Rigorous Derivation of Population Cross-Diffusion Systems from Moderately Interacting Particle Systems

Li Chen\textsuperscript{1} · Esther S. Daus\textsuperscript{2} · Alexandra Holzinger\textsuperscript{2} · Ansgar Jüngel\textsuperscript{2} \textsuperscript{1}

Received: 23 October 2020 / Accepted: 2 September 2021 / Published online: 28 September 2021 © The Author(s) 2022

Abstract
Population cross-diffusion systems of Shigesada–Kawasaki–Teramoto type are derived in a mean-field-type limit from stochastic, moderately interacting many-particle systems for multiple population species in the whole space. The diffusion term in the stochastic model depends nonlinearly on the interactions between the individuals, and the drift term is the gradient of the environmental potential. In the first step, the mean-field limit leads to an intermediate nonlocal model. The local cross-diffusion system is derived in the second step in a moderate scaling regime, when the interaction potentials approach the Dirac delta distribution. The global existence of strong solutions to the intermediate and the local diffusion systems is proved for sufficiently small initial data. Furthermore, numerical simulations on the particle level are presented.

Communicated by Pierre Degond.

The first author acknowledges support from the DFG Project CH955/3-1. The last three authors acknowledge partial support from the Austrian Science Fund (FWF), Grants F65, I3401, P30000, P33010, and W1245.

\textsuperscript{1} Li Chen
chen@math.uni-mannheim.de

\textsuperscript{2} Ansgar Jüngel
juengel@tuwien.ac.at

\textsuperscript{3} Esther S. Daus
esther.daus@tuwien.ac.at

\textsuperscript{4} Alexandra Holzinger
alexandra.holzinger@tuwien.ac.at

1 School of Business Informatics and Mathematics, University of Mannheim, 68131 Mannheim, Germany

2 Institute of Analysis and Scientific Computing, TU Wien, Wiedner Hauptstraße 8–10, 1040 Vienna, Austria
Keywords  Moderately interacting particle systems · Stochastic particle systems · Cross-diffusion system · Rigorous derivation · Shigesada–Kawasaki–Teramoto model · Mean-field limit · Population dynamics

Mathematics Subject Classification  35Q92 · 35K45 · 60J70 · 82C22

1 Introduction

The aim of this paper was to derive the population cross-diffusion system of Shigesada et al. (1979) from a stochastic, moderately interacting particle system in a mean-field-type limit. More precisely, we derive the system of equations

\[
\begin{align*}
\partial_t u_i &= \text{div}(u_i \nabla U_i) + \Delta \left( \sigma_i u_i + u_i \sum_{j=1}^{n} f(a_{ij} u_j) \right), \quad u_i(0) = u_{0,i} \quad \text{in } \mathbb{R}^d, \quad t > 0, \\
\end{align*}
\]

where \( i = 1, \ldots, n \) is the species index, \( d \geq 1 \) the space dimension, \( u = (u_1, \ldots, u_n) \) is the vector of population densities, and \( U_i = U_i(x) \) are given environmental potentials. The parameters \( \sigma_i > 0 \) are the constant diffusion coefficients in the stochastic system, and \( a_{ij} \geq 0 \) are limiting values of the interaction potentials. In the linear case \( f(s) = s \), we obtain the population model in Shigesada et al. (1979). System (1) with nonlinear functions \( f \) have also been studied in the mathematical literature; see, e.g., (Chen et al. 2018; Desvillettes et al. 2015; Lepoutre and Moussa 2017). Such systems can be formally derived from random walks on a lattice, where the nonlinearity originates from the transition rates in the random-walk model (Zamponi and Jüngel 2017, Appendix A). Assuming that the transition rates depend in a nonlinear way on the densities leads to equations similar to (1). We assume that \( f \) is smooth but possibly not globally Lipschitz continuous (including power functions). Our results are valid for functions \( f_i \) depending on the species type, but we choose the same function for all species to simplify the presentation.

This paper extends the many-particle limit of Chen et al. (2019) leading to the cross-diffusion system

\[
\partial_t u_i = \text{div} \left( \sigma_i \nabla u_i + \sum_{j=1}^{n} a_{ij} u_j \nabla u_j \right) \quad \text{in } \mathbb{R}^d, \quad t > 0, \quad i = 1, \ldots, n, \tag{2}
\]

which differs from (1) by the drift term, the nonlinear function \( f \), and the diffusion term \( \text{div} \sum_{j=1}^{n} a_{ij} u_j \nabla u_i \). System (2) is the mean-field limit of the particle system for \( N \) individuals

\[
\begin{align*}
dY_{k,i}^{N,\eta} &= -\sum_{j=1}^{n} \frac{1}{N} \sum_{\ell=1}^{N} \nabla B_{ij}^{N,\eta}(Y_{k,i}^{N,\eta} - Y_{\ell,j}^{N,\eta}) dt + \sqrt{2\sigma_i} dW_i^k(t), \\
Y_{k,i}^{N,\eta}(0) &= \xi_{i}^{k}, \quad i = 1, \ldots, n, \quad k = 1, \ldots, N, \tag{3}
\end{align*}
\]
where \( (W^k_i(t))_{t \geq 0} \) are \( d \)-dimensional Brownian motions and \( \xi^1_i, \ldots, \xi^N_i \) are independent and identically distributed (iid) random variables with the common probability density function \( u_{0,i} \). The functions

\[
B^\eta_{ij}(x) = \eta^{-d} B_{ij}\left(\frac{|x|}{\eta}\right), \quad x \in \mathbb{R}^d,
\]

are interaction potentials regularizing the delta distribution \( \delta_0 \), i.e. \( B^\eta_{ij} \to a_{ij} \delta_0 \) as \( \eta \to 0 \) in the sense of distributions.

System (1) is derived from an interacting particle system for \( n \) species with particle numbers \( N_1, \ldots, N_n \), moving in the whole space \( \mathbb{R}^d \). To simplify, we set \( N = N_i \) for all \( i = 1, \ldots, n \). The key idea of this paper is to consider interacting diffusion coefficients:

\[
dX^\eta_{k,i} = -\nabla U_i(X^\eta_{k,i})dt + \left(2\sigma_i + 2 \sum_{j=1}^n f_\eta\left(\frac{1}{N} \sum_{l=1}^N B^\eta_{ij}(X^\eta_{k,i} - X^\eta_{l,j})\right)\right)^{1/2} dW_i(t),
\]

where \( f_\eta \) is a globally Lipschitz continuous approximation of \( f \) with a Lipschitz constant smaller or equal than \( \eta^{-\alpha} \) for some small \( \alpha > 0 \). In view of (4), we can interpret the scaling parameter \( \eta \) as the interaction radius of each particle.

Equations (1) are derived from system (5) in the limit \( N \to \infty, \eta \to 0 \), with the scaling relation between \( \eta \) and \( N \) given in (9) below. First, for fixed \( \eta > 0 \), we perform a classical mean-field limit from (5) to the following auxiliary intermediate system:

\[
d\overline{X}^\eta_{k,i} = -\nabla U_i(\overline{X}^\eta_{k,i})dt + \left(2\sigma_i + 2 \sum_{j=1}^n f_\eta\left(B^\eta_{ij} * u_\eta_{j,i}(\overline{X}^\eta_{k,i})\right)\right)^{1/2} dW_i(t),
\]

where we set \( u_\eta_{j,i}(\overline{X}^\eta_{k,i}) = u_{\eta,j}(t, \overline{X}^\eta_{k,i}(t)) \) for \( j = 1, \ldots, n \). The function \( u_\eta,j \) satisfies the nonlocal cross-diffusion system

\[
\partial_t u_{\eta,j} = \text{div}(u_{\eta,j} \nabla U_i) + \Delta \left(\sigma_i u_{\eta,i} + u_{\eta,j} \sum_{j=1}^n f_\eta(B^\eta_{ij} * u_{\eta,j})\right),
\]

and will be later identified as the probability density function of \( \overline{X}^\eta_{k,i} \). Note that we consider \( N \) independent copies \( \overline{X}^\eta_{k,i}, k = 1, \ldots, N \), and the intermediate system depends on \( k \) only through the initial datum.
Then, passing to the limit \( N \to \infty, \eta \to 0 \) in (5) leads to the *macroscopic system*

\[
\frac{d\hat{X}_{k,i}}{dt} = -\nabla U_i(\hat{X}_{k,i}) + \left( 2\sigma_i + 2 \sum_{j=1}^{n} f(a_{ij} u_j(\hat{X}_{k,i})) \right)^{1/2} dW^k_i(t), \tag{8}
\]

where the functions \( u_i \) satisfy (1) and can be identified as the probability density functions of \( \hat{X}_{k,i} \). In this limit, we assume that there exists \( \delta > 0 \), depending on \( n, \min_i \sigma_i \), and \( T \), such that

\[ \eta^{-2(d+1+\alpha)} \leq \delta \log N \tag{9} \]

holds, where \( \alpha \geq 0 \) depends on the Lipschitz condition of \( f \), see Assumption (A4) below, and that the function \( f \) and its derivatives or, alternatively the initial data, are sufficiently small (see Sect. 2 for details). The main result of the paper is the error estimate

\[
\sup_{k=1,...,N} \mathbb{E} \left( \sum_{i=1}^{n} \sup_{0<s<T} |X^{N,i}_{k,i}(s) - \hat{X}_{k,i}(s)|^2 \right) \leq C(T)\eta^{2(1-\alpha)}. \tag{10}
\]

We prove this estimate for the potential \( U_i(x) = -\frac{1}{2} |x|^2 \), but more general functions are possible; see Remark 1. Note that estimate (10) implies propagation of chaos; see Remark 6. In the case \( \alpha = 0 \), our scaling (9) for the multi-species case recovers the result in Jourdain and Méléard (1998), where a single-species, moderately interacting particle system with interaction in the diffusion part was considered. Our strategy is similar to that one of Jourdain and Méléard (1998) (and based on ideas of Oelschläger 1989). Since we allow for locally Lipschitz continuous nonlinearities only, we obtain a smaller convergence rate compared to Jourdain and Méléard (1998), which in fact is natural, since we approximate the nonlinearity with functions having a Lipschitz constant of order \( \eta^{-\alpha} \). A difference to Jourdain and Méléard (1998) is that the authors assume that the diffusion matrix in the stochastic part is positive definite. We do not suppose such a condition, but we need a smallness condition on the nonlinearity for the existence proofs of systems (1) and (7).

Next, we present a brief overview on the existing literature concerning mean-field limits and moderately interacting many-particle limits in the context of diffusion equations. Mean-field limits from stochastic differential equations have been investigated since the 1980s; see the reviews (Golse 2003; Jabin and Wang 2017) and the classical works by Sznitman (1984, 1991). Oelschläger proved that in the many-particle limit, weakly interacting stochastic particle systems converge to a deterministic nonlinear process (Oelschläger 1984). Later, he generalized his approach for systems of reaction-diffusion equations (Oelschläger 1989) and porous-medium-type equations with quadratic diffusion (Oelschläger 1990), by using moderately interacting particle systems. We also refer to the recent work (Chen et al. 2018), which also includes numerical simulations. As already mentioned, moderate interactions in stochastic particle system with nonlinear diffusion coefficients were investigated for the first time.
in Jourdain and Méléard (1998). Later, Stevens derived the chemotaxis model from a many-particle system (Stevens 2000). Further works concern the mean-field limit leading to reaction-diffusion equations with nonlocal terms (Ichikawa et al. 2012), the hydrodynamic limit in a two-component system of Brownian motions to the cross-diffusion Maxwell–Stefan equations (Seo 2018), and the large population limit of point measure-valued Markov processes to nonlocal Lotka–Volterra systems with cross-diffusion (Fontbona and Méléard 2015). The latter model is similar to the nonlocal system (7). The limit from the nonlocal to the local diffusion system was shown in Moussa (2020) but only for triangular diffusion matrices. The many-particle limit from a particle system driven by Lévy noise to a fractional cross-diffusion system related to (2) was recently shown in Daus et al. (2020). Furthermore, the population system (1) was derived in Daus et al. (2019) from a time-continuous Markov chain model using the BBGKY hierarchy. This paper presents, up to our knowledge, the first rigorous derivation of the Shigesada–Kawasaki–Teramoto (SKT) model (1) from a stochastic particle system in the moderate many-particle limit.

Porous-medium-type equations can be derived from stochastic interacting particle systems by assuming interactions in the drift term (Figalli and Philipowski 2008) or in the diffusion term (Jourdain and Méléard 1998). We allow for interactions in the diffusion part but in a multi-species setting. The paper (Fontbona and Méléard 2015) is concerned with a multi-species framework too, but the authors assume bounded Lipschitz continuous interaction potentials and derive a nonlocal cross-diffusion system only. We are able to relax the assumptions and derive the local cross-diffusion system (1).

Compared to the work (Daus et al. 2019), we take the limits $N \to \infty$, $\eta \to 0$ simultaneously. However, our approach also implies the two-step limit. Indeed, we can first perform the limit $N \to \infty$ for fixed $\eta > 0$ and afterward the limit $\eta \to 0$ on the PDE level; see Lemma 9 and Theorem 3. The simultaneous limit $N \to \infty$, $\eta \to 0$, satisfying the scaling relation (9), gives a more complete picture, since we can prove the convergence in expectation for the difference of the solutions to the stochastic systems (5) and (8).

Finally, we remark that the cross-diffusion models (1) and (2) have quite different structural properties; also see (Burger et al. 2020a, b). First, system (2) has a formal gradient-flow structure for each species separately, while system (1) can be written, under the detailed-balance condition (Chen et al. 2019), only in a vector-valued gradient-flow form. Second, the segregation behavior of both models is different, i.e., segregation is stronger for the solutions to (2) than for model (1); see the numerical experiments in Sect. 7.

The paper is organized as follows. We present our assumptions and main results in Sect. 2. The existence of smooth solutions to the cross-diffusion systems (1) and (7) and an error estimate for the difference of the corresponding solutions is proved in Sects. 3 and 4, respectively. The proofs are based on Banach’s fixed-point theorem and higher-order estimations. We present the full proof since the environmental potential $U_i(x) = -\frac{1}{2}|x|^2$ is not square-integrable, which requires some care; see the arguments following (22). Section 5 is concerned with the identification of the solutions to the local and nonlocal cross-diffusion systems (1) and (7), respectively, with the probability density functions associated to the particle systems (8) and (6), respectively.
Error estimate (10), the main result of the paper, is proved in Sect. 6. In Sect. 7, we present Monte–Carlo simulations for an Euler–Maruyama discretization of system (5) and compare them to the numerical results from the particle system associated to (2). In the appendix, we recall some inequalities used in the paper.

# 2 Assumptions and Main Results

We impose the following assumptions:

(A1) **Data:** $\sigma_i \in (0, \infty)$ and $\xi_1^i, \ldots, \xi_N^i$ are independent and identically distributed (iid) square-integrable random variables with the common density function $u_{0,i}$ for $i = 1, \ldots, n$ on the probability space $(\Omega, \mathcal{F}, P)$.

(A2) **Environmental potential:** $U_i(x) = -\frac{1}{2}|x|^2$, $i = 1, \ldots, n$.

(A3) **Interaction potential:** $B_{ij} \in C^\infty_0(\mathbb{R}^d)$ satisfies $\text{supp}(B_{ij}) \subset B_1(0)$, where $B_1(0)$ is the unit ball in $\mathbb{R}^d$ and $i, j = 1, \ldots, n$.

(A4) **Nonlinearity:** $f \in W^{s+1,\infty}(\mathbb{R}; [0, \infty))$ and $f_\eta \in W^{s+1,\infty}(\mathbb{R}, [0, \infty))$ is such that $f_\eta = f$ on $[-a_\eta, a_\eta]$ and the Lipschitz constant of $f_\eta$ is less than or equal to $\eta^{-\alpha}$ for a fixed $\alpha \in [0, 1)$. Here, $s > d/2 + 1$ and $a_\eta \to \infty$ as $\eta \to 0$. If $f$ is globally Lipschitz continuous, we set $\alpha = 0$ and $f_\eta = f$.

