# NONLOCAL CROSS-DIFFUSION SYSTEMS FOR MULTI-SPECIES POPULATIONS AND NETWORKS

ANSGAR JÜNGEL, STEFAN PORTISCH, AND ANTOINE ZUREK

Abstract. Nonlocal cross-diffusion systems on the torus, arising in population dynamics and neural networks, are analyzed. The global existence of weak solutions, the weak-strong uniqueness, and the localization limit are proved. The kernels are assumed to be positive definite and in detailed balance. The proofs are based on entropy estimates coming from Shannon-type and Rao-type entropies, while the weak-strong uniqueness result follows from the relative entropy method. The existence and uniqueness theorems hold for nondifferentiable kernels. The associated local cross-diffusion system is also discussed.

## 1. Introduction

The segregation and migration in populations [31] and the dynamics of multilayer neural networks [29] can be modeled by stochastic interacting particle systems. Since the numerical approximation of such systems is very time-consuming, it is reasonable to investigate simpler macroscopic models. They can be derived from the many-particle systems in the mean-field limit, typically leading to nonlocal diffusion equations [25]. When the model involves several species, the mean-field system includes nonlocal cross-diffusion terms [12]. In this paper, we analyze nonlocal cross-diffusion systems that are motivated from [12]. We prove the global existence of weak solutions, show a weak-strong uniqueness result, and perform the localization limit.

The nonlocal cross-diffusion equations are given by

$$\partial_t u_i - \sigma\Delta u_i = \operatorname{div}(u_i\nabla p_i[u]), \quad t>0, \quad u_i(0)=u_i^0 \quad \text{in } \mathbb{T}^d,\ i=1,\ldots,n, \tag{1}$$

where $\sigma > 0$ is the diffusion coefficient, $\mathbb{T}^d$ is the $d$-dimensional torus ($d \geq 1$), $p_i[u]$ describes the multi-species connectivity of the network, given by the nonlocal operator

$$p_i[u](x) = \sum_{j=1}^n \int_{\mathbb{T}^d} K_{ij}(x-y)u_j(y)dy, \quad i=1,\ldots,n, \tag{2}$$

*Date*: April 13, 2021.

2000 *Mathematics Subject Classification.* 35K40, 35K55, 35Q92, 68T07, 92B20.

*Key words and phrases.* Cross diffusion, neural network dynamics, entropy method, localization limit, global existence of solutions, weak-strong uniqueness.

The authors have been partially supported by the Austrian Science Fund (FWF), grants P30000, P33010, F65, and W1245, and by the multilateral project of the Austrian Agency for International Cooperation in Education and Research (OeAD), grant MULT 11/2020.

$K_{ij}:\mathbb{T}^d\to\mathbb{R}$ are the kernel functions (extended periodically to $\mathbb{R}^d$), and $u=(u_1,\ldots,u_n)$ is the solution vector. In neural network theory, the kernels describe the weight of a connection between the node $x$ associated to species $i$ and node $y$ associated to species $j$. In population dynamics, $K_{ij}=\nabla V_{ij}$ is the force field associated to the interaction potential $V_{ij}$ between species $i$ and species $j$ [12]. The torus is chosen to simplify the problem; we may consider (1) also on the whole space. In population dynamics, $\mathbb{T}^d$ is the "physical" space, while it is the space of predictor variables in network theory.

Model (1)–(2) was derived in [12] from an interacting particle system. If the kernels converge, up to the factor $a_{ij}$, to the delta distribution, the nonlocal system converges to a local system, consisting of (1) with the local operator

$$p_i[u]=\sum_{j=1}^n a_{ij}u_j,\quad i=1,\ldots,n. \tag{3}$$

The localization limit was proved in [12] but for smooth solutions and in the whole space; we show it for weak solutions.

Most nonlocal models studied in the literature describe a single species. A simple example is $\partial_t u=\operatorname{div}(uv)$ with $v=\nabla(K*u)$. This corresponds to the mass continuity equation for the density $u$ with a nonlocal velocity $v$. An $L^p$ theory for this equation was provided in [6], while the Wasserstein gradient-flow structure was explored in [11]. In the neural network context, the equation can be seen as the mean-field limit of infinitely many hidden network units [29, 34].

In the multiple-species case, the mean-field limit performed in [3, 23] leads to the McKean–Vlasov–Focker–Planck equations for the probability distributions $u_i$,

$$\partial_t u_i=\sigma\Delta u_i+\operatorname{div}\sum_{j=1}^n\int_{\mathbb{T}^d}M_{ij}(x,y)u_i(x)u_j(y)dxdy,\quad i=1,\ldots,n,$$

which correspond to (1)–(2) if $M_{ij}(x,y)=\nabla K_{ij}(x-y)$. For two species and symmetrizable cross-interaction potentials (i.e. $K_{12}=\alpha K_{21}$ for some $\alpha>0$), a complete existence and uniqueness theory for measure solutions to (1)–(2) in the whole space with smooth convolution kernels was established in [18] using the Wasserstein gradient-flow theory. A nonlocal system with size exclusion was analyzed in [4], using entropy methods. In [22], a nonlocal version of the Shigesada–Kawasaki–Teramoto cross-diffusion system was analyzed, using finite differences instead of a gradient. All these works are concerned with two-species models. Compared to previous results, we allow for an arbitrary number of species and nondifferentiable kernel functions.

The existence of a unique global smooth solution to (1)–(2) in $\mathbb{R}^d$ was proved in [12], but only for small initial data. Existence and localization results for global bounded solutions to a related model, but with the same kernel for all species, were recently shown in [17]. In this paper, we study system (1)–(2) for weak solutions, allowing for large initial data and revealing a (new) double entropy structure, and prove a new weak-strong uniqueness result.

The mathematical difficulties are the cross-diffusion terms and the nonlocality, which exclude the application of standard techniques like maximum principles and regularity theory. For instance, it is well known that nonlocal diffusion operators generally do not possess regularizing effects on the solution [2]. The key of our analysis is the observation that the nonlocal system possesses, like the associated local one, *two* entropies, namely the Shannon-type entropy $H_1$ [33] and the Rao-type entropy $H_2$ [32],

$$H_1(u)=\sum_{i=1}^n\int_{\mathbb{T}^d}\pi_iu_i(\log u_i-1)dx,\quad H_2(u)=\frac12\sum_{i,j=1}^n\int_{\mathbb{T}^d}\int_{\mathbb{T}^d}\pi_iK_{ij}(x-y)u_i(x)u_j(y)dxdy,$$

if the kernels are in detailed balance and positive definite in the sense specified below and for some numbers $\pi_i>0$. A formal computation that is made rigorous below shows that the following entropy inequalities hold:

$$\frac{dH_1}{dt}+4\sigma\sum_{i=1}^n\int_{\mathbb{T}^d}\pi_i|\nabla\sqrt{u_i}|^2dx=-\sum_{i,j=1}^n\int_{\mathbb{T}^d}\int_{\mathbb{T}^d}\pi_iK_{ij}(x-y)\nabla u_i(x)\cdot\nabla u_j(y)dxdy, \tag{4}$$

$$\frac{dH_2}{dt}+\sum_{i=1}^n\int_{\mathbb{T}^d}\pi_iu_i|\nabla p_i[u]|^2dx=-\sigma\sum_{i,j=1}^n\int_{\mathbb{T}^d}\int_{\mathbb{T}^d}\pi_iK_{ij}(x-y)\nabla u_i(x)\cdot\nabla u_j(y)dxdy. \tag{5}$$

These computations are valid if $K_{ij}$ is in detailed balance, which means that there exist $\pi_1,\ldots,\pi_n>0$ such that

$$\pi_iK_{ij}(x-y)=\pi_jK_{ji}(y-x)\quad\text{for all } i,j=1,\ldots,n,\ x,y\in\mathbb{T}^d. \tag{6}$$

We recognize these identities as a generalized detailed-balance condition for the Markov chain associated to $(K_{ij}(x-y))$ (for fixed $x-y$), and in this case, $(\pi_1,\ldots,\pi_n)$ is the corresponding reversible measure. The functionals $H_1$ and $H_2$ are Lyapunov functionals if $(\pi_iK_{ij})$ is positive definite in the sense

$$\sum_{i,j=1}^n\int_{\mathbb{T}^d}\int_{\mathbb{T}^d}\pi_iK_{ij}(x-y)v_i(x)v_j(y)dxdy\geq 0\quad\text{for all } v_i,v_j\in L^2(\mathbb{T}^d). \tag{7}$$

This condition generalizes the usual definition of positive definite kernels to the multi-species case [10]. Examples of kernels that satisfy (6) and (7) are given in Remark 1. Because of the nonlocality, we cannot conclude $L^2(\mathbb{T}^d)$ estimates for $u_i$ and $\nabla u_i$ like in the local case; see [28] and Appendix B. We deduce from (4) only bounds for $u_i\log u_i$ in $L^1(\mathbb{T}^d)$ and $\sqrt{u_i}$ in $H^1(\mathbb{T}^d)$.

These bounds are not sufficient to pass to the limit in the approximate problem. In particular, we cannot identify the limit of the product $u_i\nabla p_i[u]$, since $u_i$ and $\nabla p_i[u]$ are elements in larger spaces than $L^2(\mathbb{T}^d)$. We solve this issue by exploiting the uniform $L^2(\mathbb{T}^d)$ bound for $\sqrt{u_i}\nabla p_i[u]$ from (5) and prove a "compensated compactness" lemma (see Lemma 9 in Appendix A): If $u_\varepsilon\to u$ strongly in $L^p(\mathbb{T}^d)$, $v_\varepsilon\rightharpoonup v$ weakly in $L^p(\mathbb{T}^d)$, and $u_\varepsilon v_\varepsilon\rightharpoonup w$ weakly in $L^p(\mathbb{T}^d)$ for some $1<p<2$ then $uv=w$. The estimates from (4)–(5) are the key for the proof of the global existence of weak solutions to (1)–(2).

As a second result, we prove the weak-strong uniqueness of solutions, i.e., if $u$ is a weak solution to (1)–(2) satisfying $u_i \in L^2(0,T;H^1(\mathbb{T}^d))$ and if $v$ is a "strong" solution to this problem with the same initial data, then $u(t) = v(t)$ for a.e. $t \geq 0$. The proof uses the relative entropy

$$H(u|v) = \sum_{i=1}^n \int_{\mathbb{T}^d} \pi_i \big( u_i(\log u_i - 1) - u_i \log v_i + v_i \big) dx,$$

a variant of which was used in [21] for reaction-diffusion systems in the context of renormalized solutions and later extended to Shigesada–Kawasaki–Teramoto systems [14]. The recent work [24] generalizes this approach to more general Shigesada–Kawasaki–Teramoto as well as energy-reaction-diffusion systems. Originally, the relative entropy method was devised for conservation laws to estimate the $L^2$ distance between two solutions [16, 19]. Up to our knowledge, we apply this techniques for the first time to nonlocal cross-diffusion systems. The idea is to differentiate $H(u|v)$ and to derive the inequality

$$H(u(t)|v(t)) \leq C \sum_{i=1}^n \int_0^t \|u_i - v_i\|_{L^1(\mathbb{T}^d)}^2 ds \quad \text{for } t > 0.$$

The Csiszár–Kullback–Pinsker inequality [27, Theorem A.2] allows us to estimate the relative entropy from below by $\|u_i(t) - v_i(t)\|_{L^1(\mathbb{T}^d)}^2$, up to some factor. Then Gronwall's lemma implies that $u_i(t) = v_i(t)$ for a.e. $t > 0$. The application of this inequality is different from the proof in [14, 21, 24], where the relative entropy is estimated from below by $|u_i - v_i|^2$ on the set $\{u_i \leq K\}$. The difference originates from the nonlocal terms. Indeed, if $K_{ij}$ is bounded,

$$\sum_{i,j=1}^n \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \pi_i K_{ij}(x-y)(u_i - v_i)(x)(u_j - v_j)(y) dx dy \leq C \sum_{i=1}^n \bigg( \int_{\mathbb{T}^d} |u_i - v_i| dx \bigg)^2,$$

leading to an estimate in $L^1(\mathbb{T}^d)$. In the local case, the associated estimate yields an $L^2(\mathbb{T}^d)$ estimate:

$$\sum_{i,j=1}^n \int_{\mathbb{T}^d} \pi_i a_{ij} (u_i - v_i)(x)(u_j - v_j)(x) dx \leq C \sum_{i=1}^n \int_{\mathbb{T}^d} |u_i - v_i|^2 dx.$$

As the densities $u_i$ may be only nonnegative, we cannot use $\pi_i \log u_i$ as a test function in (1) to compute $dH(u|v)/dt$. This issue is overcome by regularizing the entropy by using $\log(u_i + \varepsilon)$ for some $\varepsilon > 0$ as a test funtion and then to pass to the limit $\varepsilon \to 0$.

We observe that the uniqueness of weak solutions to cross-diffusion systems is a delicate task, and there are only few results in the literature. Most of the results are based on the fact that the total density $\sum_{i=1}^n u_i$ satisfies a simpler equation for which uniqueness can be shown; see [4, 13]. A duality method for a nonlocal version of the Shigesada–Kawasaki–Teramoto system was used in [22]. In [5], a weak-strong uniqueness result on a cross-diffusion system, based on $L^2$ estimates, was shown.

