Persisting entropy structure for nonlocal cross-diffusion systems

Helge Dietert, Ayman Moussa

Persisting entropy structure for nonlocal cross-diffusion systems

Helge Dietert* Ayman Moussa†

January 7, 2021

For cross-diffusion systems possessing an entropy (i.e. a Lyapunov functional) we study nonlocal versions and exhibit sufficient conditions to ensure that the nonlocal version inherits the entropy structure. These nonlocal systems can be understood as population models *per se* or as approximation of the classical ones. With the preserved entropy, we can rigorously link the approximating nonlocal version to the classical local system. From a modelling perspective this gives a way to prove a derivation of the model and from a PDE perspective this provides a regularisation scheme to prove the existence of solutions. A guiding example is the SKT model [18] and in this context we answer positively to a question raised by Fontbona and Méléard in [9] and thus provide a full derivation.

1. Introduction

1.1. Cross-diffusion systems with entropy structure

Our starting points are cross-diffusion systems of n species with densities $u = (u_i)_{1\le i\le n}$ solving a system

$$\partial_t u_i - \operatorname{div}\left(\sum_{j=1}^{n} a_{ij}(u)\nabla u_j\right) = 0, \qquad \text{for } i = 1, \dots, n, \tag{1}$$

*Email: helge.dietert@imj-prg.fr
Université de Paris and Sorbonne Université, CNRS, Institut de Mathématiques de Jussieu-Paris Rive Gauche (IMJ-PRG), F-75013, Paris, France
Currently on leave and working at
Institut für Mathematik, Universität Leipzig, D-04103 Leipzig, Germany
†Email: ayman.moussa@sorbonne-universite.fr
Sorbonne Université and Université de Paris, CNRS, Laboratoire Jacques-Louis Lions (LJLL), F-75005 Paris, France

on a domain Ω supplemented with boundary conditions and initial data u^{init}. Here a_{ij} are given scalar functions ($\mathbb{R}^n_{\geq 0} \to \mathbb{R}_{\geq 0}$) and the unknowns are the model densities u_i's, which are therefore expected to be non-negative. The matrix $A(u) := (a_{ij}(u))$ is called the *diffusion* matrix and is always assumed to be positive definite. As this work focuses on the entropy structure for the diffusion, we do not consider here any reaction terms.

Without any assumptions on the a_{ij}'s, the only estimate that we have on system (1) is the conservation of the overall mass, *i.e.*

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_\Omega u_i = 0,$$

for $i = 1, \dots, n$. Due to the severe non-linearity of the system, this sole control is not sufficient to obtain the existence of global solutions. Searching for a Lyapunov functional of the form

$$H(u) := \int_\Omega \sum_{i=1}^n h_i(u_i), \tag{2}$$

where $h_i \in \mathscr{C}^0(\mathbb{R}_{\geq 0}) \cap \mathscr{C}^2(\mathbb{R}_{>0})$, we find formally without boundary terms that

$$\frac{\mathrm{d}}{\mathrm{d}t} H(u) = -\int_\Omega \begin{pmatrix} \nabla u_1 \\ \vdots \\ \nabla u_n \end{pmatrix} \cdot M(u) \begin{pmatrix} \nabla u_1 \\ \vdots \\ \nabla u_n \end{pmatrix},$$

with $M : \mathbb{R}^n_{>0} \to \mathbb{R}^{n\times n}$ defined by

$$M(y) = \begin{pmatrix} h_1''(y_1) & 0 & \dots & 0 \\ 0 & h_2''(y_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & h_n''(y_n) \end{pmatrix} \begin{pmatrix} a_{11}(y) & a_{12}(y) & \dots & a_{1n}(y) \\ a_{21}(y) & a_{22}(y) & \dots & a_{2n}(y) \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}(y) & a_{n2}(y) & \dots & a_{nn}(y) \end{pmatrix}. \tag{3}$$

Hence we have a positive dissipation $I = -\mathrm{d}H/\mathrm{d}t$ if (the symmetric part of) M is positive semi-definite. This motivates the following definition, where the second part quantifies the dissipation.

Definition 1 (Entropy structure)**.** *We say that the system* (1) *has an* entropy structure *if there exist n functions $h_1, \dots, h_n \in \mathscr{C}^0(\mathbb{R}_{\geq 0}) \cap \mathscr{C}^2(\mathbb{R}_{>0})$ such that the corresponding matrix map $M : \mathbb{R}^n_{>0} \to \mathbb{R}^{n\times n}$ defined by* (3) *takes its values in the cone of positive definite matrices. We say that this entropy structure is* uniform *when there exist furthermore n functions $\alpha_1, \dots, \alpha_n : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ such that for all $(z, v) \in \mathbb{R}^n \times \mathbb{R}^n_{>0}$ it holds*

$$z^T \cdot M(v) z \geq \sum_{i=1}^n \alpha_i(v_i)^2 z_i^2. \tag{4}$$

For a given entropy structure the functions h_i's are called the entropy densities, *α_i's are the* dissipations *and the functional H defined in* (2) *is called the* entropy.

Remark 2. *From the assumed positive definiteness of the diffusion matrix A, it directly follows that for every entropy structure all the functions h_i, $i = 1, \dots, n$, are convex.*

Remark 3. *For typical examples, as the SKT system (5) below, the entropy densities h_i have diverging derivative towards the origin so that we define the matrix map M only for positive arguments. In this work we also take $\mathbb{R}_{\geq 0}$ for the range of the densities which is the most common case. In general the entropy structure can also be defined for bounded subsets of $\mathbb{R}$, cf. [11].*

Smooth solutions for the system (1) are known to exist, at least locally in time, thanks to the work of Amann [1] which gives also a criteria of explosion for such solutions. A part from the very specific case of triangular system [10], for global solutions the current literature allows only weak solutions and relies crucially on the entropy structure.

For an overview of such cross-diffusion systems we refer to Jüngel [11], which gives a list of examples in the introduction and also uses the quantified condition (4). Note that [11] allows in principle more general entropies but, apart from the volume-filling models, all examples have the additive form (2) required in this work.

A guiding example is the SKT system with densities u_1 and u_2

$$\begin{cases} \partial_t u_1 = \Delta\big((d_1 + d_{12}u_2)u_1\big) \\ \partial_t u_2 = \Delta\big((d_2 + d_{21}u_1)u_2\big) \end{cases} \tag{5}$$

with parameters $d_1, d_2, d_{12}, d_{21} \geq 0$. This system has been introduced by Shigesada, Kawasaki and Teramoto [18]. Writing the system in divergence form (1), the matrix $(a_{ij}(u))_{ij}$ reads

$$\begin{pmatrix} d_1 + d_{12}u_2 & d_{12}u_1 \\ d_{21}u_2 & d_2 + d_{21}u_1 \end{pmatrix}.$$

For non-negative solutions this matrix has non-negative trace and determinant. As remarked by Chen and Jüngel [3] the following entropy allows to symmetrize the system

$$H(u_1, u_2) := \int_{\mathbb{T}^d} \big(h_1(u_1) + h_2(u_2)\big), \tag{6}$$

with

$$h_1(z) = d_{21}\psi(z), \quad h_2(z) := d_{12}\psi(z), \quad \psi(z) = z\log(z) - z + 1. \tag{7}$$

Indeed, one checks that

$$M(u_1, u_2) = \begin{pmatrix} h_1''(u_1) & 0 \\ 0 & h_2''(u_2) \end{pmatrix} \begin{pmatrix} d_1 + d_{12}u_2 & d_{12}u_1 \\ d_{21}u_2 & d_2 + d_{21}u_1 \end{pmatrix} = d_{12}d_{21} \begin{pmatrix} \star & 1 \\ 1 & \star \end{pmatrix},$$

so that M is symmetric and still has non-negative determinant and trace. Thus M is positive semi-definite and forms with H an entropy structure again under the necessary condition that the solution is non-negative.

It is also known (see [11, 7, 12] for instance) that the previous entropy structure of the additive form (2) can be found for substantial generalization of (5) in the following general class of cross-diffusion systems

$$\begin{cases} \partial_t u_1 = \Delta\big(\mu_1(u_1, u_2)\, u_1\big), \\ \partial_t u_2 = \Delta\big(\mu_2(u_1, u_2)\, u_2\big), \end{cases} \tag{8}$$

where the non-linear functions μ_1 and μ_2 are assumed $\mathscr{C}^0(\mathbb{R}^2_{\geq 0}) \cap \mathscr{C}^1(\mathbb{R}^2_{>0})$ so that (8) can be written in divergence form (1) in order for the entropy structure to makes sense. We will propose in Section 2 a proof of existence for this type of generalisations of the SKT system, which does not follow the usual approximation procedure, whether it is by an entropic change of variable [11] or with a semi-discrete scheme [7]. For the existence of solutions the difficulty comes from the cross-diffusion effect so that we will focus on the case without self-diffusion (imposing that μ_i does not depend on u_i)

$$\begin{cases} \partial_t u_1 = \Delta\big(\mu_1(u_2)\, u_1\big), \\ \partial_t u_2 = \Delta\big(\mu_2(u_1)\, u_2\big). \end{cases} \tag{9}$$

In all these studies, one important difficulty in the use of the entropy structure to build global weak solution is the elaboration of an adequate approximation scheme which has to verify two mandatory conditions:

- compatibility with the entropy structure;
- non-negativeness of the solutions (which is needed for the entropy structure to make sense).

A natural temptation to approximate the system (8) is to smooth out the non-linearity by a mollifier. This has been first proposed by Lepoutre, Pierre, and Rolland [13] in which a regularised model was studied, but no rigorous link was established with the usual model. A few years later Fontbona and Méléard managed to produce by a stochastic derivation a convoluted regularisation of the system in Fontbona and Méléard [9]. It is important to note that their purpose was not as much to produce an adequate approximation scheme but to derive the SKT system from a particle model. However they did not manage to handle the last step of the derivation and they explicitly raised the question, whether it is possible to find in the limit of small regularisation the classical local cross-diffusion system. A partial answer in this direction is given in [16], where only the special case of triangular diffusion coefficients is handled. A recent contribution to the derivation of the system from a particle model has been made by Chen et al. [4], who perform the limit under the assumption of small cross-diffusion effects. The approach of this paper is therefore perturbative and does not allow for a treatment of the derivation in full generality as suggested by the method proposed in Fontbona and Méléard [9].

The systems obtained after the use of such mollification or convolution operator are called *nonlocal* because the diffusion rate of one species at a given point x does not depend anymore solely on the population's density at this place, but on a *space average* around it.

To the best of our knowledge, the current literature does not offer any example of persisting entropy structure for a nonlocal cross-diffusion systems. In this work we aim at giving sufficient conditions on the mollifying process for the entropy structure to persist on the nonlocal system. Our intuition takes its origin from the article [6] in which the first author of the current article exhibited an entropy structure for the SKT systems under a spatial discretization. It is a striking fact that this very discretization is everything but a nonlocal version of the SKT system. However we will see that the mechanism which allows the entropy to persist on this discretized system can be translated for nonlocal versions of it.