**Remark 1 (Discussion) Environmental potential:** The sign of $U_i$ guarantees that the populations are dispersed since the drift term becomes $-x \cdot \nabla u_i - u_i$. We have taken a quadratic potential $U_i$ to simplify the presentation. “Dispersive” potentials (i.e., potentials $U_i$ with $\Delta U_i \leq 0$) are needed in the analysis, since we cannot bound terms including $\Delta U_i$ if $\Delta U_i \geq 0$. It is possible to choose general (dispersive) potentials $U_i \in C^\infty(\mathbb{R}^d)$ such that $\nabla U_i$ is globally Lipschitz continuous, $D^k U_i \in L^\infty(\mathbb{R}^d)$ for $k = 2, \ldots, s + 2$, the Hessian $D^2 U_i$ is negative semidefinite, $\Delta U_i < 0$, and $D^k U_i$ for $k = 3, \ldots, s$ is sufficiently small in the $L^\infty(\mathbb{R}^d)$ norm. Thus, we may choose $U_i(x) = -|x|^2 + g(x)$ and $g$ is a smooth perturbation.

**Nonlinearity:** Since $f$ is not assumed to be globally Lipschitz continuous, we need to approximate the nonlinearity. The condition on the Lipschitz constant of $f_\eta$ ensures that we have a control on the growth of the Lipschitz constant of $f_\eta$ in the limit $N \to \infty$ and $\eta \to 0$. This growth condition is needed in the proof of Lemma 9; see (34) and thereafter. The condition $s > d/2 + 1$ ensures that the embedding $H^s(\mathbb{R}^d) \hookrightarrow W^{1,\infty}(\mathbb{R}^d)$ is continuous, and this embedding is needed to obtain solutions in $H^s(\mathbb{R}^d)$ and to derive the estimates.

We introduce some notation. We set

$$a_{ij} = \int_{\mathbb{R}^d} B_{ij}(|x|)dx, \quad i, j = 1, \ldots, n,$$

$$B_{ij}^\eta(x) = \eta^{-d} B_{ij}(|x|/\eta), A_{ij} = \|B_{ij}\|_{L^1(\mathbb{R}^d)} = \|B_{ij}^\eta\|_{L^1(\mathbb{R}^d)} \text{ and } A = \max_{i,j=1,\ldots,n} A_{ij}.$$ 

Let $C_s > 0$ be the constant of the continuous embedding $H^s(\mathbb{R}^d) \hookrightarrow L^\infty(\mathbb{R}^d)$ and set

$$I = [-2AC_s\|u_0\|_{H^s(\mathbb{R}^d)}, 2AC_s\|u_0\|_{H^s(\mathbb{R}^d)}]. \quad (11)$$
Then, for small $\eta > 0$ such that $a_\eta \ge 2AC_\varepsilon u_0 \| H^s(\mathbb{R}^d) \|$, we have $f_\eta = f$ on $I$.

First, we ensure that the nonlocal and local cross-diffusion systems (7) and (1), respectively, have global smooth solutions.

**Theorem 2.** (Existence for the nonlocal system) Let Assumptions (A2) and (A4) hold, $u_0 \in H^s(\mathbb{R}^d; \mathbb{R}^n)$ for $s > d/2 + 1$, and let $\eta > 0$ be such that $a_\eta \ge 2AC_\varepsilon H^s(\mathbb{R}^d)$. There exists $\varepsilon > 0$ depending on $u_0$ such that if $\| f \|_{C^{s+1}(I)} \le \varepsilon$, system (7) possesses a unique solution $u_\eta = (u_{\eta,1}, \ldots, u_{\eta,n})$ satisfying

$$
\begin{align*}
&u_{\eta,i} \in L^\infty(0, \infty; H^s(\mathbb{R}^d)) \cap L^2(0, \infty; H^{s+1}(\mathbb{R}^d)), \\
&\| u_{\eta} \|^2_{L^2(0,T; H^s(\mathbb{R}^d))} + \sigma_* \| \nabla u_{\eta} \|^2_{L^2(0,T; H^s(\mathbb{R}^d))} \le \| u_0 \|^2_{H^s(\mathbb{R}^d)},
\end{align*}
$$

where $0 < \sigma_* < \sigma_{\min} := \min_{i=1,\ldots,n} \sigma_i$.

The dependence of $\varepsilon$ on $u_0$ can be made more explicit. The proof shows that we need to choose $0 < \varepsilon < C\sigma_\min^{1/2} u_0^{-s} H^s(\mathbb{R}^d)$, where $C > 0$ is independent of $u_0$ and $\sigma_i$. Thus, if $\| f \|_{C^{s+1}(I)}$ is finite, the global existence result is valid for small initial data.

**Theorem 3.** (Existence for the local system) Let $u_0$ and $\eta$ satisfy the assumptions of Theorem 2. Then there exists $\varepsilon > 0$ depending on $u_0$ such that if $\| f \|_{C^{s+1}(I)} \le \varepsilon$, system (1) possesses a unique solution $u = (u_1, \ldots, u_n)$ satisfying

$$
\begin{align*}
&u_i \in L^\infty(0, \infty; H^s(\mathbb{R}^d)) \cap L^2(0, \infty; H^{s+1}(\mathbb{R}^d)), \quad i = 1, \ldots, n, \\
&\| u \|^2_{L^2(0,T; H^s(\mathbb{R}^d))} + \sigma_* \| \nabla u \|^2_{L^2(0,T; H^s(\mathbb{R}^d))} \le \| u_0 \|^2_{H^s(\mathbb{R}^d)},
\end{align*}
$$

where $0 < \sigma_* < \sigma_{\min}$. Moreover, with the solution $u_\eta$ from Theorem 2, it holds that for an arbitrary $T > 0$,

$$
\| u - u_\eta \|_{L^\infty(0,T; L^2(\mathbb{R}^d))} + \| \nabla (u - u_\eta) \|_{L^2(0,T; L^2(\mathbb{R}^d))} \le C(T) \eta.
$$

Next, we state an existence result for the stochastic particle systems (5), (6), and (8).

**Proposition 4.** Let Assumptions (A1)–(A4) hold and let $\eta > 0$, $N \in \mathbb{N}$. Then:

(i) There exist unique square-integrable adapted stochastic processes with continuous paths, which are strong solutions to systems (5), (6), and (8), respectively.

(ii) For each $t > 0$, the $(nNd)$-dimensional random variables $\vec{X}_\eta(t)$ and $\hat{X}(t)$ possess density functions $\vec{u}_\eta(t) \otimes^N$ and $\tilde{u}(t) \otimes^N$ with respect to the Lebesgue measure on $\mathbb{R}^{nNd}$, respectively.

The proof follows from Karatzas and Shreve (1991) and Nualart (2006). Indeed, Theorem 2.9 in (Karatzas and Shreve 1991, page 289) shows that there exist continuous square-integrable stochastic processes, which are strong solutions to (5), (6), and (8), respectively. Strong uniqueness is guaranteed by Theorem 2.5 in (Karatzas and Shreve 1991, page 287). We conclude from (Nualart 2006, Theorem 2.3.1) that $\vec{X}_\eta(t)$ and $\hat{X}(t)$ are absolutely continuous with respect to the Lebesgue measure and thus, they
possess density functions \( \bar{u}_\eta(t, x) \otimes^N \) and \( \hat{u}(t, x) \otimes^N \), respectively. We prove in Sect. 5 that the density functions \( \bar{u}_\eta \) and \( \hat{u} \) can be identified with \( u_\eta \) and \( u \), the solutions to (7) and (1), respectively.

The following theorem is our main result.

**Theorem 5** Let \( X_{k, i}^\eta \) and \( \hat{X}_{k, i} \) be the solutions to (5) and (8), respectively. Then there exist parameters \( \delta > 0 \), depending on \( n, \sigma_{\min} \), and \( \varepsilon > 0 \), depending on \( u_0 \), such that if \( \eta^{-2(d+1+\alpha)} \leq \delta \log N \) and \( \| f \|_{C^{s+1}(I)} \leq \varepsilon \),

\[
\sup_{k=1, \ldots, N} \mathbb{E} \left( \sum_{i=1}^n \sup_{0 < s < T} \left| (X_{k, i}^\eta - \hat{X}_{k, i})(s) \right|^2 \right) \leq C(T, n, \sigma_{\min}) \eta^{2(1-\alpha)},
\]

where \( \alpha \geq 0 \) is defined in Assumption (A4).

**Remark 6** It is well-known that this result implies propagation of chaos in the single-species case; see, e.g., (Jabin and Wang 2017, Section 3.1). In the multi-species case, this generalizes for fixed \( k \) to the convergence of the \( k \)-marginal distribution \( F_k(t) \) of \( (X_{N, 1}^\eta(t), \ldots, X_{N, k}^\eta(t)) \) at any time \( t > 0 \) towards the product measure \( \otimes_{\ell=1}^k u_{i\ell}(\cdot, t) \) as \( N \to \infty \), \( \eta \to 0 \), i.e.

\[
W_2^2(F_k(t), \otimes_{\ell=1}^k u_{i\ell}(\cdot, t)) \leq k C(T, n, \sigma_{\min}) \eta \to 0,
\]

where \( W_2 \) denotes the 2-Wasserstein distance. \( \square \)

## 3 Proof of Theorem 2

We prove the global existence of smooth solutions to the nonlocal system (7). Since \( \eta \) is fixed in the proof, we omit it for \( u_\eta \) to simplify the notation. We split the proof in several steps. In the first step, we prove the existence of local-in-time solutions satisfying \( \| u_i(t) \|_{H^s(\mathbb{R}^d)} \leq 2 \| u_0 \|_{H^s(\mathbb{R}^d)} \) for \( 0 < t < T(\eta) \) for some (possibly) small \( T(\eta) > 0 \). Actually, we show in the second step, that the factor 2 can be replaced by one. This uniform estimate allows us in the third step to conclude the global existence.

**Step 1: Local existence of solutions.** In this step, the smallness conditions on \( \eta \) and \( f \) are not needed. The idea is to apply the Banach fixed-point theorem on the space

\[
X_T := \left\{ v \in L^\infty(0, T; H^s(\mathbb{R}^d); \mathbb{R}^n) : \| v \|_{L^\infty(0, T; H^s(\mathbb{R}^d))} \leq 2 \| u_0 \|_{H^s(\mathbb{R}^d)} \right\},
\]

where \( T > 0 \) will be determined later in this proof. We define the fixed-point operator \( S : X_T \to X_T, S(v) = u \), where \( u \) is the unique solution to the linear problem

\[
\partial_t u_i = \text{div}(u_i \nabla U_i) + \Delta(u_i(\sigma_i + K_i(v(t, x)))) \quad u_i(0) = u_{0,i} \quad \text{in } \mathbb{R}^d, \quad t > 0,
\] (12)
with \( K_i(v) = \sum_{j=1}^n f_\eta(B_{ij}^\eta v_j) \geq 0, i = 1, \ldots, n \). We need to show that \( S \) is well defined. We infer from Young’s convolution inequality (Lemma 11) and the embedding \( H^s(\mathbb{R}^d) \hookrightarrow L^{\infty}(\mathbb{R}^d) \) that

\[
\sup_{0 < t < T} \| \nabla K_i(v) \|_{L^\infty(\mathbb{R}^d)} \leq \sum_{j=1}^n \| f_\eta' \|_{L^\infty(\mathbb{R})} \| \nabla B_{ij}^\eta \|_{L^1(\mathbb{R}^d)} \sup_{0 < t < T} \| v_j(t) \|_{L^\infty(\mathbb{R}^d)}
\]

\[
\leq C(\eta) \sum_{j=1}^n \| v_j \|_{L^\infty(0, T; H^s(\mathbb{R}^d))} < \infty,
\]

i.e., \( K_i(v) \) is globally Lipschitz continuous. Therefore, a Galerkin argument to verify higher-order regularity shows that, for given \( v \in X_T \), there exists a unique solution \( u_i \in L^\infty(0, T; H^s(\mathbb{R}^d)) \cap L^2(0, T; H^{s+1}(\mathbb{R}^d)) \) to (12). It remains to show that \( u = (u_1, \ldots, u_n) \in X_T \) for some \( T > 0 \). The estimations are not difficult, but since \( \nabla U_i \) is not square integrable, some care is needed.

First, we prove higher-order estimates for \( K_i(v) \). Let \( \alpha \in \mathbb{N}_0^d \) be a multi-index with order \( |\alpha| = m \leq s \). By Lemma 13 and Young’s convolution inequality,

\[
\int_0^T \| D^\alpha K_i(v) \|_{L^2(\mathbb{R}^d)}^2 \, dt \leq C \int_0^T \sum_{j=1}^n \| f_\eta' \|_{C^{m-1}(\mathbb{R})} \| B_{ij}^\eta * v_j \|_{L^\infty(\mathbb{R}^d)}^{2(m-1)} \| D^\alpha (B_{ij}^\eta * v_j) \|_{L^2(\mathbb{R}^d)}^2 \, dt
\]

\[
\leq C(\eta) \sum_{j=1}^n \| B_{ij}^\eta \|_{L^1(\mathbb{R}^d)}^{2m} \| v_j \|_{L^\infty(\mathbb{R}^d)}^{2(m-1)} \| D^\alpha v_j \|_{L^2(\mathbb{R}^d)}^2 \, dt
\]

\[
\leq C(\eta) \sum_{j=1}^n \int_0^T \| v_j \|_{H^s(\mathbb{R}^d)}^{2m} \, dt < \infty,
\]

where here and in the following, \( C > 0, C(\eta) > 0 \), etc. are generic constants with values changing from line to line. In a similar way, applying Lemmas 11 and 12,

\[
\sup_{0 < t < T} \| D^\alpha \nabla K_i(v) \|_{L^2(\mathbb{R}^d)}^2 \leq C \sup_{0 < t < T} \sum_{j=1}^n \| D^\alpha (f_\eta' (B_{ij}^\eta * v_j) \nabla B_{ij}^\eta * v_j) \|_{L^2(\mathbb{R}^d)}^2
\]

\[
\leq C \sup_{0 < t < T} \sum_{j=1}^n \left( \| f_\eta' (B_{ij}^\eta * v_j) \|_{L^\infty(\mathbb{R}^d)} \| \nabla B_{ij}^\eta \|_{L^1(\mathbb{R}^d)} \| D^m v_j \|_{L^2(\mathbb{R}^d)} + \| D^m (f_\eta' (B_{ij}^\eta * v_j)) \|_{L^2(\mathbb{R}^d)} \| \nabla B_{ij}^\eta \|_{L^1(\mathbb{R}^d)} \| v_j \|_{L^\infty(\mathbb{R}^d)} \right)^2 \leq C(\eta),
\]

since, according to Lemma 13, we can bound \( \sup_{0 < t < T} \| D^m (f_\eta' (B_{ij}^\eta * v_j)) \|_{L^2(\mathbb{R}^d)} \) in terms of \( \| f_\eta \|_{C^{s+1}(\mathbb{R})}, \| B_{ij}^\eta \|_{L^1(\mathbb{R}^d)} \), and \( \sup_{0 < t < T} \| v_j \|_{H^s(\mathbb{R}^d)} \), and it holds that \( \| \nabla B_{ij}^\eta \|_{L^1(\mathbb{R}^d)} \leq C(\eta) \).
We proceed with the proof of \( u \in X_T \) for some \( T > 0 \). Applying \( D^\alpha \) to (12), multiplying the resulting equation by \( D^\alpha u_i \), and integrating over \((0, \tau) \times \mathbb{R}^d\) for \( \tau < T \) yields

\[
\frac{1}{2} \int_{\mathbb{R}^d} |D^\alpha u_i(\tau)|^2 \, dx - \frac{1}{2} \int_{\mathbb{R}^d} |D^\alpha u_{0,i}|^2 \, dx + \sigma_i \int_0^\tau \int_{\mathbb{R}^d} |\nabla D^\alpha u_i|^2 \, dx \, dt = I_1 + I_2 + I_3, \tag{16}
\]

where

\[
I_1 = - \int_0^\tau \int_{\mathbb{R}^d} \nabla D^\alpha u_i \cdot D^\alpha (u_i \nabla U_i) \, dx \, dt,
\]

\[
I_2 = - \int_0^\tau \int_{\mathbb{R}^d} \nabla D^\alpha u_i \cdot D^\alpha (\nabla u_i K_i(v)) \, dx \, dt,
\]

\[
I_3 = - \int_0^\tau \int_{\mathbb{R}^d} \nabla D^\alpha u_i \cdot D^\alpha (u_i \nabla K_i(v)) \, dx \, dt.
\]

First, let \( |\alpha| = m = 0 \). Then, integrating by parts in \( I_1 \), using Young’s inequality, and observing that \( U_i(x) = -\frac{1}{2} |x|^2 \),

\[
I_1 = \frac{1}{2} \int_0^\tau \int_{\mathbb{R}^d} u_i^2 \Delta U_i \, dx \, dt = -\frac{d}{2} \int_0^\tau \int_{\mathbb{R}^d} u_i^2 \, dx \, dt \leq 0,
\]

\[
I_2 = - \int_0^\tau \int_{\mathbb{R}^d} K_i(v) |\nabla u_i|^2 \, dx \, dt \leq 0,
\]

\[
I_3 \leq \frac{\sigma_i}{2} \int_0^\tau \int_{\mathbb{R}^d} |\nabla u_i|^2 \, dx \, dt + \frac{1}{2\sigma_i} \|\nabla K_i(v)\|_{L^\infty(0,T;L^\infty(\mathbb{R}^d))}^2 \int_0^\tau \|u_i\|_{L^2(\mathbb{R}^d)}^2 \, dt,
\]

where we used \( K_i(v) \geq 0 \) for \( I_2 \). It follows from (13) that

\[
I_1 + I_2 + I_3 \leq \frac{\sigma_i}{2} \int_0^\tau \int_{\mathbb{R}^d} |\nabla u_i|^2 \, dx \, dt + C \int_0^\tau \|u_i\|_{L^2(\mathbb{R}^d)}^2 \, dt,
\]

\( \copyright \) Springer
where \( C > 0 \) depends on the \( L^\infty(0, T; H^s(\mathbb{R}^d)) \) norm of \( v \). Inserting this estimate into (16) with \( \alpha = 0 \) and applying the Gronwall inequality, we infer that

\[
\int_{\mathbb{R}^d} u_i(\tau)^2 dx + \frac{\sigma_i}{2} \int_0^\tau \int_{\mathbb{R}^d} |\nabla u_i|^2 dx dt \leq C(u_0) e^{C_T}.
\]

This shows that \( u_i \) is bounded in \( L^\infty(0, T; L^2(\mathbb{R}^d)) \) and \( L^2(0, T; H^1(\mathbb{R}^d)) \).