The bounds obtained in the proof of our existence result are independent of the kernels, such that we can perform the localization limit, our third main result. For this, we assume that $K_{ij} = B_{ij}^\varepsilon \to a_{ij}\delta_0$ as $\varepsilon \to 0$ in the sense of distributions, where $\delta_0$ is the Dirac delta distribution. Then, if $u^\varepsilon$ is a weak solution to (1)–(2), we prove that $u_i^\varepsilon \to u_i$ strongly in $L^1(\mathbb{T}^d \times (0,T))$, and the limit $u$ solves the local system (1) and (3). As a by-product, we obtain the global existence of weak solutions to this problem; see Appendix B for the precise statement.

We summarize our main results:

- global existence of weak solutions to the nonlocal system (1)–(2) for nondifferentiable positive definite kernels in detailed balance;
- weak-strong uniqueness of solutions $u \in L^2(0,T;\, H^1(\mathbb{T}^d;\mathbb{R}^n))$ to the nonlocal system;
- localization limit to the local system (1) and (3).

The paper is organized as follows. Our hypotheses and main results are made precise in Section 2. The global existence of weak solutions to the nonlocal system and some regularity results are proved in Section 3. The weak-strong uniqueness result is shown in Section 4. In Section 5, the localization limit, based on the a priori estimates of Section 3, is performed. Finally, we collect some auxiliary lemmas in Appendix A and state a global existence result for the local system (1) and (3) in Appendix B.

## 2. Main results

We collect the main theorems which are proved in the subsequent sections. We impose the following hypotheses:

(H1) Data: Let $d \geq 1$, $T > 0$, $\sigma > 0$, and $u^0 \in L^2(\mathbb{T}^d)$ satisfies $u_i^0 \geq 0$ in $\mathbb{T}^d$, $i = 1,\ldots,n$.

(H2) Regularity: $K_{ij} \in L^s(\mathbb{T}^d)$ for $i,j = 1,\ldots,n$, where $s = d/2$ if $d > 2$, $s > 1$ if $d = 2$, and $s = 1$ if $d = 1$.

(H3) Detailed balance: There exist $\pi_1,\ldots,\pi_n > 0$ such that $\pi_i K_{ij}(x-y) = \pi_j K_{ji}(y-x)$ for all $i,j = 1,\ldots,n$, $x,y \in \mathbb{T}^d$.

(H4) Positive definiteness: For all $v_1,\ldots,v_n \in L^2(\mathbb{T}^d)$, it holds that

$$\sum_{i,j=1}^n \int_{\mathbb{T}^d}\int_{\mathbb{T}^d} \pi_i K_{ij}(x-y)v_i(x)v_j(y)dxdy \geq 0.$$

We need the same diffusivity $\sigma$ for all species, since otherwise we cannot prove that the Rao-type functional $H_2$ is a Lyapunov functional. The reason is the mixing of the species in the definition of $H_2$.

**Remark 1** (Kernels satisfying Hypotheses (H2)–(H4))**.** Kernels satisfying Hypothesis (H4) with $n = 1$ can be characterized by Mercer's theorem [10, 30].

An example that satisfies Hypotheses (H2)–(H4) is given by the Gaussian kernel $B(|x-y|) = (2\pi)^{-d/2}\exp(-|x-y|^2/2)$. We define for $i,j = 1,\ldots,n$ and $x,y \in \mathbb{R}^d$,

$$K_{ij}(x-y) = B_{ij}^\varepsilon(x-y) := \frac{a_{ij}}{(2\pi\varepsilon^2)^{d/2}}\exp\bigg(-\frac{|x-y|^2}{2\varepsilon^2}\bigg),$$

where $\varepsilon > 0$ and $a_{ij} \geq 0$ are such that the matrix $(\pi_i a_{ij})$ is symmetric and positive definite for some $\pi_i > 0$. Thus, Hypothesis (H3) holds. Hypothesis (H4) can be verified as follows. The identity

$$\frac{e^{-|x-y|^2/(2\varepsilon^2)}}{(2\pi\varepsilon^2)^{d/2}} = \int_{\mathbb{R}^d} \frac{e^{-|x-z|^2/\varepsilon^2}}{(\pi\varepsilon^2)^{d/2}} \frac{e^{-|y-z|^2/\varepsilon^2}}{(\pi\varepsilon^2)^{d/2}} dz,$$

shows that

$$\begin{aligned}
\sum_{i,j=1}^n \int_{\mathbb{T}^d}\int_{\mathbb{T}^d} \pi_i K_{ij}(x-y)v_i(x)v_j(y)dxdy &= \sum_{i,j=1}^n \int_{\mathbb{T}^d}\int_{\mathbb{T}^d} \pi_i a_{ij} \frac{e^{-|x-y|^2/(2\varepsilon^2)}}{(2\pi\varepsilon^2)^{d/2}} v_i(x)v_j(y)dxdy \\
&= \sum_{i,j=1}^n \pi_i a_{ij} \int_{\mathbb{R}^d}\int_{\mathbb{T}^d} \bigg(\frac{e^{-|x-z|^2/\varepsilon^2}}{(\pi\varepsilon^2)^{d/2}} v_i(x)\bigg)dx \int_{\mathbb{T}^d} \bigg(\frac{e^{-|y-z|^2/\varepsilon^2}}{(\pi\varepsilon^2)^{d/2}} v_j(y)\bigg)dydz \\
&\geq \frac{\alpha}{(\pi\varepsilon^2)^d} \sum_{i=1}^n \int_{\mathbb{R}^d} \bigg(\int_{\mathbb{T}^d} e^{-|x-z|^2/\varepsilon^2} v_i(x)dx\bigg)^2 dz \geq 0,
\end{aligned}$$

where $\alpha > 0$ is the smallest eigenvalue of $(\pi_i a_{ij})$. This proves the positive definiteness of $K_{ij}$. Note that $B_{ij}^\varepsilon \to a_{ij}\delta_0$ as $\varepsilon \to 0$ in the sense of distributions.

We can construct further examples from the Gaussian kernel. For instance,

$$K_{ij}(x-y) = \frac{a_{ij}}{1+|x-y|^2}, \quad i,j = 1,\ldots,n, \ x,y \in \mathbb{R}^d,$$

satisfies Hypothesis (H4), since

$$\frac{1}{1+|x-y|^2} = \int_0^\infty e^{-s(1+|x-y|^2)}ds,$$

and $B(x-y) = \exp(-s(1+|x-y|^2))$ is positive definite. □

First, we show the global existence of weak solutions. Let $Q_T = \mathbb{T}^d \times (0,T)$.

**Theorem 2** (Global existence)**.** *Let Hypotheses (H1)–(H4) hold. Then there exists a global weak solution $u = (u_1, \ldots, u_n)$ to (1)–(2) satisfying $u_i \geq 0$ in $Q_T$ and*

$$\begin{aligned}
&u_i^{1/2} \in L^2(0,T;H^1(\mathbb{T}^d)), \quad u_i \in L^{1+2/d}(Q_T) \cap L^q(0,T;W^{1,q}(\mathbb{T}^d)), \\
&\partial_t u_i \in L^q(0,T;W^{-1,q}(\mathbb{T}^d)), \quad u_i \nabla p_i[u] \in L^q(Q_T),
\end{aligned} \tag{8}$$

*where $q = (d+2)/(d+1)$ and $i = 1,\ldots,n$. The initial datum in (1) is satisfied in the sense of $W^{-1,q}(\mathbb{T}^d) := W^{1,d+2}(\mathbb{T}^d)'$. Moreover, the following entropy inequalities hold:*

$$H_1(u(t)) + 4\sigma \sum_{i=1}^n \int_0^t \int_{\mathbb{T}^d} \pi_i |\nabla u_i^{1/2}|^2 dxds \leq H_1(u^0), \tag{9}$$

$$H_2(u(t)) + \sum_{i=1}^n \int_0^t \int_{\mathbb{T}^d} \pi_i u_i |\nabla p_i[u]|^2 dxds \leq H_2(u^0). \tag{10}$$

Imposing more regularity on the kernel functions, we can derive $H^1(\mathbb{T}^d)$ regularity for $u_i$, which is needed for the weak-strong uniqueness result.

**Proposition 3** (Regularity). *Let Hypotheses (H1)–(H4) hold and let $\nabla K_{ij} \in L^{d+2}(\mathbb{T}^d)$ for $i,j = 1,\ldots,n$. Then there exists a weak solution $u = (u_1,\ldots,u_n)$ to (1)–(2) satisfying $u_i \geq 0$ in $\mathbb{T}^d$ and*

$$u_i \in L^2(0,T;H^1(\mathbb{T}^d)), \quad \partial_t u_i \in L^2(0,T;H^{-1}(\mathbb{T}^d)), \quad \nabla p_i[u] \in L^\infty(0,T;L^\infty(\mathbb{T}^d)).$$

*Moreover, if additionally $\nabla K_{ij}$, $\Delta K_{ij} \in L^\infty(\mathbb{T}^d)$ and $m_0 \leq u_i^0 \leq M_0$ in $\mathbb{T}^d$ then $0 < m_0 e^{-\lambda t} \leq u_i(t) \leq M_0 e^{\lambda t}$ in $\mathbb{T}^d$ for $t > 0$, where $\lambda > 0$ depends on $\Delta K_{ij}$ and $u^0$.*

The proof of the $H^1(\mathbb{T}^d)$ regularity is based on standard $L^2$ estimates if $\nabla K_{ij} \in L^\infty(\mathbb{T}^d)$. The difficulty is the reduced regularity $\nabla K_{ij} \in L^{d+2}(\mathbb{T}^d)$ which requires some care. Indeed, the test function $u_i$ in the weak formulation of (1) leads to a cubic term, which is reduced to a subquadratic term for $\nabla u_i$ by combining the Gagliardo–Nirenberg inequality and the uniform $L^1(\mathbb{T}^d)$ bound for $u_i$.

Similar lower and upper bounds as in Proposition 3 were obtained in [17] with a different proof. Since the $L^\infty$ bounds depend on the derivatives of $K_{ij}$, they do not carry over in the localization limit to the local system. In fact, it is an open problem whether the local system (1) and (3) possesses *bounded* weak solutions. The proposition also holds for kernel functions $K_{ij}(x,y)$ that are used in neural network theory; see Remark 6.

**Theorem 4** (Weak-strong uniqueness). *Let $K_{ij} \in L^\infty(\mathbb{R}^d)$ for $i,j = 1,\ldots,n$. Let $u$ be a nonnegative weak solution to (1)–(2) satisfying (8) as well as $u_i \in L^2(0,T;H^1(\mathbb{T}^d)) \cap H^1(0,T;H^{-1}(\mathbb{T}^d))$, and let $v$ be a "strong" solution to (1)–(2) satisfying*

$$c \leq v_i \leq C \quad \text{in } Q_T, \quad \partial_t v_i \in L^2(0,T;H^{-1}(\mathbb{T}^d)), \quad v_i \in L^\infty(0,T;W^{1,\infty}(\mathbb{T}^d)),$$

*for some $C \geq c > 0$ and having the same initial data as $u$. Then $u(t) = v(t)$ in $\mathbb{T}^d$ for a.e. $t \in (0,T)$.*

For the localization limit, we choose the radial kernel

$$K_{ij}^\varepsilon(x-y) = \frac{a_{ij}}{\varepsilon^d} B\left(\frac{|x-y|}{\varepsilon}\right), \quad i,j = 1,\ldots,n, \ x,y \in \mathbb{T}^d, \tag{11}$$

where $B \in C^0(\mathbb{R})$, $\operatorname{supp}(B) \subset (-1,1)$, $\int_{\mathbb{R}} B(z)dz = 1$, and $a_{ij} \geq 0$ is such that $(\pi_i a_{ij})$ is symmetric and positive definite for some $\pi_i > 0$, $i = 1,\ldots,n$.

**Theorem 5** (Localization limit). *Let $K_{ij}^\varepsilon$ be given by (11) and satisfying Hypothesis (H4). Let $u^\varepsilon$ be the weak solution to (1)–(2), constructed in Theorem 2. Then there exists a subsequence of $(u^\varepsilon)$ that is not relabeled such that, as $\varepsilon \to 0$,*

$$u_\varepsilon \to u \quad \text{strongly in } L^2(0,T;L^{d/(d-1)}(\mathbb{T}^d)),$$

*if $d \geq 2$ and strongly in $L^2(0,T;L^r(\mathbb{T}^d))$ for any $r < \infty$ if $d = 1$. Moreover, $u$ is a nonnegative weak solution to (1) and (3).*

The existence of global weak solutions to (1) and (3) can also be proved for any bounded domain; see Appendix B.

## 3. Global existence for the nonlocal system

We prove the global existence of a nonnegative weak solution $u$ to (1)–(2) and show the regularity properties of Proposition 3. Since the proof is based on the entropy method similar to [27, Chapter 4], we sketch the standard arguments and focus on the derivation of uniform estimates.