In the following Subsections 1.2 and 1.3, we introduce the nonlocal versions of the system to which our method applies. We then state the existence and convergence results in the following Subsection 1.4. In the remainder of the paper these results are then proved.

1.2. Regularisation on the torus

The idea comes from the work by Daus, Desvillettes, and Dietert [6], where the entropy structure was understood for the linear rates SKT model in a spatial discretisation. In this paper the intuition is to relate the entropy structure to the reversibility of a Markov chain modelling an N-particle system whose mean-field limit converges (formally) to the spatially discrete system.

Briefly, the idea in [6] is that, on a particle model with discrete space variable, the entropy structure is obtained by imposing that a pair of particles is jumping together with a suitable rate. Trying to use this idea for a nonlocal approximation, we intuitively want to make pairs of particle with a given distance jump together. In order to identify the pairs, we therefore take the convolution reflected between the two species.

For $\Omega = \mathbb{T}^d$ and a convolution kernel $\rho : \mathbb{T}^d \to \mathbb{R}_{\geq 0}$ this motivates the following regularisation of (5)

$$\begin{cases} \partial_t u_1 = \Delta\big((d_1 + d_{12}\, u_2 \star \rho)u_1\big) \\ \partial_t u_2 = \Delta\big((d_2 + d_{21}\, u_1 \star \check{\rho})u_2\big) \end{cases} \tag{10}$$

where $\check{\rho}$ is the reflected convolution kernel, i.e.

$$\check{\rho}(y) = \rho(-y).$$

The key-observation is that, due to this reflected choice, for any function φ one has formally (with the translation operator $\tau_y u_2 = u_2(\cdot - y)$)

$$\int_{\mathbb{T}^d} \varphi(u_2)\Delta\big((u_1 \star \check{\rho})u_2\big) = \int_{\mathbb{T}^d} \rho(y) \left\{ \int_{\mathbb{T}^d} \varphi(\tau_y u_2)\Delta\big(u_1 \tau_y u_2\big) \right\} \mathrm{d}y.$$

In particular, for H of the form (2) we have

$$-\frac{\mathrm{d}}{\mathrm{d}t} H(u_1, u_2) = \int_{\mathbb{T}^d} \rho(y) \left\{ \int_{\mathbb{T}^d} \begin{pmatrix} \nabla u_1 \\ \nabla \tau_y u_2 \end{pmatrix} \cdot M\Big(u_1, \tau_y u_2\Big) \begin{pmatrix} \nabla u_1 \\ \nabla \tau_y u_2 \end{pmatrix} \right\} \mathrm{d}y,$$

where M is the same matrix associated to H, as in the local case. In particular, since the kernel is non-negative, the entropy structure of the local case persists in the nonlocal system in the sense that H still defines a Lyapunov functional. The previous computation can be adapted to the generalizations of the SKT system (8) with the following caution: the spatial regularisation has to be applied after the nonlinearity, without affecting the self-diffusion. We have more precisely the following proposition.

Proposition 4. *Consider $\mu_1, \mu_2 \in \mathscr{C}^0(\mathbb{R}^2_{\geq 0}) \cap \mathscr{C}^1(\mathbb{R}^2_{>0})$ and the corresponding system* (8)*. If this system has an entropy H, then for any non-negative kernel ρ of integral* 1*, any solution of the following non-local system (τ_y is the translation operator)*

$$\begin{cases} \partial_t u_1 - \Delta \left[\int_{\mathbb{T}^d} \rho(y)\, \mu_1\big(u_1, \tau_y u_2\big)\big)\, \mathrm{d}y\, u_1 \right] = 0, \\ \partial_t u_2 - \Delta \left[\int_{\mathbb{T}^d} \check{\rho}(y)\, \mu_2\big(\tau_y u_1, u_2\big)\big)\, \mathrm{d}y\, u_2 \right] = 0 \end{cases} \tag{11}$$

satisfies formally

$$\frac{\mathrm{d}}{\mathrm{d}t} H(u_1(t), u_2(t)) \leq 0.$$

If the entropy structure of the system (8) *is furthermore assumed uniform with dissipation α_1 and α_2, then we have formally*

$$\frac{\mathrm{d}}{\mathrm{d}t} H(u_1(t), u_2(t)) + D(t) \leq 0,$$

where

$$D(t) := \int_{\mathbb{T}^d} \alpha_1(u_1(t))^2 |\nabla u_1(t)|^2 + \int_{\mathbb{T}^d} \alpha_2(u_2(t))^2 |\nabla u_2(t)|^2.$$

Proof. Denoting by h_1 and h_2 the entropy densities, we find by multiplying the first equation of (11) by $h_1'(u_1)$ and integrating over $\mathbb{T}^d$ that

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} \int_{\mathbb{T}^d} h_1(u_1)\, \mathrm{d}x &= -\int_{\mathbb{T}^d} h_1''(u_1) \nabla u_1 \cdot \left\{ \int_{\mathbb{T}^d} \rho(y) \nabla(\mu_1(u_1, \tau_y u_2) u_1)\, \mathrm{d}y \right\} \mathrm{d}x \\ &= -\int_{\mathbb{T}^d} \rho(y) \left\{ \int_{\mathbb{T}^d} h_1''(u_1) \nabla u_1 \cdot \nabla(\mu_1(u_1, \tau_y u_2) u_1)\, \mathrm{d}x \right\} \mathrm{d}y, \end{aligned}$$

where τ_y is the translation operator and ∇ acts on the x (not noted) variable only. We have a similar formula for the second equation, that is

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_{\mathbb{T}^d} h_2(u_2)\, \mathrm{d}x = -\int_{\mathbb{T}^d} \check{\rho}(y) \left\{ \int_{\mathbb{T}^d} h_1''(u_2) \nabla u_2 \cdot \nabla(\mu_2(\tau_y u_1, u_2) u_2)\, \mathrm{d}x \right\} \mathrm{d}y.$$

Intuitively speaking, we want to collect the pairs $u_1(x)$ and $u_2(x-y)$ in both expressions. This motivates in the double integral in the variables x, y of the last r.h.s. the change of variable $(x,y) \mapsto (z-w,-w)$. Using that the translation commutes with differential operators, we find

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{T}^d} h_2(u_2) = -\int_{\mathbb{T}^d} \check{\rho}(-w)\left\{\int_{\mathbb{T}^d} h_2''(\tau_w u_2)\nabla\tau_w u_2 \cdot \nabla(\mu_1(u_1,\tau_w u_2)\tau_w u_2)\mathrm{d}z\right\}\mathrm{d}w.$$

Since $\check{\rho}(-w) = \rho(w)$, renaming the variables as before, we can collect both contributions as

$$\frac{\mathrm{d}}{\mathrm{d}t}H(u_1(t),u_2(t)) = -\int_{\mathbb{T}^d}\rho(y)\left\{\int_{\mathbb{T}^d}\begin{pmatrix}\nabla u_1\\ \nabla\tau_y u_2\end{pmatrix}\cdot M\Big(u_1,\tau_y u_2\Big)\begin{pmatrix}\nabla u_1\\ \nabla\tau_y u_2\end{pmatrix}\right\}\mathrm{d}y,$$

where M is given by (3), the coefficients of the matrix A being the one used to write (8) in divergence form (1). The fact that H is a Lyapunov functional and the precised dissipation in case of uniform entropy follow (for the latter, we use the normalization of ρ). $\square$

If there is no self diffusion in the generalized system (8) like in (9), the nonlocal system (11) becomes simply

$$\begin{cases}\partial_t u_1 = \Delta\big((\mu_1(u_2)\star\rho)u_1\big),\\ \partial_t u_2 = \Delta\big((\mu_2(u_1)\star\check{\rho})u_2\big).\end{cases} \tag{12}$$

Thus, compared to Fontbona and Méléard [9], the spatial regularisation is applied after the nonlinearity, while they do it the opposite way in their stochastic derivation. However, we note that in the fundamental case of the (linear) SKT system (5), we get the same system.

For these systems with the Laplace structure an other important role is played by the duality estimates, see [7, 12, 16]. An advantage of the previous scheme is that these duality estimates naturally continue to work in the nonlocal versions.

Remark 5. *In the regularisation* (11) *the rate* $(\mu_i)_{i=1,2}$ *is averaged with respect to the cross-diffusion influence but a possible nonlinear self-diffusion is not regularised. However, a nonlinear self-diffusion tends to improve the entropy-dissipation estimates and we thus focus on cases without self-diffusion.*

In particular for stochastic derivations it is still interesting to also regularise the self-diffusion. However, in a general setting this destroys the entropy structure and we need a compatibility with the entropy structure. For this consider a symmetric kernel σ*, i.e.* $\check{\sigma} = \sigma$*, and assume that* (8) *can be written as*

$$\begin{cases}\partial_t u_1 = \Delta\big((\mu_1(u_2)+\kappa_1(u_1))\,u_1\big),\\ \partial_t u_2 = \Delta\big((\mu_2(u_1)+\kappa_2(u_2))\,u_2\big),\end{cases}$$

where the system without the κ *has an entropy structure with an entropy* H *consisting* *of h_1 and h_2 and matrix map M. We then propose the regularisation*

$$\begin{cases} \partial_t u_1(x) = \Delta_x \left[\left(\int_{y\in\mathbb{T}^d} \rho(x-y)\, \mu_1\big(u_2(y)\big)\,\mathrm{d}y + \int_{y\in\mathbb{T}^d} \sigma(x-y)\, \kappa_1\big(u_1(y)\big)\,\mathrm{d}y \right) u_1(x) \right] \\ \partial_t u_2(y) = \Delta_y \left[\left(\int_{x\in\mathbb{T}^d} \rho(x-y)\, \mu_2\big(u_1(x)\big)\,\mathrm{d}x + \int_{x\in\mathbb{T}^d} \sigma(x-y)\, \kappa_2\big(u_2(x)\big)\,\mathrm{d}x \right) u_2(y) \right] \end{cases}$$

For the dissipation we then find

$$-\frac{\mathrm{d}}{\mathrm{d}t} H(u_1, u_2) = I_\mu + I_1 + I_2$$

where I_μ is the dissipation with $\kappa_1 = \kappa_2 = 0$ and thus has a good sign. The new terms are after using symmetrisation $\check{\sigma} = \sigma$

$$I_i = \frac{1}{2} \int_{x\in\Omega} \int_{y\in\mathbb{R}^d} \sigma(y) \begin{pmatrix} \nabla u_i(x) \\ \nabla u_i(x-y) \end{pmatrix} \cdot N_i \begin{pmatrix} \nabla u_i(x) \\ \nabla u_i(x-y) \end{pmatrix} \mathrm{d}y\,\mathrm{d}x$$

with

$$N_i = \begin{pmatrix} h_i''(u_i(x)) & 0 \\ 0 & h_i''(u_i(x-y)) \end{pmatrix} \begin{pmatrix} \kappa_1(u_i(x-y)) & u_i(x)\kappa_1'(u_i(x-y)) \\ u_i(x-y)\kappa_1'(u_i(x)) & \kappa_1(u_i(x)) \end{pmatrix}$$

for $i = 1, 2$.

