \to L^p(\mathbb{R}^d)\) is bounded for \(0 < s \leq \frac{1}{2}\), while the operator \((-\Delta)^{s} : W^{2,p}(\mathbb{R}^d) \to L^p(\mathbb{R}^d)\) is bounded for \(1/2 < s \leq 1\). Thus, if \(0 < s \leq \frac{1}{2},\)

\[ \|(-\Delta)^{s}u\|_p \leq C\|u\|_p\|\nabla u\|_p \quad \text{for} \quad u \in W^{1,p}(\mathbb{R}^d). \]

Replacing \(u\) by \(u_\lambda(x) = \lambda^{d/(p-2s)}u(\lambda x)\) with \(\lambda > 0\) yields

\[ \|(-\Delta)^{s}u\|_p \leq \frac{\|(-\Delta)^{s}u_\lambda\|_p}{\|\nabla u_\lambda\|_p} \leq C\|u\|_p\|\nabla u\|_p \]

We minimize the right-hand side with respect to \(\lambda\) giving the value \(\lambda_0 = 2s(1-2s)^{-1}\|u\|_p\|\nabla u\|_p^{-1}\) and therefore,

\[ \|(-\Delta)^{s}u\|_p \leq C\|u\|_p\|\nabla u\|^{2s}_p. \]

The case \(1/2 < s \leq 1\) is proved in a similar way. \(\square\)

**Lemma 25.** Let \(d \geq 2\), \(0 < s \leq \frac{1}{2},\) \(p \in (1, \infty),\) and \(q \in [p, \infty).\) If \(p < d/(2s),\) we assume additionally that \(q \leq dp/(d - 2sp).\) Then, there exists \(C > 0\) such that for all \(u \in W^{1,p}(\mathbb{R}^d),\)

\[ \|(-\Delta)^{-s}\nabla u\|_q \leq C\|u\|_p\|\nabla u\|^\theta_p \]

where \(\theta = 1 + d/p - d/q - 2s \in [0, 1].\)

**Proof.** The statement is true for \(s = 1/2\) since the operator \((-\Delta)^{-1/2}\nabla : L^q(\mathbb{R}^d) \to L^q(\mathbb{R}^d)\) is bounded for any \(q \in (1, \infty)\) [46, theorem 3, p. 96]. Then, the inequality follows from the standard Gagliardo–Nirenberg inequality.

Thus, let \(0 < s < 1/2.\) We claim that it is sufficient to prove that \((-\Delta)^{-s}\nabla : W^{1,p}(\mathbb{R}^d) \to L^q(\mathbb{R}^d)\) is bounded. Indeed, assume that

\[ \|(-\Delta)^{-s}\nabla u\|_q \leq C\|u\|_p\|\nabla u\|_p \quad \text{for} \quad u \in W^{1,p}(\mathbb{R}^d). \]

(B4)
Replacing, as in the proof of Lemma 24, \( u \) by \( u_\lambda(x) = \lambda^{d/q-1+2s}u(\lambda x) \) with \( \lambda > 0 \) yields
\[
\|(-\Delta)^{-\frac{1}{2}}\nabla u\|_q \leq C\lambda^{-\theta}\left(\|u\|_p + \lambda\|\nabla u\|_p\right),
\]
where \( \theta \) is defined in the statement of the theorem. Minimizing the right-hand side with respect to \( \lambda \) gives the value \( \lambda_0 = \theta(1-\theta)^{-1}\|u\|_p\|\nabla u\|_p^{-1} \) and therefore,
\[
\|(-\Delta)^{-\frac{1}{2}}\nabla u\|_q \leq C\|u\|_p^{\frac{1}{\theta}}\|\nabla u\|_p^{\frac{1}{p}}.
\]

It remains to show (B4). To this end, we distinguish two cases. First, let \( p < d/(2s) \). By assumption, \( p \leq q \leq r(1) := dp/(d-2sp) \). We apply the Hardy–Littlewood–Sobolev inequality (Lemma 22) to find that
\[
\|(-\Delta)^{-\frac{1}{2}}\nabla u\|_p \leq C\|\nabla u\|_p \leq C\left(\|u\|_p + \|\nabla u\|_p\right).
\]
Furthermore, by using (in this order) the boundedness of \( (-\Delta)^{-1/2} : L^p(\mathbb{R}^d) \to L^p(\mathbb{R}^d) \), lemma 2 in [46, p. 133], eq. (40) in [46, p. 135], and theorem 3 in [46, p. 135f],
\[
\|(-\Delta)^{-\frac{1}{2}}\nabla u\|_p = \|\nabla(-\Delta)^{-\frac{1}{2}}(-\Delta)^{1/2-s}u\|_p \leq C\|(-\Delta)^{1/2-s}u\|_p
\]
\[
\leq C\|(I-\Delta)^{1/2-s}u\|_p \leq C\|I-\Delta\|^1\|u\|_p \leq C\left(\|u\|_p + \|\nabla u\|_p\right). \tag{B5}
\]
These inequalities hold for any \( p \in (1, \infty) \). Now, it is sufficient to interpolate with \( 1/q = \mu/p + (1-\mu)/r(1) \):
\[
\|(-\Delta)^{-\frac{1}{2}}\nabla u\|_q \leq \|(-\Delta)^{-\frac{1}{2}}\nabla u\|_p^\mu\|(-\Delta)^{-\frac{1}{2}}\nabla u\|_r^{1-\mu} \leq C\left(\|u\|_p + \|\nabla u\|_p\right).
\]