Now, let \( |\alpha| = m \geq 1 \). Then, integrating by parts, using \( /Delta_1 U_i \leq 0 \), and applying Young’s inequality again,

\[
I_1 = \frac{1}{2} \int_0^\tau \int_{\mathbb{R}^d} (D^\alpha u_i)^2 \Delta U_i dx dt
\]

\[
- \int_0^\tau \int_{\mathbb{R}^d} \nabla D^\alpha u_i \cdot (D^\alpha (u_i \nabla U_i) - D^\alpha u_i \nabla U_i) dx dt
\]

\[
\leq \frac{\sigma_i}{4} \int_0^\tau \int_{\mathbb{R}^d} |\nabla D^\alpha u_i|^2 dx dt
\]

\[
+ \sum_{0 < |\beta| \leq |\alpha|} \int_0^\tau c_\beta \|D^{\alpha-\beta} u_i\|_{L^2(\mathbb{R}^d)}^2 \|D^\beta \nabla U_i\|_{L^\infty(\mathbb{R}^d)}^2 dt
\]

\[
\leq \frac{\sigma_i}{4} \int_0^\tau \int_{\mathbb{R}^d} |\nabla D^\alpha u_i|^2 dx dt + C \int_0^\tau \|u_i\|_{H^{m-1}(\mathbb{R}^d)}^2 dt,
\]

where we used the fact that \( D^\beta \nabla U_i \) is bounded for \( |\beta| = 1 \) and vanishes for \( |\beta| > 1 \).

It follows from integration by parts, \( K_i(v) \geq 0 \), and Lemma 14 that

\[
I_2 = - \int_0^\tau \int_{\mathbb{R}^d} \nabla D^\alpha u_i \cdot (D^\alpha (\nabla u_i K_i(v)) - \nabla D^\alpha u_i K_i(v)) dx dt
\]

\[
- \int_0^\tau \int_{\mathbb{R}^d} K_i(v) |\nabla D^\alpha u_i|^2 dx dt
\]

\[
\leq \frac{\sigma_i}{4} \int_0^\tau \int_{\mathbb{R}^d} |\nabla D^\alpha u_i|^2 dx dt + C \int_0^\tau \|D K_i(v)\|_{L^\infty(\mathbb{R}^d)} \|D^{m-1} \nabla u_i\|_{L^2(\mathbb{R}^d)}
\]

\[
+ \|D^m K_i(v)\|_{L^2(\mathbb{R}^d)} \|\nabla u_i\|_{L^\infty(\mathbb{R}^d)}^2 dx dt.
\]

We infer from estimates (13) and (14) for \( K_i(v) \) and the embedding \( H^s(\mathbb{R}^d) \hookrightarrow W^{1,\infty}(\mathbb{R}^d) \) that

\[
I_2 \leq \frac{\sigma_i}{4} \int_0^\tau \int_{\mathbb{R}^d} |\nabla D^\alpha u_i|^2 dx dt + C \int_0^\tau \|u_i\|_{H^s(\mathbb{R}^d)}^2 dt.
\]
Finally, we use Lemma 12 and estimates (13) and (15) to obtain

\[
I_3 \leq \frac{\sigma_i}{4} \int_0^\tau \int_{\mathbb{R}^d} |\nabla D^\alpha u_i|^2 \, dx \, dt \\
+ C \int_0^\tau \int_{\mathbb{R}^d} \left( \|u_i\|_{L^\infty(\mathbb{R}^d)} \|D^m \nabla K_i(v)\|_{L^2(\mathbb{R}^d)} + \|D^m u_i\|_{L^2(\mathbb{R}^d)} \|\nabla K_i(v)\|_{L^\infty(\mathbb{R}^d)} \right)^2 \, dx \, dt
\]

\[
\leq \frac{\sigma_i}{4} \int_0^\tau \int_{\mathbb{R}^d} |\nabla D^\alpha u_i|^2 \, dx \, dt + C(\eta) \int_0^\tau \|u_i\|^2_{H^s(\mathbb{R}^d)} \, dt.
\]

Inserting these estimates into (16) and summing over $|\alpha| \leq s$, we arrive at

\[
\|u_i(\tau)\|^2_{H^s(\mathbb{R}^d)} + \frac{\sigma_i}{4} \int_0^\tau \|\nabla u_i\|^2_{H^s(\mathbb{R}^d)} \, dt \leq \|u_0,i\|^2_{H^s(\mathbb{R}^d)} \\
+ C(\eta) \int_0^\tau \|u_i\|^2_{H^s(\mathbb{R}^d)} \, dt.
\]

Summing over $i = 1, \ldots, n$ and applying Gronwall’s inequality gives

\[
\|u(\tau)\|^2_{H^s(\mathbb{R}^d)} \leq \|u_0\|^2_{H^s(\mathbb{R}^d)} e^{C(\eta)\tau} \leq \|u_0\|^2_{H^s(\mathbb{R}^d)} e^{C(\eta)T}.
\]

Choosing $T > 0$ sufficiently small, we can ensure that $\|u(\tau)\|_{H^s(\mathbb{R}^d)} \leq 2\|u_0\|_{H^s(\mathbb{R}^d)}$ for all $0 < \tau < T$. This shows that $u \in X_T$, i.e., the operator is well-defined.

Next, we prove that $S : X_T \to X_T$ is a contraction. Let $v, w \in X_T$ and set $\tilde{v} = S(v)$ and $\tilde{w} = S(w)$. Taking the difference of equations (12) satisfied by $\tilde{v}_i$ and $\tilde{w}_i$, respectively, using the test function $\tilde{v}_i - \tilde{w}_i$, and integrating by parts, it follows that

\[
\frac{1}{2} \int_{\mathbb{R}^d} \(
abla (\tilde{v}_i - \tilde{w}_i)(\tau) \) \, dx + \sigma_i \int_0^\tau \int_{\mathbb{R}^d} |\nabla (\tilde{v}_i - \tilde{w}_i)|^2 \, dx \, dt = I_4 + I_5 + I_6, ~ (17)
\]

where

\[
I_4 = \frac{1}{2} \int_0^\tau \int_{\mathbb{R}^d} \Delta U_i(\tilde{v}_i - \tilde{w}_i)^2 \, dx \, dt \leq 0,
\]

\[
I_5 = -\int_0^\tau \int_{\mathbb{R}^d} \nabla((\tilde{v}_i - \tilde{w}_i)K_i(v)) \cdot \nabla(\tilde{v}_i - \tilde{w}_i) \, dx \, dt,
\]

\[
I_6 = -\int_0^\tau \int_{\mathbb{R}^d} \nabla(\tilde{w}_i(K_i(v) - K_i(w))) \cdot \nabla(\tilde{v}_i - \tilde{w}_i) \, dx \, dt.
\]

Because of $K_i(v) \geq 0$ and estimate (13) for $\nabla K_i(v)$, we find that, by Young’s inequality,
It follows again from Young’s inequality that

\[ I_5 = -\int_0^\tau \int_{\mathbb{R}^d} K_i(v) \left( \nabla (\bar{v}_i - \bar{w}_i) \right)^2 \, dx \, dt \]
\[ - \int_0^\tau \int_{\mathbb{R}^d} (\bar{v}_i - \bar{w}_i) \nabla K_i(v) \cdot \nabla (\bar{v}_i - \bar{w}_i) \, dx \, dt \]
\[ \leq \frac{\sigma_i}{4} \int_0^\tau \| \nabla (\bar{v}_i - \bar{w}_i) \|_{L^2(\mathbb{R}^d)}^2 \, dt + C(\sigma_i) \int_0^\tau \| \bar{v}_i - \bar{w}_i \|_{L^2(\mathbb{R}^d)}^2 \| \nabla K_i(v) \|_{L^\infty(\mathbb{R}^d)}^2 \, dt \]
\[ \leq \frac{\sigma_i}{4} \int_0^\tau \| \nabla (\bar{v}_i - \bar{w}_i) \|_{L^2(\mathbb{R}^d)}^2 \, dt + C(\eta) \int_0^\tau \| \bar{v}_i - \bar{w}_i \|_{L^2(\mathbb{R}^d)}^2 \, dt. \]

It follows again from Young’s inequality that

\[ I_6 \leq \frac{\sigma_i}{4} \int_0^\tau \| \nabla (\bar{v}_i - \bar{w}_i) \|_{L^2(\mathbb{R}^d)}^2 \, dt + C(\sigma_i) \int_0^\tau \| \nabla \bar{w}_i \|_{L^\infty(\mathbb{R}^d)}^2 \| K_i(v) - K_j(w) \|_{L^2(\mathbb{R}^d)}^2 \, dt \]
\[ + C(\sigma_i) \int_0^\tau \| \bar{v}_i \|_{L^2(\mathbb{R}^d)}^2 \| \nabla (K_i(v) - K_j(w)) \|_{L^2(\mathbb{R}^d)}^2 \, dt. \]  

Since \( \bar{w} \in X_T \), we have \( \| \nabla \bar{w}_i \|_{L^\infty(\mathbb{R}^d)} \leq C \| \bar{w}_i \|_{H^1(\mathbb{R}^d)} \leq C(\mu_0) \) and \( \| \bar{w}_i \|_{L^\infty(\mathbb{R}^d)} \leq C(\mu_0) \). We use the fact that \( f_\eta \) and \( f'_\eta \) are globally Lipschitz continuous:

\[ \| K_i(v) - K_j(w) \|_{L^2(\mathbb{R}^d)} \leq C(\eta) \sum_{j=1}^n \| B_{ij}^\eta \ast (v_j - w_j) \|_{L^2(\mathbb{R}^d)} \]
\[ \leq C(\eta) \| v - w \|_{L^2(\mathbb{R}^d)}, \]
\[ \| \nabla (K_i(v) - K_j(w)) \|_{L^2(\mathbb{R}^d)} \leq \sum_{j=1}^n \| (f'_\eta(B_{ij}^\eta \ast v_j) - f'_\eta(B_{ij}^\eta \ast w_j)) \ast \nabla v_j \|_{L^2(\mathbb{R}^d)} \]
\[ + \sum_{j=1}^n \| f''_\eta(B_{ij}^\eta \ast w_j) \ast \nabla B_{ij}^\eta \ast (v_j - w_j) \|_{L^2(\mathbb{R}^d)} \]
\[ \leq C(\eta) \sum_{j=1}^n \| v_j - w_j \|_{L^2(\mathbb{R}^d)} \| B_{ij}^\eta \|_{L^1(\mathbb{R}^d)} \| \nabla v_j \|_{L^\infty(\mathbb{R}^d)} \]
\[ + C(\eta) \sum_{j=1}^n \| \nabla B_{ij}^\eta \|_{L^1(\mathbb{R}^d)} \| v_j - w_j \|_{L^2(\mathbb{R}^d)} \]
\[ \leq C(\eta) \| v - w \|_{L^2(\mathbb{R}^d)}. \]

Inserting these inequalities into (18) and summarizing the estimates for \( I_4, I_5, \) and \( I_6, \)
we conclude from (17) and summation over \( i = 1, \ldots, n \) that

\[ \frac{1}{2} \left( \| \bar{v} - \bar{w} \|_{L^2(\mathbb{R}^d)}^2 + \sum_{j=1}^n \| \nabla (\bar{v}_i - \bar{w}_i) \|_{L^2(\mathbb{R}^d)}^2 \right) \]
\[ \leq C_1 \int_0^\tau \| \bar{v} - \bar{w} \|_{L^2(\mathbb{R}^d)}^2 \, dt + C_2 \tau \| v - w \|_{L^\infty(0, \tau; L^2(\mathbb{R}^d))}. \]
We apply Gronwall’s inequality and the supremum over $0 < \tau < T$ to find that
\[
\| \tilde{v} - \bar{w} \|^2_{L^\infty(0,T;L^2(\mathbb{R}^d))} \leq 2C_2e^{2C_1T} T \| v - w \|^2_{L^\infty(0,T;L^2(\mathbb{R}^d))}.
\]

Thus, choosing $T > 0$ such that $2C_2e^{2C_1T} T < 1$, we infer that $S : X_T \to X_T$ is a contraction. By Banach’s fixed-point theorem, there exists a unique solution $u \in L^\infty(0,T;H^s(\mathbb{R}^d)) \cap L^2(0,T;H^{s+1}(\mathbb{R}^d))$ to (7).

**Step 2: A priori estimates.** Let $u = u_\eta$ be the unique solution to (7). We know from Step 1 that $\| u_i(t) \|_{L^\infty(\mathbb{R}^d)} \leq C_s \| u_i(t) \|_{H^s(\mathbb{R}^d)} \leq 2C_s \| u_0 \|_{H^s(\mathbb{R}^d)}$ for any $0 < t < T$. Recall that $T = T(\eta)$ and hence we do not have uniform estimates in $\eta$ even for small $T > 0$ at this step. We show in this step the estimate $\| u_i(t) \|_{H^s(\mathbb{R}^d)} \leq \| u_0 \|_{H^s(\mathbb{R}^d)}$, which allows us to conclude that the end time $T$ can be arbitrary and actually does not depend on $\eta$. We apply $D^\alpha$ to (7) (with $|\alpha| = m \leq s$), multiply the resulting equation by $D^\alpha u_i$, and integrate over $(0,\tau) \times \mathbb{R}^d$ for $\tau < T$, similarly to the corresponding estimate in Step 1:

\[
\frac{1}{2} \int_{\mathbb{R}^d} |D^\alpha u_i(\tau)|^2 dx - \frac{1}{2} \int_{\mathbb{R}^d} |D^\alpha u_{0,i}|^2 dx
\]

\[
+ \sigma_i \int_0^\tau \int_{\mathbb{R}^d} |\nabla D^\alpha u_i|^2 dx dt = I_7 + I_8 + I_9,
\]

where

\[
I_7 = -\int_0^\tau \int_{\mathbb{R}^d} \nabla D^\alpha u_i \cdot D^\alpha (u_i \nabla U_i) dx dt,
\]

\[
I_8 = -\int_0^\tau \int_{\mathbb{R}^d} \nabla D^\alpha u_i \cdot D^\alpha (\nabla u_i K_i(u)) dx dt,
\]

\[
I_9 = -\int_0^\tau \int_{\mathbb{R}^d} \nabla D^\alpha u_i \cdot D^\alpha (u_i \nabla K_i(u)) dx dt,
\]

and we recall that $K_i(u) = \sum_{j=1}^n f_\eta(B_{ij}^\eta * u_j)$.