Hence H is still an entropy if $(N_i)_{i=1,2}$ are always positive semi-definite which gives an extra condition on the system. We note, however, that for the studied SKT system (5) this condition is always satisfied under the natural assumption that effect on the other species is of the same form as the self-diffusion effect, i.e. that it takes the form

$$\begin{cases} \partial_t u_1 = \Delta\big((d_1 + d_{12}u_2^\alpha + d_{11}u_1^\beta)u_1\big), \\ \partial_t u_2 = \Delta\big((d_2 + d_{21}u_1^\beta + d_{22}u_2^\alpha)u_2\big), \end{cases}$$

for constants $d_1, d_2, d_{11}, d_{12}, d_{21}, d_{22}, \alpha, \beta \in \mathbb{R}_{>0}$ with $\alpha\beta \leq 1$ (see, e.g., [7] for the discussion of the local case).

Remark 6. *Consider the SKT system on a regular bounded set $\Omega \in \mathbb{R}^d$ with* constant *Dirichlet boundary conditions*[1]*; that is, for $x \in \partial\Omega$ we impose for all times $t \in \mathbb{R}_+$ $(u_1, u_2)(t, x) = (b_1, b_2) \in \mathbb{R}^2_{>0}$ and we will now explain how the previous regularisation scheme on the torus can be used to design an approximation procedure for this type of boundary conditions by a penalisation method.*

The first remark is that, whenever the system has an entropy structure, each elementary functions $h_i : \mathbb{R}_{\geq 0} \to \mathbb{R}$ can be exchanged with $z \mapsto h_i(z) + \ell_i(z)$ for any affine functions ℓ_i, without changing the entropy estimate (which solely relies on second derivatives). W.l.o.g. we can therefore assume $h_i(b_i) = h_i'(b_i) = 0$. By convexity h_i reaches therefore its minimum at b_i.

[1] We need to assume constant boundary data in order to avoid boundary terms in the entropy estimate.

Now for an approximating system take N large enough such that $\overline{\Omega} \subset\subset (-N,N)^d$ and identify the hypercube $[-N,N]^d$ with the flat torus $\mathbb{T}_N^d := (\mathbb{R}/2N\mathbb{Z})^d$ and consider for $\varepsilon > 0$ the approximating system

$$\partial_t u_{1,\varepsilon} - \Delta((d_1 + d_{12} u_{2,\varepsilon} \star \rho_\varepsilon) u_{1,\varepsilon}) = -\frac{1}{\varepsilon}(u_{1,\varepsilon} - b_1)\mathbf{1}_{\mathbb{T}_N^d \setminus \Omega},$$
$$\partial_t u_{2,\varepsilon} - \Delta((d_2 + d_{21} u_{1,\varepsilon} \star \check{\rho}_\varepsilon) u_{2,\varepsilon}) = -\frac{1}{\varepsilon}(u_{2,\varepsilon} - b_2)\mathbf{1}_{\mathbb{T}_N^d \setminus \Omega}.$$

The key idea is that $h_i'(z)(z - b_i)$ is a non-negative function vanishing at only one point: multiplying respectively the equations by $h_i'(u_{i,\varepsilon})$, we recover an entropy estimate with an extra penalisation term (which has the good sign). At the limit $\varepsilon \to 0$ the penalisation forces therefore $u_i = b_i$ outside Ω and thus on $\partial\Omega$.

1.3. General regularisation scheme

In the previous subsection we considered the special case of two species on the torus and in this subsection we will discuss the case of several species on general domains Ω. In this we will see that the specific Laplace structure as in (8) is not preserved and we have the general divergence structure as in (1), see Remark 8.

For the two densities case on the torus, we used the convolution in order to define how a pair is interacting in the cross-diffusion. In the general case of n densities on a domain Ω with a possible boundary, the suitable generalisation is a kernel $K : \Omega^n \to \mathbb{R}_{\geq 0}$ between all densities and the intuitive idea is that the cross-diffusion between the densities $u_1(x_1), u_2(x_2), \dots, u_n(x_n)$ at positions $x_1, x_2, \dots, x_n \in \Omega$ happens with the intensity $K(x_1, x_2, \dots, x_n)$. The idea of using a kernel on a bounded domain has been proposed in [13], where K is the fundamental solution the (Neumann) operator $\mathrm{Id} - \delta\Delta$ with $0 < \delta \ll 1$. However, the authors kept the Laplace structure and applied the regularisation before the nonlinearity so that the entropy structure was lost, see Remark 8 below.

At the boundary such a general tuple cannot diffuse freely if we impose no-flux boundary conditions. Hence in order to rule out boundary terms we further assume that

$$K(x_1, \dots, x_n) = 0 \qquad \text{if } x_i \in \partial\Omega \text{ for } i = 1, \dots, n. \tag{13}$$

A family of kernel K^ϵ for $\epsilon > 0$ then yields an approximation of the local system if the kernel is concentrating on the diagonal as $\epsilon \to 0$, i.e. for a species $i = 1, \dots, n$, a point $x_i \in \Omega$ and a sufficiently nice test function $\phi : \Omega^n \to \mathbb{R}$ it holds that

$$\prod_{j \neq i} \int_{x_j \in \Omega} \mathrm{d}x_j K^\epsilon(x_1, \dots, x_n)\, \phi(x_1, \dots, x_n) \to \phi(x_i, \dots, x_i) \qquad \text{as } \epsilon \to 0,$$

where we introduced the notation $\prod_{j \neq i} \int_{x_j \in \Omega} \mathrm{d}x_j$ to denote the repeated integral over all coordinates x_j with $j \neq i$, i.e.

$$\prod_{j \neq i} \int_{x_j \in \Omega} \mathrm{d}x_j := \int_{x_1 \in \Omega} \mathrm{d}x_1 \cdots \int_{x_{i-1} \in \Omega} \mathrm{d}x_{i-1} \int_{x_{i+1} \in \Omega} \mathrm{d}x_{i+1} \cdots \int_{x_n \in \Omega} \mathrm{d}x_n.$$

A natural candidate of such kernels K^ϵ is a smoothing of

$$K^\epsilon(x_1,\dots,x_n) = C_\epsilon 1_{|x_i-x_j|\le\epsilon, i,j=1,\dots,n}$$

with a cutoff towards the boundary and a suitable constant C_ϵ.

With this we can state our proposed general regularisation.

Proposition 7. *Let $n\in\mathbb{N}$ be the number of densities and assume rates $(a_{ij})_{i,j=1,\dots,n}$ such that the local system (1) has an entropy H.*

For a constant $\epsilon > 0$, a domain $\Omega \in \mathbb{R}^d$ and a kernel $K : \Omega^n \to \mathbb{R}_{\ge 0}$ satisfying (13), the evolution of densities $u_1,\dots,u_n : \Omega \to \mathbb{R}_{\ge 0}$ in time t by the nonlocal system $(i = 1,\dots,n)$

$$\begin{aligned}&\partial_t u_i(x_i) - \epsilon\Delta u_i(x_i)\\ &\quad - \mathrm{div}_{x_i}\left(\prod_{k\neq i}\int_{x_k\in\Omega} \mathrm{d}x_k\, K(x_1,\dots,x_n)\sum_{j=1}^n a_{ij}\big(u_1(x_1),\dots,u_n(x_n)\big)\nabla u_j(x_j)\right)\\ &= 0\end{aligned} \tag{14}$$

supplemented in the case of boundaries with von Neumann boundary conditions

$$n\cdot\nabla u_i(x) = 0 \text{ for } x\in\partial\Omega \tag{15}$$

satisfies formally

$$\frac{\mathrm{d}}{\mathrm{d}t}H(u_1(t),\dots,u_n(t)) \le 0.$$

In the case of uniform dissipations $(\alpha_i)_i$ it holds that

$$\frac{\mathrm{d}}{\mathrm{d}t}H(u_1(t),\dots,u_n(t)) + D(t) \le 0.$$

where

$$D(t) = \sum_{i=1}^n \int_\Omega \left[\epsilon\, h_i''(u_i(x)) + \alpha_i(u_i(x))^2 w_i(x)\right]\ |\nabla u_i(x)|^2\,\mathrm{d}x$$

with the weights $(w_i)_{i=1,\dots,n}$ defined as

$$\prod_{j\neq i}\int_{x_j\in\Omega}\mathrm{d}x_j K(x_1,\dots,x_n) = w_i(x_i) \ge 0, \qquad \forall x_i\in\Omega. \tag{16}$$

Here we added a small global self-diffusion with ϵ in order to compensate that the kernel K vanishes at the boundary so that we can obtain global regularity estimates for the regularised system. By the assumption (13) the imposed von Neumann boundary conditions imply zero-flux boundary conditions.

Proof. In order to obtain the estimate, the idea is to collect the interaction in a tuple $u_1(x_1), \dots, u_n(x_n)$. We then find for the dissipation

$$
\begin{aligned}
&-\frac{\mathrm{d}}{\mathrm{d}t} H(u(t)) \\
&= -\sum_{i=1}^{n} \frac{\mathrm{d}}{\mathrm{d}t} \int_{x_i \in \Omega} h_i(x_i)\,\mathrm{d}x_i \\
&= \sum_{i=1}^{n} \int_{x_i} \nabla u_i(x_i)\, h_i''(u_i(x_i)) \\
&\qquad \cdot \left[\epsilon + \prod_{k \neq i} \int_{x_k \in \Omega} \mathrm{d}x_k\, K(x_1, \dots, x_n) \sum_{j=1}^{n} a_{ij}\big(u_1(x_1), \dots, u_n(x_n)\big) \nabla u_j(x_j) \right] \mathrm{d}x_i \\
&= \int_{x_1} \mathrm{d}x_1 \cdots \int_{x_n} \mathrm{d}x_n K(x_1, \dots, x_n) \begin{pmatrix} \nabla u_1(x_1) \\ \vdots \\ \nabla u_n(x_n) \end{pmatrix} \cdot M\big(u_1(x_1), \dots, u_n(x_n)\big) \begin{pmatrix} \nabla u_1(x_1) \\ \vdots \\ \nabla u_n(x_n) \end{pmatrix} \\
&\qquad + \epsilon \int_{x \in \Omega} \sum_{i=1}^{n} h_i''(u_i(x))\, |\nabla u_i(x)|^2 \,\mathrm{d}x,
\end{aligned}
$$

where M is the matrix from the entropy structure, Definition 1, and the boundary terms vanish due to the von Neumann boundary condition and (13).