Second, let \( p \geq d/(2s) \). We choose \( \lambda \in (0, d/(2sp)) \subset (0, 1) \) and apply the Hardy–Littlewood–Sobolev inequality:
\[
\|(-\Delta)^{-\frac{1}{2}}\nabla u\|_{r(\lambda)} = \|(-\Delta)^{-\frac{1}{2}}(-\Delta)^{\frac{1}{2}}\lambda^2\nabla u\|_{r(\lambda)} \leq C\|(-\Delta)^{\frac{1}{2}}\lambda^2\nabla u\|_p,
\]
where \( r(\lambda) = dp/(d-2sp) \). Since \( (1-\lambda)s < 1/2 \), we deduce from (B5) that
\[
\|(-\Delta)^{-\frac{1}{2}}\nabla u\|_{r(\lambda)} \leq C\left(\|u\|_p + \|\nabla u\|_p\right).
\]
Since \( r(\lambda) \to \infty \) as \( \lambda \to d/(2sp) \), the result follows. \( \square \)

**Appendix C. Parabolic regularity**

**Lemma 26** (Parabolic regularity). Let \( 1 < p < \infty \), \( T > 0 \) and let \( u \) be the (weak) solution to the heat equation
\[
\partial_t u - \Delta u = v, \quad u(0) = u^0 \quad \text{in} \quad \mathbb{R}^d,
\]
where \( v \in L^p(0,T;L^p(\mathbb{R}^d)) \) and \( u^0 \in W^{2,p}(\mathbb{R}^d) \). Then, there exists \( C > 0 \), depending on \( T \) and \( p \), such that
\[
\|\partial_t u\|_{L^p(0,T;L^p(\mathbb{R}^d))} + \|D^2 u\|_{L^p(0,T;L^p(\mathbb{R}^d))} \leq C\left(\|v\|_{L^p(0,T;L^p(\mathbb{R}^d))} + \|D^2 u^0\|_{L^p(\mathbb{R}^d)}\right). \tag{C1}
\]
Furthermore, if \( v = \text{div} w \) for some \( w \in L^p(0,T;L^p(\mathbb{R}^d;\mathbb{R}^d)) \), then,
\[
\|\nabla u\|_{L^p(0,T;L^p(\mathbb{R}^d))} \leq C\left(\|w\|_{L^p(0,T;L^p(\mathbb{R}^d))} + T^{\frac{1}{2p}}\|\nabla u^0\|_{L^p(\mathbb{R}^d)}\right). \tag{C2}
\]
Proof. We use a known result on the parabolic regularity for the equation
\[ \partial_t \bar{u} - \Delta \bar{u} = \nu, \quad \bar{u}(0) = 0 \quad \text{in} \quad \mathbb{R}^d. \] (C3)

It holds that [49]
\[ \| \partial_t \bar{u} \|_{L^p(0,T;L^p(\mathbb{R}^d))} + \| D^2 \bar{u} \|_{L^p(0,T;L^p(\mathbb{R}^d))} \leq C \| \nu \|_{L^p(0,T;L^p(\mathbb{R}^d))}. \] (C4)

We apply this result to \( \bar{u} = u - e^{t \Delta} u^0 \), where \( e^{t \Delta} u^0 \) is the solution to the homogeneous heat equation in \( \mathbb{R}^d \) with initial datum \( u^0 \). Then, \( \bar{u} \) solves (C3) and satisfies estimate (C4). Inserting the definition of \( \bar{u} \) and observing that \( \| D^2 (e^{t \Delta} u^0) \|_p \leq C \| D^2 u^0 \|_p \), we obtain (C1).

If \( \nu = \text{div} \, w \) for some \( w \in L^p(0,T;L^p(\mathbb{R}^d;\mathbb{R}^d)) \), the uniqueness of solutions to the heat equation yields \( u = e^{t \Delta} u^0 + \text{div} \, U \), where \( U \) solves
\[ \partial_t U - \Delta U = w, \quad U(0) = 0 \quad \text{in} \quad \mathbb{R}^d. \]

Then, we deduce from the regularity result of [49] with \( \bar{u} = U \) and \( \nu = w \) that
\[ \| D^2 U \|_{L^p(0,T;L^p(\mathbb{R}^d))} \leq C \| w \|_{L^p(0,T;L^p(\mathbb{R}^d))}. \]

Since \( \nabla u = e^{t \Delta} \nabla u^0 + \nabla \text{div} \, U \), inequality (C2) follows. \( \square \)