*Step 1: Solution of an approximated system.* Let $T>0$, $N\in\mathbb{N}$, $\tau=T/N$, $\delta>0$, and $m\in\mathbb{N}$ with $m>d/2$. We proceed by induction over $k$. Let $u^{k-1}\in L^2(\mathbb{T}^d;\mathbb{R}^n)$ be given. Set $u_i(w)=\exp(w_i/\pi_i)>0$. We wish to find $w^k\in H^m(\mathbb{T}^d;\mathbb{R}^n)$ to the approximated system

$$\begin{aligned}\frac{1}{\tau}\int_{\mathbb{T}^d}(u(w^k)-u^{k-1})\cdot\phi dx+\sigma\sum_{i=1}^n\int_{\mathbb{T}^d}\nabla u_i(w^k)\cdot\nabla\phi_i dx+\delta b(w^k,\phi)\\ =-\sum_{i=1}^n\int_{\mathbb{T}^d}u_i(w^k)\nabla p_i[u(w^k)]\cdot\nabla\phi_i dx,\end{aligned}\tag{12}$$

for $\phi=(\phi_1,\ldots,\phi_n)\in H^m(\mathbb{T}^d;\mathbb{R}^n)$. The bilinear form

$$b(w^k,\phi)=\int_{\mathbb{T}^d}\bigg(\sum_{|\alpha|=m}D^\alpha w^k\cdot D^\alpha\phi+w\cdot\phi\bigg)dx,$$

is coercive on $H^m(\mathbb{T}^d;\mathbb{R}^n)$, i.e. $b(w^k,w^k)\ge C\|w^k\|^2_{H^m(\mathbb{T}^d)}$ for some $C>0$, as a consequence of the generalized Poincaré–Wirtinger inequality. By the fixed-point argument on the space $L^\infty(\mathbb{T}^d;\mathbb{R}^n)$ used in [27, Section 4.4], it is sufficient to derive a uniform bound for $w_i^k$ in $H^m(\mathbb{T}^d)$, which embeddes compactly into $L^\infty(\mathbb{T}^d)$. To this end, we use the test function $\phi_i=w_i^k=\pi_i\log u_i^k$ (with $u_i^k:=u_i(w^k)$) in (12):

$$\begin{aligned}\sum_{i=1}^n\frac{\pi_i}{\tau}\int_{\mathbb{T}^d}(u_i^k-u_i^{k-1})\cdot\log u_i^k dx+4\sigma\sum_{i=1}^n\pi_i\int_{\mathbb{T}^d}|\nabla(u_i^k)^{1/2}|^2dx+\delta b(w^k,w^k)\\ =-\sum_{i=1}^n\int_{\mathbb{T}^d}u_i^k\nabla p_i[u^k]\cdot\nabla w_i^k dx=-\sum_{i=1}^n\int_{\mathbb{T}^d}\pi_i\nabla p_i[u^k]\cdot\nabla u_i^k dx,\end{aligned}$$

where we used the identity $u_i^k\nabla w_i^k=\pi_i\nabla u_i^k$. An integration by parts gives

$$\int_{\mathbb{T}^d}\nabla K_{ij}(x-y)u_j^k(y)dy=\int_{\mathbb{T}^d}\nabla K_{ij}(z)u_j^k(x-z)dz=\int_{\mathbb{T}^d}K_{ij}(x-y)\nabla u_j^k(y)dy.$$

Thus, in view of definition (2) of $p_i[u^k]$ and Hypothesis (H4),

$$\begin{aligned}\sum_{i=1}^n\frac{\pi_i}{\tau}\int_{\mathbb{T}^d}(u_i^k-u_i^{k-1})\cdot\log u_i^k dx+4\sigma\sum_{i=1}^n\pi_i\int_{\mathbb{T}^d}|\nabla(u_i^k)^{1/2}|^2dx+\delta b(w^k,w^k)\\ =-\sum_{i,j=1}^n\int_{\mathbb{T}^d}\int_{\mathbb{T}^d}\pi_iK_{ij}(x-y)\nabla u_j^k(y)\cdot\nabla u_i^k(x)dxdy\le 0.\end{aligned}$$

Taking into account the convexity of $s \mapsto s\log s$ to estimate the first integral and the coercivity of $b(w^k, w^k)$ to estimate the third term, we find that

$$\begin{aligned}\frac{1}{\tau}\sum_{i=1}^n\int_{\mathbb{T}^d}\pi_i\big(u_i^k(\log u_i^k - 1) - u_i^{k-1}(\log u_i^{k-1}-1)\big)dx \\ + 4\sigma\sum_{i=1}^n \pi_i\|\nabla(u_i^k)^{1/2}\|_{L^2(\mathbb{T}^d)}^2 + \delta C\sum_{i=1}^n\|w_i^k\|_{H^m(\mathbb{T}^d)}^2 \le 0. \qquad (13)\end{aligned}$$

This provides a uniform estimate for $w^k$ in $H^m(\mathbb{T}^d) \hookrightarrow L^\infty(\mathbb{T}^d)$ (not uniform in $\delta$), necessary to conclude the fixed-point argument and giving the existence of a solution $w^k \in H^m(\mathbb{T}^d;\mathbb{R}^n)$ to (12). This defines $u^k := u(w^k)$, finishing the induction step.

To derive further uniform estimates, we wish to use $\psi_i = \pi_i p_i[u^k]$ as a test function in (12). However, we cannot estimate the term $\delta b(w^k,\psi)$ appropriately. Therefore, we perform the limits $\delta \to 0$ and $\tau \to 0$ separately.

*Step 2: Limit* $\delta \to 0$. Let $u^\delta = (u_1^\delta,\ldots,u_n^\delta)$ with $u_i^\delta = u_i(w^k)$ be a solution to (12) and let $w_i^\delta = \pi_i\log u_i^\delta$ for $i=1,\ldots,n$. Estimate (13) and the Poincaré–Wirtinger inequality show that $(u_i^\delta)^{1/2}$ is uniformly bounded in $H^1(\mathbb{T}^d)$ and, by Sobolev's embedding, in $L^{r_1}(\mathbb{T}^d)$, where $r_1 = 2d/(d-2)$ if $d>2$, $r_1<\infty$ if $d=2$, and $r_1=\infty$ if $d=1$. Therefore, $\nabla u_i^\delta = 2(u_i^\delta)^{1/2}\nabla(u_i^\delta)^{1/2}$ is uniformly bounded in $L^{r_2}(\mathbb{T}^d)$, where $r_2 = d/(d-1)$ if $d>2$, $r_2<2$ if $d=2$, and $r_2=2$ if $d=1$. By Sobolev's embedding, $(u_i^\delta)$ is relatively compact in $L^r(\mathbb{T}^d)$ for $r<r_1/2$, and there exists a subsequence that is not relabeled such that, as $\delta\to 0$,

$$\begin{aligned} u_i^\delta &\to u_i \quad \text{strongly in } L^r(\mathbb{T}^d), \quad r<r_1/2, \\ \nabla u_i^\delta &\rightharpoonup \nabla u_i \quad \text{weakly in } L^{r_2}(\mathbb{T}^d), \\ \delta w_i^\delta &\to 0 \quad \text{strongly in } H^m(\mathbb{T}^d). \end{aligned}$$

In particular, we have, up to a subsequence, $u_i^\delta \to u_i$ a.e. and $|u_i^\delta(y)| \le g(y) \in L^1(\mathbb{T}^d)$. By dominated convergence, $p_i[u^\delta]\to p_i[u]$ a.e. Young's convolution inequality (see Lemma 7 in Appendix A) shows that, for $d>2$,

$$\begin{aligned}\|p_i[u^\delta]\|_{L^\infty(\mathbb{T}^d)} &\le \sum_{j=1}^n\bigg\|\int_{\mathbb{T}^d}K_{ij}(\cdot - y)u_j^\delta(y)dy\bigg\|_{L^\infty(\mathbb{T}^d)} \\ &\le \sum_{j=1}^n\|K_{ij}\|_{L^{d/2}(\mathbb{T}^d)}\|u_j^\delta\|_{L^{d/(d-2)}(\mathbb{T}^d)} \le C,\end{aligned}$$

In a similar way, we can prove that $(p_i[u^\delta])$ is bounded in $L^r(\mathbb{T}^d)$ for any $r<\infty$ if $d=2$ and in $L^\infty(\mathbb{T}^d)$ if $d=1$, assuming that $K_{ij}\in L^1(\mathbb{T}^d)$. Lemma 8 in Appendix A implies that $p_i[u^\delta]\to p_i[u]$ strongly in $L^r(\mathbb{T}^d)$ for any $r<\infty$. Furthermore, if $d>2$,

$$\|\nabla p_i[u^\delta]\|_{L^{r_3}(\mathbb{T}^d)} \le \sum_{j=1}^n\|K_{ij}\|_{L^{d/2}(\mathbb{T}^d)}\|\nabla u_j^\delta\|_{L^{d/(d-1)}(\mathbb{T}^d)} \le C,$$

where $r_3 = d$, and $(\nabla p_i[u^\delta])$ is bounded in $L^{r_3}(\mathbb{T}^d)$ for some $r_3 > 2$ if $d = 2$ and for $r_3 = 2$ if $d = 1$. Hence, for a subsequence,

$$\nabla p_i[u^\delta] \rightharpoonup \nabla p_i[u] \quad \text{weakly in } L^{r_3}(\mathbb{T}^d).$$

It follows that $(u_i^\delta \nabla p_i[u^\delta])$ is bounded in $L^{\min\{2,d/(d-1)\}}(\mathbb{T}^d)$ and

$$u_i^\delta \nabla p_i[u^\delta] \rightharpoonup u_i \nabla p_i[u] \quad \text{weakly in } L^{\min\{2,d/(d-1)\}}(\mathbb{T}^d).$$

Thus, we can pass to the limit $\delta \to 0$ in (12) to conclude that $u_i^k := u_i \geq 0$ for $i = 1, \ldots, n$ solves

$$\frac{1}{\tau} \int_{\mathbb{T}^d} (u^k - u^{k-1}) \cdot \phi dx + \sigma \sum_{i=1}^n \int_{\mathbb{T}^d} \nabla u_i^k \cdot \nabla \phi_i dx = -\sum_{i=1}^n \int_{\mathbb{T}^d} u_i^k \nabla p_i[u^k] \cdot \nabla \phi_i dx, \tag{14}$$

for all test functions $\phi_i \in W^{1,r_3}(\mathbb{T}^d)$. Observe that $p_i[u^k]$ is an element of the space $W^{1,r_3}(\mathbb{T}^d)$ and is an admissible test function; this will be used in the next step.

*Step 3: Uniform estimates.* We introduce the piecewise constant in time functions $u^{(\tau)}(x,t) = u^k(x)$ for $x \in \mathbb{T}^d$ and $t \in ((k-1)\tau, k\tau]$, $k = 1, \ldots, N$. At time $t = 0$, we set $u_i^{(\tau)}(\cdot, 0) = u_i^0$. Furthermore, we use the time-shift operator $(\sigma_\tau u^{(\tau)})(x,t) = u^{k-1}(x)$ for $x \in \mathbb{T}^d$, $t \in ((k-1)\tau, k\tau]$. Then, summing (14) over $k$, we obtain

$$\begin{aligned} \frac{1}{\tau} \int_0^T \int_{\mathbb{T}^d} (u^{(\tau)} - \sigma_\tau u^{(\tau)}) \cdot \phi dx dt + \sigma \sum_{i=1}^n \int_0^T \int_{\mathbb{T}^d} \nabla u_i^{(\tau)} \cdot \nabla \phi_i dx dt \\ = -\sum_{i=1}^n \int_0^T \int_{\mathbb{T}^d} u_i^{(\tau)} \nabla p_i[u^{(\tau)}] \cdot \nabla \phi_i dx dt, \end{aligned}$$

for piecewise constant functions $\phi : (0,T) \to W^{1,r_3}(\mathbb{T}^d; \mathbb{R}^n)$ and, by density, for all functions $\phi \in L^2(0,T; W^{1,r_3}(\mathbb{T}^d; \mathbb{R}^n))$. Summing the entropy inequality (13) from $k = 1, \ldots, N$, it follows that

$$H_1(u^{(\tau)}(T)) + 4\sigma \sum_{i=1}^n \int_0^T \pi_i \|\nabla (u_i^{(\tau)})^{1/2}\|_{L^2(\mathbb{T}^d)}^2 dt \leq H_1(u^0). \tag{15}$$

These bounds allow us to derive further estimates. Since the $L^1 \log L^1$ bound dominates the $L^1(\mathbb{T}^d)$ norm, we deduce from the Poincaré–Wirtinger inequality that

$$\|u_i^{(\tau)} \log u_i^{(\tau)}\|_{L^\infty(0,T;L^1(\mathbb{T}^d))} + \|(u_i^{(\tau)})^{1/2}\|_{L^2(0,T;H^1(\mathbb{T}^d))} \leq C(u^0), \quad i = 1, \ldots, n.$$