First, let $m = 0$. Arguing similarly as for $I_1$ and $I_2$, we find that $I_7 \leq 0$ and $I_8 \leq 0$. We estimate $\nabla K_i(u) = \sum_{j=1}^n f_\eta'(B_{ij}^\eta * u_j) B_{ij}^\eta * \nabla u_j$:

\[
\| \nabla K_i(u) \|_{L^2(\mathbb{R}^d)} \leq A \sum_{j=1}^n \| f_\eta'(B_{ij}^\eta * u_j) \|_{L^\infty(\mathbb{R}^d)} \| \nabla u_j \|_{L^2(\mathbb{R}^d)},
\]

recollecting that $A = \max_{i,j=1,...,n} \| B_{ij}^\eta \|_{L^1(\mathbb{R}^d)}$. This gives for $m = 0$:

\[
I_9 \leq \| u_i \|_{L^\infty(0,T;L^\infty(\mathbb{R}^d))} \int_0^\tau \| \nabla u_i \|_{L^2(\mathbb{R}^d)} \| \nabla K_i(u) \|_{L^2(\mathbb{R}^d)} dt
\]

\[
\leq C \| u_0 \|_{H^s(\mathbb{R}^d)} \sum_{j=1}^n \| f_\eta'(B_{ij}^\eta * u_j) \|_{L^\infty(0,T;L^\infty(\mathbb{R}^d))} \int_0^\tau \| \nabla u_j \|_{L^2(\mathbb{R}^d)}^2 dt.
\]
Moreover, the last integral in (22) vanishes since $\int L^\infty(\mathbb{R}^d)$ and $
abla U_i$ because of noting that the second terms in both integrals are the same (with different signs) derivative equals 0. Thus, we have

$$I_9 \leq C \varepsilon \|u_0\|_{H^1(\mathbb{R}^d)} \sum_{j=1}^n \int_0^\tau \|\nabla u_j\|_{L^2(\mathbb{R}^d)}^2 \, dt.$$ 

Inserting these estimates into (19), we conclude that

$$\|u_i(\tau)\|_{L^2(\mathbb{R}^d)}^2 + \left(\sigma_i - C \varepsilon \|u_0\|_{H^1(\mathbb{R}^d)}\right) \int_0^\tau \|\nabla u_i\|_{L^2(\mathbb{R}^d)}^2 \, dt \leq \|u_0, i\|_{L^2(\mathbb{R}^d)}^2.$$ 

Choosing $\varepsilon > 0$ sufficiently small, this gives an estimate for $u_i$ in $L^\infty(0, T; L^2(\mathbb{R}^d)) \cap L^2(0, T; H^1(\mathbb{R}^d))$.

Next, let $m \geq 1$. The estimate for $I_7$ is delicate since $\nabla U_i \notin L^2(\mathbb{R}^d)$, and the corresponding estimate for $I_1$ cannot be directly used. We split $I_7$ into two parts:

$$I_7 = \int_0^\tau \int_{\mathbb{R}^d} D^\alpha u_i \cdot \nabla u_i + u_i \Delta U_i \, dx \, dt$$

$$= \int_0^\tau \int_{\mathbb{R}^d} D^\alpha u_i \left(D^{\alpha} \nabla u_i \cdot \nabla U_i - D^\alpha \nabla u_i \cdot \nabla U_i\right) \, dx \, dt$$

$$+ \int_0^\tau \int_{\mathbb{R}^d} D^\alpha u_i \left(D^{\alpha} (u_i \Delta U_i) - D^\alpha u_i \Delta U_i\right) \, dx \, dt,$$

noting that the second terms in both integrals are the same (with different signs) because of

$$- \int_{\mathbb{R}^d} D^\alpha u_i \cdot \nabla U_i \, dx = -\frac{1}{2} \int_{\mathbb{R}^d} \nabla (D^\alpha u_i)^2 \cdot \nabla U_i \, dx = \frac{1}{2} \int_{\mathbb{R}^d} (D^\alpha u_i)^2 \Delta U_i \, dx.$$ 

Moreover, the last integral in (22) vanishes since $\Delta U_i = -d$. In the first integral of the right-hand side of (22), the first-order derivative of $U_i$ cancels, while the second-order derivative equals $\partial^2 U_i / \partial x_j \partial x_k = -\delta_{jk}$ and all higher-order derivatives of $U_i$ vanish. Then a straightforward computation leads to

$$I_7 = -d \int_0^\tau \int_{\mathbb{R}^d} (D^\alpha u_i)^2 \, dx \, dt \leq 0.$$
For the estimates of $I_8$ and $I_9$, we need a smallness condition on $f$ and its derivatives. We apply Young’s inequality and Lemma 12 to estimate the (more delicate) term $I_9$:

$$I_9 \leq \frac{\sigma_i}{4} \int_0^\tau \|\nabla D^\alpha u_i\|_{L^2(\mathbb{R}^d)}^2 \, dt + C(\sigma_i) \int_0^\tau \|D^\alpha(u_i \nabla K_i(u))\|_{L^2(\mathbb{R}^d)}^2 \, dt$$

$$\leq \frac{\sigma_i}{4} \int_0^\tau \|\nabla D^\alpha u_i\|_{L^2(\mathbb{R}^d)}^2 \, dt + C \int_0^\tau \left(\|u_i\|_{L^\infty(\mathbb{R}^d)} \|D^m \nabla K_i(u)\|_{L^2(\mathbb{R}^d)} \right)$$

$$+ \|D^m u_i\|_{L^2(\mathbb{R}^d)} \|\nabla K_i(u)\|_{L^\infty(\mathbb{R}^d)}^2 \, dt. $$

Estimate (21) shows that $f_\eta = f$ and $|f'| \leq \varepsilon$ on $I$. Then, by similar arguments leading to (20),

$$\|\nabla K_i(u)\|_{L^\infty(\mathbb{R}^d)} \leq A \sum_{j=1}^n \|f'_\eta(B_{ij}^\eta * u_j)\|_{L^\infty(\mathbb{R}^d)} \|\nabla u_j\|_{L^\infty(\mathbb{R}^d)}$$

$$\leq A \varepsilon \|\nabla u\|_{L^\infty(\mathbb{R}^d)} \leq \varepsilon AC_s \|\nabla u\|_{H^s(\mathbb{R}^d)}. $$

Moreover, using Lemma 13, the embedding $H^s(\mathbb{R}^d) \hookrightarrow W^{1,\infty}(\mathbb{R}^d)$, and $m \leq s$,

$$\|D^m \nabla K_i(u)\|_{L^2(\mathbb{R}^d)} \leq A \sum_{j=1}^n \|\nabla u_j\|_{L^\infty(\mathbb{R}^d)} \|D^m (f'_\eta(B_{ij}^\eta * u_j))\|_{L^2(\mathbb{R}^d)}$$

$$\leq C \sum_{j=1}^n \|\nabla u_j\|_{H^s(\mathbb{R}^d)} \|f''_\eta(B_{ij}^\eta * u_j)\|_{L^\infty(\mathbb{R}^d)} \|B_{ij}^\eta * u_j\|_{L^\infty(\mathbb{R}^d)} \|B_{ij}^\eta * D^m u_j\|_{L^2(\mathbb{R}^d)}$$

$$\leq \varepsilon C \|\nabla u\|_{H^s(\mathbb{R}^d)} \|u\|_{L^\infty(\mathbb{R}^d)} \|D^m u\|_{L^2(\mathbb{R}^d)} \leq \varepsilon C \|\nabla u\|_{H^s(\mathbb{R}^d)} \|u_0\|_{H^s(\mathbb{R}^d)},$$

recalling definition (11) of the interval $I$. Consequently, the estimate for $I_9$ becomes

$$I_9 \leq \frac{\sigma_i}{4} \int_0^\tau \|\nabla D^\alpha u_i\|_{L^2(\mathbb{R}^d)}^2 \, dt + C \varepsilon^2 \|u_0\|_{H^s(\mathbb{R}^d)}^2 \int_0^\tau \|\nabla u\|_{H^s(\mathbb{R}^d)}^2 \, dt. $$

The term $I_8$ is treated in a similar way, resulting in

$$I_8 \leq \frac{\sigma_i}{4} \int_0^\tau \|\nabla D^\alpha u_i\|_{L^2(\mathbb{R}^d)}^2 \, dt + C \varepsilon^2 \|u_0\|_{H^s(\mathbb{R}^d)}^2 \int_0^\tau \|\nabla u\|_{H^s(\mathbb{R}^d)}^2 \, dt. $$

Set $\sigma_{\min} = \min_{i=1,\ldots,n} \sigma_i > 0$. We conclude from (19) after summation over $|\alpha| \leq s$ and $i = 1, \ldots, n$ that

$$\|u(\tau)\|_{H^s(\mathbb{R}^d)}^2 + \left(\sigma_{\min} - C \varepsilon^2 \|u_0\|_{H^s(\mathbb{R}^d)}^s\right) \int_0^\tau \|\nabla u\|_{H^s(\mathbb{R}^d)}^2 \, dt \leq \|u_0\|_{H^s(\mathbb{R}^d)}^2. $$

Thus, for sufficiently small $\varepsilon > 0$, we arrive at the desired estimate uniform in $\eta$.

Step 3: Global existence and uniqueness. We have proved that $\|u(\tau)\|_{H^s(\mathbb{R}^d)} \leq \|u_0\|_{H^s(\mathbb{R}^d)}$ for $0 < \tau \leq T$ for some sufficiently small $T > 0$. The value for $T$ does not
not depend on the solution. Thus, we can use \( u(T) \) as an initial datum and solve the equation in \([T, 2T]\). Repeating this argument leads to a global solution. The uniqueness of a solution follows after standard estimates, based on the global Lipschitz continuity of \( f_\eta \) and \( f'_\eta \) (see the calculations for \( I_4, I_5, \) and \( I_6 \)) and choosing \( \varepsilon > 0 \) sufficiently small.

### 4 Proof of Theorem 3

We show the global existence of smooth solutions to the local system (1) and an error estimate for the difference of the solutions to (1) and (7), respectively. First, we prove that a solution \( u_\eta \) to (7) converges to a solution \( u \) to (1) in a certain sense. Then we prove the error bound in Theorem 3 by estimating the difference \( f_\eta(B^n_{ij} \ast u_{\eta,j}) - f_\eta(a_{ij}u_{\eta,j}) \).

**Step 1. Existence and uniqueness of solutions.** Let \( u_\eta \) be a smooth solution to (7) and let \( \phi \in C_0^\infty(\mathbb{R}^d) \) with \( \text{supp}(\phi) \subset B_R, \xi \in C^0([0, T]) \) be test functions, where \( B_R \) is a ball around the origin with radius \( R > 0 \). Then the weak formulation of (7) reads as

\[
\int_0^T \langle \partial_t u_{\eta,i}, \phi(t) \rangle \, dt = -\int_0^T \int_{\mathbb{R}^d} u_{\eta,i} \nabla U_i \cdot \nabla \phi(t) \, dx \, dt - \int_0^T \int_{\mathbb{R}^d} (\sigma_i \nabla u_{\eta,i} + \nabla (u_{\eta,i} K_i(u_\eta))) \cdot \nabla \phi(t) \, dx \, dr,
\]

where \( \langle \cdot, \cdot \rangle \) is the duality pairing between \( H^{-1}(\mathbb{R}^d) \) and \( H^1(\mathbb{R}^d) \) and \( K_i(u) = \sum_{j=1}^n f_\eta(B^n_{ij} \ast u_j) \). We want to perform the limit \( \eta \to 0 \). By the uniform estimate of Theorem 2, there exists a subsequence, which is not relabeled, such that \( u_\eta \rightharpoonup u \) weakly in \( L^2(0, T; H^{s+1}(\mathbb{R}^d)) \) and weakly* in \( L^\infty(0, T; H^s(\mathbb{R}^d)) \subset L^\infty(0, T; L^\infty(\mathbb{R}^d)) \) as \( \eta \to 0 \). Our aim is to prove that \( u \) is a weak solution to (1).

It follows from the proof of Lemma 7 in Chen et al. (2019) that

\[
B^n_{ij} \ast \nabla u_{\eta,j} \rightharpoonup a_{ij} \nabla u_j \quad \text{weakly in } L^2(0, T; L^2(\mathbb{R}^d)).
\]

We claim that \( f_\eta(B^n_{ij} \ast u_{\eta,j}) \to f(a_{ij}u_j) \) strongly in \( L^2(0, T; L^2(B_R)) \). First, we observe that \( u \in L^\infty(0, T; L^\infty(\mathbb{R}^d)) \). The weak formulation (23) gives

\[
\| \partial_t u_{\eta,i} \|_{L^2(0, T; H^{-1}(B_R))} \leq \| u_{\eta,i} \|_{L^2(0, T; L^2(\mathbb{R}^d))} \| \nabla U_i \|_{L^\infty(B_R)} + \sigma_i \| \nabla u_{\eta,i} \|_{L^2(0, T; L^2(\mathbb{R}^d))} + \| \nabla u_{\eta,i} \|_{L^2(0, T; L^2(\mathbb{R}^d))} \| K_i(u_\eta) \|_{L^\infty(0, T; L^\infty(\mathbb{R}^d))} + \| u_{\eta,i} \|_{L^2(0, T; L^2(\mathbb{R}^d))} \| \nabla K_i(u_\eta) \|_{L^\infty(0, T; L^\infty(\mathbb{R}^d)).}
\]

Because of

\[
\| K_i(u_\eta) \|_{L^\infty(0, T; L^\infty(\mathbb{R}^d))} \leq \sum_{j=1}^n \| f_\eta(B^n_{ij} \ast u_{\eta,j}) \|_{L^\infty(0, T; L^\infty(\mathbb{R}^d))} \leq C \| f \|_{L^\infty(I)},
\]
\[ \lVert \nabla K_i(u_\eta) \rVert_{L^\infty(0,T;L^\infty(\mathbb{R}^d))} \leq \sum_{j=1}^{n} \lVert f'_\eta(B_{ij}^\eta \ast u_{\eta,j}) \rVert_{L^\infty(0,T;L^\infty(\mathbb{R}^d))} \times \lVert B_{ij}^\eta \ast \nabla u_{\eta,j} \rVert_{L^\infty(0,T;L^\infty(\mathbb{R}^d))} \leq C \lVert f' \rVert_{L^\infty(I)} \lVert \nabla u_\eta \rVert_{L^\infty(0,T;L^\infty(\mathbb{R}^d))} \leq C \lVert u_0 \rVert_{H^s(\mathbb{R}^d)}, \]

we obtain a uniform bound for \( \partial_t u_{\eta,i} \) in \( L^2(0, T; H^{-1}(B_R)) \) (the bound might depend on \( R \)). In particular, up to a subsequence, as \( \eta \to 0 \),

\[ \partial_t u_{\eta,i} \rightharpoonup \partial_t u_i \text{ weakly in } L^2(0, T; H^{-1}(B_R)). \]

Since \( u_\eta \) is uniformly bounded in \( L^2(0, T; H^1(B_R)) \), the Aubin–Lions lemma implies the existence of a subsequence (not relabeled) such that

\[ u_{\eta,i} \to u_i \text{ strongly in } L^2(0, T; L^2(B_R)). \]

We use the Lipschitz continuity of \( f = f_\eta \) on \( I \) to infer that

\[ \lVert f_\eta(B_{ij}^\eta \ast u_{\eta,j}) - f(a_{ij} u_j) \rVert_{L^2(0,T;L^2(B_R))} \leq C \lVert B_{ij}^\eta \ast (u_{\eta,j} - u_j) + B_{ij}^\eta \ast u_j - a_{ij} u_j \rVert_{L^2(0,T;L^2(B_R))} \leq C \lVert B_{ij}^\eta \rVert_{L^1(\mathbb{R}^d)} \lVert u_{\eta,j} - u_j \rVert_{L^2(0,T;L^2(B_R))} + \lVert B_{ij}^\eta \ast u_j - a_{ij} u_j \rVert_{L^2(0,T;L^2(B_R))} \to 0. \]

This shows the claim. In a similar way, it follows from the Lipschitz continuity of \( f'_\eta \) that \( f'_\eta(B_{ij}^\eta \ast u_{\eta,j}) \to f'(a_{ij} u_j) \) strongly in \( L^2(0, T; L^2(B_R)) \).

The previous convergences allow us to perform the limit \( \eta \to 0 \) in (23), leading to

\[ \int_0^T \langle \partial_t u_i, \phi \rangle \eta(t) dt = - \int_0^T \int_{\mathbb{R}^d} u_i \nabla U_i \cdot \nabla \phi(t) dx dt - \int_0^T \int_{\mathbb{R}^d} \nabla F_i(u) \cdot \nabla \phi(t) dx dt, \]

where \( F_i(u) = u_i (\sigma_i + \sum_{j=1}^{n} f(a_{ij} u_j)). \) Moreover, \( u_i(0) = u_{0,i} \) in \( B_R \) for any \( R > 0 \). Thus, \( u \) is a weak solution to (1). Standard estimates show that \( u \) is the unique solution, again choosing \( \varepsilon > 0 \) sufficiently small.

