RIGOROUS DERIVATION OF POPULATION CROSS-DIFFUSION SYSTEMS FROM MODERATELY INTERACTING PARTICLE SYSTEMS

LI CHEN, ESTHER S. DAUS, ALEXANDRA HOLZINGER, AND ANSGAR JÜNGEL

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.

1. Introduction

The aim of this paper is to derive the population cross-diffusion system of Shigesada, Kawasaki, and Teramoto [27] from a stochastic, moderately interacting particle system in a mean-field-type limit. More precisely, we derive the system of equations

$$\partial_t u_i = \operatorname{div}(u_i \nabla U_i) + \Delta\bigg(\sigma_i u_i + u_i \sum_{j=1}^n f(a_{ij}u_j)\bigg), \quad u_i(0) = u_{0,i} \quad \text{in } \mathbb{R}^d,\ t>0, \tag{1}$$

where $i = 1,\ldots,n$ is the species index, $u = (u_1,\ldots,u_n)$ 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 [27]. System (1) with nonlinear functions f have also been studied in the literature; see, e.g., [6, 9, 18]. 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.

Date: October 23, 2020.

2000 *Mathematics Subject Classification.* 35Q92, 35K45, 60J70, 82C22.

Key words and phrases. Moderately interacting particle systems, stochastic particle systems, cross-diffusion system, rigorous derivation, Shigesada–Kawasaki–Teramoto model, mean-field limit, population dynamics.

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.

This paper extends the many-particle limit of [4] leading to the cross-diffusion system

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

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

$$\begin{aligned} \mathrm{d}Y_{k,i}^{N,\eta} &= -\sum_{j=1}^n \frac{1}{N} \sum_{\ell=1}^N \nabla B_{ij}^\eta \big(Y_{k,i}^{N,\eta} - Y_{\ell,j}^{N,\eta}\big)\mathrm{d}t + \sqrt{2\sigma_i}\mathrm{d}W_i^k(t), \\ Y_{k,i}^{N,\eta}(0) &= \xi_i^k, \quad i = 1, \ldots, n,\ k = 1, \ldots, N, \end{aligned} \tag{3}$$

where $(W_i^k(t))_{t\ge 0}$ are d-dimensional Brownian motions and $\xi_i^1, \ldots, \xi_i^N$ are independent and identically distributed (iid) random variables with the common probability density function $u_{0,i}$. The functions

$$B_{ij}^\eta(x) = \eta^{-d} B_{ij}\left(\frac{|x|}{\eta}\right), \quad x \in \mathbb{R}^d, \tag{4}$$

are interaction potentials regularizing the delta distribution δ_0, i.e. $B_{ij}^\eta \to a_{ij}\delta_0$ 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:

$$\begin{aligned} \mathrm{d}X_{k,i}^{N,\eta} &= -\nabla U_i(X_{k,i}^{N,\eta})\mathrm{d}t + \left(2\sigma_i + 2\sum_{j=1}^n f_\eta\left(\frac{1}{N}\sum_{\substack{\ell=1 \\ (\ell,j)\neq(k,i)}}^N B_{ij}^\eta(X_{k,i}^{N,\eta} - X_{\ell,j}^{N,\eta})\right)\right)^{1/2} \mathrm{d}W_i^k(t), \\ X_{k,i}^{N,\eta}(0) &= \xi_i^k, \quad i = 1, \ldots, n,\ k = 1, \ldots, N, \end{aligned} \tag{5}$$

where f_η 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 η 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 η 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:

$$\begin{aligned} \mathrm{d}\overline{X}_{k,i}^\eta &= -\nabla U_i(\overline{X}_{k,i}^\eta)\mathrm{d}t + \left(2\sigma_i + 2\sum_{j=1}^n f_\eta\big(B_{ij}^\eta * u_{\eta,j}(\overline{X}_{k,i}^\eta)\big)\right)^{1/2} \mathrm{d}W_i^k(t), \\ \overline{X}_{k,i}^\eta(0) &= \xi_i^k, \quad i = 1, \ldots, n,\ k = 1, \ldots, N, \end{aligned} \tag{6}$$

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

$$
\begin{aligned}
&\partial_t u_{\eta,i} = \operatorname{div}(u_{\eta,i}\nabla U_i) + \Delta\bigg(\sigma_i u_{\eta,i} + u_{\eta,i}\sum_{j=1}^n f_\eta(B^\eta_{ij} * u_{\eta,j})\bigg), \\
&u_{\eta,i}(0) = u_i^0 \text{ in } \mathbb{R}^d, \quad i = 1, \ldots, n,
\end{aligned} \tag{7}
$$

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*

$$
\begin{aligned}
&\mathrm{d}\widehat{X}_{k,i} = -\nabla U_i(\widehat{X}_{k,i})\mathrm{d}t + \bigg(2\sigma_i + 2\sum_{j=1}^n f(a_{ij}u_j(\widehat{X}_{k,i}))\bigg)^{1/2}\mathrm{d}W_i^k(t), \\
&\widehat{X}^\eta_{k,i}(0) = \xi_i^k, \quad i = 1, \ldots, n, \ k = 1, \ldots, N,
\end{aligned} \tag{8}
$$

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

$$
\eta^{-2(d+1+\alpha)} \le \delta \log N \tag{9}
$$

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

$$
\sup_{k=1,\ldots,N} \mathbb{E}\bigg(\sum_{i=1}^n \sup_{0<s<T} \big|X^{N,\eta}_{k,i}(s) - \widehat{X}_{k,i}(s)\big|^2\bigg) \le C(T)\eta^{2(1-\alpha)}. \tag{10}
$$

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 [15], where a single-species, moderately interacting particle system with interaction in the diffusion part was considered. Since we allow for not globally Lipschitz continuous functions, our convergence rate is smaller, but for $\alpha \to 0$ we recover the rate in [15].

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 [12, 14] and the classical works by Sznitman [29, 30]. Oelschläger proved that in the many-particle limit, weakly interacting stochastic particle systems converge to a deterministic nonlinear process [23]. Later, he generalized his approach for systems of reaction-diffusion equations [24] and porous-medium-type equations with quadratic diffusion [25], by using moderately interacting particle systems. We also refer to the recent work [5], 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 [15]. Later, Stevens derived the chemotaxis model from a many-particle system [28]. Further works concern the mean-field limit leading to reaction-diffusion equations with nonlocal terms [13], the hydrodynamic limit in a two-component system of Brownian motions to the cross-diffusion Maxwell–Stefan equations [26], and the large population limit of point measure-valued Markov processes to nonlocal Lotka–Volterra systems with cross diffusion [11]. The latter model is similar to the nonlocal system (7). The limit from the nonlocal to the local diffusion system was shown in [21] 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 [8]. Furthermore, the population system (1) was derived in [7] 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 [10] or in the diffusion term [15]. We allow for interactions in the diffusion part but in a multi-species setting. The paper [11] 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 [7], 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 afterwards 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 [2, 3]. First, system (2) has a formal gradient-flow structure for each species separately, while system (1) can be written, under the detailed-balance condition [4], 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 Section 7.

The paper is organized as follows. We present our assumptions and main results in Section 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 Sections 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 Section 6. In Section 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 $\xi_i^1,\ldots,\xi_i^N$ 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:** $B_{ij} \in C_0^\infty(\mathbb{R}^d)$ satisfies $\operatorname{supp}(B_{ij}) \subset B_1(0)$, where $B_1(0)$ is the unit ball in $\mathbb{R}^d$ and $i,j=1,\ldots,n$.

(A4) **Nonlinearity:** $f \in W^{s+1,\infty}_{\rm loc}(\mathbb{R};[0,\infty))$ 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_η 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. It is possible to choose general potentials $U_i \in C^\infty(\mathbb{R}^d)$ such that ∇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^2U_i is negative semidefinite, $\Delta U_i < 0$, and D^kU_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_η ensures that we have a control on the growth of the Lipschitz constant of f_η 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

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

$B_{ij}^\eta(x) = \eta^{-d}B_{ij}(|x|/\eta)$, $A_{ij} = \|B_{ij}\|_{L^1(\mathbb{R}^d)} = \|B_{ij}^\eta\|_{L^1(\mathbb{R}^d)}$ 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

$$I = [-2AC_s\|u_0\|_{H^s(\mathbb{R}^d)}, 2AC_s\|u_0\|_{H^s(\mathbb{R}^d)}]. \tag{11}$$

Then, for small $\eta > 0$ such that $a_\eta \ge 2AC_s\|u_0\|_{H^s(\mathbb{R}^d)}$, 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 $a_\eta \ge 2AC_s\|u_0\|_{H^s(\mathbb{R}^d)}$. 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*

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

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

The dependence of ε 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 σ_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* η *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*

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

where $0 < \sigma_* < \sigma_{\min}$. *Moreover, with the solution* u_η *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,* $\eta > 0$, $N \in \mathbb{N}$, *and let* $\xi_1^k, \ldots, \xi_n^k$ *for* $k = 1, \ldots, N$ *be iid random variables with values in* $\mathbb{R}^d$ *and with density function* $u_0 = (u_{0,1}, \ldots, u_{0,n})$. *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* $\overline{X}^\eta(t)$ *and* $\widehat{X}(t)$ *possess density functions* $\overline{u}_\eta(t)^{\otimes N}$ *and* $\widehat{u}(t)^{\otimes N}$ *with respect to the Lebesgue measure on* $\mathbb{R}^{nNd}$, *respectively.*

The proof follows from [16] and [22]. Indeed, Theorem 2.9 in [16, 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 [16, page 287]. We conclude from [22, Theorem 2.3.1] that $\overline{X}_\eta(t)$ and $\widehat{X}(t)$ are absolutely continuous with respect to the Lebesgue measure and thus, they possess density functions $\overline{u}_\eta(t,x)^{\otimes N}$ and $\widehat{u}(t,x)^{\otimes N}$, respectively. We prove in Section 5 that the density functions $\overline{u}_\eta$ and $\widehat{u}$ can be identified with u_η and u, the solutions to (7) and (1), respectively.

The following theorem is our main result.

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

$$\sup_{k=1,\ldots,N} \mathbb{E}\bigg(\sum_{i=1}^n \sup_{0<s<T} \big|(X_{k,i}^{N,\eta} - \widehat{X}_{k,i})(s)\big|^2 \bigg) \le C(T,n,\sigma_{\min})\eta^{2(1-\alpha)},$$

where $\alpha \ge 0$ *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., [14, 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 $(X_{j_1,i_1}^{N,\eta}(t),\ldots,X_{j_k,i_k}^{N,\eta}(t))$ 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.

$$W_2^2\bigg(F_k(t), \bigotimes_{\ell=1}^k u_{i_\ell}(\cdot,t) \bigg) \le kC(T,n,\sigma_{\min})\eta \to 0,$$

where W_2 denotes the 2-Wasserstein distance. □

3. Proof of Theorem 2

We prove the global existence of smooth solutions to the nonlocal system (7). Since η is fixed in the proof, we omit it for u_η to simplify the notation. We split the proof in several steps.