This implies, by the Gagliardo–Nirenberg inequality with $\theta = d/(d+2)$, that

$$\begin{aligned} \|u_i^{(\tau)}\|_{L^{1+2/d}(Q_T)}^{1+2/d} &= \int_0^T \|(u_i^{(\tau)})^{1/2}\|_{L^{2+4/d}(\mathbb{T}^d)}^{2+4/d} dt \\ &\leq C \int_0^T \|(u_i^{(\tau)})^{1/2}\|_{H^1(\mathbb{T}^d)}^{\theta(2d+4)/d} \|(u_i^{(\tau)})^{1/2}\|_{L^2(\mathbb{T}^d)}^{(1-\theta)(2d+4)/d} dt \\ &\leq C \|u_i^{(\tau)}\|_{L^\infty(0,T;L^1(\mathbb{T}^d))}^{2/d} \int_0^T \|(u_i^{(\tau)})^{1/2}\|_{H^1(\mathbb{T}^d)}^2 dt \leq C(u^0), \end{aligned} \tag{16}$$

and by Hölder's inequality with $q=(d+2)/(d+1)$,

$$\begin{aligned}\|\nabla u_i^{(\tau)}\|_{L^q(Q_T)} &= 2\|(u_i^{(\tau)})^{1/2}\nabla(u_i^{(\tau)})^{1/2}\|_{L^q(Q_T)} \\ &\le 2\|(u_i^{(\tau)})^{1/2}\|_{L^{2+4/d}(Q_T)}\|\nabla(u_i^{(\tau)})^{1/2}\|_{L^2(Q_T)} \le C. \end{aligned} \tag{17}$$

These bounds are not sufficient to pass to the limit $\tau\to 0$. To derive further estimates, we use the test function $\phi_i = \pi_i p_i[u^k]\in W^{1,r_3}(\mathbb{T}^d)$ in (14):

$$\begin{aligned}&\frac{1}{\tau}\sum_{i,j=1}^n\int_{\mathbb{T}^d}\int_{\mathbb{T}^d}\pi_i(u_i^k(x)-u_i^{k-1}(x))K_{ij}(x-y)u_j^k(y)dxdy \\ &+\sigma\sum_{i,j=1}^n\int_{\mathbb{T}^d}\int_{\mathbb{T}^d}\pi_iK_{ij}(x-y)\nabla u_i^k(x)\cdot\nabla u_j^k(y)dxdy = -\sum_{i=1}^n\int_{\mathbb{T}^d}\pi_iu_i^k|\nabla p_i[u^k]|^2dx. \end{aligned} \tag{18}$$

Because of the positive definiteness of $\pi_iK_{ij}$, the second term on the left-hand side is nonnegative. Exploiting the symmetry and positive definiteness of $\pi_iK_{ij}$ (Hypotheses (H3)–(H4)), the first integral can be estimated from below as

$$\begin{aligned}&\frac{1}{2\tau}\sum_{i,j=1}^n\int_{\mathbb{T}^d}\int_{\mathbb{T}^d}\pi_iK_{ij}(x-y)\Big(u_i^k(x)u_j^k(y)-u_i^{k-1}(x)u_j^{k-1}(y) \\ &\qquad + (u_i^k(x)-u_i^{k-1}(x))(u_j^k(y)-u_j^{k-1}(y))\Big)dxdy \\ &\ge \frac{1}{2\tau}\sum_{i,j=1}^n\int_{\mathbb{T}^d}\int_{\mathbb{T}^d}\pi_iK_{ij}(x-y)\big(u_i^k(x)u_j^k(y)-u_i^{k-1}(x)u_j^{k-1}(y)\big)dxdy \\ &= \frac{1}{\tau}(H_2(u^k)-H_2(u^{k-1})).\end{aligned}$$

Therefore, we infer from (18) that

$$H_2(u^k)+\tau\sum_{i=1}^n\int_{\mathbb{T}^d}\pi_iu_i^k|\nabla p_i[u^k]|^2dx \le H_2(u^{k-1}),$$

and summing this inequality from $k=1,\ldots,N$, we have

$$H_2(u^{(\tau)}(T))+\sum_{i=1}^n\pi_i\int_0^T\big\|(u_i^{(\tau)})^{1/2}\nabla p_i[u^{(\tau)}]\big\|^2_{L^2(\mathbb{T}^d)}dt \le H_2(u^0). \tag{19}$$

The previous bound allows us to derive an estimate for the discrete time derivative. Indeed, estimates (16) and (19) imply that

$$u_i^{(\tau)}\nabla p_i[u^{(\tau)}] = (u_i^{(\tau)})^{1/2}\cdot(u_i^{(\tau)})^{1/2}\nabla p_i[u^{(\tau)}],$$

is uniformly bounded in $L^q(Q_T)$, where $q=(d+2)/(d+1)$. Let $\phi\in L^{q'}(0,T;W^{2,(d+2)/2}(\mathbb{T}^d))$, where $q'=d+2$. Then $1/q+1/q'=1$ and

$$\frac{1}{\tau}\bigg|\int_0^T\int_{\mathbb{T}^d}(u^{(\tau)}-\sigma_\tau u^{(\tau)})\cdot\phi dxdt\bigg|$$

$$\leq \sigma \sum_{i=1}^n \|u_i^{(\tau)}\|_{L^{1+2/d}(Q_T)} \|\Delta\phi_i\|_{L^{(d+2)/2}(Q_T)} + \sum_{i=1}^n \|u_i^{(\tau)}\nabla p_i[u^{(\tau)}]\|_{L^q(Q_T)} \|\nabla\phi_i\|_{L^{q'}(Q_T)}$$
$$\leq C\|\phi\|_{L^{q'}(0,T;W^{2,(d+2)/2}(\mathbb{T}^d))}.$$

We conclude that

$$\tau^{-1}\|u^{(\tau)} - \sigma_\tau u^{(\tau)}\|_{L^q(0,T;W^{2,(d+2)/2}(\mathbb{T}^d)')} \leq C. \tag{20}$$

*Step 4: Limit $\tau \to 0$.* Estimates (15) and (20) allow us to apply the Aubin–Lions compactness lemma in the version of [7] to conclude the existence of a subsequence that is not relabeled such that, as $\tau \to 0$,

$$u_i^{(\tau)} \to u_i \quad \text{strongly in } L^2(0,T;L^{d/(d-1)}(\mathbb{T}^d)),\ i = 1,\ldots,n,$$

if $d \geq 2$ and strongly in $L^2(0,T;L^r(\mathbb{T}^d))$ for any $r < \infty$ if $d = 1$. Strictly speaking, the version of [7] holds for the continuous time derivative, but the technique of [20] shows that the conclusion of [7] also holds for the discrete time derivative. Then, maybe for another subsequence, $u_i^{(\tau)} \to u_i$ a.e. in $Q_T$, and we deduce from (16) that $u_i^{(\tau)} \to u_i$ strongly in $L^r(Q_T)$ for all $r < 1 + 2/d$ (see Lemma 8 in Appendix A). Furthermore, we obtain from (15), (17), (19), and (20) the convergences

$$\begin{aligned}
\nabla u_i^{(\tau)} &\rightharpoonup \nabla u_i \quad \text{weakly in } L^q(Q_T),\ i = 1,\ldots,n, \\
\tau^{-1}(u^{(\tau)} - \sigma_\tau u^{(\tau)}) &\rightharpoonup \partial_t u_i \quad \text{weakly in } L^q(0,T;W^{2,(d+2)/2}(\mathbb{T}^d)'), \\
(u_i^{(\tau)})^{1/2}\nabla p_i[u^{(\tau)}] &\rightharpoonup z_i \quad \text{weakly in } L^2(Q_T),
\end{aligned}$$

where $z_i \in L^2(Q_T)$ and $q = (d+2)/(d+1)$. Since $u_i^{(\tau)} \geq 0$, we infer that $u_i \geq 0$ in $Q_T$.

*Step 5: Identification of the limit.* We need to identify $z_i$ with $u_i^{1/2}\nabla p_i[u]$. We show first that

$$\nabla p_i[u^{(\tau)}] \rightharpoonup \nabla p_i[u] \quad \text{weakly in } L^q(Q_T).$$

Indeed, it follows from the strong convergence of $(u_i^{(\tau)})$ that (up to a subsequence) $K_{ij}(x-y) \times u_j^{(\tau)}(y,t) \to K_{ij}(x-y)u_j(y,t)$ for a.e. $(y,t) \in Q_T$ and for a.e. $x \in \mathbb{T}^d$. Hence, because of the uniform bounds, $p_i[u^{(\tau)}] \to p_i[u]$ a.e. in $Q_T$. We deduce from Young's convolution inequality and the uniform bound for $\nabla u_i^{(\tau)}$ in $L^q(Q_T)$ that $\nabla p_i[u^{(\tau)}]$ is uniformly bounded in $L^q(Q_T)$. (Here, we only need $K_{ij} \in L^1(\mathbb{T}^d)$. The estimate for $\nabla p_i[u^{(\tau)}]$ is better under Hypothesis (H2), but the time regularity cannot be improved.) Therefore,

$$\nabla p_i[u^{(\tau)}] \rightharpoonup \nabla p_i[u] \quad \text{weakly in } L^q(Q_T).$$

When $d = 2$, we have the convergences $\nabla p_i[u^{(\tau)}] \rightharpoonup \nabla p_i[u]$ weakly in $L^{4/3}(Q_T)$ and $(u_i^{(\tau)})^{1/2} \to u_i^{1/2}$ strongly in $L^4(Q_T)$, which is sufficient to pass to the limit in the product and to identify it with $z_i$. However, this argument fails when $d > 2$, and we need a more sophisticated proof. The div-curl lemma does not apply, since the exponents of the Lebesgue spaces, in which the convergences of $(u_i^{(\tau)})^{1/2}$ and $\nabla p_i[u^{(\tau)}]$ take place, are not conjugate for $d > 2$. Also the generalization [9, Theorem 2.1] to nonconjugate exponents cannot be used for general $d$.

Our idea is to exploit the fact that the product converges in a space better than $L^1$. Then Lemma 9 in Appendix A immediately implies that

$$(u_i^{(\tau)})^{1/2}\nabla p_i[u^{(\tau)}] \rightharpoonup u_i^{1/2}\nabla p_i[u] \quad \text{weakly in } L^q(Q_T).$$

In fact, estimate (19) implies that this convergence holds in $L^2(Q_T)$. Then the strong convergence of $(u_i^{(\tau)})^{1/2}$ in $L^2(Q_T)$ gives

$$u_i^{(\tau)}\nabla p_i[u^{(\tau)}] \rightharpoonup u_i\nabla p_i[u] \quad \text{weakly in } L^1(Q_T).$$

In view of the uniform bounds for $(u_i^{(\tau)})^{1/2}$ in $L^{2+4/d}(Q_T)$ and of $(u_i^{(\tau)})^{1/2}\nabla p_i[u^{(\tau)}]$ in $L^2(Q_T)$, the product $u_i^{(\tau)}\nabla p_i[u^{(\tau)}]$ is uniformly bounded in $L^q(Q_T)$. Thus, the previous weak convergence also holds in $L^q(Q_T)$.

*Step 6: End of the proof.* The convergences of the previous step allow us to pass to the limit $\tau \to 0$ in (14), yielding

$$\int_0^T \langle \partial_t u_i, \phi_i\rangle dt + \sigma\int_0^T\int_{\mathbb{T}^d} \nabla u_i \cdot \nabla\phi_i dxdt = -\int_0^T\int_{\mathbb{T}^d} u_i\nabla p_i[u]\cdot\nabla\phi_i dxdt,$$

for all smooth test functions. Because of $\nabla u_i$, $u_i\nabla p_i[u] \in L^q(Q_T)$, a density argument shows that the weak formulation holds for all $\phi \in L^{q'}(0,T;W^{1,q'}(\mathbb{T}^d))$, recalling that $q' = d+2$. Then $\partial_t u_i$ lies in the space $L^q(0,T;W^{-1,q}(\mathbb{T}^d))$, where $W^{-1,q}(\mathbb{T}^d) := W^{1,q'}(\mathbb{T}^d)'$. The proof that $u(\cdot,0)$ satisfies the initial datum can be done exactly as in [26, p. 1980]. Finally, using the convexity of $H_1$ and the lower semi-continuity of convex functions, the entropy inequalities (15) and (19) become (9) and (10), respectively, in the limit $\tau \to 0$. This ends the proof of Theorem 2.