**Step 2: Convergence rate.** We take the difference of (7) and (1), multiply the resulting equation by \( u_{\eta,i} - u_i \), integrate over \((0, \tau) \times \mathbb{R}^d\) for any \( \tau > 0 \), and integrate by parts:

\[
\frac{1}{2} \int_{\mathbb{R}^d} (u_{\eta,i} - u_i)(\tau)^2 dx + \sigma_i \int_0^\tau \int_{\mathbb{R}^d} |\nabla (u_{\eta,i} - u_i)|^2 dx dt = \frac{1}{2} \int_0^\tau \int_{\mathbb{R}^d} \Delta U_i (u_{\eta,i} - u_i)^2 dx dt - \int_0^\tau \int_{\mathbb{R}^d} \nabla \sum_{j=1}^{n} (u_{\eta,i} f_\eta(B_{ij}^\eta \ast u_{\eta,j}) - u_i f(a_{ij} u_j)) \cdot \nabla (u_{\eta,i} - u_i) dx dt. \tag{24}
\]
The first integral on the right-hand side is nonpositive since $\Delta U_i = -d$. We split the second integral into three parts:

$$
- \int_0^\tau \int_{\mathbb{R}^d} \sum_{j=1}^n \nabla (u_{\eta,i} f_\eta (B_{ij}^\eta * u_{\eta,j}) - u_i f (a_{ij} u_{\eta,j})) \cdot \nabla (u_{\eta,i} - u_i) \, dx \, dt = J_1 + J_2 + J_3, \tag{25}
$$

where

$$
J_1 = - \int_0^\tau \int_{\mathbb{R}^d} \sum_{j=1}^n \nabla (u_{\eta,i} - u_i) f_\eta (B_{ij}^\eta * u_{\eta,j}) \cdot \nabla (u_{\eta,i} - u_i) \, dx \, dt,
$$

$$
J_2 = - \int_0^\tau \int_{\mathbb{R}^d} \sum_{j=1}^n \nabla \left( u_i \left( f_\eta (B_{ij}^\eta * u_{\eta,j}) - f_\eta (a_{ij} u_{\eta,j}) \right) \right) \cdot \nabla (u_{\eta,i} - u_i) \, dx \, dt,
$$

$$
J_3 = - \int_0^\tau \int_{\mathbb{R}^d} \sum_{j=1}^n \nabla \left( u_i \left( f_\eta (a_{ij} u_{\eta,j}) - f (a_{ij} u_{\eta,j}) \right) \right) \cdot \nabla (u_{\eta,i} - u_i) \, dx \, dt.
$$

We start with the estimate of $J_1$. The families $(B_{ij}^\eta * u_{\eta,j})$ and $(B_{ij}^\eta * \nabla u_{\eta,j})$ are bounded in $L^\infty (0, T; L^\infty (\mathbb{R}^d))$. Using $\|f_\eta\|_{L^\infty (I)} = \|f\|_{L^\infty (I)} \leq \varepsilon$ and Young’s inequality, we have

$$
J_1 \leq \|f_\eta (B_{ij}^\eta * u_{\eta,j})\|_{L^\infty (0, T; L^\infty (\mathbb{R}^d))} \int_0^\tau \|\nabla (u_{\eta,i} - u_i)\|_{L^2 (\mathbb{R}^d)}^2 \, dt + \int_0^\tau \|u_{\eta,i} - u_i\|_{L^2 (\mathbb{R}^d)} \|f_\eta' (B_{ij}^\eta * u_{\eta,j})\|_{L^\infty (0, T; L^\infty (\mathbb{R}^d))} \times \|B_{ij}^\eta * \nabla u_{\eta,j}\|_{L^\infty (0, T; L^\infty (\mathbb{R}^d))} \|\nabla (u_{\eta,i} - u_i)\|_{L^2 (\mathbb{R}^d)} \, dt
$$

$$
\leq \left( \frac{\sigma_i}{4} + \varepsilon \right) \int_0^\tau \|\nabla (u_{\eta,i} - u_i)\|_{L^2 (\mathbb{R}^d)}^2 \, dt + C (\sigma_i) \int_0^\tau \|u_{\eta,i} - u_i\|_{L^2 (\mathbb{R}^d)}^2 \, dt. \tag{26}
$$

Next, we estimate $J_2 = J_{21} + J_{22}$, where

$$
J_{21} = - \int_0^\tau \int_{\mathbb{R}^d} \nabla u_i \sum_{j=1}^n \left( f_\eta (B_{ij}^\eta * u_{\eta,j}) - f_\eta (a_{ij} u_{\eta,j}) \right) \cdot \nabla (u_{\eta,i} - u_i) \, dx \, dt,
$$

$$
J_{22} = - \int_0^\tau \int_{\mathbb{R}^d} u_i \sum_{j=1}^n \left( f_\eta' (B_{ij}^\eta * u_{\eta,j}) B_{ij}^\eta * \nabla u_{\eta,j} - f_\eta' (a_{ij} u_{\eta,j}) a_{ij} \nabla u_{\eta,j} \right) \cdot \nabla (u_{\eta,i} - u_i) \, dx \, dt.
$$
It follows that

\[
J_{21} \leq \|\nabla u_i\|_{L^\infty(0,T;L^\infty(\mathbb{R}^d))} \sum_{j=1}^{n} \int_0^T \|f_\eta(B_{ij}^\eta * u_{\eta,j}) - f_\eta(a_{ij}u_{\eta,j})\|_{L^2(\mathbb{R}^d)} \times \|\nabla (u_{\eta,i} - u_i)\|_{L^2(\mathbb{R}^d)} \, dt \\
\leq \frac{\sigma_i}{8} \sum_{j=1}^{n} \int_0^T \|\nabla (u_{\eta,i} - u_i)\|_{L^2(\mathbb{R}^d)}^2 + C \sum_{j=1}^{n} \int_0^T \|f_\eta(B_{ij}^\eta * u_{\eta,j}) - f_\eta(a_{ij}u_{\eta,j})\|_{L^2(\mathbb{R}^d)}^2 \, dt.
\]

Since both $B_{ij}^\eta * u_{\eta,j}$ and $u_{\eta,j}$ are uniformly bounded in $L^\infty(0, T; L^\infty(\mathbb{R}^d))$, we can choose $\eta > 0$ sufficiently small such that $f = f_\eta$ on $I$. On that interval, $f$ is Lipschitz continuous uniformly in $\eta$. We use this information in

\[
\left| \int_{\mathbb{R}^d} (f_\eta(B_{ij}^\eta * u_{\eta,j}) - f_\eta(a_{ij}u_{\eta,j})) g(x) \, dx \right| \\
\leq C \int_{\mathbb{R}^d} |B_{ij}^\eta * u_{\eta,j} - a_{ij}u_{\eta,j}| |g(x)| \, dx,
\]

where $g \in L^2(\mathbb{R}^d)$. Recalling that supp$(B_{ij}^\eta) \subset B_\eta(0)$ and $a_{ij} = \int_{B_\eta} B_{ij}^\eta \, dx$, we obtain

\[
\left| \int_{\mathbb{R}^d} (f_\eta(B_{ij}^\eta * u_{\eta,j}) - f_\eta(a_{ij}u_{\eta,j})) g(x) \, dx \right| \\
\leq C \int_{\mathbb{R}^d} \left| \int_{B_\eta} B_{ij}^\eta(y)(u_{\eta,j}(x - y) - u_{\eta,j}(x)) \, dy \right| |g(x)| \, dx \\
\leq C \int_{\mathbb{R}^d} \int_{B_\eta} |B_{ij}^\eta(y)| \left( \int_0^1 |\nabla u_{\eta,j}(x - ry)| \eta \, dr \right) \, dy |g(x)| \, dx \\
= C \eta \int_0^1 \int_{B_\eta} |B_{ij}^\eta(y)| \left( \int_{\mathbb{R}^d} |\nabla u_{\eta,j}(x - ry)| |g(x)| \, dx \right) \, dy \\
\leq C \eta \int_0^1 \int_{B_\eta} |B_{ij}^\eta(y)| \|\nabla u_{\eta,j}\|_{L^2(\mathbb{R}^d)} \|g\|_{L^2(\mathbb{R}^d)} \, dy \\
\leq C \eta \int_{B_\eta} |B_{ij}^\eta(y)| \|\nabla u_{\eta,j}\|_{L^2(\mathbb{R}^d)} \|g\|_{L^2(\mathbb{R}^d)} \leq C \eta \|g\|_{L^2(\mathbb{R}^d)}.
\]

By duality, we find that

\[
J_{21} \leq \frac{\sigma_i}{8} \int_0^T \|\nabla (u_{\eta,i} - u_i)\|_{L^2(\mathbb{R}^d)}^2 \, dt + C \eta^2.
\]

The integral $J_{22}$ is split into $J_{22} = J_{221} + J_{222}$, where

\[
J_{221} = -\int_0^T \int_{\mathbb{R}^d} u_i \sum_{j=1}^{n} f_\eta' (B_{ij}^\eta * u_{\eta,j}) (B_{ij}^\eta * \nabla u_{\eta,j} - a_{ij} \nabla u_{\eta,j}) \cdot \nabla (u_{\eta,i} - u_i) \, dx \, dt,
\]
\[ J_{222} = - \int_0^\tau \int_{\mathbb{R}^d} u_i \sum_{j=1}^n \left( f'_\eta(B_{ij}^\eta * u_{\eta,j}) - f'_\eta(a_{ij}u_{\eta,j}) \right) a_{ij} \nabla u_{\eta,j} \cdot \nabla (u_{\eta,i} - u_i) \, dx \, dt. \]

We infer from the uniform boundedness of \( B_{ij}^\eta * u_{\eta,j} \) in \( L^\infty(0, T; L^\infty(\mathbb{R}^d)) \) and the fact that \( f'_\eta = f' \) on \( I \) for sufficiently small \( \eta > 0 \) that

\[
J_{221} \leq \frac{\sigma_i}{16} \int_0^\tau \| \nabla (u_{\eta,i} - u_i) \|_{L^2(\mathbb{R}^d)}^2 \, dt + C \sum_{j=1}^n \int_0^\tau \| B_{ij}^\eta * \nabla u_{\eta,j} - a_{ij} \nabla u_{\eta,j} \|_{L^2(\mathbb{R}^d)}^2 \, dt \\
\leq \frac{\sigma_i}{16} \int_0^\tau \| \nabla (u_{\eta,i} - u_i) \|_{L^2(\mathbb{R}^d)}^2 \, dt + C \eta^2 \sum_{j=1}^n \int_0^\tau \| D^2 u_{\eta,j} \|_{L^2(\mathbb{R}^d)}^2 \, dt,
\]

where we estimated the difference \( B_{ij}^\eta * \nabla u_{\eta,j} - a_{ij} \nabla u_{\eta,j} \) similarly as for \( J_{21} \). Furthermore, the Lipschitz continuity of \( f'_\eta = f' \) on \( I \) leads to

\[
J_{222} \leq C \sum_{j=1}^n \int_0^\tau \| u_i \|_{L^\infty(\mathbb{R}^d)} \| B_{ij}^\eta * u_{\eta,j} \\
- a_{ij} u_{\eta,j} \|_{L^2(\mathbb{R}^d)} \| \nabla u_{\eta,j} \|_{L^\infty(\mathbb{R}^d)} \| \nabla (u_{\eta,i} - u_i) \|_{L^2(\mathbb{R}^d)} \, dt \\
\leq \frac{\sigma_i}{16} \int_0^\tau \| \nabla (u_{\eta,i} - u_i) \|_{L^2(\mathbb{R}^d)}^2 \, dt + C \eta^2 \sum_{j=1}^n \int_0^\tau \| \nabla u_{\eta,j} \|_{L^2(\mathbb{R}^d)}^2 \, dt.
\]

Summarizing these estimates, we infer that

\[
J_{22} \leq \frac{\sigma_i}{8} \int_0^\tau \| \nabla (u_{\eta,i} - u_i) \|_{L^2(\mathbb{R}^d)}^2 \, dt + C \eta^2,
\]

and combining the estimate for \( J_{21} \) and \( J_{22} \),

\[
J_2 \leq \frac{\sigma_i}{4} \int_0^\tau \| \nabla (u_{\eta,i} - u_i) \|_{L^2(\mathbb{R}^d)}^2 \, dt + C \eta^2. \tag{27}
\]

It remains to estimate \( J_3 = J_{31} + J_{32} \), where

\[
J_{31} = - \int_0^\tau \int_{\mathbb{R}^d} \sum_{j=1}^n \left( f_\eta(a_{ij}u_{\eta,j}) - f(a_{ij}u_j) \right) \nabla u_i \cdot \nabla (u_{\eta,i} - u_i) \, dx \, dt,
\]

\[
J_{32} = - \int_0^\tau \int_{\mathbb{R}^d} u_i \sum_{j=1}^n \left( f'_\eta(a_{ij}u_{\eta,j})a_{ij} \nabla u_{\eta,j} - f'(a_{ij}u_j)a_{ij} \nabla u_j \right) \cdot \nabla (u_{\eta,i} - u_i) \, dx \, dt.
\]

Similar arguments as above yield

\[
J_{31} \leq \frac{\sigma_i}{8} \int_0^\tau \| \nabla (u_{\eta,i} - u_i) \|_{L^2(\mathbb{R}^d)}^2 \, dt + C \int_0^\tau \| \nabla u_i \|_{L^\infty(\mathbb{R}^d)} \| u_{\eta} - u \|_{L^2(\mathbb{R}^d)}^2 \, dt.
\]
The second term $J_{32}$ is again split into two parts, $J_{32} = J_{321} + J_{322}$, where

$$J_{321} = \int_0^\tau \int_{\mathbb{R}^d} u_i \sum_{j=1}^n (f'_\eta(a_{ij} u_{n,j}) - f'_\eta(a_{ij} u_j)) a_{ij} \nabla u_{n,j} \cdot \nabla(u_{n,i} - u_i) \, dx \, dt,$$

$$J_{322} = \int_0^\tau \int_{\mathbb{R}^d} u_i \sum_{j=1}^n a_{ij} (f'_\eta(a_{ij} u_{j}) - f'(a_{ij} u_j) \nabla u_j) \cdot \nabla(u_{n,i} - u_i) \, dx \, dt.$$

Using the Lipschitz continuity again, $f'_\eta = f'$ on $I$, and $|f'| \leq \varepsilon$, we deduce that

$$J_{321} \leq C \|u_i\|_{L^\infty(0,T;L^\infty(\mathbb{R}^d))} \int_0^\tau \sum_{j=1}^n \|\nabla u_{n,j}\|_{L^\infty(\mathbb{R}^d)} \|u_{n,j}\|_2 \|u_{n,i} - u_i\|_2 \, dt,$$

$$J_{322} \leq C \int_0^\tau \sum_{j=1}^n \|f'(a_{ij} u_j)\|_{L^\infty(\mathbb{R}^d)} \|\nabla(u_{n,j} - u_j)\|_2 \|\nabla(u_{n,i} - u_i)\|_2 \, dt,$$

This shows that

$$J_{32} \leq \left( \frac{\sigma_i}{8} + C\varepsilon \right) \int_0^\tau \|\nabla(u_{n,i} - u_i)\|^2_{L^2(\mathbb{R}^d)} \, dt + C \int_0^\tau \|u_{n} - u\|^2_{L^2(\mathbb{R}^d)} \, dt.$$
This is the desired estimate for the last integral in (24). We conclude for sufficiently small $\varepsilon > 0$ and after summation over $i = 1, \ldots, n$ that

$$
\| (u_\eta - u)(\tau) \|_{L^2(\mathbb{R}^d)}^2 + \sigma_{\min} C \int_0^\tau \| \nabla (u_\eta - u) \|_{L^2(\mathbb{R}^d)}^2 \, dt \\
\leq C \int_0^\tau \| u_\eta - u \|_{L^2(\mathbb{R}^d)}^2 \, dt + C \eta^2.
$$

The proof ends after applying Gronwall’s inequality.

5 Links Between the SDEs and PDEs

We show that the density function $\hat{u}$ from Proposition 4 coincides with the unique weak solution $u$ to (1).