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

$$X_T := \big\{ v \in L^\infty(0,T;H^s(\mathbb{R}^d;\mathbb{R}^n)) : \|v\|_{L^\infty(0,T;H^s(\mathbb{R}^d))} \le 2\|u_0\|_{H^s(\mathbb{R}^d)} \big\},$$

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

$$\partial_t u_i = \operatorname{div}(u_i \nabla U_i) + \Delta\big(u_i(\sigma_i + K_i(v(t,x)))\big), \quad u_i(0)=u_{0,i} \quad \text{in } \mathbb{R}^d,\ t>0, \tag{12}$$

with $K_i(v) = \sum_{j=1}^n f_\eta(B_{ij}^\eta * v_j) \ge 0$, $i=1,\ldots,n$. 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

$$\begin{aligned} \sup_{0<t<T} \|\nabla K_i(v)\|_{L^\infty(\mathbb{R}^d)} &\le \sum_{j=1}^n \|f_\eta'\|_{L^\infty(\mathbb{R})} \|\nabla B_{ij}^\eta\|_{L^1(\mathbb{R}^d)} \sup_{0<t<T} \|v_j(t)\|_{L^\infty(\mathbb{R}^d)} \\ &\le C(\eta) \sum_{j=1}^n \|v_j\|_{L^\infty(0,T;H^s(\mathbb{R}^d))} < \infty, \end{aligned} \tag{13}$$

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 ∇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 $|\alpha| = m \le s$. By Lemma 13 and Young's convolution inequality,

$$\begin{aligned}\int_0^T \|D^\alpha K_i(v)\|_{L^2(\mathbb{R}^d)}^2 \mathrm{d}t &\le C\int_0^T \sum_{j=1}^n \|f_\eta'\|_{C^{m-1}(\mathbb{R})}^2 \|B_{ij}^\eta * v_j\|_{L^\infty(\mathbb{R}^d)}^{2(m-1)} \|D^\alpha(B_{ij}^\eta * v_j)\|_{L^2(\mathbb{R}^d)}^2 \mathrm{d}t \\ &\le C(\eta)\int_0^T \sum_{j=1}^n \|B_{ij}^\eta\|_{L^1(\mathbb{R}^d)}^{2m} \|v_j\|_{L^\infty(\mathbb{R}^d)}^{2(m-1)} \|D^\alpha v_j\|_{L^2(\mathbb{R}^d)}^2 \mathrm{d}t \\ &\le C(\eta)\sum_{j=1}^n \int_0^T \|v_j\|_{H^s(\mathbb{R}^d)}^{2m}\mathrm{d}t < \infty, \end{aligned} \tag{14}$$

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,

$$\begin{aligned}\sup_{0<t<T}\|D^\alpha \nabla K_i(v)\|_{L^2(\mathbb{R}^d)}^2 &\le C\sup_{0<0<T}\sum_{j=1}^n \big\|D^\alpha\big(f_\eta'(B_{ij}^\eta * v_j)\nabla B_{ij}^\eta * v_j\big)\big\|_{L^2(\mathbb{R}^d)}^2 \\ \le C\sup_{0<t<T}\sum_{j=1}^n \Big(&\|f_\eta'(B_{ij}^\eta * v_j)\|_{L^\infty(\mathbb{R}^d)}\|\nabla B_{ij}^\eta\|_{L^1(\mathbb{R}^d)}\|D^m v_j\|_{L^2(\mathbb{R}^d)} \\ &+ \|D^m(f_\eta'(B_{ij}^\eta * v_j))\|_{L^2(\mathbb{R}^d)}\|\nabla B_{ij}^\eta\|_{L^1(\mathbb{R}^d)}\|v_j\|_{L^\infty(\mathbb{R}^d)}\Big)^2 \le C(\eta), \end{aligned} \tag{15}$$

since, according to Lemma 13, we can bound $\sup_{0<t<T}\|D^m(f_\eta'(B_{ij} * v_j))\|_{L^2(\mathbb{R}^d)}$ 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 $\|\nabla B_{ij}^\eta\|_{L^1(\mathbb{R}^d)} \le C(\eta)$.

We proceed with the proof of $u \in X_T$ for some $T > 0$. Applying D^α to (12), multiplying the resulting equation by $D^\alpha u_i$, and integrating over $(0,\tau)\times\mathbb{R}^d$ for $\tau < T$ yields

$$\frac12\int_{\mathbb{R}^d}|D^\alpha u_i(\tau)|^2\mathrm{d}x - \frac12\int_{\mathbb{R}^d}|D^\alpha u_{0,i}|^2\mathrm{d}x + \sigma_i\int_0^\tau\int_{\mathbb{R}^d}|\nabla D^\alpha u_i|^2\mathrm{d}x\mathrm{d}t = I_1 + I_2 + I_3, \tag{16}$$

where

$$\begin{aligned} I_1 &= -\int_0^\tau\int_{\mathbb{R}^d}\nabla D^\alpha u_i\cdot D^\alpha(u_i\nabla U_i)\mathrm{d}x\mathrm{d}t, \\ I_2 &= -\int_0^\tau\int_{\mathbb{R}^d}\nabla D^\alpha u_i\cdot D^\alpha(\nabla u_i K_i(v))\mathrm{d}x\mathrm{d}t, \\ I_3 &= -\int_0^\tau\int_{\mathbb{R}^d}\nabla D^\alpha u_i\cdot D^\alpha(u_i\nabla K_i(v))\mathrm{d}x\mathrm{d}t. \end{aligned}$$

First, let $|\alpha| = m = 0$. Then, integrating by parts in I_1, using Young's inequality, and observing that $U_i(x) = -\frac12|x|^2$,

$$\begin{aligned} I_1 &= \frac12\int_0^\tau\int_{\mathbb{R}^d}u_i^2\Delta U_i\mathrm{d}x\mathrm{d}t = -\frac{d}{2}\int_0^\tau\int_{\mathbb{R}^d}u_i^2\mathrm{d}x\mathrm{d}t \le 0, \\ I_2 &= -\int_0^\tau\int_{\mathbb{R}^d}K_i(v)|\nabla u_i|^2\mathrm{d}x\mathrm{d}t \le 0, \end{aligned}$$

$$I_3 \le \frac{\sigma_i}{2}\int_0^\tau\int_{\mathbb{R}^d}|\nabla u_i|^2\mathrm{d}x\mathrm{d}t + \frac{1}{2\sigma_i}\|\nabla K_i(v)\|^2_{L^\infty(0,T;L^\infty(\mathbb{R}^d))}\int_0^\tau \|u_i\|^2_{L^2(\mathbb{R}^d)}\mathrm{d}t,$$

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

$$I_1+I_2+I_3 \le \frac{\sigma_i}{2}\int_0^\tau\int_{\mathbb{R}^d}|\nabla u_i|^2\mathrm{d}x\mathrm{d}t + C\int_0^\tau \|u_i\|^2_{L^2(\mathbb{R}^d)}\mathrm{d}t,$$

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

$$\int_{\mathbb{R}^d} u_i(\tau)^2\mathrm{d}x + \frac{\sigma_i}{2}\int_0^\tau\int_{\mathbb{R}^d}|\nabla u_i|^2\mathrm{d}x\mathrm{d}t \le C(u_0)e^{C\tau}.$$

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 $|\alpha|=m\ge 1$. Then, integrating by parts, using $\Delta U_i\le 0$, and applying Young's inequality again,

$$\begin{aligned}
I_1 &= \frac{1}{2}\int_0^\tau\int_{\mathbb{R}^d}(D^\alpha u_i)^2\Delta U_i\mathrm{d}x\mathrm{d}t - \int_0^\tau\int_{\mathbb{R}^d}\nabla D^\alpha u_i\cdot\big(D^\alpha(u_i\nabla U_i) - D^\alpha u_i\nabla U_i\big)\mathrm{d}x\mathrm{d}t \\
&\le \frac{\sigma_i}{4}\int_0^\tau\int_{\mathbb{R}^d}|\nabla D^\alpha u_i|^2\mathrm{d}x\mathrm{d}t + \sum_{0<|\beta|\le|\alpha|}\int_0^\tau c_\beta\|D^{\alpha-\beta}u_i\|^2_{L^2(\mathbb{R}^d)}\|D^\beta\nabla U_i\|^2_{L^\infty(\mathbb{R}^d)}\mathrm{d}t \\
&\le \frac{\sigma_i}{4}\int_0^\tau\int_{\mathbb{R}^d}|\nabla D^\alpha u_i|^2\mathrm{d}x\mathrm{d}t + C\int_0^\tau\|u_i\|^2_{H^{m-1}(\mathbb{R}^d)}\mathrm{d}t,
\end{aligned}$$

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, $K_i(v)\ge 0$, and Lemma 14 that

$$\begin{aligned}
I_2 &= -\int_0^\tau\int_{\mathbb{R}^d}\nabla D^\alpha u_i\cdot\big(D^\alpha(\nabla u_i K_i(v)) - \nabla D^\alpha u_i K_i(v)\big)\mathrm{d}x\mathrm{d}t \\
&\quad - \int_0^\tau\int_{\mathbb{R}^d}K_i(v)|\nabla D^\alpha u_i|^2\mathrm{d}x\mathrm{d}t \\
&\le \frac{\sigma_i}{4}\int_0^\tau\int_{\mathbb{R}^d}|\nabla D^\alpha u_i|^2\mathrm{d}x\mathrm{d}t + C\int_0^\tau\big(\|DK_i(v)\|_{L^\infty(\mathbb{R}^d)}\|D^{m-1}\nabla u_i\|_{L^2(\mathbb{R}^d)} \\
&\quad + \|D^m K_i(v)\|_{L^2(\mathbb{R}^d)}\|\nabla u_i\|_{L^\infty(\mathbb{R}^d)}\big)^2\mathrm{d}x\mathrm{d}t.
\end{aligned}$$

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

$$I_2 \le \frac{\sigma_i}{4}\int_0^\tau\int_{\mathbb{R}^d}|\nabla D^\alpha u_i|^2\mathrm{d}x\mathrm{d}t + C\int_0^\tau\|u_i\|^2_{H^s(\mathbb{R}^d)}\mathrm{d}t.$$

Finally, we use Lemma 12 and estimates (13) and (15) to obtain

$$\begin{aligned}
I_3 &\le \frac{\sigma_i}{4}\int_0^\tau\int_{\mathbb{R}^d}|\nabla D^\alpha u_i|^2\mathrm{d}x\mathrm{d}t + C\int_0^\tau\int_{\mathbb{R}^d}\big(\|u_i\|_{L^\infty(\mathbb{R}^d)}\|D^m\nabla K_i(v)\|_{L^2(\mathbb{R}^d)} \\
&\quad + \|D^m u_i\|_{L^2(\mathbb{R}^d)}\|\nabla K_i(v)\|_{L^\infty(\mathbb{R}^d)}\big)^2\mathrm{d}x\mathrm{d}t
\end{aligned}$$

$$\leq \frac{\sigma_i}{4}\int_0^\tau\int_{\mathbb{R}^d}|\nabla D^\alpha u_i|^2\mathrm{d}x\mathrm{d}t + C(\eta)\int_0^\tau \|u_i\|^2_{H^s(\mathbb{R}^d)}\mathrm{d}t.$$

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

$$\|u_i(\tau)\|^2_{H^s(\mathbb{R}^d)} + \frac{\sigma_i}{4}\int_0^\tau \|\nabla u_i\|^2_{H^s(\mathbb{R}^d)}\mathrm{d}t \leq \|u_{0,i}\|^2_{H^s(\mathbb{R}^d)} + C(\eta)\int_0^\tau \|u_i\|^2_{H^s(\mathbb{R}^d)}\mathrm{d}t.$$

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 $\bar v = S(v)$ and $\bar w = S(w)$. Taking the difference of equations (12) satisfied by $\bar v_i$ and $\bar w_i$, respectively, using the test function $\bar v_i - \bar w_i$, and integrating by parts, it follows that

$$\frac12\int_{\mathbb{R}^d}(\bar v_i - \bar w_i)(\tau)^2\mathrm{d}x + \sigma_i\int_0^\tau\int_{\mathbb{R}^d}|\nabla(\bar v_i - \bar w_i)|^2\mathrm{d}x\mathrm{d}t = I_4 + I_5 + I_6, \tag{17}$$

where

$$\begin{aligned} I_4 &= \frac12\int_0^\tau\int_{\mathbb{R}^d}\Delta U_i(\bar v_i - \bar w_i)^2\mathrm{d}x\mathrm{d}t \leq 0, \\ I_5 &= -\int_0^\tau\int_{\mathbb{R}^d}\nabla\big((\bar v_i - \bar w_i)K_i(v)\big)\cdot\nabla(\bar v_i - \bar w_i)\mathrm{d}x\mathrm{d}t, \\ I_6 &= -\int_0^\tau\int_{\mathbb{R}^d}\nabla\big(\bar w_i(K_i(v) - K_i(w))\big)\cdot\nabla(\bar v_i - \bar w_i)\mathrm{d}x\mathrm{d}t. \end{aligned}$$

Because of $K_i(v)\geq 0$ and estimate (13) for $\nabla K_i(v)$, we find that, by Young's inequality,

$$\begin{aligned} I_5 &= -\int_0^\tau\int_{\mathbb{R}^d}K_i(v)|\nabla(\bar v_i - \bar w_i)|^2\mathrm{d}x\mathrm{d}t - \int_0^\tau\int_{\mathbb{R}^d}(\bar v_i - \bar w_i)\nabla K_i(v)\cdot\nabla(\bar v_i - \bar w_i)\mathrm{d}x\mathrm{d}t \\ &\leq \frac{\sigma_i}{4}\int_0^\tau \|\nabla(\bar v_i - \bar w_i)\|^2_{L^2(\mathbb{R}^d)}\mathrm{d}t + C(\sigma_i)\int_0^\tau \|\bar v_i - \bar w_i\|^2_{L^2(\mathbb{R}^d)}\|\nabla K_i(v)\|^2_{L^\infty(\mathbb{R}^d)}\mathrm{d}t \\ &\leq \frac{\sigma_i}{4}\int_0^\tau \|\nabla(\bar v_i - \bar w_i)\|^2_{L^2(\mathbb{R}^d)}\mathrm{d}t + C(\eta)\int_0^\tau \|\bar v_i - \bar w_i\|^2_{L^2(\mathbb{R}^d)}\mathrm{d}t. \end{aligned}$$