**Proof of Proposition 3.** The proof of the $H^1(\mathbb{T}^d)$ regularity requires an approximate scheme different from that one used in the proof of Theorem 2. Given $u^{k-1} \in L^2(\mathbb{T}^d;\mathbb{R}^n)$ with $u_i^{k-1} \geq 0$, we wish to find $u^k \in H^1(\mathbb{T}^d;\mathbb{R}^n)$ such that

$$\frac{1}{\tau}\int_{\mathbb{T}^d}(u_i^k - u_i^{k-1})\phi_i dx + \sigma\int_{\mathbb{T}^d}\nabla u_i^k\cdot\nabla\phi_i dx + \int_{\mathbb{T}^d}\frac{(u_i^k)^+}{1+\delta(u_i^k)^+}\nabla p_i[u^k]\cdot\nabla\phi_i dx = 0, \tag{21}$$

for $\phi_i \in H^1(\mathbb{T}^d)$, where $\delta > 0$ and $z^+ = \max\{0,z\}$. Since $\nabla K_{ij} \in L^{d+2}(\mathbb{T}^d)$, $\nabla p_i[u^k]$ can be bounded in $L^{d+2}(\mathbb{T}^d)$ in terms of the $L^1(\mathbb{T}^d)$ norm of $u^k$. Thus, the last term on the left-hand side is well defined. The existence of a solution to this discrete scheme is proved by a fixed-point argument, and the main step is the derivation of uniform estimates. First, we observe that the test function $(u_i^k)^- = \min\{0,u_i^k\}$ yields

$$\frac{1}{\tau}\int_{\mathbb{T}^d}(u_i^k - u_i^{k-1})(u_i^k)^- dx + \sigma\int_{\mathbb{T}^d}|\nabla(u_i^k)^-|^2 dx = -\int_{\mathbb{T}^d}\frac{(u_i^k)^+}{1+\delta(u_i^k)^+}\nabla p_i[u^k]\cdot\nabla(u_i^k)^- dx = 0,$$

and consequently, $(u_i^k)^- = 0$ in $\mathbb{T}^d$. Thus, $u_i^k \geq 0$ and we can remove the plus sign in (21). Second, we use the test function $u_i^k$ in (21) and sum over $i = 1,\ldots,n$:

$$\frac{1}{\tau}\sum_{i=1}^n\int_{\mathbb{T}^d}(u_i^k - u_i^{k-1})u_i^k dx + \sigma\sum_{i=1}^n\int_{\mathbb{T}^d}|\nabla u_i^k|^2 dx = -\sum_{i=1}^n\int_{\mathbb{T}^d}\frac{u_i^k}{1+\delta u_i^k}\nabla p_i[u^k]\cdot\nabla u_i^k dx. \tag{22}$$

The first integral becomes

$$\sum_{i=1}^n\int_{\mathbb{T}^d}(u_i^k-u_i^{k-1})u_i^kdx\geq\frac12\sum_{i=1}^n\int_{\mathbb{T}^d}\big((u_i^k)^2-(u_i^{k-1})^2\big)dx.$$

The right-hand side in (22) is estimated by Hölder's inequality and Young's convolution inequality:

$$\begin{aligned}-\sum_{i=1}^n\int_{\mathbb{T}^d}\frac{u_i^k}{1+\delta u_i^k}\nabla p_i[u^k]\cdot\nabla u_i^kdx&\leq\sum_{i=1}^n\|u_i^k\|_{L^{2+4/d}(\mathbb{T}^d)}\|\nabla p_i[u^k]\|_{L^{d+2}(\mathbb{T}^d)}\|\nabla u_i^k\|_{L^2(\mathbb{T}^d)}\\&\leq C_K\sum_{i,j=1}^n\|u_i^k\|_{L^{2+4/d}(\mathbb{T}^d)}\|u_j^k\|_{L^1(\mathbb{T}^d)}\|\nabla u_i^k\|_{L^2(\mathbb{T}^d)},\end{aligned}$$

where $C_K>0$ depends on the $L^{d+2}(\mathbb{T}^d)$ norm of $\nabla K_{ij}$. Taking the test function $\phi_i=1$ in (21) shows that $\|u_i^k\|_{L^1(\mathbb{T}^d)}=\|u_i^0\|_{L^1(\mathbb{T}^d)}$ is uniformly bounded. This allows us to reduce the cubic expression on the right-hand side of the previous inequality to a quadratic one. This is the key idea of the proof. Combining the previous arguments, (21) becomes

$$\begin{aligned}\frac12\sum_{i=1}^n\int_{\mathbb{T}^d}(u_i^k)^2dx-\frac12\sum_{i=1}^n\int_{\mathbb{T}^d}(u_i^{k-1})^2dx+\tau\sigma\sum_{i=1}^n\int_{\mathbb{T}^d}|\nabla u_i^k|^2dx&\\\leq\tau C\sum_{i=1}^n\|u_i^k\|_{L^{2+4/d}(\mathbb{T}^d)}\|\nabla u_i^k\|_{L^2(\mathbb{T}^d)}&\\\leq\frac12\tau\sigma\sum_{i=1}^n\|\nabla u_i^k\|_{L^2(\mathbb{T}^d)}^2+\tau C\sum_{i=1}^n\|u_i^k\|_{L^{2+4/d}(\mathbb{T}^d)}^2.&\end{aligned}$$

The Gagliardo–Nirenberg and Poincaré–Wirtinger inequalities show that

$$\begin{aligned}\|u_i^k\|_{L^{2+4/d}(\mathbb{T}^d)}^2&\leq C\|u_i^k\|_{H^1(\mathbb{T}^d)}^{2\theta}\|u_i^k\|_{L^1(\mathbb{T}^d)}^{2(1-\theta)}\\&\leq C\big(\|\nabla u_i^k\|_{L^2(\mathbb{T}^d)}+\|u_i^k\|_{L^1(\mathbb{T}^d)}\big)^{2\theta}\|u_i^k\|_{L^1(\mathbb{T}^d)}^{2(1-\theta)}\\&\leq C(u^0)\|\nabla u_i^k\|_{L^2(\mathbb{T}^d)}^{2\theta}+C(u^0),\end{aligned}$$

where $\theta=d(d+4)/(d+2)^2<1$. We deduce from Young's inequality that for any $\varepsilon>0$,

$$\|u_i^k\|_{L^{2+4/d}(\mathbb{T}^d)}^2\leq\varepsilon\|\nabla u_i^k\|_{L^2(\mathbb{T}^d)}^2+C(\varepsilon).$$

Therefore, choosing $\varepsilon>0$ sufficiently small, we infer that

$$\frac12\sum_{i=1}^n\int_{\mathbb{T}^d}(u_i^k)^2dx-\frac12\sum_{i=1}^n\int_{\mathbb{T}^d}(u_i^{k-1})^2dx+\frac14\tau\sigma\sum_{i=1}^n\int_{\mathbb{T}^d}|\nabla u_i^k|^2dx\leq C. \tag{23}$$

This provides a uniform $H^1(\mathbb{T}^d)$ estimate for $u^k$. Defining the fixed-point operator as a mapping from $L^2(\mathbb{T}^d)$ to $L^2(\mathbb{T}^d)$, the compact embedding $H^1(\mathbb{T}^d)\hookrightarrow L^2(\mathbb{T}^d)$ implies the compactness of this operator (see [27, Chapter 4] for details). This shows that (21) possesses a weak solution $u^k\in H^1(\mathbb{T}^d)$.

In order to pass to the limit $(\delta,\tau)\to 0$, we need uniform estimates for the piecewise constant in time functions $u_i^{(\tau)}$, using the same notation as in the proof of Theorem 2. Estimate (23) provides uniform bounds for $u_i^{(\tau)}$ in $L^\infty(0,T;L^2(\mathbb{T}^d))$ and $L^2(0,T;H^1(\mathbb{T}^d))$. By the Gagliardo–Nirenberg inequality, $(u_i^{(\tau)})$ is bounded in $L^{2+4/d}(Q_T)$. By Young's convolution inequality,

$$\sup_{t\in(0,T)}\|\nabla p_i[u^{(\tau)}(t)]\|_{L^\infty(\mathbb{T}^d)} \le \sum_{j=1}^n \|\nabla K_{ij}\|_{L^{d+2}(\mathbb{T}^d)} \sup_{t\in(0,T)}\|u_j^{(\tau)}\|_{L^q(\mathbb{T}^d)} \le C,$$

where $q=(d+2)/(d+1)$. Thus, $(\nabla p_i[u^{(\tau)}])$ is bounded in $L^\infty(0,T;L^\infty(\mathbb{T}^d))$. From these estimates, we can derive a uniform bound for the discrete time derivative. Therefore, by the Aubin–Lions lemma [20], up to a subsequence, as $(\delta,\tau)\to 0$,

$$u_i^{(\tau)}\to u_i \quad \text{strongly in } L^2(Q_T),$$

and this convergence even holds in $L^r(Q_T)$ for any $r<2+4/d$. We can show as in the proof of Theorem 2 that $p_i[u^{(\tau)}]\to p_i[u]$ a.e. and consequently, for a subsequence, $\nabla p_i[u^{(\tau)}]\rightharpoonup \nabla p_i[u]$ weakly in $L^2(Q_T)$. We infer that

$$u_i^{(\tau)}\nabla p_i[u^{(\tau)}]\rightharpoonup u_i\nabla p_i[u] \quad \text{weakly in } L^1(Q_T).$$

Omitting the details, it follows that $u=(u_1,\ldots,u_n)$ is a weak solution to (1)–(2) satisfying $u_i\in L^2(0,T;H^1(\mathbb{T}^d))$ for $i=1,\ldots,n$.

Next, we show the lower and upper bounds for $u_i$. Define $M(t)=M_0e^{\lambda t}$, where $\lambda>0$ will be specified later. Since $\nabla K_{ij}\in L^\infty$, we can apply the Young convolution inequality and estimate $\nabla p_i[u]$ in $L^\infty(\mathbb{T}^d)$ in terms of the $L^1(\mathbb{T}^d)$ bounds for $u_j$. Then, with the test function $e^{-\lambda t}(u_i-M)^+(t)=e^{-\lambda t}\max\{0,(u_i-M)(t)\}$ in the weak formulation of (21), we deduce from

$$\partial_t u_i e^{-\lambda t}(u_i-M)^+ = \frac12\partial_t\big\{e^{-\lambda t}[(u_i-M)^+]^2\big\} + \frac{\lambda}{2}e^{-\lambda t}[(u_i-M)^+]^2 + \lambda e^{-\lambda t}M(u_i-M)^+,$$

that

$$\begin{aligned}
&\frac12\int_{\mathbb{T}^d} e^{-\lambda t}(u_i-M)^+(t)^2dx + \sigma\int_0^t\int_{\mathbb{T}^d} e^{-\lambda s}|\nabla(u_i-M)^+|^2dxds \\
&\quad = -\int_0^t\int_{\mathbb{T}^d} e^{-\lambda s}(u_i-M)\nabla p_i[u]\cdot\nabla(u_i-M)^+dxds - \frac{\lambda}{2}\int_0^t\int_{\mathbb{T}^d} e^{-\lambda s}[(u_i-M)^+]^2dxds \\
&\qquad - \int_0^t\int_{\mathbb{T}^d} e^{-\lambda s}M\nabla p_i[u]\cdot\nabla(u_i-M)^+dxds - \lambda\int_0^t\int_{\mathbb{T}^d} e^{-\lambda s}M(u_i-M)^+dxds.
\end{aligned}$$

We write $(u_i-M)\nabla(u_i-M)^+ = \frac12\nabla[(u_i-M)^+]^2$ and integrate by parts in the first and third terms of the right-hand side:

$$\begin{aligned}
&\frac12\int_{\mathbb{T}^d} e^{-\lambda s}(u_i-M)^+(t)^2dx + \sigma\int_0^t\int_{\mathbb{T}^d} e^{-\lambda s}|\nabla(u_i-M)^+|^2dxds \\
&\quad \le \frac12\big(\|\Delta p_i[u]\|_{L^\infty(0,T;L^\infty(\mathbb{T}^d))} - \lambda\big)\int_0^t\int_{\mathbb{T}^d} e^{-\lambda s}[(u_i-M)^+]^2dxds
\end{aligned}$$

$$+\big(\|\Delta p_i[u]\|_{L^\infty(0,T;L^\infty(\mathbb{T}^d))}-\lambda\big)\int_0^t\int_{\mathbb{T}^d}e^{-\lambda s}M(u_i-M)^+dxds.$$

By Young's convolution inequality and the regularity assumptions on $K_{ij}$,

$$\|\Delta p_i[u]\|_{L^\infty(0,T;L^\infty(\mathbb{T}^d))}\le C\sum_{j=1}^n\|u_j\|_{L^\infty(0,T;L^1(\mathbb{T}^d))}\le C_0.$$

Therefore, choosing $\lambda\ge C_0$, it follows that

$$\int_{\mathbb{T}^d}e^{-\lambda t}(u_i-M)^+(t)^2dx\le 0,$$

and we infer that $e^{-\lambda t}(u_i-M)^+(t)=0$ and $u_i(t)\le M(t)=M_0e^{\lambda t}$ for $t>0$. The inequality $u_i(t)\ge m(t):=m_0e^{-\lambda t}$ is proved in the same way, using the test function $e^{-\lambda t}(u_i-m)^-=e^{-\lambda t}\min\{0,u_i-m\}$.