By the assumed sign of the matrix M and the lower bound by α_i, respectively, the result follows. □

Remark 8. *The previous regularisation* (10) *for two species in the simple setting* $\Omega = \mathbb{R}^d$ *or* $\Omega = \mathbb{T}^d$ *is exactly recovered by setting* $K(x_1, x_2) = \rho(x_1 - x_2)$ *and dropping the normal diffusion with* ϵ*.*

This leaves the question whether the Laplace structure of a system of a the form (8) *can be preserved in the nonlocal version. Applying the regularisation procedure for a general kernel* K*, we can rewrite the regularised evolution in the Laplace structure if*

$$\nabla_x K(x, y) = -\nabla_y K(x, y).$$

This, however, is only true if K *has a convolution structure and thus does not work for domains with boundaries. Indeed we find for the two species system* (8) *the regu-*

larisation

$$\begin{cases} \partial_t u_1(x) - \epsilon\Delta u_1 - \Delta_x \left[\int_{y\in\Omega} K(x,y)\,\mu_1\big(u_1(x),u_2(y)\big)\,\mathrm{d}y\,u_1(x)\right] \\ \qquad\qquad = -\nabla_x \left[\int_{y\in\Omega} \big[(\partial_x K)(x,y) + (\partial_y K)(x,y)\big]\mu_1\big(u_1(x),u_2(y)\big)\,\mathrm{d}y\,u_1(x)\right], \\ \partial_t u_2(y) - \epsilon\Delta u_2 - \Delta_y \left[\int_{x\in\Omega} K(x,y)\,\mu_2\big(u_1(x),u_2(y)\big)\,\mathrm{d}x\,u_2(y)\right] \\ \qquad\qquad = -\nabla_y \left[\int_{x\in\Omega} \big[(\partial_x K)(x,y) + (\partial_y K)(x,y)\big]\mu_2\big(u_1(x),u_2(y)\big)\,\mathrm{d}x\,u_2(y)\right], \end{cases} \tag{17}$$

which contains corrector terms for the defect of the convolution structure on the RHS.

For the linear rate SKT system, this matches the regularisation following the discrete structure in [6], where we identified the entropy with the reversibility of a corresponding Markov chain, see Appendix A.

1.4. Results

A strong motivation of this regularisation is the derivation of cross-diffusion models from a particle model and indeed we find for the linear rate SKT model that Fontbona and Méléard [9] derived our approximate system. We leave a general derivation from particle models for future work.

Another future direction is the study of the gradient flow structure. Formally, the original local system often has a gradient flow structure which in most studies is only used in the form of the dissipation inequality inequality (an exception is Zinsl and Matthes [20]). Having found a regularisation, we plan for future work to investigate the gradient flow formulation of the nonlocal system and the limit towards the local system. Such limits of gradient flows are an active field and we only mention [17, 2, 14] as starting points.

In this work we show by PDE methods the existence of solutions for the approximation system, where we see the effectiveness of the regularisation. Subsequently, we prove the limit of the approximating system and thus provide an alternative to other existence results based on the boundedness-by-entropy method [11, 5].

Our first result shows that the regularisation (12) for (9) is sufficient to find solutions satisfying the entropy-dissipation inequality. It will be clear from the proof below that the diffusivity μ_i's could be assumed sublinear, instead of being controlled by the entropy densities. The self-diffusion could be included *via* the more general approximation (11) (for which there is a similar existence result) but we have chosen to avoid it to simplify the presentation.

Theorem 9. *Consider the generalised SKT system* (9) *with* $\mu_1, \mu_2 \in \mathscr{C}^0(\mathbb{R}_{\geq 0}) \cap \mathscr{C}^1(\mathbb{R}_{>0})$*. Assume that it admits a uniform entropy structure with entropy* H*, entropy densities* $h_1, h_2 \in \mathscr{C}^0(\mathbb{R}_{\geq 0}) \cap \mathscr{C}^2(\mathbb{R}_{>0})$ *and dissipations* $\alpha_1, \alpha_2 : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$*. Assume* *furthermore that for two positive constants* δ, A *such that for all* $z \in \mathbb{R}_{\geq 0}$

$$\delta \leq \mu_1(z) \leq \mathrm{A}(1 + h_2(z)) \quad \textit{and} \quad \delta \leq \mu_2(z) \leq \mathrm{A}(1 + h_1(z)). \tag{18}$$

Fix $\rho \in \mathscr{C}^2(\mathbb{T}^d)$ *non-negative having integral* 1 *over* $\mathbb{T}^d$*, and a bounded initial data* $u_1^{\mathrm{init}}, u_2^{\mathrm{init}}$ *satisfying for some positive constant* γ

$$\gamma \leq u_i^{\mathrm{init}} \leq \gamma^{-1},$$

so that $H^{\mathrm{init}} := H(u_1^{\mathrm{init}}, u_2^{\mathrm{init}}) < +\infty$*. Then, there exists positive functions*

$$u_1, u_2 \in \mathscr{C}^0([0,T]; \mathrm{L}^2(\mathbb{T}^d)) \cap \mathrm{L}^2(0,T; \mathrm{H}^1(\mathbb{T}^d)) \cap \mathrm{L}^\infty(0,T; \mathrm{L}^\infty(\mathbb{T}^d)), \tag{19}$$

such that (u_1, u_2) *is a distributional solution to the system* (12) *initiated by* $(u_1^{\mathrm{init}}, u_2^{\mathrm{init}})$*. This solution* (u_1, u_2) *satisfies furthermore the following estimates for* $i = 1, 2$*:*

- conservation of the mass: $u_i \in \mathscr{C}^0([0,T]; \mathrm{L}^1(\mathbb{T}^d))$ *and for* $t \in [0,T]$

$$\int_{\mathbb{T}^d} u_i(t) = \int_{\mathbb{T}^d} u_i^{\mathrm{init}}. \tag{20}$$

- entropy estimate: $h_i(u_i) \in \mathscr{C}^0([0,T]; \mathrm{L}^1(\mathbb{T}^d))$ *and for* $t \in [0,T]$

$$H(u_1(t), u_2(t)) + \int_0^t D(s)\,\mathrm{d}s \leq H^{\mathrm{init}}, \tag{21}$$

where

$$D(t) := \int_{\mathbb{T}^d} \alpha_1(u_1(t))^2 |\nabla u_1(t)|^2 + \alpha_2(u_2(t))^2 |\nabla u_2(t)|^2.$$

- maximum principle:

$$\gamma \exp\left(-\mathrm{A}\mathrm{B}_{T,\mathrm{init}} \|\Delta\rho\|_{\mathrm{L}^\infty(\mathbb{T}^d)}\right) \leq u_i \leq \gamma^{-1} \exp\left(\mathrm{A}\mathrm{B}_{T,\mathrm{init}} \|\Delta\rho\|_{\mathrm{L}^\infty(\mathbb{T}^d)}\right), \tag{22}$$

where

$$\mathrm{B}_{T,\mathrm{init}} := T(1 + H^{\mathrm{init}}).$$

- duality estimate:

$$\int_{Q_T} \left(\left[\mu_1(u_2) \star \rho\right] u_1 + \left[\mu_2(u_1) \star \check{\rho}\right] u_2\right) (u_1 + u_2) \\ \lesssim_d (1 + 2\mathrm{A}\mathrm{B}_{T,\mathrm{init}}) \left(\int_{\mathbb{T}^d} (u_1^{\mathrm{init}})^2 + \int_{\mathbb{T}^d} (u_2^{\mathrm{init}})^2\right), \tag{23}$$

where the constant behind $\lesssim_d$ *depends only on the dimension* d*.*

Remark 10. *The upper-bound in assumption* (18) *is natural for many cross-diffusion systems. For instance if* μ_1 *and* μ_2 *are given by power-laws (as in [7]), the entropy densities are precisely given by the same exponents (with an exception for the linear case). See also Remark 12.*

As a by-product of this existence result, we recover the (known) existence of weak solutions to the generalized SKT system (9).

Theorem 11. *Consider the assumptions of Theorem 9, for a sequence of non-negative functions* $(\rho_n)_n \in \mathscr{C}^2(\mathbb{T}^d)^{\mathbb{N}}$ *which converges weakly towards the Dirac mass, with dissipation rates* α_1 *and* α_2 *vanishing on a set of measure* 0*. Assume furthermore that the diffusivities are strictly subquadratic or controlled by the entropy densities, that is*

$$\lim_{z\to+\infty} \frac{\mu_1(z)}{h_2(z)\vee z^2} + \frac{\mu_2(z)}{h_1(z)\vee z^2} = 0. \tag{24}$$

Then, the corresponding sequence of solutions $(u_{1,n}, u_{2,n})_n$ *given by Theorem 9 converges (up to a subsequence) in* $\mathrm{L}^1(Q_T)$ *towards a weak global solution* (u_1, u_2) *of the SKT system which satisfies for a.e.* $t \in [0,T]$ *the conservation of the mass* (20)*, the entropy estimate* (21) *and the following duality estimate*

$$\int_{Q_T} \big(\mu_1(u_2)u_1 + \mu_2(u_1)u_2\big)(u_1+u_2)$$
$$\lesssim_d (1 + 2\mathrm{AB}_{T,\mathrm{init}}) \int_{\mathbb{T}^d} (u_1^{\mathrm{init}})^2 + \int_{\mathbb{T}^d} (u_2^{\mathrm{init}})^2. \tag{25}$$

Remark 12. *The assumption* (24) *is crucial to avoid any concentration in the non-linearities of the system. However, in practice (see for instance the power-law case in [7]) the control of gradients of the entropy estimate gives raise (by Sobolev embedding) to another estimate on* $\mu_1(u_2)$ *and* $\mu_2(u_1)$*.*

In a general setting we described the regularisation scheme (14), for which we can state the following existence result.