**Theorem 7** Let the assumptions of Theorem 3 hold. Let $\hat{X}_i$ for $i = 1, \ldots, n$ be the square-integrable process solving (8) with density function $\hat{u}_i$ and let $u_i$ be the unique weak solution to (1). Then $\hat{u} = (\hat{u}_1, \ldots, \hat{u}_n)$ solves the linear equation

$$
\partial_t \hat{u}_i = \text{div}(\hat{u}_i \nabla U_i) + \Delta \left( \sigma_i \hat{u}_i + \hat{u}_i \sum_{j=1}^n f(a_{ij} u_j) \right) \quad \text{in } \mathbb{R}^d, \ i = 1, \ldots, n, \quad (29)
$$

in the weak integrable sense, i.e.

$$
\int_{\mathbb{R}^d} \hat{u}_i(t) \phi(t) \, dx - \int_{\mathbb{R}^d} u_{0,i} \phi(0) \, dx - \int_0^t \int_{\mathbb{R}^d} \hat{u}_i \partial_t \phi \, dx \, ds \\
= - \int_0^t \int_{\mathbb{R}^d} \hat{u}_i \nabla U_i \cdot \nabla \phi \, dx \, dt + \int_0^t \int_{\mathbb{R}^d} \hat{u}_i \left( \sigma_i + \sum_{j=1}^n f(a_{ij} u_j) \right) \Delta \phi \, dx \, ds
$$

for all $\phi \in C_0^\infty([0, \infty) \times \mathbb{R}^d)$ and $t > 0$, where we assume that the initial datum $\hat{u}_i(0) = u_{0,i}$ fulfills

$$
\int_{\mathbb{R}^d} u_{0,i}(x) \, dx = 1, \quad \int_{\mathbb{R}^d} u_{0,i}(x) |x|^2 \, dx < \infty. \quad (30)
$$

Additionally, $\hat{u} = u$ in $(0, \infty) \times \mathbb{R}^d$, $u_i \geq 0$, and (30) is fulfilled for $u_i$ instead of $u_{0,i}$ for almost all $t > 0$ and all $i = 1, \ldots, n$.

**Proof** Since $\hat{X}_{k,i}$ depends on $k$ only via the initial data $\xi_i^k$ with the same law $u_{0,i}$, we can omit the index $k$. Let $\phi \in C_0^\infty([0, \infty) \times \mathbb{R}^d)$ and set $F_i(u) = \sigma_i + \sum_{j=1}^n f(a_{ij} u_j)$. By Itô’s lemma, we obtain

$$
\phi(t, \hat{X}_i(t)) = \phi(0, \xi_i) + \int_0^t \partial_t \phi(s, \hat{X}_i(s)) \, ds - \int_0^t \nabla U_i(s) \cdot \nabla \phi(s, \hat{X}_i(s)) \, ds
$$
\[ + \int_0^t F_i(u(\hat{X}_i(s))) \Delta \phi(s, \hat{X}_i(s))\,ds \]
\[ + \int_0^t F_i(u(\hat{X}_i(s)))^{1/2} \nabla \phi(s, X(s)) \cdot dW_i(s). \]  
(31)

We claim that the density function \( \hat{u}_i : [0, \infty) \to \mathcal{P}_2(\mathbb{R}^d) \), where \( \mathcal{P}_2(\mathbb{R}^d) \) is the space of all density functions with finite second moment, is continuous with respect to the 2-Wasserstein distance \( W_2 \). Indeed, since \( \hat{X}_i \) is square-integrable, we have \( \hat{u}_i(t) \in \mathcal{P}_2(\mathbb{R}^d) \) for almost all \( t > 0 \) and the limit \( s \to t \) in the Wasserstein distance leads to

\[ W_2(\hat{u}_i(t), \hat{u}_i(s)) = \inf \left\{ \left( \mathbb{E}(|Y_t - Y_s|^2) \right)^{1/2} : \text{Law}(Y_t) = \hat{u}_i(t), \ \text{Law}(Y_s) = \hat{u}_i(s) \right\} \]
\[ \leq \left( \mathbb{E}(|\hat{X}_i(t) - \hat{X}_i(s)|^2) \right)^{1/2} \to 0, \]

using the facts that \( \hat{X}_i \) is continuous in time and has bounded second moments. This shows the claim. We conclude that the point evaluation \( \hat{u}_i(t) \) is well defined.

The previous argumentation shows that we can apply the expectation to (31) to obtain

\[ \int_{\mathbb{R}^d} \hat{u}_i(t) \phi(t)\,dx = \int_{\mathbb{R}^d} u_{0,i} \phi(0)\,dx + \int_0^t \int_{\mathbb{R}^d} \hat{u}_i(s) \partial_t \phi(s)\,dx\,ds \]
\[ - \int_0^t \int_{\mathbb{R}^d} \hat{u}_i(s) \nabla U_i \cdot \nabla \phi(s)\,dx\,ds + \int_0^t \int_{\mathbb{R}^d} \hat{u}_i(s) F_i(u(s)) \Delta \phi(s)\,dx\,ds. \]

This is the very weak formulation of (29), showing the first part of the theorem.

Next, we verify that the solution to (29) is unique. More precisely, we take \( u_0 = 0 \) and show that \( \hat{u}_i(t) = 0 \) for almost all \( t > 0 \). The statement is usually proved by a duality argument. However, the coefficients of the dual problem associated to (29) are not regular enough such that we need to regularize it. As the proof is rather standard but tedious, we only sketch the arguments. Let \( \chi_k \) be a family of mollifiers and consider the regularized dual backward problem on the ball \( B_R \) around the origin with radius \( R > 0 \):

\[ \partial_t w_{k,R} - \nabla U_i \cdot \nabla w_{k,R} + (\chi_k * F_i(u)) \Delta w_{k,R} = 0 \quad \text{in } B_R, \ 0 < s < t, \]
\[ w_{k,R} = 0 \quad \text{on } \partial B_R, \quad w_{k,R}(t) = g \in C_0^\infty(B_R) \quad \text{in } B_R. \]

We extend the unique smooth solution \( w_{k,R} \) to the whole space by setting \( w_{k,R} = 0 \) on \( \mathbb{R}^d \setminus B_R \). Since the extension may be not smooth, we choose a cut-off function \( \psi_R \in C^\infty(\mathbb{R}^d) \) and use \( w_{k,R} \psi_R \) as an admissible test function in the very weak formulation of (29). Standard estimations give bounds for \( w_{k,R} \) uniform in \( k \) and \( R \). Then, passing to the limit \( k \to \infty, R \to \infty \) in the weak formulation shows that \( \int_{\mathbb{R}^d} g(x) \hat{u}_i(s, x)\,dx = 0 \), and since \( g \) was arbitrary, we conclude that \( \hat{u}_i(s) = 0 \) for \( 0 < s < t \).

The weak solution \( u \) to (1) is also a very weak solution to (29). Therefore, by the previous uniqueness result, \( \hat{u} = u \).
Similar arguments lead to the following result that relates the solutions $\bar{u}_\eta$ and $u_\eta$.

**Theorem 8** Let the assumptions of Theorem 2 hold and let $\eta > 0$. Let $\bar{X}_{k,i}^\eta$ for $i = 1, \ldots, n$ and $k = 1, \ldots, N$ be the square-integrable process solving (6) with density function $\bar{u}_{\eta,i}$. Then $\bar{u}_\eta = (\bar{u}_{\eta,1}, \ldots, \bar{u}_{\eta,n})$ solves the linear problem

$$\partial_t \bar{u}_{\eta,i} = \text{div}(\bar{u}_{\eta,i} \nabla U_i) + \Delta \left( \sigma_i \bar{u}_{\eta,i} + \bar{u}_{\eta,i} \sum_{j=1}^n f_\eta(B_{ij}^\eta * u_{\eta,j}) \right)$$

in $\mathbb{R}^d$, $i = 1, \ldots, n$, with initial datum $\bar{u}_{\eta,i}(0) = u_{0,i}$, which fulfils (30), where $u_{\eta,i}$ is the unique weak solution to (7). Then $\bar{u}_\eta = u_\eta$ in $(0, \infty) \times \mathbb{R}^d$, $u_{\eta,i} \geq 0$, and

$$\int_{\mathbb{R}^d} u_{\eta,i}(x,t)dx = 1, \quad \int_{\mathbb{R}^d} u_{\eta,i}(x,t)|x|^2dx < \infty$$

for almost all $t > 0$ and all $i = 1, \ldots, n$.

**6 Proof of Theorem 5**

The proof is split into two parts. We estimate first the square mean error of the difference $X_{k,i}^{N,\eta} - \bar{X}_{k,i}^\eta$, where $\bar{X}_{k,i}^\eta$ is the solution to the intermediate system (6). In fact, this error bound is a generalization of a result due to Sznitman (1991). Essential for this step are the facts that the Lipschitz constant of $B_{ij}^\eta$ is of order $\eta^{-d-1}$, while the Lipschitz constant of $f_\eta$ is of order $\eta^{-\alpha}$. Second, we estimate the square mean error of the difference $\bar{X}_{k,i}^\eta - \bar{X}_{k,i}$, based on an estimate of $f_\eta(B_{ij}^\eta * u_j) - f_\eta(a_{ij}u_j)$ in $L^2$, which is of the order of $\eta^{1-\alpha}$.

**Lemma 9** Let $X_{k,i}^{N,\eta}$ and $\bar{X}_{k,i}^\eta$ be the solutions to (5) and (8), respectively, in the sense of Proposition 4. Under the assumptions of Theorem 5, there exists $\delta > 0$, depending on $n$, $\sigma_{\text{min}}$, and $T$, such that if $\eta^{-2(d+1+\alpha)} \leq 2 \log N$, where $\alpha > 0$ is fixed in Assumption (A4), we have

$$\sup_{k=1,\ldots,N} \mathbb{E} \left( \sum_{i=1}^n \sup_{0<s<T} \left| (X_{k,i}^{N,\eta} - \bar{X}_{k,i}^\eta)(s) \right|^2 \right) \leq C(T, n, \sigma_{\text{min}}) N^{-1+(T+1)} C(n, T, \sigma_{\text{min}})^\delta,$$

where $C(T, n, \sigma_{\text{min}}) > 0$ is a positive constant.

**Proof** The process $D_{k,i}^{N,\eta} := X_{k,i}^{N,\eta} - \bar{X}_{k,i}^\eta$ solves

$$D_{k,i}^{N,\eta}(s) = E_{1,i}(s) + E_{2,i}(s), \quad 0 \leq s \leq T,$$

where

$$E_{1,i}(s) = -\int_0^s \left( \nabla U_i(X_{k,i}^{N,\eta}(t)) - \nabla U_i(\bar{X}_{k,i}^\eta(t)) \right) dt,$$

and

$$E_{2,i}(s) = \int_0^s \left( \nabla U_i(X_{k,i}^{N,\eta}(t)) - \nabla U_i(\bar{X}_{k,i}^\eta(t)) \right) dt.$$
\[ E_{2,i}(s) = \int_0^s (E_{21}(t) - E_{22}(t))dW_i^k(t), \]
\[ E_{21}(t) = \left(2\sigma_i + 2\sum_{j=1}^n f_\eta \left(\frac{1}{N} \sum_{\ell=1, (i, j) \neq (k, i)} B_{ij}^\eta(X_{k,i}^N, \eta(t) - X_{\ell,i}^N, \eta(t))\right)\right)^{1/2}, \]
\[ E_{22}(t) = \left(2\sigma_i + 2\sum_{j=1}^n f_\eta(B_{ij}^\eta \ast u_{\eta,j}(t, \mathcal{X}_{k,i}^\eta(t)))\right)^{1/2}. \]

We use the global Lipschitz continuity of \( \nabla U_i \) and the Fubini theorem to estimate the first term:
\[
\mathbb{E}\left(\sup_{0 < s < T} |E_{1,i}(s)|^2\right) \leq CT \mathbb{E} \int_0^T \left|\left(X_{k,i}^N, \eta - \mathcal{X}_{k,i}^\eta\right)(s)\right|^2 ds \\
\leq CT \int_0^T \mathbb{E}\left(\sup_{0 < s < t} \left|\left(X_{k,i}^N, \eta - \mathcal{X}_{k,i}^\eta\right)(s)\right|^2\right) dt.
\]

Summing over \( i = 1, \ldots, n \) and taking the supremum over \( k = 1, \ldots, N \) leads to
\[
\sup_{k=1, \ldots, N} \mathbb{E}\left(\sum_{i=1}^n \sup_{0 < s < T} |E_{1,i}(s)|^2\right) \\
\leq CT \int_0^T \sup_{k=1, \ldots, N} \mathbb{E}\left(\sup_{0 < s < t} \left|\left(X_{k,i}^N, \eta - \mathcal{X}_{k,i}^\eta\right)(s)\right|^2\right) dt. \tag{33}
\]

Next, we apply the Burkholder–Davis–Gundy inequality (Karatzas and Shreve 1991, Theorem 3.28) to the second term \( E_{2,i} \) and use the Lipschitz continuity of \( x \mapsto (2\sigma_i + x)^{1/2} \) for \( x \geq 0 \):
\[
\mathbb{E}\left(\sup_{0 < s < T} |E_{2,i}(s)|^2\right) \leq C \mathbb{E} \int_0^T \left(E_{21}(t) - E_{22}(t)\right)^2 dt \\
\leq C \mathbb{E} \int_0^T \left[\sum_{j=1}^n f_\eta \left(\frac{1}{N} \sum_{\ell=1, (\ell, j) \neq (k, i)} B_{ij}^\eta(X_{k,i}^N, \eta(t) - X_{\ell,i}^N, \eta(t))\right) \right. \\
\left. - \sum_{j=1}^n f_\eta(B_{ij}^\eta \ast u_{\eta,j}(t, \mathcal{X}_{k,i}^\eta(t)))\right]^2 dt \\
= C \mathbb{E} \int_0^T \left[\sum_{j=1}^n (L_j^1(t) + L_j^2(t) + L_j^3(t))\right]^2 dt \\
\leq C(n) \mathbb{E} \int_0^T \sum_{j=1}^n (L_j^1(t)^2 + L_j^2(t)^2 + L_j^3(t)^2) dt, \tag{34}
\]

\( \copyright \) Springer
where

\[ L_j^1(t) = f_\eta \left( \frac{1}{N} \sum_{\ell=1}^{N} B_{ij}^\eta (X_{k,i}^{N,\eta}(t) - X_{\ell,j}^{N,\eta}(t)) \right) \]

\[ - f_\eta \left( \frac{1}{N} \sum_{\ell=1}^{N} B_{ij}^\eta (\overline{X}_{k,i}^{\eta}(t) - X_{\ell,j}^{N,\eta}(t)) \right) , \]

\[ L_j^2(t) = f_\eta \left( \frac{1}{N} \sum_{\ell=1}^{N} B_{ij}^\eta (\overline{X}_{k,i}^{\eta}(t) - X_{\ell,j}^{N,\eta}(t)) \right) \]

\[ - f_\eta \left( \frac{1}{N} \sum_{\ell=1}^{N} B_{ij}^\eta (\overline{X}_{k,i}^{\eta}(t) - \overline{X}_{\ell,j}^{\eta}(t)) \right) , \]

\[ L_j^3(t) = f_\eta \left( \frac{1}{N} \sum_{\ell=1}^{N} B_{ij}^\eta (\overline{X}_{k,i}^{\eta}(t) - \overline{X}_{\ell,j}^{\eta}(t)) \right) - f_\eta (B_{ij}^\eta * u_{\eta,j}(t, \overline{X}_{k,i}^{\eta}(t))). \]

We estimate these three terms separately. By construction, the Lipschitz constant of \( f_\eta \) can be estimated by \( L_f \leq \eta^{-\alpha} \). Moreover, the Lipschitz constant of \( B_{ij}^\eta (x) = \eta^{-d} B_{ij}(|x|/\eta) \) is computed by \( L_B = \max_{i,j=1,...,n} \| \nabla B_{ij}^\eta \|_{L^\infty(\mathbb{R}^d)} \leq C \eta^{-d-1} \). This shows that

\[ \left| L_j^1(t) \right| \leq L_f \left| \frac{1}{N} \sum_{\ell=1}^{N} \left( B_{ij}^\eta (X_{k,i}^{N,\eta}(t) - X_{\ell,j}^{N,\eta}(t)) - B_{ij}^\eta (\overline{X}_{k,i}^{\eta}(t) - X_{\ell,j}^{N,\eta}(t)) \right) \right| \]

\[ \leq L_f L_B \left| X_{k,i}^{N,\eta}(t) - \overline{X}_{k,i}^{\eta}(t) \right| \leq C \eta^{-d-1-\alpha} \left| X_{k,i}^{N,\eta}(t) - \overline{X}_{k,i}^{\eta}(t) \right|. \]