**Remark 6.** Proposition 3 holds true for functions $K_{ij}(x,y)$ that are not convolution kernels. We need the regularity $\nabla_xK_{ij}\in L_y^\infty L_x^{d+2}\cap L_x^\infty L_y^{d+2}$ to apply the Young inequality for kernels; see [35, Theorem 0.3.1]. For the lower and upper bounds of the solution, we additionally need the regularity $\nabla_xK_{ij}$, $\Delta_xK_{ij}\in L_x^\infty L_y^\infty$. □

## 4. Weak-strong uniqueness for the nonlocal system

In this section, we prove Theorem 4. Let $u$ be a nonnegative weak solution to (1)–(2) and $v$ be a positive "strong" solution to (1)–(2) with the regularity properties specified in the theorem. Then, for $0<\varepsilon<1$, we define the regularized relative entropy density

$$h_\varepsilon(u|v)=\sum_{i=1}^n\pi_i\big((u_i+\varepsilon)(\log(u_i+\varepsilon)-1)-(u_i+\varepsilon)\log v_i+v_i\big),$$

and the associated relative entropy

$$H_\varepsilon(u|v)=\int_{\mathbb{T}^d}h_\varepsilon(u|v)dx.$$

*Step 1: Preparations.* We compute

$$\frac{\partial h_\varepsilon}{\partial u_i}(u|v)=\pi_i\log(u_i+\varepsilon)-\pi_i\log v_i,\quad \frac{\partial h_\varepsilon}{\partial v_i}(u|v)=-\pi_i\frac{u_i+\varepsilon}{v_i}+\pi_i.$$

The second function is an admissible test function for the weak formulation of (1), satisfied by $v_i$, since $\nabla u_i\in L^2(Q_T)$ and $\nabla v_i\in L^\infty(Q_T)$. Strictly speaking, the first function is not an admissible test function for the weak formulation of (1), satisfied by $u_i$, since it needs test functions in $W^{1,d+2}(\mathbb{T}^d)$. However, the nonlocal term becomes with this test function

$$\int_{\mathbb{T}^d}\int_{\mathbb{T}^d}K_{ij}(x-y)\nabla u_j(y)\cdot\frac{\nabla u_i(x)}{u_i(x)+\varepsilon}dxdy,$$

which is finite since $K_{ij}$ is essentially bounded and $\nabla u_i \cdot \nabla u_j \in L^1(Q_T)$. Thus, by a suitable approximation argument, the following computation can be made rigorous. We find that

$$\begin{aligned}
\frac{d}{dt}H_\varepsilon(u|v) &= \sum_{i=1}^n \left( \left\langle \partial_t u_i, \frac{\partial h_\varepsilon}{\partial u_i}(u|v) \right\rangle + \left\langle \partial_t v_i, \frac{\partial h_\varepsilon}{\partial v_i}(u|v) \right\rangle \right) \\
&= -\sigma \sum_{i=1}^n \int_{\mathbb{T}^d} \left( \nabla u_i \cdot \nabla \frac{\partial h_\varepsilon}{\partial u_i}(u|v) + \nabla v_i \cdot \nabla \frac{\partial h_\varepsilon}{\partial v_i}(u|v) \right) dx \\
&\quad - \sum_{i=1}^n \int_{\mathbb{T}^d} \left( u_i \nabla p_i[u] \cdot \nabla \frac{\partial h_\varepsilon}{\partial u_i}(u|v) + v_i \nabla p_i[v] \cdot \nabla \frac{\partial h_\varepsilon}{\partial v_i}(u|v) \right) dx \\
&= -\sigma \sum_{i=1}^n \int_{\mathbb{T}^d} \pi_i \left| \frac{\nabla u_i}{\sqrt{u_i+\varepsilon}} - \sqrt{u_i+\varepsilon}\frac{\nabla v_i}{v_i} \right|^2 dx \\
&\quad - \sum_{i=1}^n \int_{\mathbb{T}^d} \pi_i \left( \frac{u_i}{u_i+\varepsilon}\nabla p_i[u] \cdot \nabla u_i - \frac{u_i}{v_i}\nabla p_i[u] \cdot \nabla v_i - \nabla p_i[v] \cdot \nabla u_i \right. \\
&\quad \left. + \frac{u_i+\varepsilon}{v_i}\nabla p_i[v] \cdot \nabla v_i \right) dx.
\end{aligned}$$

The first integral is nonpositive. Thus, an integration over $(0,t)$ gives

$$\begin{aligned}
&H_\varepsilon(u(t)|v(t)) - H_\varepsilon(u(0)|v(0)) \\
&\quad \le -\sum_{i=1}^n \int_0^t \int_{\mathbb{T}^d} \pi_i (u_i+\varepsilon)\nabla(p_i[u]-p_i[v]) \cdot \nabla \log \frac{u_i+\varepsilon}{v_i} dxds \\
&\qquad + \varepsilon \sum_{i=1}^n \int_0^t \int_{\mathbb{T}^d} \pi_i \nabla p_i[u] \cdot \nabla \log \frac{u_i+\varepsilon}{v_i} dxds =: I_1 + I_2. 
\end{aligned} \tag{24}$$

*Step 2: Estimation of $I_1$ and $I_2$.* Inserting the definition of $p_i$,

$$\begin{aligned}
\nabla(p_i[u]-p_i[v])(x) &= \sum_{j=1}^n \int_{\mathbb{T}^d} K_{ij}(x-y)\nabla(u_j-v_j)(y)dy \\
&= \sum_{j=1}^n \int_{\mathbb{T}^d} K_{ij}(x-y)\bigg( (u_j+\varepsilon)(y)\nabla \log \frac{u_j+\varepsilon}{v_j}(y) \\
&\quad + (u_j-v_j)(y)\nabla \log v_j(y) + \varepsilon \nabla \log v_j(y) \bigg) dy,
\end{aligned}$$

leads to

$$\begin{aligned}
I_1 = -\sum_{i,j=1}^n \int_0^t \int_{\mathbb{T}^d}\int_{\mathbb{T}^d} \pi_i K_{ij}(x-y)\bigg( &(u_i+\varepsilon)(x)(u_j+\varepsilon)(y)\nabla \log \frac{u_j+\varepsilon}{v_j}(y) \\
\times \nabla \log \frac{u_i+\varepsilon}{v_i}(x) + (u_i+\varepsilon)(x)&(u_j-v_j)(y)\nabla \log v_j(y) \cdot \nabla \log \frac{u_i+\varepsilon}{v_i}(x) \bigg) dxdyds
\end{aligned}$$

$$
\begin{aligned}
&- \varepsilon \sum_{i,j=1}^n \int_0^t \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \pi_i K_{ij}(x-y)(u_i+\varepsilon)(x)\nabla \log v_j(y) \cdot \nabla \log \frac{u_i+\varepsilon}{v_i}(x) dxdyds \\
&=: I_{11} + I_{12}.
\end{aligned}
$$

Setting

$$
U_i = (u_i+\varepsilon)\nabla \log \frac{u_i+\varepsilon}{v_i}, \quad V_i = \frac{1}{2}(u_i - v_i)\nabla \log v_i,
$$

we can formulate the first integral as

$$
\begin{aligned}
I_{11} &= -\sum_{i,j=1}^n \int_0^t \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \pi_i K_{ij}(x-y)\big(U_i(x)\cdot U_j(y) + 2U_i(x)\cdot V_j(y)\big) dxdyds \\
&= -\sum_{i,j=1}^n \int_0^t \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \pi_i K_{ij}(x-y)(U_i+V_i)(x)\cdot (U_j+V_j)(y) dxdyds \\
&\quad + \sum_{i,j=1}^n \int_0^t \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \pi_i K_{ij}(x-y) V_i(x)\cdot V_j(y) dxdyds \\
&\le \frac{1}{4}\sum_{i,j=1}^n \int_0^t \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \pi_i K_{ij}(x-y)(u_i-v_i)(x)(u_j-v_j)(y)\nabla \log v_i(x)\cdot \nabla \log v_j(y) dxdyds \\
&\le \frac{1}{4} \max_{i,j=1,\ldots,n} \|\pi_i K_{ij}\|_{L^\infty(\mathbb{T}^d)} \max_{k=1,\ldots,n} \|\nabla \log v_k\|_{L^\infty(Q_T)}^2 \\
&\quad \times \sum_{i,j=1}^n \int_0^t \int_{\mathbb{T}^d} |(u_i-v_i)(x)| dx \int_{\mathbb{T}^d} |(u_j-v_j)(y)| dyds \\
&\le C \sum_{i=1}^n \int_0^t \bigg( \int_{\mathbb{T}^d} |u_i - v_i| dx \bigg)^2 ds,
\end{aligned}
$$

using the symmetry and positive definiteness of $\pi_i K_{ij}$ as well as the regularity assumptions on $K_{ij}$ and $\nabla \log v_i$. The second integral $I_{12}$ is estimated as

$$
\begin{aligned}
I_{12} &= -\varepsilon \sum_{i,j=1}^n \int_0^t \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \pi_i K_{ij}(x-y)\nabla \log v_j(y) \cdot \big(\nabla u_i - (u_i+\varepsilon)\nabla \log v_i\big)(x) dxdyds \\
&\le \varepsilon C \sum_{i,j=1}^n \|\nabla \log v_j\|_{L^\infty(Q_T)} \int_0^t \int_{\mathbb{T}^d} \big(|\nabla u_i| + (u_i+1)|\nabla \log v_i|\big) dxds \\
&\le \varepsilon C \sum_{i=1}^n \big(\|\nabla u_i\|_{L^1(Q_T)} + \|u_i\|_{L^1(Q_T)} + 1\big) \le \varepsilon C.
\end{aligned}
$$

We conclude that

$$
I_1 \le C \sum_{i=1}^n \int_0^t \bigg( \int_{\mathbb{T}^d} |u_i - v_i| dx \bigg)^2 ds + \varepsilon C.
$$

It remains to estimate $I_2$. Here we need the improved regularity $\nabla u_i \in L^2(Q_T)$. Inserting the definition of $p_i[u]$, we have

$$I_2 = \varepsilon \sum_{i,j=1}^n \int_0^t \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} \pi_i K_{ij}(x-y) \nabla u_j(y) \cdot \nabla \log \frac{u_i + \varepsilon}{v_i}(x) dx dy ds.$$

Since

$$\varepsilon |\nabla u_j(y) \cdot \nabla \log(u_i + \varepsilon)(x)| = 2\varepsilon \left| \nabla u_j(y) \cdot \frac{\nabla \sqrt{u_i + \varepsilon}}{\sqrt{u_i + \varepsilon}}(x) \right| \le 2\sqrt{\varepsilon} |\nabla u_j(y)| |\nabla \sqrt{u_i(x)}|,$$

we find that

$$I_2 \le C \sum_{i,j=1}^n \big( \varepsilon \|\nabla u_j\|_{L^1(Q_T)} + \sqrt{\varepsilon} \|\nabla u_j\|_{L^2(Q_T)} \|\nabla \sqrt{u_i}\|_{L^2(Q_T)} \big) \le \sqrt{\varepsilon} C.$$

We summarize the estimates for $I_1$ and $I_2$ and conclude from (24) that

$$H_\varepsilon(u(t)|v(t)) - H_\varepsilon(u(0)|v(0)) \le C \sum_{i=1}^n \int_0^t \left( \int_{\mathbb{T}^d} |u_i - v_i| dx \right)^2 ds + \sqrt{\varepsilon} C. \tag{25}$$

*Step 3: Limit $\varepsilon \to 0$.* We perform first the limit in $H_\varepsilon(u(t)|v(t))$. Since $u_i \in L^2(0,T; H^1(\mathbb{T}^d)) \cap H^1(0,T; H^{-1}(\mathbb{T}^d)) \hookrightarrow C^0([0,T]; L^2(\mathbb{T}^d))$, we have

$$\big| (u_i + \varepsilon)(\log(u_i + \varepsilon) - 1) \big| \le u_i(\log u_i + 1) + C \in L^\infty(0,T; L^1(\mathbb{T}^d)).$$

Therefore, by dominated convergence, as $\varepsilon \to 0$,

$$\sum_{i=1}^n \int_{\mathbb{T}^d} \pi_i (u_i(t) + \varepsilon)(\log(u_i(t) + \varepsilon) - 1) dx \to \sum_{i=1}^n \int_{\mathbb{T}^d} \pi_i u_i(t)(\log u_i(t) - 1) dx,$$

and this convergence holds for a.e. $t \in (0,T)$. Furthermore, in view of the bound for $\log v_i$,

$$\sum_{i=1}^n \pi_i(-(u_i + \varepsilon) \log v_i + v_i) \le C(v) \left( \sum_{i=1}^n u_i + 1 \right) \in L^\infty(0,T; L^1(\mathbb{T}^d)),$$

and we can again use dominated convergence:

$$\sum_{i=1}^n \int_{\mathbb{T}^d} \pi_i \big( -(u_i(t) + \varepsilon) \log v_i(t) + v_i(t) \big) dx \to \sum_{i=1}^n \int_{\mathbb{T}^d} \pi_i \big( -u_i(t) \log v_i(t) + v_i(t) \big) dx.$$

This shows that for a.e. $t \in (0,T)$,

$$H_\varepsilon(u(t)|v(t)) \to H(u(t)|v(t)) \quad \text{as } \varepsilon \to 0, \text{ where}$$
$$H(u|v) = \sum_{i=1}^n \int_{\mathbb{T}^d} \pi_i \big( u_i(\log u_i - 1) - u_i \log v_i + v_i \big) dx,$$

and $H_\varepsilon(u(0)|v(0)) = H_\varepsilon(u^0|u^0) \to 0$. Then we deduce from (25) in the limit $\varepsilon \to 0$ that

$$H(u(t)|v(t)) \le C \sum_{i=1}^n \int_0^t \|u_i - v_i\|_{L^1(\mathbb{T}^d)}^2 ds. \tag{26}$$

Taking the test function $\phi_i = 1$ in the weak formulation of (1), we find that $\int_{\mathbb{T}^d} u_i^0 dx = \int_{\mathbb{T}^d} u_i(t)dx$ for all $t > 0$. Since $u$ and $v$ have the same initial data, it follows that $\int_{\mathbb{T}^d} u_i(t)dx = \int_{\mathbb{T}^d} v_i(t)dx$ for all $t > 0$. Thus, by the classical Csiszár–Kullback–Pinsker inequality [27, Theorem A.2], we have

$$\begin{aligned} H(u|v) &= \sum_{i=1}^n \int_{\mathbb{T}^d} \pi_i u_i \log \frac{u_i}{v_i} dx + \sum_{i=1}^n \int_{\mathbb{T}^d} \pi_i (v_i - u_i) dx \\ &= \sum_{i=1}^n \int_{\mathbb{T}^d} \pi_i u_i \log \frac{u_i}{v_i} dx \geq C(u^0) \sum_{i=1}^n \|u_i - v_i\|_{L^1(\mathbb{T}^d)}^2. \end{aligned}$$

We infer from (26) that

$$\sum_{i=1}^n \|(u_i - v_i)(t)\|_{L^1(\mathbb{T}^d)}^2 \leq C \int_0^t \sum_{i=1}^n \|u_i - v_i\|_{L^1(\mathbb{T}^d)}^2 ds.$$

Gronwall's inequality implies that $\|(u_i - v_i)(t)\|_{L^1(\mathbb{T}^d)} = 0$ and hence $u_i(t) = v_i(t)$ in $\mathbb{T}^d$ for a.e. $t > 0$ and $i = 1, \ldots, n$.