Theorem 13. *Consider a cross-diffusion system* (1) *for* n *species and rates* $a_{ij} \in \mathscr{C}^0(\mathbb{R}^n_{\geq 0})$*,* $i,j = 1,\ldots,n$*. Assume that it admits a uniform entropy structure with entropy* H*, entropy densities* $h_1,\ldots,h_n \in \mathscr{C}^0(\mathbb{R}_{\geq 0})\cap\mathscr{C}^2(\mathbb{R}_{>0})$ *and dissipations* $\alpha_1,\ldots,\alpha_n : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$*.*

For $i \neq j$ *define* $\tilde{a}_{ij} : \mathbb{R}^n_{\geq 0} \to \mathbb{R}$ *by*

$$\partial_j \tilde{a}_{ij}(v) = a_{ij}(v), \quad v \in \mathbb{R}^n_{\geq 0}, \qquad \text{and} \qquad \tilde{a}_{ij}(v) = 0, \quad \text{if } v_j = 0.$$

Suppose that $v \mapsto \tilde{a}_{ij}(v)$ *is continuously differentiable with respect to* v_i *and that there exists a constant* A *such that for all* $v \in \mathbb{R}^n_{\geq 0}$ *and* $i = 1,\ldots,n$

$$a_{ii}(v) \leq A\,(1 + h_1(v_1) + \cdots + h_n(v_n)),$$
$$\frac{\tilde{a}_{ij}(v)}{v_i} \leq A\,(1 + h_1(v_1) + \cdots + h_n(v_n)),$$

and
$$\partial_i \tilde a_{ij}(v) \le A\,(1 + h_1(v_1) + \cdots + h_n(v_n)).$$

Let $\Omega \in \mathbb{R}^d$ be a domain with piecewise $\mathscr{C}^1$ boundary and $K \in \mathscr{C}^2_c(\Omega^n)$ be a nonnegative kernel satisfying (13)*. Further fix bounded initial data $u^{\mathrm{init}} = (u_1^{\mathrm{init}}, \dots, u_n^{\mathrm{init}})$ satisfying for some positive constant γ*

$$\gamma \le u_i^{\mathrm{init}} \le \gamma^{-1},$$

so that $H^{\mathrm{init}} := H(u^{\mathrm{init}}) < +\infty$. Then, there exists positive functions

$$u_1, \dots, u_n \in \mathscr{C}^0([0,T]; \mathrm{L}^2(\mathbb{T}^d)) \cap \mathrm{L}^2(0,T; \mathrm{H}^1(\mathbb{T}^d)) \cap \mathrm{L}^\infty(0,T; \mathrm{L}^\infty(\mathbb{T}^d)), \tag{26}$$

such that $(u_1, \dots, u_n)$ is a distributional solution to the system (17) *with initial data u^{init} and von Neumann boundary data* (15)*. Furthermore, the solution $u = (u_1, \dots, u_n)$ satisfies furthermore the following estimates for $i = 1, \dots, n$:*

- conservation of the mass: $u_i \in \mathscr{C}^0([0,T]; \mathrm{L}^1(\Omega))$ *and for* $t \in [0,T]$
$$\int_{\mathbb{T}^d} u_i(t) = \int_{\mathbb{T}^d} u_i^{\mathrm{init}}. \tag{27}$$

- entropy estimate: $h_i(u_i) \in \mathscr{C}^0([0,T]; \mathrm{L}^1(\Omega))$ *and for* $t \in [0,T]$
$$H(u(t))) + \int_0^t D(s)\,\mathrm{d}s \le H^{\mathrm{init}}, \tag{28}$$
$$D(t) = \sum_{i=1}^n \int_\Omega \left[\epsilon\, h_i''(u_i(x)) + \alpha_i(u_i(x))^2 w_i(x)\right]\ |\nabla u_i(x)|^2\,\mathrm{d}x$$
with the weights $(w_i)_{i=1,\dots,n}$ defined in (16).

- maximum principle:
$$\gamma \exp\left(-MT\right) \le u_i \le \gamma^{-1} \exp\left(MT\right), \tag{29}$$
where
$$M = 2A\,\max(\|K\|_\infty, \|\nabla K\|_\infty, \|\nabla^2 K\|_\infty)\,(|\Omega| + H^{\mathrm{init}}). \tag{30}$$

- ϵ regularity:
$$\sup_{t\in[0,T]} \|u_i(t,\cdot)\|^2_{\mathrm{L}^2(\Omega)} + \epsilon \int_0^T \|\nabla u_i(t,\cdot)\|^2_{\mathrm{L}^2(\Omega)} \le \exp\left[TM\Big(2 + \frac{1}{\epsilon}\Big)\right] \|u_i^{\mathrm{init}}\|^2_{\mathrm{L}^2(\Omega)}.$$

Under the assumption that the dissipation is big enough, one can conclude that the approximations converge to the local version. In the setting of their time-discretisation approximation scheme, [5] discusses possible conditions for such a convergence. Nevertheless, they need to treat the SKT case separately. As the SKT case is the motivating example, we therefore focus on the SKT case, where we replace the duality estimate with a positive self-diffusion. For n species with densities $u = (u_1, \dots, u_n)$ the SKT system corresponds to the evolution

$$\partial_t u_i = \Delta \left(d_i u_i + \sum_{j=1}^{n} d_{ij} u_j u_i \right) \tag{31}$$

with constants $d_1, \dots, d_n \geq 0$ and $(d_{ij})_{ij} \geq 0$. Further suppose that there exist weights $\pi_1, \dots, \pi_n \geq 0$ such that the diffusion coefficients satisfy the detailed balance condition

$$\pi_i d_{ij} = \pi_j d_{ji}, \qquad \text{for } i, j = 1, \dots, n, \tag{32}$$

see [6] for a discussion on the condition. Then the evolution (31) has an entropy structure with

$$h_i(z) = \pi_i \left(z \log z - z + 1 \right)$$

and dissipation

$$\alpha_i(z) = \pi_i d_{ii}.$$

Theorem 14. *Given a bounded domain $\Omega \subset \mathbb{R}^d$ with $\mathscr{C}^1$ boundary and an increasing sequence of sets $(A_m)_{m\in\mathbb{N}}$ with $A_m \subset\subset \Omega$ and $A_m \uparrow \Omega$ as $m \to \infty$. Suppose that there exists a constant c and extension operators $\mathcal{E}_m : W^{1,p}(A_m) \to W^{1,p}(\mathbb{R}^d)$ for $p = 2+2/d$ such that $\|\mathcal{E}_m(f)\|_{W^{1,p}(\mathbb{R}^d)} \leq c\|f\|_{W^{1,p}(A_m)}$.*

Assume a corresponding sequence of non-negative regularisation kernels $(K^m)_{m\in\mathbb{N}}$ in $\mathscr{C}_c^2(\Omega^n)$ for n densities satisfying (13) with weights w_i^m, $i = 1, \dots, n$, as in (16). Suppose that the weights always map to $[0, 1]$ and

$$w_i^m(x) = 1, \qquad \textit{for } x \in A_m \textit{ and } i = 1, \dots, n.$$

Moreover, suppose that K^m concentrates along the diagonal, i.e.

$$K^m(x_1, \dots, x_n) = 0, \quad \textit{if } |x_i - x_j| \geq \frac{1}{m} \textit{ for some } i, j = 1, \dots, n.$$

Consider the SKT system (31) for n densities with constants $d_1, \dots, d_n \geq 0$ and $(d_{ij})_{ij} \geq 0$ and weights $\pi_1, \dots, \pi_n \geq 0$ satisfying (32) and $d_{ii} > 0$ for $i = 1, \dots, n$ with initial data $u^{\text{init}} = (u_1^{\text{init}}, \dots, u_n^{\text{init}})$ with

$$\gamma \leq u_i \leq \gamma^{-1}$$

for $i = 1, \dots, n$ and a constant $\gamma \in \mathbb{R}_{>0}$.

Then there exists a sequence of $(\epsilon_m)_{m\in\mathbb{N}}$ with $\epsilon_m \downarrow 0$ as $m \to \infty$ such that the approximating solutions $(u^m)_m$ as constructed in Theorem 13 converge along a subsequence to u in $\mathrm{L}^q([0, T) \times \Omega; \mathbb{R}^n_{\geq 0})$ with $q = 2 + (1/2d)$. The limit u is a non-negative weak solution to (31) with the no-flux boundary conditions satisfying the entropy-dissipation inequality, i.e. for $\phi \in \mathscr{C}^\infty([0, T] \times \Omega)$ with $\phi(T, \cdot) \equiv 0$ and $i = 1, \dots, n$ it holds

$$-\int_0^T \int_\Omega u_i \, \partial_t \phi + \int_0^T \int_\Omega \left(\sum_{j=1}^{n} a_{ij}(u) \nabla u_j \right) \cdot \nabla \phi = \int_\Omega u_i^{\text{init}} \phi(0, \cdot),$$

where

$$a_{ij}(u) = \begin{cases} d_i + 2d_{ii}u_i + \sum_{j\neq i} d_{ij}u_j & \textit{if } i = j, \\ d_{ij}u_i & \textit{otherwise.} \end{cases}$$

Remark 15. *A sequence of such sets A_m can be constructed for locally Lipschitz domains. For this locally write the boundary as a graph of $d-1$ variables and locally then such a sequence can be constructed. For the construction of extensions we refer to the treatment of [8, Section 5.4] and [19, Section VI].*

2. The convolution scheme on the flat torus

This section is dedicated to the proofs of Theorem 9, Theorem 11 on the torus. For the domain, we introduce the notation

$$Q_T := [0,T) \times \mathbb{T}^d,$$

and start by recalling some useful results about the Kolmogorov equation, that is

$$\partial_t z - \Delta(\mu z) = G, \tag{33}$$

$$z(0,\cdot) = z^{\mathrm{init}}, \tag{34}$$

where G, μ and z^{init} are given and z is the unknown. Solutions will be understood in the following sense:

Definition 16. *Given a measurable function $\mu : Q_T \to \mathbb{R}$, and two square-integrable functions $G, z^{\mathrm{init}} : Q_T \to \mathbb{R}$ we say that $z \in \mathrm{L}^1(Q_T)$ is a distributional solution of* (33) *–* (34) *if $z\mu$ is integrable on Q_T and for all test function $\varphi \in \mathscr{D}(Q_T)$ there holds*

$$-\int_{Q_T} z(\partial_t\varphi + \mu\Delta\varphi) = \int_{\mathbb{T}^d} z^{\mathrm{init}}\varphi(0,\cdot) + \int_{Q_T} G\varphi.$$

In the second paragraph we proceed to the proof of Theorem 9. The last paragraph deals with the asymptotic limit stated in Theorem 11.

2.1. Reminder on the Kolmogorov equation

The following result is directly extracted from [16], more precisely merging results obtained in Theorem 3, Proposition 2 and Proposition 3 therein.

Theorem 17. *Fix $\mu \in \mathrm{L}^\infty(Q_T)$ such that $\inf_{Q_T}\mu > 0$. For any $z^{\mathrm{init}} \in \mathrm{L}^2(\mathbb{T}^d)$ there exists a unique solution z to* (33) *–* (34) *in the sense of Definition 16. This solution belongs to $\mathrm{L}^2(Q_T)$ and satisfies*

- maximum principle*: if G and z^{init} are non-negative, then so is z;*

- duality estimate*: $\mu^{1/2} z \in \mathrm{L}^2(Q_T)$ and*

$$\int_{Q_T} \mu z^2 \lesssim_d \left(1 + \int_{Q_T} \mu\right) \left(\int_{\mathbb{T}^d} (z^{\mathrm{init}})^2 + T \int_{Q_T} G^2\right),$$

where the constant behind $\lesssim_d$ depends only on the dimension;

- sequential stability*: for fixed G and z^{init} as above, the map $\mu \mapsto z$, restricted to those μ who are bounded and positively lower-bounded, is continuous in the $\mathrm{L}^1(Q_T)$ topology for the argument μ and the $\mathrm{L}^2(Q_T)$ topology for the image z.*

We will use two corollaries of the previous theorem.