It follows again from Young's inequality that

$$\begin{aligned} I_6 &\leq \frac{\sigma_i}{4}\int_0^\tau \|\nabla(\bar v_i - \bar w_i)\|^2_{L^2(\mathbb{R}^d)}\mathrm{d}t + C(\sigma_i)\int_0^\tau \|\nabla\bar w_i\|^2_{L^\infty(\mathbb{R}^d)}\|K_i(v) - K_i(w)\|^2_{L^2(\mathbb{R}^d)}\mathrm{d}t \\ &\quad + C(\sigma_i)\int_0^\tau \|\bar w_i\|^2_{L^\infty(\mathbb{R}^d)}\|\nabla(K_i(v) - K_i(w))\|^2_{L^2(\mathbb{R}^d)}\mathrm{d}t. \end{aligned} \tag{18}$$

Since $\bar{w} \in X_T$, we have $\|\nabla \bar{w}_i\|_{L^\infty(\mathbb{R}^d)} \le C\|\bar{w}_i\|_{H^s(\mathbb{R}^d)} \le C(u_0)$ and $\|\bar{w}_i\|_{L^\infty(\mathbb{R}^d)} \le C(u_0)$. We use the fact that f_η and f'_η are globally Lipschitz continuous:

$$\|K_i(v) - K_i(w)\|_{L^2(\mathbb{R}^d)} \le C(\eta)\sum_{j=1}^n \|B^\eta_{ij} * (v_j - w_j)\|_{L^2(\mathbb{R}^d)} \le C(\eta)\|v - w\|_{L^2(\mathbb{R}^d)},$$

$$\begin{aligned}\|\nabla(K_i(v) - K_i(w))\|_{L^2(\mathbb{R}^d)} &\le \sum_{j=1}^n \|(f'_\eta(B^\eta_{ij} * v_j) - f'_\eta(B^\eta_{ij} * w_j))B^\eta_{ij} * \nabla v_j\|_{L^2(\mathbb{R}^d)} \\ &\quad + \sum_{j=1}^n \|f'_\eta(B^\eta_{ij} * w_j)\nabla B^\eta_{ij} * (v_j - w_j)\|_{L^2(\mathbb{R}^d)} \\ &\le C(\eta)\sum_{j=1}^n \|v_j - w_j\|_{L^2(\mathbb{R}^d)}\|B^\eta_{ij}\|_{L^1(\mathbb{R}^d)}\|\nabla v_j\|_{L^\infty(\mathbb{R}^d)} \\ &\quad + C(\eta)\sum_{j=1}^n \|\nabla B^\eta_{ij}\|_{L^1(\mathbb{R}^d)}\|v_j - w_j\|_{L^2(\mathbb{R}^d)} \\ &\le C(\eta)\|v - w\|_{L^2(\mathbb{R}^d)}.\end{aligned}$$

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

$$\begin{aligned}&\frac{1}{2}\|(\bar{v} - \bar{w})(\tau)\|^2_{L^2(\mathbb{R}^d)} + \sum_{i=1}^n \frac{\sigma_i}{4}\int_0^\tau \|\nabla(\bar{v}_i - \bar{w}_i)\|^2_{L^2(\mathbb{R}^d)}\mathrm{d}t \\ &\quad \le C_1\int_0^\tau \|\bar{v} - \bar{w}\|^2_{L^2(\mathbb{R}^d)}\mathrm{d}t + C_2\tau\|v - w\|^2_{L^\infty(0,\tau;L^2(\mathbb{R}^d))}.\end{aligned}$$

We apply Gronwall's inequality and the supremum over $0 < \tau < T$ to find that

$$\|\bar{v} - \bar{w}\|^2_{L^\infty(0,T;L^2(\mathbb{R}^d))} \le C_2 e^{C_1 T} T\|v - w\|^2_{L^\infty(0,T;L^2(\mathbb{R}^d))}.$$

Thus, choosing $T > 0$ such that $C_2 e^{C_1 T} T < 1$, 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: Uniform estimates. Let $u = u_\eta$ be the unique solution to (7). We know from Step 1 that $\|u_i(t)\|_{L^\infty(\mathbb{R}^d)} \le 2C_s\|u_i(t)\|_{H^s(\mathbb{R}^d)} \le 2C_s\|u_0\|_{H^s(\mathbb{R}^d)}$ for any $0 < t < T$. Recall that $T = T(\eta)$ and hence we do not have uniform estimates in η even for small $T > 0$ at this step. We apply D^α to (7) (with $|\alpha| = m \le s$), 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:

$$\frac{1}{2}\int_{\mathbb{R}^d} |D^\alpha u_i(\tau)|^2\mathrm{d}t - \frac{1}{2}\int_{\mathbb{R}^d} |D^\alpha u_{0,i}|^2\mathrm{d}t + \sigma_i\int_0^\tau\int_{\mathbb{R}^d} |\nabla D^\alpha u_i|^2\mathrm{d}x\mathrm{d}t = I_7 + I_8 + I_9, \tag{19}$$

where

$$\begin{aligned} I_7 &= -\int_0^\tau\int_{\mathbb{R}^d}\nabla D^\alpha u_i\cdot D^\alpha(u_i\nabla U_i)\mathrm{d}x\mathrm{d}t, \\ I_8 &= -\int_0^\tau\int_{\mathbb{R}^d}\nabla D^\alpha u_i\cdot D^\alpha(\nabla u_i K_i(u))\mathrm{d}x\mathrm{d}t, \\ I_9 &= -\int_0^\tau\int_{\mathbb{R}^d}\nabla D^\alpha u_i\cdot D^\alpha(u_i\nabla K_i(u))\mathrm{d}x\mathrm{d}t, \end{aligned}$$

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 $I_7\le 0$ and $I_8\le 0$. We estimate $\nabla K_i(u)=\sum_{j=1}^n f_\eta'(B_{ij}^\eta * u_j)B_{ij}^\eta * \nabla u_j$:

$$\|\nabla K_i(u)\|_{L^2(\mathbb{R}^d)}\le A\sum_{j=1}^n\|f_\eta'(B_{ij}^\eta * u_j)\|_{L^\infty(\mathbb{R}^d)}\|\nabla u_j\|_{L^2(\mathbb{R}^d)}, \tag{20}$$

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

$$\begin{aligned} I_9 &\le \|u_i\|_{L^\infty(0,\tau;L^\infty(\mathbb{R}^d))}\int_0^\tau\|\nabla u_i\|_{L^2(\mathbb{R}^d)}\|\nabla K_i(u)\|_{L^2(\mathbb{R}^d)}\mathrm{d}t \\ &\le C\|u_0\|_{H^s(\mathbb{R}^d)}\sum_{j=1}^n\|f_\eta'(B_{ij}^\eta * u_j)\|_{L^\infty(0,\tau;L^\infty(\mathbb{R}^d))}\int_0^\tau\|\nabla u_j\|_{L^2(\mathbb{R}^d)}^2\mathrm{d}t. \end{aligned}$$

From this point on, we will need the smallness condition on f_η and f_η'. Because of

$$\|B_{ij}^\eta * u_j(t)\|_{L^\infty(\mathbb{R}^d)}\le\|B_{ij}^\eta\|_{L^1(\mathbb{R}^d)}C_s\|u_j(t)\|_{H^s(\mathbb{R}^d)}\le 2AC_s\|u_0\|_{H^s(\mathbb{R}^d)}, \tag{21}$$

where $C_s>0$ is the constant of the embedding $H^s(\mathbb{R}^d)\hookrightarrow L^\infty(\mathbb{R}^d)$, $(B_{ij}^\eta * u_j(t))(x)$ lies in the interval $I=[-2AC_s\|u_0\|_{H^s(\mathbb{R}^d)},2AC_s\|u_0\|_{H^s(\mathbb{R}^d)}]$ for $0<t<T$ and $x\in\mathbb{R}^d$. On this interval, $f_\eta=f$ if $\eta>0$ is sufficiently small. From now on, we use $f\le\varepsilon$ and $|f'|\le\varepsilon$ on I for a small $\varepsilon>0$. Thus, we have

$$I_9\le C\varepsilon\|u_0\|_{H^s(\mathbb{R}^d)}\int_0^\tau\|\nabla u_i\|_{L^2(\mathbb{R}^d)}^2\mathrm{d}t.$$

Inserting these estimates into (19), we conclude that

$$\|u_i(\tau)\|_{L^2(\mathbb{R}^d)}^2+\big(\sigma_i-C\varepsilon\|u_0\|_{H^s(\mathbb{R}^d)}\big)\int_0^\tau\|\nabla u_i\|_{L^2(\mathbb{R}^d)}^2\mathrm{d}t\le\|u_{0,i}\|_{L^2(\mathbb{R}^d)}^2.$$

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

$$I_7=\int_0^\tau\int_{\mathbb{R}^d}D^\alpha u_i D^\alpha(\nabla u_i\cdot\nabla U_i+u_i\Delta U_i)\mathrm{d}x\mathrm{d}t$$

$$= \int_0^\tau \int_{\mathbb{R}^d} D^\alpha u_i \big(D^\alpha(\nabla u_i \cdot \nabla U_i) - D^\alpha \nabla u_i \cdot \nabla U_i\big)\mathrm{d}x\mathrm{d}t$$

$$+ \int_0^\tau \int_{\mathbb{R}^d} D^\alpha u_i \big(D^\alpha(u_i \Delta U_i) - D^\alpha u_i \Delta U_i\big)\mathrm{d}x\mathrm{d}t, \tag{22}$$

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

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

Moreover, the last integral in (22) vanishes since $\Delta U_i = -1$. 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 straightforard computation leads to

$$I_7 = -d \int_0^\tau \int_{\mathbb{R}^d} (D^\alpha u_i)^2 \mathrm{d}x\mathrm{d}t \le 0.$$

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:

$$\begin{aligned} I_9 &\le \frac{\sigma_i}{4}\int_0^\tau \|\nabla D^\alpha u_i\|^2_{L^2(\mathbb{R}^d)}\mathrm{d}t + C(\sigma_i)\int_0^\tau \|D^\alpha(u_i \nabla K_i(u))\|^2_{L^2(\mathbb{R}^d)}\mathrm{d}t \\ &\le \frac{\sigma_i}{4}\int_0^\tau \|\nabla D^\alpha u_i\|^2_{L^2(\mathbb{R}^d)}\mathrm{d}t + C\int_0^\tau \big(\|u_i\|_{L^\infty(\mathbb{R}^d)}\|D^m \nabla K_i(u)\|_{L^2(\mathbb{R}^d)} \\ &\quad + \|D^m u_i\|_{L^2(\mathbb{R}^d)}\|\nabla K_i(u)\|_{L^\infty(\mathbb{R}^d)}\big)^2 \mathrm{d}t. \end{aligned}$$

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

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

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

$$\begin{aligned} \|D^m \nabla K_i(u)\|_{L^2(\mathbb{R}^d)} &\le A\sum_{j=1}^n \|\nabla u_j\|_{L^\infty(\mathbb{R}^d)} \|D^m(f'_\eta(B^\eta_{ij} * u_j))\|_{L^2(\mathbb{R}^d)} \\ &\le C\sum_{j=1}^n \|\nabla u_j\|_{H^s(\mathbb{R}^d)}\|f''\|_{C^{m-1}(I)}\|B^\eta_{ij} * u_j\|^{m-1}_{L^\infty(\mathbb{R}^d)}\|B^\eta_{ij} * D^m u_j\|_{L^2(\mathbb{R}^d)} \\ &\le \varepsilon C \|\nabla u\|_{H^s(\mathbb{R}^d)}\|u\|^{m-1}_{L^\infty(\mathbb{R}^d)}\|D^m u\|_{L^2(\mathbb{R}^d)} \le \varepsilon C\|\nabla u\|_{H^s(\mathbb{R}^d)}\|u_0\|^s_{H^s(\mathbb{R}^d)}, \end{aligned}$$

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

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

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

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

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

$$\|u(\tau)\|_{H^s(\mathbb{R}^d)}^2 + \big(\sigma_{\min} - C\varepsilon^2 \|u_0\|_{H^s(\mathbb{R}^d)}^s\big) \int_0^\tau \|\nabla u\|_{H^s(\mathbb{R}^d)}^2 \mathrm{d}t \le \|u_0\|_{H^s(\mathbb{R}^d)}^2.$$

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

Step 3: Global existence and uniqueness. We have proved that $\|u(\tau)\|_{H^s(\mathbb{R}^d)} \le \|u_0\|_{H^s(\mathbb{R}^d)}$ for $0 < \tau \le T$ for some sufficiently small $T > 0$. The value for T does 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_η and f'_η (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.