## 5. Localization limit

We prove Theorem 5. Let $u^\varepsilon$ be the nonnegative weak solution to (1)–(2) with kernel (11), constructed in Theorem 2. The entropy inequalities (9) and (10) as well as the proof of Theorem 2 show that all estimates are independent of $\varepsilon$. (More precisely, the right-hand side of (10) depends on $K_{ij}^\varepsilon$, but in view of [8, Theorem 4.22], the right-hand side can be bounded uniformly in $\varepsilon$.) Therefore, for $i = 1, \ldots, n$ (see (15)–(17), (19)–(20)),

$$\begin{aligned} \|u_i^\varepsilon \log u_i^\varepsilon\|_{L^\infty(0,T;L^1(\mathbb{T}^d))} + \|u_i^\varepsilon\|_{L^{1+2/d}(Q_T)} + \|u_i^\varepsilon\|_{L^q(0,T;W^{1,q}(\mathbb{T}^d))} &\leq C, \\ \|(u_i^\varepsilon)^{1/2}\|_{L^2(0,T;H^1(\mathbb{T}^d))} + \|\partial_t u_i^\varepsilon\|_{L^q(0,T;W^{1,d+2}(\mathbb{T}^d)')} + \|(u_i^\varepsilon)^{1/2} \nabla p_i^\varepsilon[u^\varepsilon]\|_{L^2(Q_T)} &\leq C, \end{aligned}$$

where $C > 0$ is independent of $\varepsilon$, $q = (d+2)/(d+1)$, and $p_i^\varepsilon[u_i^\varepsilon] = \sum_{j=1}^n \int_{\mathbb{T}^d} K_{ij}^\varepsilon(x - y) u_j^\varepsilon(y) dy$. We infer from the Aubin–Lions lemma in the version of [7, 20] that there exists a subsequence (not relabeled) such that, as $\varepsilon \to 0$,

$$u_i^\varepsilon \to u_i \quad \text{strongly in } L^2(0,T;L^{d/(d-1)}(\mathbb{T}^d)), \quad i = 1, \ldots, n, \tag{27}$$

if $d \geq 2$ and strongly in $L^2(0,T;L^r(\mathbb{T}^d))$ for any $r < \infty$ if $d = 1$. Moreover,

$$\nabla u_i^\varepsilon \rightharpoonup \nabla u_i \quad \text{weakly in } L^q(Q_T), \quad i = 1, \ldots, n, \tag{28}$$

$$\partial_t u_i^\varepsilon \rightharpoonup \partial_t u_i \quad \text{weakly in } L^q(0,T;W^{1,d+2}(\mathbb{T}^d)'), \tag{29}$$

$$(u_i^\varepsilon)^{1/2} \nabla p_i^\varepsilon[u^\varepsilon] \rightharpoonup z_i \quad \text{weakly in } L^2(Q_T), \tag{30}$$

where $z_i \in L^2(Q_T)$ for $i = 1, \ldots, n$.

As in Section 3, the main difficulty is the identification of $z_i$ with $u_i^{1/2} \nabla p_i[u]$, where $p_i[u] := \sum_{j=1}^n a_{ij} \nabla u_j$. Since the kernel functions also depend on $\varepsilon$, the proof is different from that one in Section 3. We claim that

$$\nabla p_i^\varepsilon[u^\varepsilon] \rightharpoonup \nabla p_i[u] \quad \text{weakly in } L^q(Q_T). \tag{31}$$

Indeed, let $\phi \in L^{q'}(Q_T;\mathbb{R}^n)$, where $q' = d+2$ satisfies $1/q + 1/q' = 1$. We compute

$$\begin{aligned}
\bigg|\int_0^T\int_{\mathbb{T}^d}(\nabla p_i^\varepsilon[u^\varepsilon] - \nabla p_i[u])\cdot\phi dxdt\bigg| \\
&= \bigg|\sum_{j=1}^n\int_0^T\int_{\mathbb{T}^d}\bigg(\int_{\mathbb{T}^d}K_{ij}^\varepsilon(x-y)\nabla u_j^\varepsilon(y,t)dy\bigg)\cdot\phi(x,t)dxdt \\
&\quad - \sum_{j=1}^n\int_0^T\int_{\mathbb{T}^d}a_{ij}\nabla u_j(y,t)\cdot\phi(y,t)dydt\bigg| \\
&\le \sum_{j=1}^n\bigg|\int_0^T\int_{\mathbb{T}^d}\bigg(\int_{\mathbb{T}^d}K_{ij}^\varepsilon(x-y)\phi(x,t)dx - a_{ij}\phi(y,t)\bigg)\cdot\nabla u_j^\varepsilon(y,t)dydt\bigg| \\
&\quad + \sum_{j=1}^n a_{ij}\bigg|\int_0^T\int_{\mathbb{T}^d}\nabla(u_j^\varepsilon - u_j)(y,t)\cdot\phi(y,t)dydt\bigg| \\
&\le \sum_{j=1}^n\bigg\|\int_{\mathbb{T}^d}K_{ij}^\varepsilon(\cdot-y)\phi(y)dy - a_{ij}\phi\bigg\|_{L^{q'}(Q_T)}\|\nabla u_j^\varepsilon\|_{L^q(Q_T)} \\
&\quad + \sum_{j=1}^n a_{ij}\bigg|\int_0^T\int_{\mathbb{T}^d}\nabla(u_j^\varepsilon - u_j)(y,t)\cdot\phi(y,t)dydt\bigg|.
\end{aligned}$$

Since $B$ has compact support in $\mathbb{R}$, we can apply the proof of [8, Theorem 4.22] to infer that the first term on the right-hand side, formulated as the convolution $K_{ij}^\varepsilon * \phi - a_{ij}\phi$ (slightly abusing the notation), converges to zero strongly in $L^{q'}(\mathbb{R}^d)$ as $\varepsilon \to 0$. Thus, taking into account the weak convergence (28), convergence (31) follows.

Because of the convergences (27), (30), and (31), we can apply Lemma 9 in Appendix A to infer that $z_i = u_i^{1/2}\nabla p_i[u]$. Therefore,

$$u_i^\varepsilon\nabla p_i^\varepsilon[u^\varepsilon] \rightharpoonup u_i\nabla p_i[u] \quad \text{weakly in } L^1(Q_T).$$

Estimate (30) shows that the convergence holds in $L^q(Q_T)$. This convergence as well as (28) and (29) allow us to perform the limit $\varepsilon \to 0$ in the weak formulation of (1), proving that $u$ solves (1) and (3).

## Appendix A. Auxiliary results

We recall the Young convolution inequality (the proof in [8, Theorem 4.33] also applies to the torus).

**Lemma 7** (Young's convolution inequality)**.** *Let $1 \le p \le q \le \infty$ be such that $1 + 1/q = 1/p + 1/r$ and let $K \in L^r(\mathbb{T}^d)$ (extended periodically to $\mathbb{R}^d$). Then for any $v \in L^p(\mathbb{T}^d)$,*

$$\bigg\|\int_{\mathbb{T}^d}K(\cdot-y)v(y)dy\bigg\|_{L^q(\mathbb{T}^d)} \le \|K\|_{L^r(\mathbb{T}^d)}\|v\|_{L^p(\mathbb{T}^d)}.$$

The next result is a consequence of Vitali's lemma and is well known. We recall it for the convenience of the reader.

**Lemma 8.** *Let $\Omega \subset \mathbb{R}^d$ $(d \geq 1)$ be a bounded domain, $1 < p < \infty$, and $u_\varepsilon$, $u \in L^1(\mathbb{T}^d)$ be such that $(u_\varepsilon)$ is bounded in $L^p(\Omega)$ and $u_\varepsilon \to u$ a.e. in $\Omega$. Then $u_\varepsilon \to u$ strongly in $L^r(\Omega)$ for all $1 \leq r < p$.*

*Proof.* We have for any $M > 0$,

$$\int_{\{u_\varepsilon \geq M\}} |u_\varepsilon|^r dx = \int_{\{u_\varepsilon \geq M\}} |u_\varepsilon|^p |u_\varepsilon|^{-(p-r)} dx \leq M^{-(p-r)} \int_\Omega |u_\varepsilon|^p dx \leq CM^{-(p-r)} \to 0,$$

as $M \to \infty$. Thus, $(u_\varepsilon)$ is uniformly integrable. Since convergence a.e. implies convergence in measure, we conclude with Vitali's convergence theorem. □

The following lemma specifies conditions under which the limit of the product of two converging sequences can be identified.

**Lemma 9.** *Let $p > 1$ and let $u_\varepsilon \geq 0$, $u_\varepsilon \to u$ strongly in $L^p(\mathbb{T}^d)$, $v_\varepsilon \rightharpoonup v$ weakly in $L^p(\mathbb{T}^d)$, and $u_\varepsilon v_\varepsilon \rightharpoonup w$ weakly in $L^p(\mathbb{T}^d)$ as $\varepsilon \to 0$. Then $w = uv$.*

The lemma is trivial if $p \geq 2$. We apply it in Section 3 with $1 < p < 2$. Note that the strong convergence of $(u_\varepsilon)$ cannot be replaced by weak convergence. A simple counterexample is given by $u_\varepsilon(x) = \exp(2\pi \mathrm{i} x/\varepsilon) \rightharpoonup 0$, $v_\varepsilon(x) = \exp(-2\pi \mathrm{i} x/\varepsilon) \rightharpoonup 0$ weakly in $L^2(-1,1)$, but $u_\varepsilon v_\varepsilon \equiv 1 \neq 0 \cdot 0$.

*Proof.* We define the truncation function $T_1 \in C^2([0,\infty))$ satisfying $T_1(s) = s$ for $0 \leq s \leq 1$, $T_1(s) = 2$ for $s > 3$, and $T_1$ is nondecreasing and concave in the interval $[1,3]$. Furthermore, we set $T_k(s) = kT_1(s/k)$ for $s \geq 0$ and $k \in \mathbb{N}$. The strong convergence of $(u_\varepsilon)$ implies the existence of a subsequence that is not relabeled such that $u_\varepsilon \to u$ a.e. Hence, $T_k(u_\varepsilon) \to T_k(u)$ a.e. Since $T_k$ is bounded for fixed $k \in \mathbb{N}$, we conclude by dominated convergence that $T_k(u_\varepsilon) \to T_k(u)$ strongly in $L^r(\mathbb{T}^d)$ for any $r < \infty$. Because of the uniqueness of the limit, the convergence holds for the whole sequence. Thus, $T_k(u_\varepsilon)v_\varepsilon \rightharpoonup T_k(u)v$ weakly in $L^1(\mathbb{T}^d)$. Writing $\overline{z_\varepsilon}$ for the weak limit of a sequence $(z_\varepsilon)$ (if it exists), this result means that $\overline{T_k(u_\varepsilon)v_\varepsilon} = T_k(u)v$ and the assumption translates to $\overline{u_\varepsilon v_\varepsilon} = w$. Consequently, $w - T_k(u)v = \overline{(u_\varepsilon - T_k(u_\varepsilon))v_\varepsilon}$. Then we can estimate

$$\begin{aligned}
\|w - T_k(u)v\|_{L^1(\mathbb{T}^d)} &\leq \sup_{0<\varepsilon<1} \int_{\mathbb{T}^d} |u_\varepsilon - T_k(u_\varepsilon)||v_\varepsilon| dx \leq C \sup_{0<\varepsilon<1} \int_{\{|u_\varepsilon|>k\}} |u_\varepsilon||v_\varepsilon| dx \\
&\leq \frac{C}{k^{p-1}} \int_{\{|u_\varepsilon|>k\}} |u_\varepsilon|^p |v_\varepsilon| dx \leq \frac{C}{k^{p-1}} \int_{\mathbb{T}^d} (1 + |u_\varepsilon v_\varepsilon|^p) dx \leq \frac{C}{k^{p-1}}.
\end{aligned}$$

This shows that $T_k(u)v \to w$ strongly in $L^1(\mathbb{T}^d)$ and (for a subsequence) a.e. as $k \to \infty$. Since $T_k(u)v = uv$ in $\{|u| \leq k\}$ for any $k \in \mathbb{N}$ and $\mathrm{meas}\{|u| > k\} \leq \|u\|_{L^1(\mathbb{T}^d)}/k \to 0$, we infer in the limit $k \to \infty$ that $w = uv$ in $\mathbb{T}^d$. □

## Appendix B. Local cross-diffusion system

The existence of global weak solutions to the local system (1) and (3) in any bounded polygonal domain was shown in [28] by analyzing a finite-volume scheme. For completeness, we state the assumptions and the theorem and indicate how the result can be proved using the techniques of Section 3. We assume that $\Omega\subset\mathbb{R}^d$ ($d\ge 1$) is a bounded domain, $T>0$, and $u^0\in L^2(\Omega)$ satisfies $u_i^0\ge 0$ in $\Omega$ for $i=1,\ldots,n$. We set $Q_T=\Omega\times(0,T)$.