Corollary 18. *Consider the assumptions of Theorem 17, with $G = 0$. If furthermore z^{init} is bounded with $\gamma \le z^{\mathrm{init}} \le \gamma^{-1}$ for some positive constant γ and if $\Delta\mu \in \mathrm{L}^1(0, T; \mathrm{L}^\infty(\mathbb{T}^d))$, then $z \in \mathrm{L}^\infty(Q_T)$ with the estimate*

$$\gamma \exp\left(-\int_0^T \|(\Delta\mu)^-(s)\|_{\mathrm{L}^\infty(\mathbb{T}^d)} \mathrm{d}s\right) \le z \le \gamma^{-1} \exp\left(\int_0^T \|(\Delta\mu)^+(s)\|_{\mathrm{L}^\infty(\mathbb{T}^d)} \, \mathrm{d}s\right),$$

where the exponents $^+$ and $^-$ refer to (respectively) the positive and negative parts.

Proof. Define

$$\begin{aligned}
\Phi(t) &:= \gamma^{-1} \exp\left(\int_0^t \|(\Delta\mu)^+(s)\|_{\mathrm{L}^\infty(\mathbb{T}^d)} \, \mathrm{d}s\right) - z, \\
\Psi(t) &:= z - \gamma \exp\left(-\int_0^t \|(\Delta\mu)^-(s)\|_{\mathrm{L}^\infty(\mathbb{T}^d)} \, \mathrm{d}s\right),
\end{aligned}$$

which satisfy

$$\begin{aligned}
(\partial_t \Phi - \Delta(\mu\Phi))(t, x) &= (\Phi + z)(t, x)(\|(\Delta\mu)^+(t)\|_{\mathrm{L}^\infty(\mathbb{T}^d)} - \Delta\mu(t, x)) \ge 0, \\
(\partial_t \Psi - \Delta(\mu\Psi))(t, x) &= (z - \Psi)(t, x)(\Delta\mu(t, x) + \|(\Delta\mu)^-(t)\|_{\mathrm{L}^\infty(\mathbb{T}^d)}) \ge 0.
\end{aligned}$$

The conclusion follows using the maximum principle of Theorem 17, since Φ and Ψ are initially non-negative. □

Corollary 19. *Consider the assumptions of Theorem 17, with $G = 0$. If furthermore $\Delta\mu \in \mathrm{L}^1(0, T; \mathrm{L}^\infty(Q_T))$, then $z \in \mathrm{L}^\infty(0, T; \mathrm{L}^2(\mathbb{T}^d)) \cap \mathrm{L}^2(0, T; \mathrm{H}^1(\mathbb{T}^d))$ with the following estimate for a.e. $t \in [0, T]$*

$$\int_{\mathbb{T}^d} z(t)^2 + \int_0^t \int_{\mathbb{T}^d} \mu |\nabla z|^2 \le \exp\left(\int_0^t \|(\Delta\mu)^+(s)\|_{\mathrm{L}^\infty(\mathbb{T}^d)} \, \mathrm{d}s\right) \int_{\mathbb{T}^d} (z^{\mathrm{init}})^2. \tag{35}$$

Proof. Let's first assume that μ and the initial data are smooth. In that case, we can rewrite the Kolmogorov equation (33) as standard parabolic equation, and we get the smoothness of the solution z. In this situation, we can rigorously multiply the equation by z and integrating by parts, to get

$$\begin{aligned}\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{T}^d} z(t)^2 + \int_{\mathbb{T}^d}\mu(t)|\nabla z(t)|^2 &= -\int_{\mathbb{T}^d} z(t)\nabla z(t)\cdot\nabla\mu(t)\\ &= \frac{1}{2}\int_{\mathbb{T}^d} z(t)^2\Delta\mu(t) \le \frac{1}{2}\|(\Delta\mu)^+(t)\|_{\mathrm{L}^\infty(\mathbb{T}^d)}\int_{\mathbb{T}^d} z(t)^2.\end{aligned}$$

We have thus

$$\begin{aligned}\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}&\left\{\exp\left(-\int_0^t\|(\Delta\mu)^+(s)\|_{\mathrm{L}^\infty(\mathbb{T}^d)}\,\mathrm{d}s\right)\int_{\mathbb{T}^d} z(t)^2\right\}\\ &\qquad + \exp\left(-\int_0^t\|(\Delta\mu)^+(s)\|_{\mathrm{L}^\infty(\mathbb{T}^d)}\,\mathrm{d}s\right)\int_{\mathbb{T}^d}\mu(t)|\nabla z(t)|^2 \le 0,\end{aligned}$$

and we infer after time integration the stated estimate. For the moment, we only established the estimate in the case of smooth data. Replacing μ and z^{init} by smooth approximations $(\mu_n)_n$ and $(z_n^{\mathrm{init}})_n$, approaching them in $\mathrm{L}^1(Q_T)$ and $\mathrm{L}^2(\mathbb{T}^d)$ respectively, with furthermore $\|(\Delta\mu_n)^+\|_{\mathrm{L}^\infty(Q_T)} \le \|(\Delta\mu)^+\|_{\mathrm{L}^\infty(Q_T)}$, we get a sequence $(z_n)_n$ which, by the sequential stability of Theorem 17, approaches z in $\mathrm{L}^2(Q_T)$. The usual semi-continuity argument for weak convergence allows to obtain that $z \in \mathrm{L}^\infty(0,T;\mathrm{L}^2(\mathbb{T}^d))\cap \mathrm{L}^2(0,T;\mathrm{H}^1(\mathbb{T}^d))$, with the estimate (35) being satisfied for a.e. t. □

2.2. Proof of Theorem 9

Proof. We start by proving the four *a priori* estimates, under the assumption of positivity and regularity (19).

- *conservation of the mass:* since $u_i \in \mathscr{C}^0([0,T];\mathrm{L}^2(\mathbb{T}^d))$, it also belongs to $\mathscr{C}^0([0,T];\mathrm{L}^1(\mathbb{T}^d))$ and this is sufficient (*via* a density argument) to use $\mathbf{1}_{\mathbb{T}^d}$ as test function which allows to recover (20).

- *entropy estimate:* the h_i's and μ_i's are locally Lipschitz, so boundedness of the u_i's and their belonging to $\mathscr{C}^0([0,T];\mathrm{L}^2(\mathbb{T}^d))$ imply $h_i(u_i),\mu_i(u_i) \in \mathscr{C}^0([0,T];\mathrm{L}^1(\mathbb{T}^d))$, for $i = 1,2$. With the same type of arguments we recover $h_i'(u_i) \in \mathrm{L}^2(0,T;\mathrm{H}^1(\mathbb{T}^d))$. This is sufficient to justify the following formula for all $t \in [0,T]$, by density of smooth functions,

$$\int_0^t\int_{\mathbb{T}^d} h_i'(u_i)\partial_t u_i = \int_{\mathbb{T}^d} h_i(u_i(t)) - \int_{\mathbb{T}^d} h_i(u_i^{\mathrm{init}}).$$

  Similarly, we have that (with the analogous formula for the other species)

$$-\int_0^t\int_{\mathbb{T}^d} h_1'(u_1)\Delta\big(\big[\mu_1(u_2)\star\rho\big]u_1\big) = \int_0^t\int_{\mathbb{T}^d} h_1''(u_1)\nabla u_1\cdot\nabla\big(\big[\mu_1(u_2)\star\rho\big]u_1\big),$$

  which is sufficient to reproduce rigorously the computation done in the proof Proposition 4 and integrate it time to get (21).

- *maximum principle:* we have (using assumption (18) and the entropy estimate)

$$\int_0^T \|\Delta\left(\mu_1(u_2)\star\rho\right)(s)\|_{\mathrm{L}^\infty(\mathbb{T}^d)}\,\mathrm{d}s \le \|\mu_1(u_2)\|_{\mathrm{L}^1(Q_T)}\|\Delta\rho\|_{\mathrm{L}^\infty(\mathbb{T}^d)}$$
$$\le \mathrm{A}T(1+H^{\mathrm{init}})\|\Delta\rho\|_{\mathrm{L}^\infty(\mathbb{T}^d)}$$

  and likewise for $\Delta(\mu_2(u_1)\star\check{\rho})$ so that (22) follows directly from Corollary 18.

- *duality estimate:* the function $z := u_1 + u_2$ solves the following Kolmogorov equation

$$\partial_t z - \Delta(\mu z) = 0,$$
$$z(0,\cdot) = u_1^{\mathrm{init}} + u_2^{\mathrm{init}},$$

  with

$$\mu := \frac{(\mu_1(u_2)\star\rho)u_1 + (\mu_2(u_1)\star\check{\rho})u_2}{u_1+u_2},$$

  where μ is well-defined thanks to the positivity of the u_i's and furthermore bounded. The duality estimate of Theorem 17 applies to get

$$\int_{Q_T}\left(\left[\mu_1(u_2)\star\rho\right]u_1 + \left[\mu_2(u_1)\star\check{\rho}\right]u_2\right)(u_1+u_2)$$
$$\lesssim_d \left(1+\int_{Q_T}\mu\right)\left(\int_{\mathbb{T}^d}(u_1^{\mathrm{init}})^2 + \int_{\mathbb{T}^d}(u_2^{\mathrm{init}})^2\right). \quad (36)$$

  To recover (23), simply notice that $\mu \le \mu_1(u_2)\star\rho + \mu_2(u_1)\star\check{\rho}$ so that using the normalization of ρ and assumption (18),

$$\int_{Q_T}\mu \le \int_{Q_T}\mu_1(u_2) + \int_{Q_T}\mu_2(u_1) \le \mathrm{A}\left(\int_{Q_T}(2+h_1(u_1)+h_2(u_2))\right) \le 2\mathrm{A}B_{T,\mathrm{init}},$$

  where we used the entropy estimate and the constant $\mathrm{B}_{T,\mathrm{init}} := T(1+H^{\mathrm{init}})$.

These estimates have been proven for a positive solution with regularity (19) whose existence has been assumed. We now construct a solution by a fixed-point argument.