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

$$\begin{aligned} \int_0^T \langle \partial_t u_{\eta.i}, \phi \rangle \zeta(t) \mathrm{d}t = &- \int_0^T \int_{\mathbb{R}^d} u_{\eta,i} \nabla U_i \cdot \nabla \phi \zeta(t) \mathrm{d}x \mathrm{d}t \\ &- \int_0^T \int_{\mathbb{R}^d} \big(\sigma_i \nabla u_{\eta,i} + \nabla(u_{\eta,i} K_i(u_\eta))\big) \cdot \nabla \phi \zeta(t) \mathrm{d}x \mathrm{d}t, \end{aligned} \tag{23}$$

where $\langle \cdot, \cdot \rangle$ is the duality pairing between $H^{-1}(\mathbb{R}^d)$ and $H^1(\mathbb{R}^d)$ and $K_i(u) = \sum_{j=1}^n f_\eta(B_{ij}^\eta * u_j)$. 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 [4] that

$$B_{ij}^\eta * \nabla u_{\eta,j} \rightharpoonup a_{ij} \nabla u_j \quad \text{weakly in } L^2(0,T; L^2(\mathbb{R}^d)).$$

We claim that $f_\eta(B_{ij}^\eta * u_{\eta,j}) \to f(a_{ij} u_j)$ 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

$$\begin{aligned} \|\partial_t u_{\eta,i}\|_{L^2(0,T;H^{-1}(B_R))} \le\ & \|u_{\eta,i}\|_{L^2(0,T;L^2(\mathbb{R}^d))} \|\nabla U_i\|_{L^\infty(B_R)} + \sigma_i \|\nabla u_{\eta,i}\|_{L^2(0,T;L^2(\mathbb{R}^d))} \\ &+ \|\nabla u_{\eta,i}\|_{L^2(0,T;L^2(\mathbb{R}^d))} \|K_i(u_\eta)\|_{L^\infty(0,T;L^\infty(\mathbb{R}^d))} \\ &+ \|u_{\eta,i}\|_{L^2(0,T;L^2(\mathbb{R}^d))} \|\nabla K_i(u_\eta)\|_{L^\infty(0,T;L^\infty(\mathbb{R}^d))}. \end{aligned}$$

Because of

$$\|K_i(u_\eta)\|_{L^\infty(0,T;L^\infty(\mathbb{R}^d))} \le \sum_{j=1}^n \|f_\eta(B_{ij}^\eta * u_{\eta,j})\|_{L^\infty(0,T;L^\infty(\mathbb{R}^d))} \le C\|f\|_{L^\infty(I)},$$

$$\begin{aligned}\|\nabla K_i(u_\eta)\|_{L^\infty(0,T;L^\infty(\mathbb{R}^d))} &\le \sum_{j=1}^n \|f_\eta'(B_{ij}^\eta * u_{\eta,j})\|_{L^\infty(0,T;L^\infty(\mathbb{R}^d))}\|B_{ij}^\eta * \nabla u_{\eta,j}\|_{L^\infty(0,T;L^\infty(\mathbb{R}^d))} \\ &\le C\|f'\|_{L^\infty(I)}\|\nabla u_\eta\|_{L^\infty(0,T;L^\infty(\mathbb{R}^d))} \le C\|u_0\|_{H^s(\mathbb{R}^d)},\end{aligned}$$

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 $\eta\to 0$,

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

Since u_η 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

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

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

$$\begin{aligned}&\|f_\eta(B_{ij}^\eta * u_{\eta,j}) - f(a_{ij}u_j)\|_{L^2(0,T;L^2(B_R))} \\ &\quad\le C\|B_{ij}^\eta * (u_{\eta,j}-u_j) + B_{ij}^\eta * u_j - a_{ij}u_j\|_{L^2(0,T;L^2(B_R))} \\ &\quad\le C\|B_{ij}^\eta\|_{L^1(\mathbb{R}^d)}\|u_{\eta,j}-u_j\|_{L^2(0,T;L^2(B_R))} + \|B_{ij}^\eta * u_j - a_{ij}u_j\|_{L^2(0,T;L^2(B_R))} \to 0.\end{aligned}$$

This shows the claim. In a similar way, it follows from the Lipschitz continuity of f_η' that $f_\eta'(B_{ij}^\eta * u_{\eta,j}) \to f'(a_{ij}u_j)$ 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

$$\int_0^T \langle \partial_t u_i, \phi\rangle \zeta(t)\mathrm{d}t = -\int_0^T\int_{\mathbb{R}^d} u_i\nabla U_i\cdot\nabla\phi\zeta(t)\mathrm{d}x\mathrm{d}t - \int_0^T\int_{\mathbb{R}^d}\nabla F_i(u)\cdot\nabla\phi\zeta(t)\mathrm{d}x\mathrm{d}t,$$

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:

$$\begin{aligned}\frac12\int_{\mathbb{R}^d}(u_{\eta,i}-u_i)(\tau)^2\mathrm{d}x + \sigma_i\int_0^\tau\int_{\mathbb{R}^d}|\nabla(u_{\eta,i}-u_i)|^2\mathrm{d}x\mathrm{d}t = \frac12\int_0^\tau\int_{\mathbb{R}^d}\Delta U_i(u_{\eta,i}-u_i)^2\mathrm{d}x\mathrm{d}t \\ -\int_0^\tau\int_{\mathbb{R}^d}\nabla\sum_{j=1}^n\big(u_{\eta,i}f_\eta(B_{ij}^\eta * u_{\eta,j}) - u_if(a_{ij}u_j)\big)\cdot\nabla(u_{\eta,i}-u_i)\mathrm{d}x\mathrm{d}t. \qquad (24)\end{aligned}$$

The first integral on the right-hand side is nonpositive since $\Delta U_i = -d$. We split the second integral into three parts:

$$-\int_0^\tau\int_{\mathbb{R}^d}\sum_{j=1}^n\nabla\big(u_{\eta,i}f_\eta(B_{ij}^\eta * u_{\eta,j}) - u_if(a_{ij}u_j)\big)\cdot\nabla(u_{\eta,i}-u_i)\mathrm{d}x\mathrm{d}t = J_1+J_2+J_3, \qquad (25)$$

where

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

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

$$J_3 = -\int_0^\tau \int_{\mathbb{R}^d} \sum_{j=1}^n \nabla\big(u_i\big(f_\eta(a_{ij}u_{\eta,j}) - f(a_{ij}u_j)\big)\big) \cdot \nabla(u_{\eta,i} - u_i)\mathrm{d}x\mathrm{d}t.$$

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 $\|f_\eta\|_{L^\infty(I)} = \|f\|_{L^\infty(I)} \le \varepsilon$ and Young's inequality, we have

$$\begin{aligned} J_1 &\le \|f_\eta(B_{ij}^\eta * u_{\eta,j})\|_{L^\infty(0,T;L^\infty(\mathbb{R}^d))} \int_0^\tau \|\nabla(u_{\eta,i} - u_i)\|_{L^2(\mathbb{R}^d)}^2 \mathrm{d}t \\ &\quad + \int_0^\tau \|u_{\eta,i} - u_i\|_{L^2(\mathbb{R}^d)} \|f_\eta'(B_{ij}^\eta * u_{\eta,j})\|_{L^\infty(0,T;L^\infty(\mathbb{R}^d))} \\ &\quad \times \|B_{ij}^\eta * \nabla u_{\eta,j}\|_{L^\infty(0,T;L^\infty(\mathbb{R}^d))} \|\nabla(u_{\eta,i} - u_i)\|_{L^2(\mathbb{R}^d)}\mathrm{d}t \\ &\le \left(\frac{\sigma_i}{4} + \varepsilon\right) \int_0^\tau \|\nabla(u_{\eta,i} - u_i)\|_{L^2(\mathbb{R}^d)}^2 \mathrm{d}t + C(\sigma_i) \int_0^\tau \|u_{\eta,i} - u_i\|_{L^2(\mathbb{R}^d)}^2 \mathrm{d}t. \end{aligned} \tag{26}$$

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

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

$$J_{22} = -\int_0^\tau \int_{\mathbb{R}^d} u_i \sum_{j=1}^n \big(f_\eta'(B_{ij}^\eta * u_{\eta,j})B_{ij}^\eta * \nabla u_{\eta,j} - f_\eta'(a_{ij}u_{\eta,j})a_{ij}\nabla u_{\eta,j}\big) \cdot \nabla(u_{\eta,i} - u_i)\mathrm{d}x\mathrm{d}t.$$

It follows that

$$\begin{aligned} J_{21} &\le \|\nabla u_i\|_{L^\infty(0,T;L^\infty(\mathbb{R}^d))} \sum_{j=1}^n \int_0^\tau \|f_\eta(B_{ij}^\eta * u_{\eta,j}) - f_\eta(a_{ij}u_{\eta,j})\|_{L^2(\mathbb{R}^d)} \|\nabla(u_{\eta,i} - u_i)\|_{L^2(\mathbb{R}^d)}\mathrm{d}t \\ &\le \frac{\sigma_i}{8} \int_0^\tau \|\nabla(u_{\eta,i} - u_i)\|_{L^2(\mathbb{R}^d)}^2 \mathrm{d}t + C \sum_{j=1}^n \int_0^\tau \|f_\eta(B_{ij}^\eta * u_{\eta,j}) - f_\eta(a_{ij}u_{\eta,j})\|_{L^2(\mathbb{R}^d)}^2 \mathrm{d}t. \end{aligned}$$

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 η. We use this information in

$$\left| \int_{\mathbb{R}^d} \big(f_\eta(B_{ij}^\eta * u_{\eta,j}) - f_\eta(a_{ij}u_{\eta,j})\big)g(x)\mathrm{d}x \right| \le C \int_{\mathbb{R}^d} \big|B_{ij}^\eta * u_{\eta,j} - a_{ij}u_{\eta,j}\big| |g(x)|\mathrm{d}x,$$

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

$$\begin{aligned} &\left| \int_{\mathbb{R}^d} \big(f_\eta(B_{ij}^\eta * u_{\eta,j}) - f_\eta(a_{ij}u_{\eta,j})\big)g(x)\mathrm{d}x \right| \\ &\qquad \le C \int_{\mathbb{R}^d} \left| \int_{B_\eta} B_{ij}^\eta(y)\big(u_{\eta,j}(x-y) - u_{\eta,j}(x)\big)\mathrm{d}y \right| |g(x)|\mathrm{d}x \end{aligned}$$

$$
\begin{aligned}
&\le C\int_{\mathbb{R}^d}\int_{B_\eta}|B_{ij}^\eta(y)|\bigg(\int_0^1|\nabla u_{\eta,j}(x-ry)|\eta\mathrm{d}r\bigg)\mathrm{d}y|g(x)|\mathrm{d}x \\
&= C\eta\int_0^1\int_{B_\eta}|B_{ij}^\eta(y)|\bigg(\int_{\mathbb{R}^d}|\nabla u_{\eta,j}(x-ry)||g(x)|\mathrm{d}x\bigg)\mathrm{d}y\mathrm{d}r \\
&\le C\eta\int_0^1\int_{B_\eta}|B_{ij}^\eta(y)|\|\nabla u_{\eta,j}(\cdot-ry)\|_{L^2(\mathbb{R}^d)}\|g\|_{L^2(\mathbb{R}^d)}\mathrm{d}y\mathrm{d}r \\
&\le C\eta\int_{B_\eta}|B_{ij}^\eta(y)|\mathrm{d}y\|\nabla u_{\eta,j}\|_{L^2(\mathbb{R}^d)}\|g\|_{L^2(\mathbb{R}^d)}\le C\eta\|g\|_{L^2(\mathbb{R}^d)}.
\end{aligned}
$$

By duality, we find that

$$
J_{21}\le\frac{\sigma_i}{8}\int_0^\tau\|\nabla(u_{\eta,i}-u_i)\|_{L^2(\mathbb{R}^d)}^2\mathrm{d}t+C\eta^2.
$$