Therefore, by Fubini’s theorem,

\[ \mathbb{E} \int_0^T \sum_{j=1}^n \left| L_j^1(t) \right|^2 dt \leq C(n) \eta^{-2(d+1+\alpha)} \mathbb{E} \int_0^T \left| X_{k,i}^{N,\eta}(t) - \overline{X}_{k,i}^{\eta}(t) \right|^2 dt \]

\[ \leq C(n) \eta^{-2(d+1+\alpha)} \int_0^T \sup_{k=1,...,N} \mathbb{E} \left( \sup_{0<s<t} \left| X_{k,i}^{N,\eta}(t) - \overline{X}_{k,i}^{\eta}(t) \right|^2 \right) dt. \quad (35) \]

We can estimate the second term \( L_j^2(t) \) in a similar way, leading to

\[ \mathbb{E} \int_0^T \sum_{j=1}^n L_j^2(t)^2 dt \leq C(n) \eta^{-2(d+1+\alpha)} \]
\[ \times \int_0^T \sup_{\ell=1,\ldots,N} \mathbb{E} \left( \sup_{0<s<t} \sum_{j=1}^n \left| X_{\ell,j}^{\eta}(t) - \overline{X}_{\ell,j}^{\eta}(t) \right|^2 \right) dt. \]  

(36)

The third term \( L_3^3(t) \) has to be treated in a different way. First, we use the Lipschitz continuity of \( f_\eta \) to find that

\[ L_3^3(t) \leq \frac{C(n)}{N \eta^\alpha} \sum_{\ell=1}^N \left( B_{ij}^\eta (\overline{X}_{k,i}^\eta(t) - \overline{X}_{\ell,j}^\eta(t)) - B_{ij}^\eta * u_{\eta,j}(\overline{X}_{k,i}^\eta(t)) - \frac{1}{\eta^\alpha} B_{i\ell}(0) \right). \]

This implies that

\[ \mathbb{E} \int_0^T \sum_{j=1}^n L_3^3(t)^2 dt \leq \frac{C(n,T)}{N^2 \eta^{2(d+\alpha)}} \]

\[ + \frac{C(n)}{N^2 \eta^{2\alpha}} \sum_{j=1}^n \int_0^T \mathbb{E} \left( \sum_{\ell=1}^N \left( B_{ij}^\eta (\overline{X}_{k,i}^\eta(t) - \overline{X}_{\ell,j}^\eta(t)) - B_{ij}^\eta * u_{\eta,j}(\overline{X}_{k,i}^\eta(t)) \right) \right)^2 dt. \]

(37)

It remains to estimate the expectation. To this end, we introduce

\[ D_{(k,i),(\ell,j)}(t) := B_{ij}^\eta (\overline{X}_{k,i}^\eta(t) - \overline{X}_{\ell,j}^\eta(t)) - B_{ij}^\eta * u_{\eta,j}(t, \overline{X}_{k,i}^\eta(t)), \quad (\ell, j) \neq (k, i). \]

The processes \( \overline{X}_{k,i}^\eta \) and \( \overline{X}_{\ell,j}^\eta \) are independent, since for \( i = j \), we are considering \( N \) independent copies of the same process and for \( i \neq j \), the equation fulfilled by \( \overline{X}_{k,i}^\eta \) does not depend on the process \( \overline{X}_{\ell,j}^\eta \). If \( (k, i) \neq (\ell, j) \), \( (k, i) \neq (m, j) \), and \( \ell \neq m \), the processes \( D_{(k,i),(\ell,j)}(t) \) and \( D_{(k,i),(m,j)}(t) \) are orthogonal, since

\[ \mathbb{E} \left( D_{(k,i),(\ell,j)}(t) D_{(k,i),(m,j)}(t) \right) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} B_{ij}^\eta(x-y)B_{ij}^\eta(x-z)u_{\eta,j}(t,y)u_{\eta,j}(t,z) dy dz \]

\[ - 2 \int_{\mathbb{R}^d} B_{ij}^\eta(x-y)u_{\eta,j}(t,y)(B_{ij}^\eta * u_{\eta,j})(t,y) dy \]

\[ + (B_{ij}^\eta * u_{\eta,j})(t,x)(B_{ij}^\eta * u_{\eta,j})(t,x)u_{\eta,i}(t,x) dx = 0. \]

Together with \( \mathbb{E}(D_{(k,i),(\ell,j)}) = 0 \), this shows that the processes \( D_{(k,i),(\ell,j)} \) are uncorrelated.

However, if \( (k, i) \neq (\ell, j) \), \( (k, i) \neq (m, j) \), and \( \ell = m \), the expectation does not vanish:

\[ \mathbb{E}(D_{(k,i),(\ell,j)}(t)^2) = \int_{\mathbb{R}^d} \left( (B_{ij}^\eta * u_{\eta,j})(t,x)(B_{ij}^\eta * u_{\eta,j})(t,x) + \int_{\mathbb{R}^d} (B_{ij}^\eta(x-y))^2 u_{\eta,j}(t,y) \right) \]

\[ \mathbb{E} \] Springer
\[-2B^\eta_{ij}(x-y)u_{\eta,j}(t,y)(B^\eta_{ij} * u_{\eta,j}(t,x)) \, dy \] \[= \int_{\mathbb{R}^d} (((B^\eta_{ij})^2 * u_{\eta,j})(t,x) - (B^\eta_{ij} * u_{\eta,j})(t,x)^2)u_{\eta,i}(t,x) \, dx . \]

This expression is independent of the particle index \( k \) and \( \ell \), it depends only on the species numbers \( i \) and \( j \). The case \((k, i) = (\ell, j)\) can be treated in a similar way with the difference that, since \( D_{(k,i),(k,i)}(t) = \eta^{-d} B_{ii}(0) - B^\eta_{ii} * u_{\eta,\eta}(X^\eta_{k,i}(t)) \), we obtain for \( \mathbb{E}(D_{(k,i),(k,i)}(t)D_{(k,i),(m,j)}(t)) \) an additional term of order \( \eta^{-2d} \). Hence, we infer from (37) and the previous computation that

\[
\mathbb{E} \int_0^T \sum_{j=1}^n L_j(t)^2 \, dt - \frac{C(n,T)}{N^2 \eta^{2(d+\alpha)}} = \frac{C(n)}{N^2 \eta^{2d}} \sum_{j=1}^n \sum_{i=1}^N \int_0^T \mathbb{E}(D_{(k,i),(\ell,j)}(t)^2) \, dt \leq \frac{C(n)}{N^2 \eta^{2d}} \sum_{j=1}^n \int_0^T \left( \|B^\eta_{ij}\|^2_{L^2(\mathbb{R}^d)} \|u_{\eta,j}\|_{L^\infty(\mathbb{R}^d)} + \|B^\eta_{ij}\|_{L^1(\mathbb{R}^d)} \|u_{\eta,j}\|_{L^2(\mathbb{R}^d)} \left( 1 + \frac{1}{\eta^{2d}} \right) \right) \, dt \leq \frac{C(T,n)}{N \eta^{2(d+\alpha)}} , \tag{38}
\]

recalling that \( \|B^\eta_{ij}\|_{L^2(\mathbb{R}^d)} \leq C \eta^{-d/2} \) and \( \|B^\eta_{ij}\|_{L^1(\mathbb{R}^d)} = A_{ij} \leq A \) and choosing \( \eta < 1 \).

Inserting estimates (35), (36), and (38) for \( L^m_j(t) \) \((m = 1, 2, 3)\) into (34), we conclude that

\[
\sup_{k=1,\ldots,N} \mathbb{E} \left( \sum_{i=1}^n \sup_{0<s<T} |E_{2,i}(s)|^2 \right) \leq \frac{C(T,n)}{N \eta^{2(d+\alpha)}} + C(n, \sigma_{\min}) \eta^{-2(d+\alpha)} \int_0^T \sup_{k=1,\ldots,N} \mathbb{E} \left( \sup_{0<s<T} |X^N_{k,i}(t) - \overline{X}^\eta_{k,i}(t)| \right) \, dt .
\]

We infer from (32), estimate (33), and the previous estimate for \( E_{2,i} \) that

\[
S(T) := \sup_{k=1,\ldots,N} \mathbb{E} \left( \sum_{i=1}^n \sup_{0<s<T} |D^N_{k,i}(s)|^2 \right) \leq \frac{C(T,n)}{N \eta^{2(d+\alpha)}} + C(n, \sigma_{\min}) \eta^{-2(d+\alpha)} + T \int_0^T S(t) \, dt .
\]
Note that the function $S$ is continuous because of the continuity of the paths of $X_{k,i}^{N,\eta}$ and $\overline{X}_{k,i}^{\eta}$. Therefore, by Gronwall’s inequality, we have

$$S(T) \leq \frac{C(T, n)}{N \eta^{2(d+\alpha)}} \exp\left(C(n, T, \sigma_{\min})\eta^{-2(d+1+\alpha)}T\right).$$

We choose $\delta > 0$ such that $C(n, T, \sigma_{\min})T\delta < 1$ and $\eta > 0$ such that $\eta^{-2(d+1+\alpha)} \leq \delta \log N$. Then

$$S(T) \leq \frac{1}{N} C(T, n) \exp\left(C(n, T, \sigma_{\min})T\delta \log N\right) = C(T, n)N^{-1+C(n,T,\sigma_{\min})T\delta}. $$

This finishes the proof. \qed

Next, we prove an error estimate for the difference $X_{k,i}^{\eta} - \hat{X}_{k,i}$.

**Lemma 10** Let $X_{k,i}^{\eta}$ and $\hat{X}_{k,i}$ be the solutions to (6) and (8) in the sense of Proposition 4. Under the assumptions of Theorem 5, it holds for small $\eta > 0$ that

$$\sup_{k=1,...,N} \mathbb{E}\left( \sum_{i=1}^{n} \sup_{0<s<T} \left| (X_{k,i}^{\eta} - \hat{X}_{k,i})(s) \right|^2 \right) \leq C(T, \sigma_{\min})\eta^{2(1-\alpha)}.$$

**Proof** Since we are considering $N$ independent copies, we can omit the particle index $k$. Set $D_{i}^{\eta}(s) := X_{k,i}^{\eta}(s) - \hat{X}_{k,i}(s)$. Then, similarly as in the proof of Lemma 9, $D_{i}^{\eta}(s) = D_{1}(s) + D_{2}(s)$, where

$$D_{1}(s) = -\int_{0}^{s} (\nabla U_{i}(X_{i}^{\eta}(t)) - \nabla U_{i}(\hat{X}_{i}(t)))dt,$$

$$D_{2}(s) = \int_{0}^{s} \left[ \left( 2\sigma_{i} + 2 \sum_{j=1}^{n} f_{\eta}(B_{ij}^{\eta} * u_{\eta,j}(X_{i}^{\eta})) \right)^{1/2} 
- \left( 2\sigma_{i} + 2 \sum_{j=1}^{n} f(a_{ij}u_{j}(\hat{X}_{i})) \right)^{1/2} \right] dW_{i}(t).$$

We infer from the Lipschitz continuity of $\nabla U_{i}$ and Fubini’s theorem that

$$\mathbb{E}\left( \sup_{0<s<T} |D_{1}(s)|^{2} \right) \leq CT \mathbb{E}\left( \int_{0}^{T} |X_{i}^{\eta}(s) - \hat{X}_{i}(s)|^{2} ds \right) \leq CT \int_{0}^{T} \mathbb{E}\left( \sup_{0<s<T} |D_{i}^{\eta}(s)|^{2} \right) dt. \quad (39)$$
Similarly as in the proof of Lemma 9, we use for $D_2$ the Burkholder–Davis–Gundy inequality and the Lipschitz continuity of $x \mapsto (2\sigma_i + x)^{1/2}$ on $[0, \infty)$ to obtain

$$
\mathbb{E}\left( \sup_{0 < s < T} |D_2(s)|^2 \right) \leq C \mathbb{E} \int_0^T \left( \sum_{j=1}^n \left( f(a_{ij} u_j(\hat{X}_i)) - f_\eta(B_{ij}^\eta * u_{\eta,j}(\bar{X}_i)) \right) \right)^2 dt
$$

$$
\leq C(n)(D_{21} + D_{22} + D_{23} + D_{24}),
$$

(40)

where

$$
D_{21} = \sum_{j=1}^n \mathbb{E} \int_0^T \left( f(a_{ij} u_j(\hat{X}_i)) - f_\eta(a_{ij} u_j(\hat{X}_i)) \right)^2 dt,
$$

$$
D_{22} = \sum_{j=1}^n \mathbb{E} \int_0^T \left( f_\eta(a_{ij} u_j(\hat{X}_i)) - f_\eta(B_{ij}^\eta * u_j(\hat{X}_i)) \right)^2 dt,
$$

$$
D_{23} = \sum_{j=1}^n \mathbb{E} \int_0^T \left( f_\eta(B_{ij}^\eta * u_j(\hat{X}_i)) - f_\eta(B_{ij}^\eta * u_j(\bar{X}_i)) \right)^2 dt,
$$

$$
D_{24} = \sum_{j=1}^n \mathbb{E} \int_0^T \left( f_\eta(B_{ij}^\eta * u_j(\bar{X}_i)) - f_\eta(B_{ij}^\eta * u_{\eta,j}(\bar{X}_i)) \right)^2 dt.
$$

The first expression $D_{21}$ vanishes if $\eta > 0$ is sufficiently small, since then $f = f_\eta$ on the range of $a_{ij} u_j(\hat{X}_i)$. Using

$$
\|a_{ij} u_j - B_{ij}^\eta * u_j\|_{L^2(0,T;L^2(\mathbb{R}^d))} \leq C \eta \|\nabla u_j\|_{L^2(0,T;L^2(\mathbb{R}^d))} \leq C \eta,
$$

which was shown in the proof of Theorem 3, and the Lipschitz continuity of $f_\eta$ with Lipschitz constant less or equal $\eta^{-\alpha}$, we find that

$$
D_{22} = \sum_{j=1}^n \int_0^T \int_{\mathbb{R}^d} \left( f_\eta(a_{ij} u_j) - f_\eta(B_{ij}^\eta * u_j) \right)^2 u_i dx dt
$$

$$
\leq \eta^{-2\alpha} \sum_{j=1}^n \|u_i\|_{L^\infty(0,T;L^\infty(\mathbb{R}^d))} \|a_{ij} u_j - B_{ij}^\eta * u_j\|^2_{L^2(0,T;L^2(\mathbb{R}^d))} \leq C(n)\eta^{2(1-\alpha)}.
$$

Thanks to the uniform boundedness of the family $B_{ij}^\eta * u_j$, we can choose $\eta > 0$ sufficiently small, say $\eta \leq \eta^*$ for some $\eta^* > 0$, such that $f(B_{ij}^\eta * u_j) = f_\eta(B_{ij}^\eta * u_j)$ for $0 < \eta \leq \eta^*$. Then, using Young’s convolution inequality and the uniform estimate $\|\nabla u_j\|_{L^\infty(0,T;L^\infty(\mathbb{R}^d))} \leq C \|u_0\|_{H^1(\mathbb{R}^d)}$ from Theorem 3, the third term $D_{23}$ is estimated as

$$
D_{23} \leq C(\eta^*) \sum_{j=1}^n \|\nabla(B_{ij}^\eta * u_j)\|_{L^\infty(0,T;L^\infty(\mathbb{R}^d))} \int_0^T \mathbb{E}(|\hat{X}_i(t) - \bar{X}_i^\eta(t)|^2) dt
$$
Finally, it follows from the error estimate for \( u - u_\eta \) from Theorem 3 that

\[
D_{24} \leq C \sum_{j=1}^{n} \int_0^T \int_{\mathbb{R}^d} |B_{ij}^\eta \ast u_j - B_{ij}^\eta \ast u_{\eta,j}|^2 u_{\eta,i} \, dx \, dt
\]

\[
\leq C \sum_{j=1}^{n} \|u_{\eta,i}\|_{L^\infty(0,T;L^\infty(\mathbb{R}^d))} \int_0^T \|B_{ij}^\eta\|_{L^1(\mathbb{R}^d)}^2 \|u_j - u_{\eta,j}\|_{L^2(\mathbb{R}^d)}^2 \, dt
\]

\[
\leq C(T) \eta^2.
\]

Inserting the estimates for \( D_{21}, \ldots, D_{24} \) into (40), we conclude that

\[
\mathbb{E} \left( \sup_{0<s<T} |D_2(s)|^2 \right) \leq C(T, n) \eta^{2(1-\alpha)} + C(T) \int_0^T \mathbb{E} \left( \sup_{0<s<T} |D_1^\eta(s)|^2 \right) dt.
\]

Together with estimate (39) for \( D_1(s) \) and recalling that \( D_1^\eta = D_1 + D_2 \), we arrive at

\[
\mathbb{E} \left( \sup_{0<s<T} |D_1^\eta(s)|^2 \right) \leq C(T, n) \eta^{2(1-\alpha)} + C(T) \int_0^T \mathbb{E} \left( \sup_{0<s<T} |D_1^\eta(s)|^2 \right) dt.
\]

The proof is finished after applying Gronwall’s inequality and summing over \( i = 1, \ldots, n \). \( \square \)

Theorem 5 now follows from Lemmas 9 and 10 and the triangle inequality:

\[
\sup_{k=1,\ldots,N} \mathbb{E} \left( \sum_{i=1}^{n} \sup_{0<s<T} \left| X_{\eta,i}^{k,N}(s) - \hat{X}_i^k(s) \right|^2 \right)
\]

\[
\leq 2 \sup_{k=1,\ldots,N} \mathbb{E} \left( \sum_{i=1}^{n} \sup_{0<s<T} \left| X_{\eta,i}^{k,N}(s) - \bar{X}_{\eta,i}^k(s) \right|^2 \right)
\]

\[
+ 2 \sup_{k=1,\ldots,N} \mathbb{E} \left( \sum_{i=1}^{n} \sup_{0<s<T} \left| \bar{X}_{\eta,i}^k(s) - \hat{X}_i^k(s) \right|^2 \right)
\]

\[
\leq C_1 N^{-1+C_2\delta} + C_3 \eta^{2(1-\alpha)}.
\]

The condition \( \log N \geq \delta^{-1} \eta^{-2(d+1+\alpha)} \) is equivalent to \( N^{-1+C_2\delta} \leq \exp((-\delta^{-1} + C_2) \eta^{-2(d+1+\alpha)}) \). We choose \( \delta > 0 \) such that \( -\delta^{-1} + C_2 < 0 \) and observe that  

\( \square \) Springer
exponential decay is always faster than algebraic decay to conclude that $\exp((-\delta^{-1} + C_2)\eta^{-2(\delta+1+\alpha)}) \leq \eta^{2(1-\alpha)}$. This yields

$$\sup_{k=1,\ldots,N} \mathbb{E}\left( \sum_{i=1}^{n} \sup_{0<s<t} \left| X_{\eta,i}^k(s) - \hat{X}_i^k(s) \right|^2 \right) \leq C_4 \eta^{2(1-\alpha)},$$

finishing the proof.