On the set $E := \mathrm{L}^1(Q_T)\times\mathrm{L}^1(Q_T)$ we define the map $\Theta : E \to E$ which sends (u_1,u_2) to the solutions $(u_1^\bullet, u_2^\bullet)$ (in the sense of Definition 16) of

$$\begin{cases}\partial_t u_1^\bullet = \Delta\left(\left[\mu_1(u_2^+\wedge M)\star\rho\right]u_1^\bullet\right),\\ \partial_t u_2^\bullet = \Delta\left(\left[\mu_2(u_1^+\wedge M)\star\check{\rho}\right]u_2^\bullet\right),\end{cases}$$

where the cutoff constant $M>0$ will be fixed later on. By continuity of μ_1 and μ_2 we have

$$\max(\mu_1(x),\mu_2(x)) \le C, \qquad \forall x\in[0,M]. \quad (37)$$

In particular, Theorem 17 applies and ensures that the previous map is well-defined. Moreover, (37) implies for $(u_1, u_2) \in E$ that

$$|\Delta\big(\mu_1(u_2^+ \wedge M) \star \rho\big)| = |\mu_1(u_2^+ \wedge M) \star \Delta\rho| \leq C\|\Delta\rho\|_{\mathrm{L}^1(\mathbb{T}^d)},$$

and likewise for $\Delta\big(\mu_2(u_1^+ \wedge M) \star \rho\big)$. Hence we infer from Corollaries 18 and 19 that the images $u_1^\bullet$ and $u_2^\bullet$ are non-negative and uniformly bounded in $\mathrm{L}^\infty(Q_T)$ and $\mathrm{L}^2(0,T;\mathrm{H}^1(\mathbb{T}^d))$. Moreover by the equations, the time derivatives are also uniformly bounded in $\mathrm{L}^2(0,T;\mathrm{H}^{-1}(\mathbb{T}^d))$. That means that there exists a constant c such that

$$\Theta(E) \subset K := \Big\{(v_1, v_2) \in E :$$
$$\max(\|v_i\|_{\mathrm{L}^\infty(Q_T)}, \|v_i\|_{\mathrm{L}^2(0,T;\mathrm{H}^1(\mathbb{T}^d))}, \|\partial_t v_i\|_{\mathrm{L}^2(0,T;\mathrm{H}^{-1}(\mathbb{T}^d))}) \leq c \text{ for } i = 1, 2\Big\}.$$

Then by the Aubin-Lions lemma the convex set K is also compact in E. Hence Schauder's fixed-point theorem applies and ensures that there exists a fixed point $(u_1, u_2) \in K$, solving therefore (the u_i's are non-negative)

$$\begin{cases} \partial_t u_1 = \Delta\big(\big[\mu_1(u_2 \wedge M) \star \rho\big] u_1\big), \\ \partial_t u_2 = \Delta\big(\big[\mu_2(u_1 \wedge M) \star \check{\rho}\big] u_2\big). \end{cases} \tag{38}$$

The bounds obtained for elements in K also ensure that u_1 and u_2 both belong to $\mathscr{C}^0([0,T];\mathrm{L}^2(\mathbb{T}^d))$ so that we have the required regularity (19). In order to conclude we just need to fix a constant M such that the corresponding saturation vanishes. For this purpose, we consider

$$M = 2\gamma^{-1}\exp(\mathrm{A}\mathrm{B}_{T,\mathrm{init}}\|\Delta\rho\|_\infty),$$

where the constants A and $\mathrm{B}_{T,\mathrm{init}}$ are defined in the statement of Theorem 9 and we recall that $\gamma > 0$ is such that

$$\gamma \leq u_i^{\mathrm{init}} \leq \gamma^{-1}.$$

We now define

$$t^\star := \sup\Big\{t \in [0,T] : \max_{i=1,2} \|u_i\|_{\mathrm{L}^\infty([0,t]\times\mathbb{T}^d)} \leq \frac{3M}{4}\Big\}.$$

By Corollary 18, we have $t^\star > 0$ and up to any $t \in (0, t^\star)$ the cutoff M has been irrelevant. Thus, for $t \in (0, t^\star)$ all the a priori estimates apply and in particular the entropy estimate which implies

$$\|\mu_1(u_2)\|_{\mathrm{L}^1([0,t]\times\mathbb{T}^d)} \leq \mathrm{A}\int_0^t \big(1 + \int_{\mathbb{T}^d} h_2(u_2)\big) \leq \mathrm{A}t(1 + H^{\mathrm{init}}) \leq \mathrm{A}\mathrm{B}_{T,\mathrm{init}},$$

with a similar estimate for the other species. This in turn implies by Corollary 18 that for $t < t^\star$

$$\max_{i=1,2} \|u_i\|_{\mathrm{L}^\infty([0,t]\times\mathbb{T}^d)} \leq \frac{M}{2},$$

which proves that $t^\star = T$ by the usual continuity argument and our fixed-point (u_1, u_2) is the required solution. □

2.3. From non-local to local SKT

We start with a compactness tool already used in [12] that we adapt slightly to our setting. The proofs are only included for the reader's convenience.

Lemma 20. *Fix $\alpha : \mathbb{R}_{>0} \to \mathbb{R}_{\geq 0}$ having a negligible set of zeros. Consider a sequence of positive functions $(w_n)_n \in \mathrm{W}^{1,1}(Q_T)$ such that*

(i) $(w_n)_n$ *bounded in* $\mathrm{L}^2(Q_T)$;

(ii) $(\partial_t w_n)_n$ *bounded in* $\mathrm{L}^1(0,T;\mathrm{H}^{-m}(\mathbb{T}^d))$ *for some integer* m;

(iii) $(\alpha(w_n)\nabla w_n)_n$ *bounded in* $\mathrm{L}^2(Q_T)$.

Then $(w_n)_n$ *admits an a.e. converging subsequence.*

Proof. By assumption (iii) the sequence $(\nabla F(w_n))_n$ is bounded in $\mathrm{L}^2(Q_T)$, where $F : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is defined by

$$F(z) := \int_0^z 1 \wedge \alpha.$$

F is an increasing (because $\alpha > 0$ a.e.) 1-Lipschitz function vanishing at 0. In particular, we infer from (i) the same bound for $(F(w_n))_n$. Up to a subsequence we can thus assume that $(w_n)_n$ and $(F(w_n))_n$ respectively converge weakly to w and $\widetilde{w}$ in $\mathrm{L}^2(Q_T)$. Using (ii) we thus infer from [15, Proposition 3] that (up to a subsequence),

$$\int_{Q_T} w_n F(w_n) \underset{N+\infty}{\longrightarrow} \int_{Q_T} w\widetilde{w}. \tag{39}$$

At this stage we use the Minty-Browder or Leray-Lions trick: one first establishes that

$$\begin{aligned}\int_{Q_T} &\overbrace{(F(w_n) - F(w))(w_n - w)}^{:=h_n} \\ &= \int_{Q_T} F(w_n)w_n + \int_{Q_T} F(w)w - \int_{Q_T} F(w_n)w - \int_{Q_T} F(w)w_n \underset{N+\infty}{\longrightarrow} 0\end{aligned}$$

by exploiting the $\mathrm{L}^2(Q_T)$ weak convergences $(w_n)_n \rightharpoonup_n w$, $(F(w_n))_n \rightharpoonup_n \widetilde{w}$, together with (39). Then, since F is increasing, we have $h_n \geq 0$ so that the previous convergence may be seen as the convergence of $(h_n)_n$ to 0 in $\mathrm{L}^1(Q_T)$. In particular, up to some subsequence, we get that $(h_n)_n$ converges a.e. to 0 which in turn implies (increasingness of F) that $(w_n)_n \to w$. □

Proof of Theorem 11. Using the duality estimate of Theorem 9 we first have

$$\int_{Q_T} \big(\big[\mu_1(u_{2,n}) \star \rho_n\big]u_{1,n} + \big[\mu_2(u_{1,n}) \star \check{\rho}_n\big]u_{2,n}\big)(u_{1,n} + u_{2,n}) \lesssim_{d,\mathrm{init}} 1, \tag{40}$$

where the constant depends on the dimension and initial data but is uniform in n. In particular, both species satisfy (since μ_1 and μ_2 are positively lower-bounded) assumptions (*i*) and (*ii*) of Lemma 20. Using the entropy estimate of Theorem 9, we have also for both species that $(\alpha_i(u_{i,n})\nabla u_n)_n$ bounded in $\mathrm{L}^2(Q_T)$, which validates assumption (*iii*) of the lemma since the dissipation rates α_i are assumed a.e. positive on $\mathbb{R}_{>0}$. We infer therefore from the previous lemma that, up to a subsequence (that we do not label), both $(u_{1,n})_n$ and $(u_{2,n})_n$ both converge a.e. respectively to some u_1 and u_2. We need now to pass to the limit (in $\mathscr{D}'(Q_T)$) in the products

$$\big[\mu_1(u_{2,n}) \star \rho_n\big]u_{1,n} \text{ and } \big[\mu_2(u_{1,n}) \star \check{\rho}_n\big]u_{2,n}.$$

W.l.o.g. we can focus on the first one. Since $(u_{2,n})_n$ converges (to u_2) a.e., so does $(\mu_1(u_{2,n})_n$ (to $\mu_1(u_2)$), by continuity of μ_1. Using assumption (18) and the entropy estimate of Theorem 9, we have of $(\mu_1(u_{2,n}))_n$ bounded in $\mathrm{L}^\infty(0,T;\mathrm{L}^1(\mathbb{T}^d))$. This is not sufficient to prevent possible concentration in the space variable: we use the growth assumption (24), to establish the uniform integrability of $(\mu_1(u_{2,n}))_n$. Indeed, since μ_1 is continuous, the sequence $c_k := \inf\{z \geq 0 \,:\, \mu_1(z) \geq R\}$ diverges to $+\infty$ and we have

$$\int_{Q_T} \mu_1(u_{2,n})\mathbf{1}_{\mu_1(u_{2,n})\geq R} \leq \int_{Q_T} \mu_1(u_{2,n})\mathbf{1}_{u_{2,n}\geq c_R} \leq \sup_{z\geq c_R} \Phi(z) \int_{Q_T} h_2(u_{2,n}) \vee u_{2,n}^2,$$

where $\Phi(z) := \frac{\mu_1(z)}{h_2(z)\vee z^2}$ goes to 0 as $z \to +\infty$, by assumption (24). Since μ_2 is positively lower-bounded (thanks to assumption (18)), we infer from (40) a bound for $(u_{2,n})_n$ in $\mathrm{L}^2(Q_T)$ (we use here the non-negativity of all the involved functions). Using the entropy estimate and the previous inequalities we therefore infer

$$\lim_{R\to+\infty} \sup_n \int_{Q_T} \mu_1(u_{2,n})\mathbf{1}_{\mu_1(u_{2,n})\geq R} = 0,$$

which the stated uniform integrability. We thus infer from Vitali's convergence theorem that $(\mu_1(u_{2,n}))_n$ converges to $\mu_1(u_2)$ in $\mathrm{L}^1(Q_T)$. The sequence $(\mu_1(u_{2,n}) \star \rho_n)_n$ shares the same behaviour. In particular, $(\mu_1(u_{2,n}) \star \rho_n)_n$ is also uniformly integrable and adding a subsequence if necessary, we can assume that it converges a.e. towards $\mu_1(u_2)$. Now, to conclude we write

$$w_n := \big[\mu_1(u_{2,n}) \star \rho_n\big]u_{1,n} = \big[\mu_1(u_{2,n}) \star \rho_n\big]^{1/2} \big[\mu_1(u_{2,n}) \star \rho_n\big]^{1/2} u_{1,n}.$$

As already noticed, $(w_n)_n$ converges a.e. to the expected limit $\mu_1(u_2)u_1$. The previous writing together with the duality estimate (40) and the Cauchy-Schwarz inequality shows that $(w_n)_n$ is bounded in $\mathrm{L}^1(Q_T)$. Even better, $(w_n)_n$ is the product of of one L^2-uniformly integrable sequence with an L^2-bounded one, therefore $(w_n)_n$ is uniformly integrable and the Vitali convergence theorem applies once more to get the convergence of $(w_n)_n$ towards $\mu_1(u_2)u_1$.