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

$$
\begin{aligned}
J_{221}&=-\int_0^\tau\int_{\mathbb{R}^d}u_i\sum_{j=1}^n f_\eta'(B_{ij}^\eta*u_{\eta,j})\big(B_{ij}^\eta*\nabla u_{\eta,j}-a_{ij}\nabla u_{\eta,j}\big)\cdot\nabla(u_{\eta,i}-u_i)\mathrm{d}x\mathrm{d}t, \\
J_{222}&=-\int_0^\tau\int_{\mathbb{R}^d}u_i\sum_{j=1}^n\big(f_\eta'(B_{ij}^\eta*u_{\eta,j})-f_\eta'(a_{ij}u_{\eta,j})\big)a_{ij}\nabla u_{\eta,j}\cdot\nabla(u_{\eta,i}-u_i)\mathrm{d}x\mathrm{d}t.
\end{aligned}
$$

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

$$
\begin{aligned}
J_{221}&\le\frac{\sigma_i}{16}\int_0^\tau\|\nabla(u_{\eta,i}-u_i)\|_{L^2(\mathbb{R}^d)}^2\mathrm{d}t+C\int_0^T\|B_{ij}^\eta*\nabla u_{\eta,j}-a_{ij}\nabla u_{\eta,j}\|_{L^2(\mathbb{R}^d)}^2\mathrm{d}t \\
&\le\frac{\sigma_i}{16}\int_0^\tau\|\nabla(u_{\eta,i}-u_i)\|_{L^2(\mathbb{R}^d)}^2\mathrm{d}t+C\eta^2\int_0^\tau\|D^2u_{\eta,j}\|_{L^2(\mathbb{R}^d)}^2\mathrm{d}t,
\end{aligned}
$$

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

$$
\begin{aligned}
J_{222}&\le C\int_0^\tau\|u_i\|_{L^\infty(\mathbb{R}^d)}\|B_{ij}^\eta*u_{\eta,j}-a_{ij}u_{\eta,j}\|_{L^2(\mathbb{R}^d)}\|\nabla u_{\eta,j}\|_{L^\infty(\mathbb{R}^d)}\|\nabla(u_{\eta,i}-u_i)\|_{L^2(\mathbb{R}^d)}\mathrm{d}t \\
&\le\frac{\sigma_i}{16}\int_0^\tau\|\nabla(u_{\eta,i}-u_i)\|_{L^2(\mathbb{R}^d)}^2\mathrm{d}t+C\eta^2\int_0^\tau\|\nabla u_{\eta,j}\|_{L^2(\mathbb{R}^d)}^2\mathrm{d}t.
\end{aligned}
$$

Summarizing these estimates, we infer that

$$
J_{22}\le\frac{\sigma_i}{8}\int_0^\tau\|\nabla(u_{\eta,i}-u_i)\|_{L^2(\mathbb{R}^d)}^2\mathrm{d}t+C\eta^2,
$$

and combining the estimate for J_{21} and J_{22},

$$
J_2\le\frac{\sigma_i}{4}\int_0^\tau\|\nabla(u_{\eta,i}-u_i)\|_{L^2(\mathbb{R}^d)}^2\mathrm{d}t+C\eta^2. \tag{27}
$$

It remains to estimate $J_3 = J_{31} + J_{32}$, where

$$J_{31} = -\int_0^\tau \int_{\mathbb{R}^d} \sum_{j=1}^n \big(f_\eta(a_{ij}u_{\eta,j}) - f(a_{ij}u_j)\big)\nabla u_i \cdot \nabla(u_{\eta,i} - u_i)\mathrm{d}x\mathrm{d}t,$$

$$J_{32} = -\int_0^\tau \int_{\mathbb{R}^d} u_i \sum_{j=1}^n \big(f_\eta'(a_{ij}u_{\eta,j})a_{ij}\nabla u_{\eta,j} - f'(a_{ij}u_j)a_{ij}\nabla u_j\big) \cdot \nabla(u_{\eta,i} - u_i)\mathrm{d}x\mathrm{d}t.$$

Similar arguments as above yield

$$\begin{aligned} J_{31} &\le \frac{\sigma_i}{8}\int_0^\tau \|\nabla(u_{\eta,i} - u_i)\|_{L^2(\mathbb{R}^d)}^2\mathrm{d}t + C\int_0^\tau \|\nabla u_i\|_{L^\infty(\mathbb{R}^d)}^2 \|u_\eta - u\|_{L^2(\mathbb{R}^d)}^2\mathrm{d}t \\ &\le \frac{\sigma_i}{8}\int_0^\tau \|\nabla(u_{\eta,i} - u_i)\|_{L^2(\mathbb{R}^d)}^2\mathrm{d}t + C\int_0^\tau \|u_\eta - u\|_{L^2(\mathbb{R}^d)}^2\mathrm{d}t. \end{aligned}$$

The second term J_{32} is again split into two parts, $J_{32} = J_{321} + J_{322}$, where

$$J_{321} = -\int_0^\tau \int_{\mathbb{R}^d} u_i \sum_{j=1}^n \big(f_\eta'(a_{ij}u_{\eta,j}) - f_\eta'(a_{ij}u_j)\big)a_{ij}\nabla u_{\eta,j} \cdot \nabla(u_{\eta,i} - u_i)\mathrm{d}x\mathrm{d}t,$$

$$J_{322} = -\int_0^\tau \int_{\mathbb{R}^d} u_i \sum_{j=1}^n a_{ij}\big(f_\eta'(a_{ij}u_j)\nabla u_{\eta,j} - f'(a_{ij}u_j)\nabla u_j\big) \cdot \nabla(u_{\eta,i} - u_i)\mathrm{d}x\mathrm{d}t.$$

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

$$\begin{aligned} J_{321} &\le C\|u_i\|_{L^\infty(0,T;L^\infty(\mathbb{R}^d))}\int_0^\tau \sum_{j=1}^n \|\nabla u_{\eta,j}\|_{L^\infty(\mathbb{R}^d)}\|u_{\eta,j} - u_j\|_{L^2(\mathbb{R}^d)}\|\nabla(u_{\eta,i} - u_i)\|_{L^2(\mathbb{R}^d)}\mathrm{d}t \\ &\le \frac{\sigma_i}{8}\int_0^\tau \|\nabla(u_{\eta,i} - u_i)\|_{L^2(\mathbb{R}^d)}^2\mathrm{d}t + C\int_0^\tau \|u_\eta - u\|_{L^2(\mathbb{R}^d)}^2, \\ J_{322} &\le C\int_0^\tau \sum_{j=1}^n \|f'(a_{ij}u_j)\|_{L^\infty(\mathbb{R}^d)}\|\nabla(u_{\eta,j} - u_j)\|_{L^2(\mathbb{R}^d)}\|\nabla(u_{\eta,i} - u_i)\|_{L^2(\mathbb{R}^d)}\mathrm{d}t \\ &\le C\varepsilon\int_0^\tau \|\nabla(u_\eta - u)\|_{L^2(\mathbb{R}^d)}^2\mathrm{d}t. \end{aligned}$$

This shows that

$$J_{32} \le \left(\frac{\sigma_i}{8} + C\varepsilon\right)\int_0^\tau \|\nabla(u_{\eta,i} - u_i)\|_{L^2(\mathbb{R}^d)}^2\mathrm{d}t + C\int_0^\tau \|u_\eta - u\|_{L^2(\mathbb{R}^d)}^2.$$

Summarizing the estimate for J_{31} and J_{32}, we arrive at

$$J_3 \le \left(\frac{\sigma_i}{4} + C\varepsilon\right)\int_0^\tau \|\nabla(u_{\eta,i} - u_i)\|_{L^2(\mathbb{R}^d)}^2\mathrm{d}t + C\int_0^\tau \|u_\eta - u\|_{L^2(\mathbb{R}^d)}^2\mathrm{d}t. \tag{28}$$

Finally, putting together the estimates (26), (27), and (28), we infer from (25) that

$$\bigg|\int_0^\tau\int_{\mathbb{R}^d}\sum_{j=1}^n\nabla\big(u_{\eta,i}f_\eta(B_{ij}^\eta*u_{\eta,j})-u_if(a_{ij}u_j)\big)\cdot\nabla(u_{\eta,i}-u_i)\mathrm{d}x\mathrm{d}t\bigg|$$
$$\le\left(\frac{3\sigma_i}{4}+C\varepsilon\right)\int_0^\tau\|\nabla(u_{\eta,j}-u_j)\|^2_{L^2(\mathbb{R}^d)}\mathrm{d}t+C\int_0^\tau\|u_\eta-u\|^2_{L^2(\mathbb{R}^d)}\mathrm{d}t+C\eta^2.$$

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

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

The proof ends after applying Gronwall's inequality.

5. Links between the SDEs and PDEs

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

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

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

in the weak integrable sense, i.e.

$$\int_{\mathbb{R}^d}\widehat{u}_i(t)\phi(t)\mathrm{d}x-\int_{\mathbb{R}^d}u_{0,i}\phi(0)\mathrm{d}x-\int_0^t\int_{\mathbb{R}^d}\widehat{u}_i\partial_t\phi\mathrm{d}x\mathrm{d}s$$
$$=-\int_0^t\int_{\mathbb{R}^d}\widehat{u}_i\nabla U_i\cdot\nabla\phi\mathrm{d}x\mathrm{d}t+\int_0^t\int_{\mathbb{R}^d}\widehat{u}_i\bigg(\sigma_i+\sum_{j=1}^n f(a_{ij}u_j)\bigg)\Delta\phi\mathrm{d}x\mathrm{d}s$$

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

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

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

Proof. Since $\widehat{X}_{k,i}$ depends on k only via the initial data ξ_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

$$\phi(t,\widehat{X}_i(t))=\phi(0,\xi_i)+\int_0^t\partial_t\phi(s,\widehat{X}_i(s))\mathrm{d}s-\int_0^t\nabla U_i(s)\cdot\nabla\phi(s,\widehat{X}_i(s))\mathrm{d}s$$

$$+\int_0^t F_i\big(u(\widehat{X}_i(s))\big)\Delta\phi(s,\widehat{X}_i(s))\mathrm{d}s+\int_0^t F_i\big(u(\widehat{X}_i(s))\big)^{1/2}\nabla\phi(s,X(s))\cdot\mathrm{d}W_i(s). \tag{31}$$

We claim that the density function $\widehat{u}_i : [0,\infty) \to \mathcal{P}_2(\mathbb{R}^d)$, 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 $\widehat{X}_i$ is square-integrable, we have $\widehat{u}_i(t) \in \mathcal{P}_2(\mathbb{R}^d)$ for almost all $t > 0$ and the limit $s \to t$ in the Wasserstein distance leads to

$$\begin{aligned} W_2(\widehat{u}_i(t),\widehat{u}_i(s)) &= \inf\big\{\big(\mathbb{E}(|Y_t-Y_s|^2)\big)^{1/2} : \ \mathrm{Law}(Y_t)=\widehat{u}_i(t),\ \mathrm{Law}(Y_s)=\widehat{u}_i(s)\big\} \\ &\le \big(\mathbb{E}(|\widehat{X}_i(t)-\widehat{X}_i(s)|^2)\big)^{1/2} \to 0, \end{aligned}$$

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

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

$$\begin{aligned} \int_{\mathbb{R}^d}\widehat{u}_i(t)\phi(t)\mathrm{d}x &= \int_{\mathbb{R}^d} u_{0,i}\phi(0)\mathrm{d}x + \int_0^t\int_{\mathbb{R}^d}\widehat{u}_i(s)\partial_t\phi(s)\mathrm{d}x\mathrm{d}s \\ &\quad - \int_0^t\int_{\mathbb{R}^d}\widehat{u}_i(s)\nabla U_i\cdot\nabla\phi(s)\mathrm{d}x\mathrm{d}s + \int_0^t\int_{\mathbb{R}^d}\widehat{u}_i(s)F_i(u(s))\Delta\phi(s)\mathrm{d}x\mathrm{d}s. \end{aligned}$$

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 $\widehat{u}_i(t) = 0$ 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 χ_k be a family of mollifiers and consider the regularized dual backward problem on the ball B_R around the origin with radius $R > 0$:

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

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 $\int_{\mathbb{R}^d} g(x)\widehat{u}_i(s,x)\mathrm{d}x = 0$, and since g was arbitrary, we conclude that $\widehat{u}_i(s) = 0$ for $0 < s < t$.

The weak solution u to (1) is also a very weak solution to (29). Therefore, by the previous uniqueness result, $\widehat{u} = u$. □

Similar arguments lead to the following result that relates the solutions $\overline{u}_\eta$ and u_η.