**Theorem 10** (Existence for the local system)**.** *Let $\sigma>0$, $a_{ij}\ge 0$, and let the matrix $(u_ia_{ij})$ be positively stable for all $u_i>0$, $i=1,\ldots,n$. Assume that there exist $\pi_1,\ldots,\pi_n>0$ such that $\pi_ia_{ij}=\pi_ja_{ji}$ for all $i,j=1,\ldots,n$. Then there exists a global weak solution to* (1) *and* (3)*, satisfying $u_i\ge 0$ in $Q_T$ and*

$$u_i\in L^2(0,T;H^1(\Omega))\cap L^{2+4/d}(Q_T),\quad \partial_tu_i\in L^q(0,T;W^{-1,q}(\Omega)),$$

*for $i=1,\ldots,n$, where $q=(d+2)/(d+1)$. The initial datum in* (1) *is satisfied in the sense of $W^{-1,q}(\Omega)$. Moreover, the following entropy inequalities are satisfied:*

$$\begin{aligned}&\frac{dH_1}{dt}+4\sigma\sum_{i=1}^n\int_\Omega\pi_i|\nabla\sqrt{u_i}|^2dx+\alpha\sum_{i=1}^n\int_\Omega|\nabla u_i|^2dx\le 0,\\ &\frac{dH_2^0}{dt}+\sum_{i=1}^n\int_\Omega\pi_iu_i|\nabla p_i[u]|^2dx+\alpha\sigma\sum_{i=1}^n\int_\Omega|\nabla u_i|^2dx\le 0,\end{aligned}\tag{32}$$

*where $\alpha>0$ is the smallest eigenvalue of $(\pi_ia_{ij})$ and $H_2^0(u):=\frac12\sum_{i,j=1}^n\int_\Omega\pi_ia_{ij}u_iu_jdx\ge 0$.*

We call a matrix *positively stable* if all eigenvalues have a positive real part. This condition means that (1) is parabolic in the sense of Petrovskii, which is a minimal condition to ensure the local solvability [1]. Inequalities (4)–(5) and (32) reveal a link between the entropy structures of the nonlocal and local systems. This link was explored recently in detail for related systems in [17].

*Proof.* If $\Omega$ is the torus, the theorem is a consequence of the localization limit (Theorem 5). If $\Omega$ is a bounded domain, the result can be proved by using the techniques of the proof of Theorem 2. In fact, the proof is simpler since the problem is local. The entropy identities are (formally)

$$\begin{aligned}&\frac{dH_1}{dt}+4\sigma\sum_{i=1}^n\int_\Omega\pi_i|\nabla\sqrt{u_i}|^2dx=-\sum_{i,j=1}^n\int_\Omega\pi_ia_{ij}\nabla u_i\cdot\nabla u_jdx,\\ &\frac{dH_2^0}{dt}+\sum_{i=1}^n\int_\Omega\pi_iu_i|\nabla p_i[u]|^2dx=-\sigma\sum_{i,j=1}^n\int_\Omega\pi_ia_{ij}\nabla u_i\cdot\nabla u_jdx.\end{aligned}\tag{33}$$

We claim that the matrix $(\pi_ia_{ij})$ is positive definite. Let $A_1:=\operatorname{diag}(u_i/\pi_i)$ and $A_2:=(\pi_ia_{ij})$. Then $A_1$ is symmetric and positive definite; by our assumptions, $A_2$ is symmetric and $A_1A_2=(u_ia_{ij})$ is positively stable. Therefore, by [15, Prop. 3], $A_2$ is positive definite. We infer that the right-hand sides in (33) are nonpositive, and we derive estimates for an approximate family of $u_i$ in $L^\infty(0,T;L^2(\Omega))$ and $L^2(0,T;H^1(\Omega))$. By the Gagliardo–Nirenberg inequality, this yields bounds for $u_i$ in $L^{2+4/d}(Q_T)$. Consequently, $u_i\nabla p_i[u]$ is bounded in $L^q(Q_T)$, where $q=(d+2)/(d+1)$ (we can even choose $q=4(d+2)/(3d+4)$), and $\partial_t u_i$ is bounded in $L^q(0,T;W^{-1,q}(\Omega))$. These estimates are sufficient to deduce from the Aubin–Lions lemma the relative compactness for the approximate family of $u_i$ in $L^2(Q_T)$. The limit in the approximate problem, similar to (12), shows that the limit satisfies (1) and (3). Finally, using the lower semicontinuity of convex functions and the norm, the weak limit in the entropy inequalities leads to (32). □

## References


[1] H. Amann. Nonhomogeneous linear and quasilinear elliptic and parabolic boundary value problems. In: H. J. Schmeisser and H. Triebel (eds.), Funct. Spaces Differ. Op. Nonlinear Anal., pp. 9–126. Teubner, Wiesbaden, 1993.

[2] F. Andreu-Vaillo, J. Mazón, J. Rossi, and J. Toledo-Melero. *Nonlocal Diffusion Problems*. Amer. Math. Soc., Providence, 2010.

[3] J. Baladron, D. Fasoli, O. Faugeras, and J. Touboul. Mean-field description and propagation of chaos in networks of Hodgkin-Huxley and FitzHugh-Nagumo neurons. *J. Math. Neurosci.* 2 (2012), no. 10, 50 pages.

[4] J. Berendsen, M. Burger, and J.-F. Pietschmann. On a cross-diffusion model for multiple species with nonlocal interaction and size exclusion. *Nonlin. Anal.* 159 (2017), 10–39.

[5] J. Berendsen, M. Burger, V. Ehrlacher, and J.-F. Pietschmann. Uniqueness of strong solutions and weak-strong stability in a system of cross-diffusion equations. *J. Evol. Eqs.* 20 (2020), 459–483.

[6] A. Bertozzi, T. Laurent, and J. Rosado. $L^p$ theory for the multidimensional aggregation equation. *Commun. Pure Appl. Math.* 64 (2011), 45–83.

[7] M. Braukhoff, X. Chen, and A. Jüngel. Corrigendum: Cross diffusion preventing blow-up in the two-dimensional Keller–Segel model. *SIAM J. Math. Anal.* 52 (2020), 2198–2200.

[8] H. Brézis. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Springer, New York, 2011.

[9] M. Briane and J. Casado-Díaz. A new div-curl result. Applications to the homogenization of elliptic systems and to the weak continuity of the Jacobian. *J. Diff. Eqs.* 260 (2016), 5678–5725.

[10] J. Buescu, A. C. Paixao, F. Garcia, and I. Lourtie. Positive-definiteness, integral equations and Fourier transforms. *J. Integral Eqs. Appl.* 16 (2004), 33–52.

[11] J. A. Carrillo, M. Di Francesco, A. Figalli, T. Laurent, and D. Slepčev. Global-in-time weak measure solutions and finite-time aggregation for nonlocal interaction equations. *Duke Math. J.* 156 (2011), 229–271.

[12] L. Chen, E. S. Daus, and A. Jüngel. Rigorous mean-field limit and cross diffusion. *Z. Angew. Math. Phys.* 70 (2019), no. 122, 21 pages.

[13] X. Chen and A. Jüngel. A note on the uniqueness of weak solutions to a class of cross-diffusion systems. *J. Evol. Eqs.* 18 (2018), 805–820.

[14] X. Chen and A. Jüngel. Weak-strong uniqueness of renormalized solutions to reaction-cross-diffusion systems. *Math. Models Meth. Appl. Sci.* 29 (2019), 237–270.

[15] X. Chen and A. Jüngel. When do cross-diffusion systems have an entropy structure? *J. Diff. Eqs.* 278 (2021), 60–72.

[16] C. M. Dafermos. The second law of thermodynamics and stability. *Arch. Ration. Mech. Anal.* 70 (1979), 167–179.

[17] H. Dietert and A. Moussa. Persisting entropy structure for nonlocal cross-diffusion systems. Submitted for publication, 2021. arXiv:2101.02893.

[18] M. Di Francesco and S. Fagioli. Measure solutions for non-local interaction PDEs with two species. *Nonlinearity* 26 (2013), 2777–2808.

[19] R. DiPerna. Uniqueness of solutions to hyperbolic conservation laws. *Indiana Univ. Math. J.* 28 (1979), 137–188.

[20] M. Dreher and A. Jüngel. Compact families of piecewise constant functions in $L^p(0,T;B)$. *Nonlin. Anal.* 75 (2012), 3072–3077.

[21] J. Fischer. Weak-strong uniqueness of solutions to entropy-dissipating reaction-diffusion equations. *Nonlin. Anal.* 159 (2017), 181–207.

[22] G. Galiano and J. Velasco. Well-posedness of a cross-diffusion population model with nonlocal diffusion. *SIAM J. Math. Anal.* 51 (2019), 2884–2902.

[23] D. Gandolfo, R. Rodriguez, and H. C. Tuckwell. Mean field analysis of large-scale interacting populations of stochastic conductance-based spiking neurons using the Klimontovich method. *J. Stat. Phys.* 166 (2017), 1310–1333.

[24] K. Hopf. Weak-strong uniqueness for energy-reaction-diffusion systems. Submitted for publication, 2021. arXiv:2102.02491.

[25] P.-E. Jabin and Z. Wang. Mean field limit for stochastic particle systems. In: N. Bellomo, P. Degond, and E. Tadmor (eds.). *Active Particles*, Vol. 1, pp. 379–402. Birkhäuser, Cham, 2017.

[26] A. Jüngel. The boundedness-by-entropy method for cross-diffusion systems. *Nonlinearity* 28 (2015), 1963-2001.

[27] A. Jüngel. *Entropy Methods for Diffusive Partial Differential Equations*. Springer Briefs Math., Springer, 2016.

[28] A. Jüngel and A. Zurek. A finite-volume scheme for a cross-diffusion model arising from interacting many-particle population systems. In: R. Klöfkorn, E. Keilegavlen, F. Radu, and J. Fuhrmann (eds.). *Finite Volumes for Complex Applications IX – Methods, Theoretical Aspects, Examples*. Springer, Cham, 2020, pp. 223-231.

[29] S. Mei, T. Misiakiewicz, and A. Montanari. Mean-field theory of two-layers neural networks: dimension-free bounds and kernel limit. *Proc. Machine Learning Res.* 99 (2019), 1–77.

[30] J. Mercer. Functions of positive and negative type and their connection with the theory of integral equations. *Phil. Trans. Roy. Soc. London A* 209 (1909), 415–446.

[31] D. Morale, V. Capasso, and K. Oelschläger. An interacting particle system modelling aggregation behavior: from individuals to populations. *J. Math. Biol.* 50 (2005), 49–66.

[32] C. Rao. Diversity and dissimilarity coefficients: a unified approach. *Theor. Popul. Biol.* 21 (1982), 24–43.

[33] C. Shannon. A mathematical theory of communication. *Bell Sys. Tech. J.* 27 (1948), 379–423.

[34] J. Sirignano and K. Spiliopoulos. Mean field analysis of neural networks: a law of large numbers. *SIAM J. Appl. Math.* 80 (2020), 725–752.

[35] C. Sogge. *Fourier Integrals in Classical Analysis*. Cambridge University Press, Cambridge, 1993.

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

Institute for Analysis and Scientific Computing, Vienna University of Technology, Wiedner Hauptstrasse 8–10, 1040 Wien, Austria
*E-mail address*: stefan.portisch@tuwien.ac.at

Institute for Analysis and Scientific Computing, Vienna University of Technology, Wiedner Hauptstrasse 8–10, 1040 Wien, Austria
*E-mail address*: antoine.zurek@tuwien.ac.at