7 Numerical Tests

In this section, we perform some numerical simulations of the particle system (5) in one space dimension, without environmental potential, and with linear function $f(x) = x$. We are interested in the numerical comparison of the solutions to the particle systems (3) and (5) in terms of the segregation behavior. We explore the ability of both systems to model the segregation of the species. Numerical tests for the associated cross-diffusion systems (1) and (2) are work in progress.

We discretize the particle systems (3) and (5) by the Euler–Maruyama scheme. Let $M \in \mathbb{N}$ and introduce the time steps $0 < t_1 < \cdots < t_M = T$ with $\Delta t_m = t_{m+1} - t_m$. We approximate $X_{\eta,i}^k(t_m)$ by $x_{m,i}^k$ and $Y_{\eta,i}^k(t_m)$ by $y_{m,i}^k$, defined by, respectively,

$$x_{m+1,i}^k = x_m^k + \left( 2\sigma_i + \frac{2}{N} \sum_{j=1}^{N} \sum_{\ell=1}^{N} B_{ij}^\eta (x_m^j - y_m^\ell) \right)^{1/2} \sqrt{\Delta t_m} w_m,$$

$$y_{m+1,i}^k = y_m^k - \sum_{j=1}^{N} \frac{1}{N} \sum_{\ell=1}^{N} \nabla B_{ij}^\eta (y_m^j - y_m^\ell) \Delta t_m + \sqrt{2\sigma_i \Delta t_m} z_m,$$

with initial conditions $x_0^i = \xi_i^k$ and $y_0^i = \xi_i^k$, where $\xi_i^k$ are iid random variables and $w_m$ and $z_m$ are normally distributed. It is well known that the solutions to the Euler–Maruyama scheme converge to the associated stochastic processes in the strong sense; see, e.g., (Kloeden and Platen 1992, Theorem 9.6.2).

The numerical scheme is implemented in MATLAB using the parallel computing toolbox to accelerate the simulations. The interaction potential is given by $B(x) = \exp(-1/(1-x^2))$ for $|x| \leq 1$ and $B(x) = 0$ else. Then $B_{ij}^\eta (x) = \eta^{-1} B(x/\eta)$. The numerical parameters are $\Delta t = 1/100$, $\eta = 2$, $N = 5000$ particles, $n_{sim} = 500$ simulations.

7.1 Two Species: Nonsymmetric Case

We consider a nonsymmetric diffusion matrix with $a_{11} = 0$, $a_{12} = 355$, $a_{21} = 25$, $a_{22} = 0$, and $\sigma_1 = 1$, $\sigma_2 = 2$. The initial data are Gaussian distributions with mean $-1$ (for species $i = 1$) and 1 (for species $i = 2$) and variance 2. Figure 1 shows the approximate densities of both species (histogram) for systems (5) and (3) at time $t = 2$. We observe a segregation of the densities in both models. In the population
system (5), species 1 develops two clusters because of the very different “population pressure” parameters $a_{12} = 355$ and $a_{21} = 25$, while species 2 develops only one cluster around $x = 0$; see Fig. 1 left. The segregation effect is stronger in the particle system (3) in the sense that both species avoid each other as far as possible; see Fig. 1 right. This is not surprising since the diffusion of system (5) is generally larger than that one of system (3). The numerical results confirm the segregation property defined in Bertsch et al. (1985). Indeed, this work considers the cross-diffusion system (3) with $\sigma_1 = \sigma_2 = 0$ and $a_{11} = a_{12} = a_{21} = a_{22} = 1$. It was proved that the two species are segregated for all times if they do so initially. Here, segregation means that the intersection of the supports of the densities is empty.

### 7.2 Two Species: Symmetric Case

We investigate the symmetric case by choosing $a_{11} = a_{22} = 0$, $a_{12} = a_{21} = 355$, and, as before, $\sigma_1 = 1$, $\sigma_2 = 2$. The initial data are chosen as in the previous example. In this example, we expect that cross-diffusion dominates self-diffusion. We present the approximate densities for different times in Fig. 2. In both models, the species have the tendency to segregate. As expected, the segregation in the particle system (3) is stronger than in system (5) corresponding to the SKT model.

### 7.3 Three Species

Our third numerical experiment illustrates the segregation behavior in case of three interacting species with coefficients $\sigma_1 = 1$, $\sigma_2 = 2$, $\sigma_3 = 3$ and

$$
(a_{ij}) = \begin{pmatrix} 0 & 355 & 355 \\ 25 & 0 & 25 \\ 355 & 0 & 0 \end{pmatrix}.
$$

Similar as in the two-species case, the initial data are overlapping normal distributions with means $-1, 2,$ and $-3$, respectively, and variance 2. The approximate densities at
Fig. 2 Symmetric case: densities of particle system (5) corresponding to the SKT population model (left) and particle system (3) (right) for different times $t = 0.01, 0.15, 2$. Solid blue line: species 1; dashed red line: species 2.

$t = 2$ are shown in Fig. 3. We observe that the approximate densities of particle model (3) show a much clearer component-wise segregation behavior than the stochastic particle model (5), which corresponds to the SKT system, where the diffusion effects are much stronger. This may be explained by the fact that, on the PDE level, the gradient-flow structure of model (2) can be written species-wise, whereas the SKT model (1) (with $f(x) = x$) only possesses a vector-valued gradient-flow structure.
Fig. 3 Three-species case: densities of particle system (5) corresponding to the SKT population model (left) and particle system (3) (right) at time $t = 2$. Solid blue line: species 1; dashed red line: species 2; dash-dotted black line: species 3

Fig. 4 Densities of particle system (5) corresponding to the SKT population model with $f(s) = s^3$ (left) and $f(s) = s$ (right) at time $t = 2$. Solid blue line: species 1; dashed red line: species 2. The right figure is the same as in Fig. 2 (bottom right) but with the range $x = -50, \ldots, 50$

7.4 Cubic Nonlinearity

For our last experiment, we compare the numerical results for the cubic nonlinearity $f(s) = s^3$ with the linear case imposed in the previous examples. The parameters are the same as in Sect. 7.2. The numerical simulations are performed without using approximating functions $f_\eta$. This may be justified by the fact that the simulations deal with (relatively) small time scales and with compactly supported initial data. We observe in Fig. 4 that the cubic nonlinearity causes more clustering than the linear case $f(s) = s$. The simulations suggest that in the cubic case, diffusion happens on a faster time scale than segregation, while in the linear case, the particles diffuse slower and hence they form bigger but fewer clusters.

Funding Open access funding provided by TU Wien (TUW) and Austrian Science Fund (FWF).

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence,
Appendix A: Auxiliary Results

For the convenience of the reader, we recall some well-known estimates used in this paper.

Lemma 11 (Young’s convolution inequality, (Lieb and Loss 2001, Formula (7), page 107)) Let \( 1 \leq p, q, r \leq \infty \) be such that \( \frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r} \) and let \( f \in L^p(\mathbb{R}^d) \), \( g \in L^q(\mathbb{R}^d) \). Then \( f * g \in L^r(\mathbb{R}^d) \) and
\[
\|f * g\|_{L^r(\mathbb{R}^d)} \leq \|f\|_{L^p(\mathbb{R}^d)} \|g\|_{L^q(\mathbb{R}^d)}.
\]

Lemma 12 (Moser-type estimate I, (Majda 1984, Prop. 2.1(A))) Let \( s \in \mathbb{N} \) and \( \alpha \in \mathbb{N}_0^n \) with \( |\alpha| = s \). Then there exists a constant \( C > 0 \) such that for all \( f, g \in H^s(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d) \),
\[
\|D^\alpha (fg)\|_{L^2(\mathbb{R}^d)} \leq C \left( \|f\|_{L^\infty(\mathbb{R}^d)} \|D^\alpha g\|_{L^2(\mathbb{R}^d)} + \|D^\alpha f\|_{L^2(\mathbb{R}^d)} \|g\|_{L^\infty(\mathbb{R}^d)} \right).
\]

Lemma 13 (Moser-type estimate II, (Majda 1984, Prop. 2.1(C))) Let \( s \in \mathbb{N} \) and \( \alpha \in \mathbb{N}_0^n \) with \( |\alpha| = s \). Then there exists a constant \( C > 0 \) such that for smooth \( g : \mathbb{R} \to \mathbb{R} \) and \( u \in H^s(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d) \),
\[
\|D^\alpha g(u)\|_{L^2(\mathbb{R}^d)} \leq C \|g'\|_{C^{s-1}(\mathbb{R})} \|u\|_{L^\infty(\mathbb{R}^d)}^{s-1} \|D^\alpha u\|_{L^2(\mathbb{R}^d)}.
\]

Lemma 14 (Moser-type commutator inequality, (Majda 1984, Prop. 2.1(B))) Let \( s \in \mathbb{N} \) and \( \alpha \in \mathbb{N}_0^n \) with \( |\alpha| = s \). Then there exists \( C > 0 \) such that for all \( f \in H^s(\mathbb{R}^d) \cap W^{1,\infty}(\mathbb{R}^d) \) and \( g \in H^{s-1}(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d) \),
\[
\|D^\alpha (fg) - f D^\alpha (g)\|_{L^2(\mathbb{R}^d)} \leq C \left( \|Df\|_{L^\infty(\mathbb{R}^d)} \|D^{s-1}g\|_{L^2(\mathbb{R}^d)} + \|D^s f\|_{L^2(\mathbb{R}^d)} \|g\|_{L^\infty(\mathbb{R}^d)} \right),
\]
where \( D^s = \sum_{|\alpha| = s} D^\alpha \).

References

Bertsch, M., Gurtin, M., Hilhorst, D., Peletier, L.: On interacting populations that disperse to avoid crowding: preservation of segregation. J. Math. Biol. 23, 1–13 (1985)
Burger, M., Carrillo, J.A., Pietschmann, J.-F., Schmidtchen, M.: Segregation effects and gap formation in cross-diffusion models. Interfaces Free Bound. 22, 175–203 (2020a)
Burger, M., Pietschmann, J.-F., Ranetbauer, H., Schmeiser, C., Wolfram, M.-T.: Mean-field models for segregation dynamics. Eur. J. Appl. Math. To appear in arXiv:1808.04069 (2020b)
Chen, L., Göttlich, S., Knapp, S.: Modeling of a diffusion with aggregation: rigorous derivation and numerical simulation. ESAIM: Math. Mod. Num. Anal. 53, 567–593 (2018a)

Chen, X., Daus, E.S., Jüngel, A.: Global existence analysis of cross-diffusion population systems for multiple species. Arch. Ration. Mech. Anal. 227, 715–747 (2018b)

Chen, L., Daus, E.S., Jüngel, A.: Rigorous mean-field limits and cross diffusion. Z. Angew. Math. Phys. 70(122), 21 (2019)

Daus, E.S., Desvillettes, L., Dietert, H.: About the entropic structure of detailed balanced multi-species cross-diffusion equations. J. Differ. Equ. 266, 3861–3882 (2019)

Daus, E.S., Ptashnyk, M., Raithel, C.: Derivation of a fractional cross-diffusion system as the limit of a stochastic many-particle system driven by Lévy noise. Submitted for publication. arXiv:2006.00277 (2020)

Desvillettes, L., Lepoutre, T., Moussa, A., Trescases, A.: On the entropic structure of reaction-cross-diffusion systems. Commun. Partial Differ. Equ. 40, 1705–1747 (2015)

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

Fontbona, J., Méléard, S.: Non local Lotka–Volterra system with cross-diffusion in an heterogeneous medium. J. Math. Biol. 70, 829–854 (2015)

Golse, F.: The mean-field limit for the dynamics of large particle systems. Journées Equations aux dérivées partielles 1–47 (2003)

Ichikawa, K., Rouzimaimaiti, M., Suzuki, T.: Reaction diffusion equation with non-local term arises as a mean field limit of the master equation. Discrete Cont. Dyn. Syst. Ser. S 5, 115–126 (2012)

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

Jourdain, B., Méléard, S.: Propagation of chaos and fluctuations for a moderate model with smooth initial data. Ann. Inst. H. Poincaré Probab. Stat. 34, 727–766 (1998)

Karatzas, I., Shreve, S.: Brownian Motion and Stochastic Calculus, 2nd edn. Springer, New York (1991)

Kloeden, P., Platen, E.: Numerical Solution of Stochastic Differential Equations. Springer, Berlin (1992)

Lepoutre, T., Moussa, A.: Entropic structure and duality for multiple species cross-diffusion systems. Nonlinear Anal. 159, 298–315 (2017)

Lieb, E.H., Loss, M.: Analysis, 2nd edn. American Mathematical Society, Providence (2001)

Majda, A.: Compressible Fluid Flow and Systems of Conservation Laws in Several Space Variables. Springer, New York (1984)

Moussa, A.: From non-local to classical SKT systems: triangular case with bounded coefficients. SIAM J. Math. Anal. 52, 42–64 (2020)

Nualart, D.: The Malliavin Calculus and Related Topics. Springer, Berlin (2006)

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

Oelschläger, K.: On the derivation of reaction-diffusion equations as limit dynamics of systems of moderately interacting stochastic processes. Prob. Theory Relat. Fields 82, 565–586 (1989)

Oelschläger, K.: Large systems of interacting particles and the porous medium equation. J. Differ. Equ. 88, 294–346 (1990)

Seo, I.: Scaling limit of two-component interacting Brownian motions. Ann. Prob. 46, 2038–2063 (2018)

Shigesada, N., Kawasaki, K., Teramoto, E.: Spatial segregation of interacting species. J. Theor. Biol. 79, 83–99 (1979)

Stevens, A.: The derivation of chemotaxis equations as limit dynamics of moderately interacting stochastic many-particle systems. SIAM J. Appl. Math. 61, 183–212 (2000). (Erratum: 61 (2000), 2200-2200.)

Sznitman, A.S.: Nonlinear reflecting diffusion process, and the propagation of chaos and fluctuations associated. J. Funct. Anal. 56, 311–336 (1984)

Sznitman, A.S.: Topics in propagation of chaos. In: P.L. Hennequin (ed.), École d’Été de Probabilités de Saint-Flour XIX-1989, Lecture Notes Mathematics 1464. Berlin, Springer (1991)

Zamponi, N., Jüngel, A.: Analysis of degenerate cross-diffusion population models with volume filling. Ann. Inst. H. Poincaré Anal. Nonlinear 34, 1–29 (2017)

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.