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

Then $\overline{u}_\eta = (\overline{u}_{\eta,1}, \ldots, \overline{u}_{\eta,n})$ solves the linear problem

$$\partial_t \overline{u}_{\eta,i} = \operatorname{div}(\overline{u}_{\eta,i} \nabla U_i) + \Delta\bigg(\sigma_i \overline{u}_{\eta,i} + \overline{u}_{\eta,i} \sum_{j=1}^n f_\eta(B^\eta_{ij} * u_{\eta,j})\bigg) \quad \text{in } \mathbb{R}^d,\ i = 1, \ldots, n,$$

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

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

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 $X^{N,\eta}_{k,i} - \overline{X}^\eta_{k,i}$, where $\overline{X}^\eta_{k,i}$ is the solution to the intermediate system (6), and then the square mean error of the difference $\overline{X}^\eta_{k,i} - \widehat{X}_{k,i}$.

Lemma 9. *Let $X^{N,\eta}_{k,i}$ and $\overline{X}^\eta_{k,i}$ 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, $\sigma_{\min}$, and T, such that if $\eta^{-2(d+1+\alpha)} \leq \delta \log N$, where $\alpha \geq 0$ is fixed in Assumption (A4), we have*

$$\sup_{k=1,\ldots,N} \mathbb{E}\bigg(\sum_{i=1}^n \sup_{0<s<T} \big|(X^{N,\eta}_{k,i} - \overline{X}^\eta_{k,i})(s)\big|^2\bigg) \leq C(T,n,\sigma_{\min}) N^{-1+(T+1)C(n,\sigma_{\min})\delta},$$

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

Proof. The process $D^{N,\eta}_{k,i} := X^{N,\eta}_{k,i} - \overline{X}^\eta_{k,i}$ solves

$$D^{N,\eta}_{k,i}(s) = E_{1,i}(s) + E_{2,i}(s), \quad 0 \leq s \leq T, \tag{32}$$

where

$$E_{1,i}(s) = -\int_0^s \big(\nabla U_i(X^{N,\eta}_{k,i}(t)) - \nabla U_i(\overline{X}^\eta_{k,i}(t))\big)\mathrm{d}t,$$

$$E_{2,i}(s) = \int_0^s (E_{21}(t) - E_{22}(t))\mathrm{d}W^k_i(t),$$

$$E_{21}(t) = \Bigg(2\sigma_i + 2\sum_{j=1}^n f_\eta\bigg(\frac{1}{N}\sum_{\substack{\ell=1\\(\ell,j)\neq(k,i)}}^N B^\eta_{ij}(X^{N,\eta}_{k,i}(t) - X^{N,\eta}_{\ell,j}(t))\bigg)\Bigg)^{1/2},$$

$$E_{22}(t) = \bigg(2\sigma_i + 2\sum_{j=1}^n f_\eta\big(B^\eta_{ij} * u_{\eta,j}(t, \overline{X}^\eta_{k,i}(t))\big)\bigg)^{1/2}.$$

We use the global Lipschitz continuity of ∇U_i and the Fubini theorem to estimate the first term:

$$
\begin{aligned}
\mathbb{E}\Big(\sup_{0<s<T}|E_{1,i}(s)|^2\Big) &\le C\mathbb{E}\int_0^T \big|(X_{k,i}^{N,\eta}-\overline{X}_{k,i}^{\eta})(s)\big|^2\mathrm{d}s \\
&\le C\int_0^T \mathbb{E}\Big(\sup_{0<s<t}|(X_{k,i}^{N,\eta}-\overline{X}_{k,i}^{\eta})(s)|^2\Big)\mathrm{d}t.
\end{aligned}
$$

Summing over $i=1,\ldots,n$ and taking the supremum over $k=1,\ldots,N$ leads to

$$
\sup_{k=1,\ldots,N}\mathbb{E}\bigg(\sum_{i=1}^n\sup_{0<s<T}|E_{1,i}(s)|^2\bigg) \le C\int_0^T\sup_{k=1,\ldots,N}\mathbb{E}\Big(\sup_{0<s<t}|(X_{k,i}^{N,\eta}-\overline{X}_{k,i}^{\eta})(s)|^2\Big)\mathrm{d}t. \tag{33}
$$

Next, we apply the Burkholder–Davis–Gundy inequality [16, 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 $x\ge 0$:

$$
\begin{aligned}
\mathbb{E}\Big(\sup_{0<s<T}|E_{2,i}(s)|^2\Big) &\le C\mathbb{E}\int_0^T (E_{21}(t)-E_{22}(t))^2\mathrm{d}t \\
&\le C\mathbb{E}\int_0^T\bigg[\sum_{j=1}^n f_\eta\bigg(\frac{1}{N}\sum_{\substack{\ell=1\\(\ell,j)\neq(k,i)}}^N B_{ij}^\eta(X_{k,i}^{N,\eta}(t)-X_{\ell,j}^{N,\eta}(t))\bigg) \\
&\qquad -\sum_{j=1}^n f_\eta\big(B_{ij}^\eta * u_{\eta,j}(t,\overline{X}_{k,i}^{\eta}(t))\big)\bigg]^2\mathrm{d}t \\
&= C\mathbb{E}\int_0^T\bigg[\sum_{j=1}^n(L_j^1(t)+L_j^2(t)+L_j^3(t))\bigg]^2\mathrm{d}t \\
&\le C(n)\mathbb{E}\int_0^T\sum_{j=1}^n\big(L_j^1(t)^2+L_j^2(t)^2+L_j^3(t)^2\big)\mathrm{d}t, \qquad (34)
\end{aligned}
$$

where

$$
\begin{aligned}
L_j^1(t) &= f_\eta\bigg(\frac{1}{N}\sum_{\substack{\ell=1\\(\ell,j)\neq(k,i)}}^N B_{ij}^\eta(X_{k,i}^{N,\eta}(t)-X_{\ell,j}^{N,\eta}(t))\bigg) - f_\eta\bigg(\frac{1}{N}\sum_{\substack{\ell=1\\(\ell,j)\neq(k,i)}}^N B_{ij}^\eta(\overline{X}_{k,i}^{\eta}(t)-X_{\ell,j}^{N,\eta}(t))\bigg), \\
L_j^2(t) &= f_\eta\bigg(\frac{1}{N}\sum_{\substack{\ell=1\\(\ell,j)\neq(k,i)}}^N B_{ij}^\eta(\overline{X}_{k,i}^{\eta}(t)-X_{\ell,j}^{N,\eta}(t))\bigg) - f_\eta\bigg(\frac{1}{N}\sum_{\substack{\ell=1\\(\ell,j)\neq(k,i)}}^N B_{ij}^\eta(\overline{X}_{k,i}^{\eta}(t)-\overline{X}_{\ell,j}^{\eta}(t))\bigg), \\
L_j^3(t) &= f_\eta\bigg(\frac{1}{N}\sum_{\substack{\ell=1\\(\ell,j)\neq(k,i)}}^N B_{ij}^\eta(\overline{X}_{k,i}^{\eta}(t)-\overline{X}_{\ell,j}^{\eta}(t))\bigg) - f_\eta\big(B_{ij}^\eta * u_{\eta,j}(t,\overline{X}_{k,i}^{\eta}(t))\big).
\end{aligned}
$$

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

$$
\begin{aligned}
|L_j^1(t)| &\le L_f\bigg|\frac{1}{N}\sum_{\substack{\ell=1\\(\ell,j)\neq(k,i)}}^N \big(B_{ij}^\eta(X_{k,i}^{N,\eta}(t)-X_{\ell,j}^{N,\eta}(t)) - B_{ij}^\eta(\overline{X}_{k,i}^\eta(t)-X_{\ell,j}^{N,\eta}(t))\big)\bigg| \\
&\le L_f L_B\big|X_{k,i}^{N,\eta}(t)-\overline{X}_{k,i}^\eta(t)\big| \le \eta^{-d-1-\alpha}\big|X_{k,i}^{N,\eta}(t)-\overline{X}_{k,i}^\eta(t)\big|.
\end{aligned}
$$

Therefore, by Fubini's theorem,

$$
\begin{aligned}
\mathbb{E}\int_0^T\sum_{j=1}^n|L_j^1(t)|^2\mathrm{d}t &\le C(n)\eta^{-2(d+1+\alpha)}\mathbb{E}\int_0^T\big|X_{k,i}^{N,\eta}(t)-\overline{X}_{k,i}^\eta(t)\big|^2\mathrm{d}t \\
&\le C(n)\eta^{-2(d+1+\alpha)}\int_0^T\sup_{k=1,\ldots,N}\mathbb{E}\Big(\sup_{0<s<t}\big|X_{k,i}^{N,\eta}(t)-\overline{X}_{k,i}^\eta(t)\big|^2\Big)\mathrm{d}t. \qquad (35)
\end{aligned}
$$

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

$$
\mathbb{E}\int_0^T\sum_{j=1}^n L_j^2(t)^2\mathrm{d}t \le C(n)\eta^{-2(d+1+\alpha)}\int_0^T\sup_{\ell=1,\ldots,N}\mathbb{E}\bigg(\sup_{0<s<t}\sum_{j=1}^n\big|X_{\ell,j}^{N,\eta}(t)-\overline{X}_{\ell,j}^\eta(t)\big|^2\bigg)\mathrm{d}t. \qquad (36)
$$

The third term $L_j^3(t)$ has to be treated in a different way. First, we use the Lipschitz continuity of f_η to find that

$$
L_j^3(t) \le \frac{C(n)}{N\eta^\alpha}\bigg|\sum_{\ell=1}^N\big(B_{ij}^\eta(\overline{X}_{k,i}^\eta-\overline{X}_{\ell,j}^\eta) - B_{ij}^\eta * u_{\eta,j}(\overline{X}_{k,i}^\eta)\big) - \frac{1}{\eta^d}B_{ii}(0)\bigg|.
$$

This implies that

$$
\begin{aligned}
\mathbb{E}\int_0^T\sum_{j=1}^n L_j^3(t)^2\mathrm{d}t &\le \frac{C(n,T)}{N^2\eta^{2(d+\alpha)}} \\
&\quad + \frac{C(n)}{N^2\eta^{2\alpha}}\sum_{j=1}^n\int_0^T\mathbb{E}\bigg(\sum_{\ell=1}^N\Big(B_{ij}^\eta\big(\overline{X}_{k,i}^\eta(t)-\overline{X}_{\ell,j}^\eta(t)\big) - B_{ij}^\eta * u_{\eta,j}(\overline{X}_{k,i}^\eta)\Big)\bigg)^2\mathrm{d}t. \qquad (37)
\end{aligned}
$$

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

$$
\mathbb{E}\big(D_{(k,i),(\ell,j)}(t)D_{(k,i),(m,j)}(t)\big) = \int_{\mathbb{R}^d}\bigg(\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}B_{ij}^\eta(x-y)B_{ij}^\eta(x-z)u_{\eta,j}(t,y)u_{\eta,j}(t,z)\mathrm{d}y\mathrm{d}z
$$

$$- 2\int_{\mathbb{R}^d} B_{ij}^\eta(x-y)u_{\eta,j}(t,y)(B_{ij}^\eta * u_{\eta,j})(t,y)\mathrm{d}y$$
$$+ (B_{ij}^\eta * u_{\eta,j})(t,x)(B_{ij}^\eta * u_{\eta,j})(t,x)\bigg)u_{\eta,i}(t,x)\mathrm{d}x = 0.$$

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:

$$\begin{aligned}\mathbb{E}\big(D_{(k,i),(\ell,j)}(t)^2\big) &= \int_{\mathbb{R}^d}\bigg((B_{ij}^\eta * u_{\eta,j})(t,x)(B_{ij}^\eta * u_{\eta,j})(t,x) + \int_{\mathbb{R}^d}\Big(B_{ij}^\eta(x-y)^2 u_{\eta,j}(t,y) \\ &\quad - 2B_{ij}^\eta(x-y)u_{\eta,j}(t,y)(B_{ij}^\eta * u_{\eta,j}(t,x)\Big)\mathrm{d}y\bigg)u_{\eta,i}(t,x)\mathrm{d}x \\ &= \int_{\mathbb{R}^d}\big(((B_{ij}^\eta)^2 * u_{\eta,j})(t,x) - (B_{ij}^\eta * u_{\eta,j})(t,x)^2\big)u_{\eta,i}(t,x)\mathrm{d}x.\end{aligned}$$

This expression is independent of the particle index k and ℓ, 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 $D_{(k,i),(k,i)}(t) = \eta^{-d}B_{ii}(0) - B_{ii}^\eta * u_{i,\eta}(\overline{X}_{k,i}^\eta(t))$, we obtain for $\mathbb{E}(D_{(k,i),(k,i)}(t)D_{(k,i),(m,j)}(t))$ an additional term of order η^{-2d}. Hence, we infer from (37) and the previous computation that

$$\begin{aligned}\mathbb{E}\int_0^T \sum_{j=1}^n L_j^3(t)^2\mathrm{d}t - \frac{C(n,T)}{N^2\eta^{2(d+\alpha)}} &= \frac{C(n)}{N^2\eta^{2\alpha}}\sum_{j=1}^n\sum_{\ell=1}^N\int_0^T \mathbb{E}\big(D_{(k,i),(\ell,j)}(t)^2\big)\mathrm{d}t \\ &\le \frac{C(n)}{N\eta^{2\alpha}}\|u_{\eta,i}\|_{L^\infty(0,T;L^\infty(\mathbb{R}^d))} \\ &\quad\times\sum_{j=1}^n\int_0^T\bigg(\|(B_{ij}^\eta)^2 * u_{\eta,j}\|_{L^1(\mathbb{R}^d)} + \|B_{ij}^\eta * u_{\eta,j}\|_{L^2(\mathbb{R}^d)}^2\Big(1+\frac{1}{\eta^{2d}}\Big)\bigg)\mathrm{d}t \\ &\le \frac{C(n)}{N\eta^{2\alpha}}\sum_{j=1}^n\int_0^T\bigg(\|B_{ij}^\eta\|_{L^2(\mathbb{R}^d)}^2\|u_{\eta,j}\|_{L^\infty(\mathbb{R}^d)} + \|B_{ij}^\eta\|_{L^1(\mathbb{R}^d)}^2\|u_{\eta,j}\|_{L^2(\mathbb{R}^d)}^2\Big(1+\frac{1}{\eta^{2d}}\Big)\bigg)\mathrm{d}t \\ &\le \frac{C(T,n)}{N\eta^{2(d+\alpha)}}, \qquad (38)\end{aligned}$$

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

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

$$\begin{aligned}\sup_{k=1,\ldots,N}\mathbb{E}\bigg(\sum_{i=1}^n\sup_{0<s<T}|E_{2,i}(s)|^2\bigg) &\le \frac{C(T,n)}{N\eta^{2(d+\alpha)}} \\ &\quad + C(n,\sigma_{\min})\eta^{-2(d+1+\alpha)}\int_0^T\sup_{k=1,\ldots,N}\mathbb{E}\Big(\sup_{0<s<t}\big|X_{k,i}^{N,\eta}(t) - \overline{X}_{k,i}^\eta(t)\big|^2\Big)\mathrm{d}t.\end{aligned}$$

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

$$\begin{aligned} S(T) &:= \sup_{k=1,\ldots,N} \mathbb{E}\bigg(\sum_{i=1}^n \sup_{0<s<t} |D_{k,i}^{N,\eta}(s)|^2 \bigg) \\ &\le \frac{C(T,n)}{N\eta^{2(d+\alpha)}} + C(n,\sigma_{\min})(\eta^{-2(d+1+\alpha)}+1)\int_0^T S(t)\mathrm{d}t. \end{aligned}$$

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

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

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

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

This finishes the proof. □

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

Lemma 10. *Let $\overline{X}_{k,i}^\eta$ and $\widehat{X}_{k,i}$ 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*

$$\sup_{k=1,\ldots,N} \mathbb{E}\bigg(\sum_{i=1}^n \sup_{0<s<T} \big|(\overline{X}_{k,i}^\eta - \widehat{X}_{k,i})(s)\big|^2 \bigg) \le C(T,\sigma_{\min})\eta^{2(1-\alpha)}.$$

Proof. Since we are considering N independent copies, we can omit the particle index k. Set $D_i^\eta(s) := \overline{X}_{k,i}^\eta(s) - \widehat{X}_{k,i}(s)$. Then, similarly as in the proof of Lemma 9, $D_i^\eta(s) = D_1(s) + D_2(s)$, where

$$\begin{aligned} D_1(s) &= -\int_0^s \big(\nabla U_i(\overline{X}_i^\eta(s)) - \nabla U_i(\widehat{X}_i(s))\big)\mathrm{d}t, \\ D_2(s) &= \int_0^s \Bigg[\bigg(2\sigma_i + 2\sum_{j=1}^n f_\eta\big(B_{ij}^\eta * u_{\eta,j}(\overline{X}_i^\eta)\big)\bigg)^{1/2} \\ &\qquad - \bigg(2\sigma_i + 2\sum_{j=1}^n f\big(a_{ij}u_j(\widehat{X}_i)\big)\bigg)^{1/2}\Bigg]\mathrm{d}W_i(t). \end{aligned}$$

We infer from the Lipschitz continuity of ∇U_i and Fubini's theorem that

$$\mathbb{E}\Big(\sup_{0<s<T}|D_1(s)|^2\Big) \le C\mathbb{E}\bigg(\int_0^T \big|\overline{X}_i^\eta(s) - \widehat{X}_i(s)\big|^2\mathrm{d}s\bigg) \le C\int_0^T \mathbb{E}\Big(\sup_{0<s<t}|D_i^\eta(s)|^2\Big)\mathrm{d}t. \tag{39}$$

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

$$\begin{aligned} \mathbb{E}\Big(\sup_{0<s<T}|D_2(s)|^2\Big) &\le C\mathbb{E}\int_0^T \bigg(\sum_{j=1}^n \big(f(a_{ij}u_j(\widehat{X}_i)) - f_\eta(B_{ij}^\eta * u_{\eta,j}(\overline{X}_i^\eta))\big)\bigg)^2 \mathrm{d}t \\ &\le C(n)(D_{21}+D_{22}+D_{23}+D_{24}), \end{aligned} \tag{40}$$

where

$$\begin{aligned} D_{21} &= \sum_{j=1}^n \mathbb{E}\int_0^T \big(f(a_{ij}u_j(\widehat{X}_i)) - f_\eta(a_{ij}u_j(\widehat{X}_i))\big)^2 \mathrm{d}t, \\ D_{22} &= \sum_{j=1}^n \mathbb{E}\int_0^T \big(f_\eta(a_{ij}u_j(\widehat{X}_i)) - f_\eta(B_{ij}^\eta * u_j(\widehat{X}_i))\big)^2 \mathrm{d}t, \\ D_{23} &= \sum_{j=1}^n \mathbb{E}\int_0^T \big(f_\eta(B_{ij}^\eta * u_j(\widehat{X}_i)) - f_\eta(B_{ij}^\eta * u_j(\overline{X}_i^\eta))\big)^2 \mathrm{d}t, \\ D_{24} &= \sum_{j=1}^n \mathbb{E}\int_0^T \big(f_\eta(B_{ij}^\eta * u_j(\overline{X}_i^\eta)) - f_\eta(B_{ij}^\eta * u_{\eta,j}(\overline{X}_i^\eta))\big)^2 \mathrm{d}t. \end{aligned}$$

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

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

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

$$\begin{aligned} D_{22} &= \sum_{j=1}^n \int_0^T \int_{\mathbb{R}^d} \big(f_\eta(a_{ij}u_j) - f_\eta(B_{ij}^\eta * u_j)\big)^2 u_i \mathrm{d}x\mathrm{d}t \\ &\le \eta^{-2\alpha}\sum_{j=1}^n \|u_i\|_{L^\infty(0,T;L^\infty(\mathbb{R}^d))}\|a_{ij}u_j - B_{ij}^\eta * u_j\|^2_{L^2(0,T;L^2(\mathbb{R}^d))} \le C(n)\eta^{2(1-\alpha)}. \end{aligned}$$

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

$$\begin{aligned} D_{23} &\le C(\eta^*)\sum_{j=1}^n \|\nabla(B_{ij}^\eta * u_j)\|_{L^\infty(0,T;L^\infty(\mathbb{R}^d)}\int_0^T \mathbb{E}\big(|\widehat{X}_i(t) - \overline{X}_i^\eta(t)|^2\big)\mathrm{d}t \\ &\le C\sum_{j=1}^n \|\nabla u_j\|_{L^\infty(0,T;L^\infty(\mathbb{R}^d)}\int_0^T \mathbb{E}\big(|\widehat{X}_i(t) - \overline{X}_i^\eta(t)|^2\big)\mathrm{d}t \end{aligned}$$

$$\leq C\int_0^T \mathbb{E}\Big(\sup_{0<s<t}|D_i^\eta(s)|^2\Big)\mathrm{d}t.$$

Finally, it follows from the error estimate for $u-u_\eta$ from Theorem 3 that

$$\begin{aligned}
D_{24} &\leq C\sum_{j=1}^n\int_0^T\int_{\mathbb{R}^d}|B_{ij}^\eta * u_j - B_{ij}^\eta * u_{\eta,j}|^2 u_{\eta,i}\mathrm{d}x\mathrm{d}t \\
&\leq C\sum_{j=1}^n \|u_{\eta,i}\|_{L^\infty(0,T;L^\infty(\mathbb{R}^d))}\int_0^T \|B_{ij}^\eta\|_{L^1(\mathbb{R}^d)}^2\|u_j-u_{\eta,j}\|_{L^2(\mathbb{R}^d)}^2\mathrm{d}t \\
&\leq C(T)\eta^2.
\end{aligned}$$

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

$$\mathbb{E}\Big(\sup_{0<s<T}|D_2(s)|^2\Big)\leq C(T,n)\eta^{2(1-\alpha)} + C\int_0^T\mathbb{E}\Big(\sup_{0<s<t}|D_i^\eta(s)|^2\Big)\mathrm{d}t.$$

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

$$\mathbb{E}\Big(\sup_{0<s<T}|D_i^\eta(s)|^2\Big)\leq C(T,n)\eta^{2(1-\alpha)} + C\int_0^T\mathbb{E}\Big(\sup_{0<s<t}|D_i^\eta(s)|^2\Big)\mathrm{d}t.$$

The proof is finished after applying Gronwall's inequality and summing over $i=1,\ldots,n$. □

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

$$\begin{aligned}
&\sup_{k=1,\ldots,N}\mathbb{E}\bigg(\sum_{i=1}^n\sup_{0<s<t}\big|X_{\eta,i}^{k,N}(s)-\widehat{X}_i^k(s)\big|^2\bigg) \\
&\qquad\leq \sup_{k=1,\ldots,N}\mathbb{E}\bigg(\sum_{i=1}^n\sup_{0<s<t}\big|X_{\eta,i}^{k,N}(s)-\overline{X}_{\eta,i}^k(s)\big|^2\bigg) \\
&\qquad\quad + \sup_{k=1,\ldots,N}\mathbb{E}\bigg(\sum_{i=1}^n\sup_{0<s<t}\big|\overline{X}_{\eta,i}^k(s)-\widehat{X}_i^k(s)\big|^2\bigg) \\
&\qquad\leq C_1N^{-1+C_2\delta}+C_3\eta^{2(1-\alpha)}.
\end{aligned}$$

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

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

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). 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 $\triangle t_m = t_{m+1} - t_m$. We approximate $X_{k,i}^{N,\eta}(t_m)$ by $x_m^{k,i}$ and $Y_{k,i}^{N,\eta}(t_m)$ by $y_m^{k,i}$, defined by, respectively,

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

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

with initial conditions $x_0^{i,k} = \xi_i^k$ and $y_0^{i,k} = \xi_i^k$, where ξ_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., [17, 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 $|x| \le 1$ and $B(x) = 0$ else. Then $B_{ij}^\eta(x) = \eta^{-1}B(x/\eta)$. The numerical parameters are $\triangle t = 1/100$, $\eta = 2$, $N = 5000$ particles, $n_{\rm sim} = 500$ 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 Figure 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 Figure 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 [1]. 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

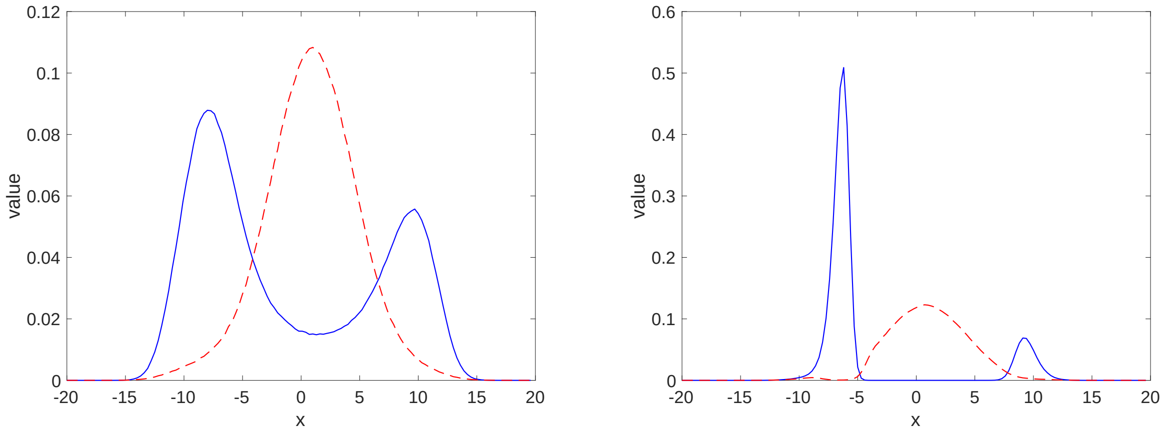

FIGURE 1. Nonsymmetric case: Densities of particle system (5) corresponding to the SKT population model (left) and particle system (3) (right) at time $t = 2$. Solid blue line: species 1; Dashed red line: species 2.

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 Figure 2. In both models, the species have the tendency to segregate. Because of the symmetric diffusion matrix, we observe only one cluster for each species. 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 behaviour 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 $t = 2$ are shown in Figure 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 posseses a vector-valued gradient-flow structure.

APPENDIX A. AUXILIARY RESULTS

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

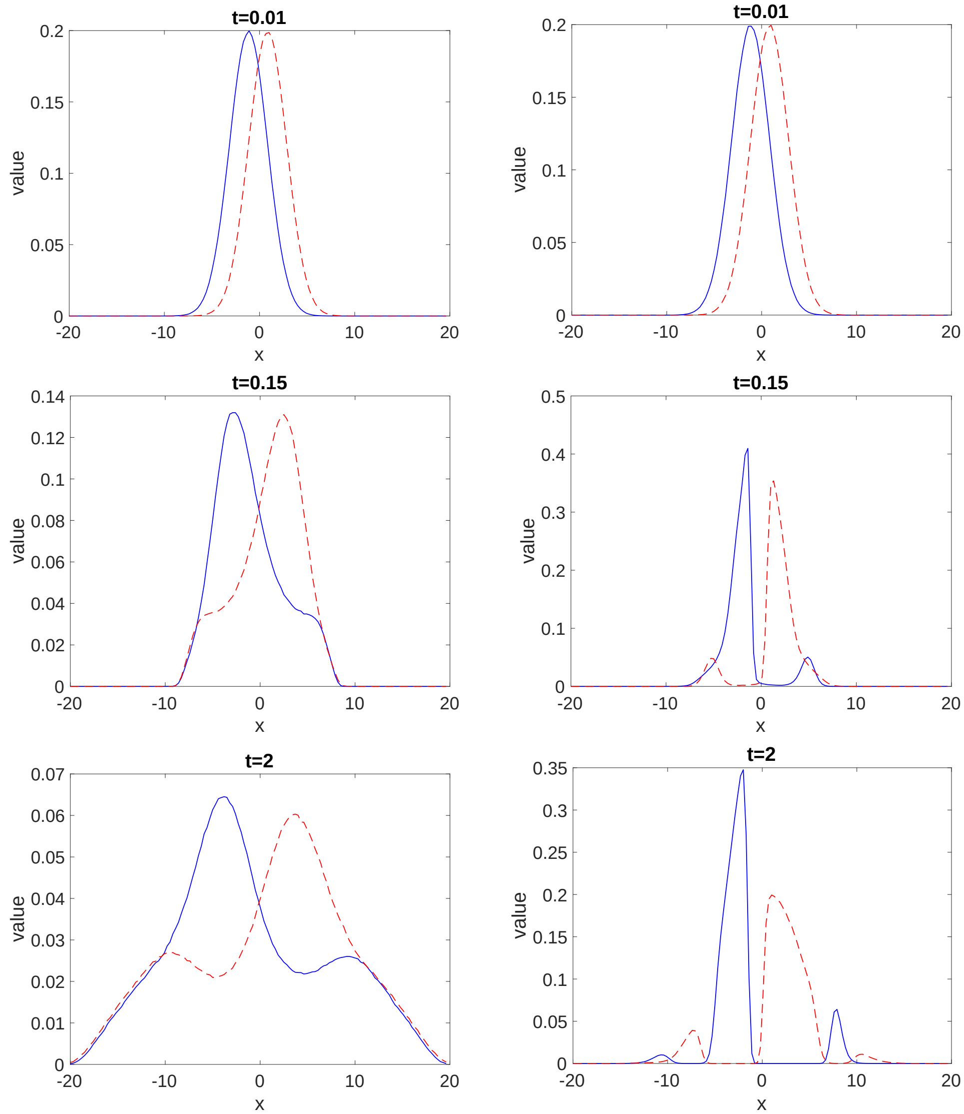

FIGURE 2. Symmetric case: Densities of particle system (5) corresponding to the SKT population model (left) and particle system (3) (right) for different times $t = 0.01,\ 0.15,\ 2$. Solid blue line: species 1; Dashed red line: species 2.

Lemma 11 (Young's convolution inequality, [19, Formula (7), page 107])**.** *Let* $1 \leq p, q, r \leq \infty$ *be such that* $1/p + 1/q = 1 + 1/r$ *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*

$$\|f * g\|_{L^r(\mathbb{R}^d)} \leq \|f\|_{L^p(\mathbb{R}^d)} \|g\|_{L^q(\mathbb{R}^d)}.$$

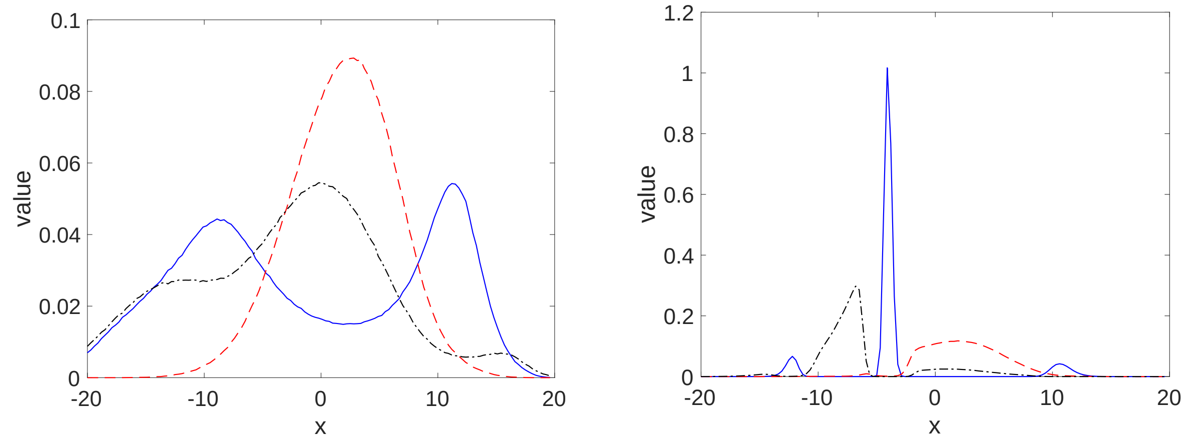

FIGURE 3. Three-species case: Densities of particle system (5) corresponding to the SKT population model (left) and particle system (3) (right) at time $t = 2$. Solid blue line: species 1; dashed red line: species 2; dash-dotted black line: species 3.

Lemma 12 (Moser-type estimate I, [20, 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, $g \in H^s(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$,*

$$\|D^\alpha(fg)\|_{L^2(\mathbb{R}^d)} \le C\big(\|f\|_{L^\infty(\mathbb{R}^d)}\|D^s g\|_{L^2(\mathbb{R}^d)} + \|D^s f\|_{L^2(\mathbb{R}^d)}\|g\|_{L^\infty(\mathbb{R}^d)}\big).$$

Lemma 13 (Moser-type estimate II, [20, 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 $u \in H^s(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$,*

$$\|D^\alpha g(u)\|_{L^2(\mathbb{R}^d)} \le C\|g'\|_{C^{s-1}(\mathbb{R})}\|u\|^{s-1}_{L^\infty(\mathbb{R}^d)}\|D^\alpha u\|_{L^2(\mathbb{R}^d)}.$$

Lemma 14 (Moser-type commutator inequality, [20, 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 $g \in H^{s-1}(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$,*

$$\|D^\alpha(fg) - fD^\alpha(g)\|_{L^2(\mathbb{R}^d)} \le C\big(\|Df\|_{L^\infty(\mathbb{R}^d)}\|D^{s-1}g\|_{L^2(\mathbb{R}^d)} + \|D^s f\|_{L^2(\mathbb{R}^d)}\|g\|_{L^\infty(\mathbb{R}^d)}\big),$$

where $D^s = \sum_{|\alpha|=s} D^\alpha$.

References

[1] M. Bertsch, M. Gurtin, D. Hilhorst, and L. Peletier. On interacting populations that disperse to avoid crowding: preservation of segregation. *J. Math. Biol.* 23 (1985), 1–13.
[2] M. Burger, J. A. Carrillo, J.-F. Pietschmann, and M. Schmidtchen. Segregation effects and gap formation in cross-diffusion models. *Interfaces Free Bound.* 22 (2020), 175–203.
[3] M. Burger, J.-F. Pietschmann, H. Ranetbauer, C. Schmeiser, and M.-T. Wolfram. Mean-field models for segregation dynamics. To appear in *Europ. J. Appl. Math.*, 2020. arXiv:1808.04069.
[4] L. Chen, E. S. Daus, and A. Jüngel. Rigorous mean-field limits and cross diffusion. *Z. Angew. Math. Phys.* 70 (2019), no. 122, 21 pages.

[5] L. Chen, S. Göttlich, and S. Knapp. Modeling of a diffusion with aggregation: rigorous derivation and numerical simulation. *ESAIM: Math. Mod. Num. Anal.* 53 (2018), 567–593.

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

[7] E. S. Daus, L. Desvillettes, and H. Dietert. About the entropic structure of detailed balanced multi-species cross-diffusion equations. *J. Diff. Eqs.* 266 (2019), 3861–3882.

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

[9] L. Desvillettes, T. Lepoutre, A. Moussa, and A. Trescases. On the entropic structure of reaction-cross-diffusion systems. *Commun. Partial Diff. Eqs.* 40 (2015), 1705–1747.

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

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

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

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

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

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

[16] I. Karatzas and S. Shreve. *Brownian Motion and Stochastic Calculus*. Second edition. Springer, New York, 1991.

[17] P. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*. Springer, Berlin, 1992.

[18] T. Lepoutre and A. Moussa. Entropic structure and duality for multiple species cross-diffusion systems. *Nonlin. Anal.* 159 (2017), 298–315.

[19] E. H. Lieb and M. Loss. *Analysis*. Second edition. Amer. Math. Soc., Providence, 2001.

[20] A. Majda. *Compressible Fluid Flow and Systems of Conservation Laws in Several Space Variables*. Springer, New York, 1984.

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

[22] D. Nualart. *The Malliavin Calculus and Related Topics*. Springer, Berlin, 2006.

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

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

[25] K. Oelschläger. Large systems of interacting particles and the porous medium equation. *J. Diff. Eqs.* 88 (1990), 294–346.

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

[27] N. Shigesada, K. Kawasaki, and E. Teramoto. Spatial segregation of interacting species. *J. Theor. Biol.* 79 (1979), 83–99.

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

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

[30] A. S. Sznitman. Topics in propagation of chaos. In: P. L. Hennequin (ed.), *École d'Été de Probabilités de Saint-Flour* XIX-1989, Lect. Notes Math. 1464. Berlin, Springer, 1991.

University of Mannheim, School of Business Informatics and Mathematics, 68131 Mannheim, Germany
E-mail address: chen@math.uni-mannheim.de

Institute of Analysis and Scientific Computing, TU Wien, Wiedner Hauptstrasse 8–10, 1040 Wien, Austria
E-mail address: esther.daus@tuwien.ac.at

Institute of Analysis and Scientific Computing, TU Wien, Wiedner Hauptstrasse 8–10, 1040 Wien, Austria
E-mail address: alexandra.holzinger@tuwien.ac.at

Institute of Analysis and Scientific Computing, TU Wien, Wiedner Hauptstrasse 8–10, 1040 Wien, Austria
E-mail address: juengel@tuwien.ac.at