The previous reasoning (which applies to both species) allows to pass to the limit the equations. As for the estimates, they are all obtained using Fatou's lemma. ☐

3. General regularised scheme on a domain

In this section, we study the general regularisation scheme introduced in Proposition 7 and prove the corresponding results Theorem 13 and Theorem 14.

3.1. Existence of regularised solutions

We start with proving the existence of solutions for the regularised scheme, i.e. Theorem 13.

The advantage of the regularisation is that the cross-diffusion terms are controllable and we thus rewrite the evolution as

$$\partial_t u_i(x_i) - \mathrm{div}_{x_i}\left[\left(\epsilon + \prod_{k\neq i}\int_{x_k\in\Omega} \mathrm{d}x_k\, K(x_1,\dots,x_n)\, a_{ii}(u_1(x_1),\dots,u_n(x_n))\right)\nabla u_i(x_i)\right]$$
$$= \mathrm{div}_{x_i}\left[\prod_{k\neq i}\int_{x_k\in\Omega} \mathrm{d}x_k\, K(x_1,\dots,x_n) \sum_{j\neq i} a_{ij}(u_1(x_1),\dots,u_n(x_n))\nabla u_j(x_j)\right].$$

For the cross-diffusion terms, the $\tilde{a}_{ij}$ in Theorem 13 are defined such that

$$a_{ij}(u_1(x_1),\dots,u_n(x_n))\nabla u_j(x_j) = \nabla_{x_j}\tilde{a}_{ij}(u_1(x_1),\dots,u_n(x_n))$$

so that the partial derivative can formally be integrated by parts onto the kernel K, where no boundary terms appear due to (15). Hence the evolution can be rewritten as

$$\partial_t u_i - \nabla\big((\epsilon + \bar{a}_i[u])\nabla u_i\big) + \bar{b}_i[u]\nabla u_i + \bar{c}_i[u]u_i = 0, \tag{41}$$

with von Neumann boundary conditions and

$$\bar{a}_i[u](x_i) = \prod_{k\neq i}\int_{x_k\in\Omega} \mathrm{d}x_k\, K(x_1,\dots,x_n)\, a_{ii}(u_1(x_1),\dots,u_n(x_n)), \tag{42}$$

$$\bar{b}_i[u](x_i) = \sum_{j\neq i}\prod_{k\neq i}\int_{x_k\in\Omega} \mathrm{d}x_k\, \partial_j K(x_1,\dots,x_n)\, \partial_i\tilde{a}_{ij}(u_1(x_1),\dots,u_n(x_n)), \tag{43}$$

$$\bar{c}_i[u](x_i) = \sum_{j\neq i}\prod_{k\neq i}\int_{x_k\in\Omega} \mathrm{d}x_k\, \partial_{ij} K(x_1,\dots,x_n)\, \frac{\tilde{a}_{ij}(u_1(x_1),\dots,u_n(x_n))}{u_i(x_i)}. \tag{44}$$

The assumptions of Theorem 13 then imply for $x_i \in \Omega$ that

$$\begin{aligned} |\bar{a}_i[u](x_i)| &\le A\,\|K\|_\infty(|\Omega| + H(u)),\\ |\bar{b}_i[u](x_i)| &\le A\,\|\nabla K\|_\infty(|\Omega| + H(u)),\\ |\bar{c}_i[u](x_i)| &\le A\,\|\nabla^2 K\|_\infty(|\Omega| + H(u)). \end{aligned} \tag{45}$$

This is enough to prove the existence of solutions by a Galerkin scheme.

Proof of Theorem 13. Let $\sigma \in \mathscr{C}_c^\infty(\mathbb{R}^n)$ a non-negative mollification kernel with $\operatorname{supp}\sigma \subset B_1$ and $\int \sigma \mathrm{d}x = 1$ and define

$$\sigma^m(x) = m^d\, \sigma(mx).$$

Extending $\bar{a}_i$, $\bar{b}_i$, $\bar{c}_i$ with zero outside Ω, we consider for $m \in \mathbb{N}$ the following system

$$\begin{aligned} \partial_t u_i^m - \nabla\big((\epsilon + ((\bar{a}_i[u^m] \wedge M) \star \sigma^m)\nabla u_i^m\big) \\ + ((\bar{b}_i[u^m] \wedge M) \star \sigma^m)\nabla u_i^m + ((\bar{c}_i[u^m] \wedge M) \star \sigma^m) u_i^m = 0, \end{aligned} \tag{46}$$

with von Neumann boundary conditions, $i = 1, \dots, n$ and the constant M as in (30).

By a standard Galerkin scheme (e.g. taking the von Neumann eigenvectors of the Laplacian on Ω), the system (46) has a solution u_i^m with initial data u_i^{init} and has any H^k, $k \in \mathbb{N}$, regularity after an arbitrary short time. Hence we can apply the maximum principle for parabolic equations and find as in Corollary 18 that

$$\gamma \exp(-TM) \le u_i^m \le \gamma^{-1} \exp(TM)\,.$$

Furthermore, each u_i^m is preserving the mass. Finally, we can test (46) against u_i^m to find the followig estimate independent of m:

$$\sup_{t\in[0,T]} \|u_i^m(t,\cdot)\|^2_{\mathrm{L}^2(\Omega)} + \epsilon \int_0^T \|\nabla u_i^m(t,\cdot)\|^2_{\mathrm{L}^2(\Omega)} \le \exp\left[TM\Big(2 + \frac{1}{\epsilon}\Big)\right] \|u_i^{\text{init}}\|^2_{\mathrm{L}^2(\Omega)},$$

where $i = 1, \dots, n$ and we used the cutoff with M. Note that here the RHS is bounded by assumption. Hence we find for a constant $C(T)$ independent of m that

$$\|\partial_t u_i^m\|_{\mathrm{L}^2(0,T,\mathrm{H}^{-1}(\Omega))} \le C(T)$$

for $i = 1, \dots, n$.

By Aubin-Lions lemma we can therefore find a subsequence (relabelling with m) and

$$u_i \in \mathscr{C}^0([0,T]; \mathrm{L}^2(\Omega)) \cap \mathrm{L}^2(0,T; \mathrm{H}^1(\Omega)) \cap \mathrm{L}^\infty(0,T; \mathrm{L}^\infty(\Omega)),$$

for $i = 1, \dots, n$ such that u_i^m converges almost everywhere to u_i and ∇u_i^m converges L^2 weakly to ∇u_i. Moreover, it holds that

$$\gamma \exp(-TM) \le u_i \le \gamma^{-1} \exp(TM)\,.$$

The convergence implies that for $\phi \in \mathscr{C}^\infty(\Omega_T)$ with $\phi(T,\cdot) \equiv 0$ it holds that

$$\begin{aligned} -\int_0^T \int_\Omega u_i \partial_t \phi + \int_0^T \int_\Omega (\epsilon + \bar{a}_i[u] \wedge M) \nabla u_i \cdot \nabla \phi + \int_0^T \int_\Omega (\bar{b}_i[u] \wedge M) \cdot \nabla u_i\, \phi \\ + \int_0^T \int_\Omega (\bar{c}_i[u] \wedge M) u_i\, \phi = \int_\Omega u_i^{\text{init}} \phi^{\text{init}}, \end{aligned}$$

i.e. $u = (u_1, \dots, u_n)$ is a weak solution with von Neumann boundary data. Moreover, by the continuity this implies directly the conservation of mass.

Until a time $T^* \le T$ for which

$$\sup_{t\in[0,T^*]} \sup_{x\in\Omega} \max(\bar{a}_i[u], |\bar{b}_i[u]|, \bar{c}_i[u]) \le M,$$

the cutoff M is not applied and we have a weak solution of (41). As in the Laplace case on the torus in Section 2, the proven regularity is sufficient to justify rigorously the formal entropy estimate as in Proposition 7.

The assumptions (45) then imply that at time T^* it holds that

$$\sup_{x\in\Omega} \max(\bar{a}_i[u], |\bar{b}_i[u]|, \bar{c}_i[u]) \le \frac{M}{2}$$

and thus by continuity $T^* = T$ and we have constructed the claimed solution. $\square$

3.2. Limit for the SKT system

Having constructed the nonlocal approximation, we now prove Theorem 14.

The assumption of the extension operator allows to find a uniform Gagliardo-Nirenberg inequality.

Lemma 21. *Assume the setup of Theorem 14. Then there exists a uniform c for the Gagliardo-Nirenberg inequality*

$$\|f\|^p_{\mathrm{L}^p(A_m)} \le c\left(\|f\|^{(1-\theta)p}_{\mathrm{L}^1(A_m)} \|\nabla f\|^{\theta p}_{\mathrm{L}^2(A_m)} + \|f\|^p_{\mathrm{L}^1(A_m)}\right) \qquad \forall f : A_m \to \mathbb{R},$$

holds on all A_m, where $m \in \mathbb{N}$, $\theta = 2d(p-1)/(p(d+2))$ and $p = 2 + 2/d$.

Proof. By the extension operator $\mathcal{E}_m$ and the Gagliardo-Nirenberg inequality on $\mathbb{R}^d$ we find (note that $\theta p = 2$)

$$\begin{aligned} \|f\|^p_{\mathrm{L}^p(A_m)} &\le \|\mathcal{E}_m(f)\|^p_{\mathrm{L}^p(\mathbb{R}^d)} \\ &\lesssim \|\mathcal{E}_m(f)\|^{(1-\theta)p}_{\mathrm{L}^1(\mathbb{R}^d)} \|\nabla\mathcal{E}_m(f)\|^2_{\mathrm{L}^2(\mathbb{R}^d)} \\ &\lesssim \|f\|^{(1-\theta)p}_{\mathrm{L}^1(\mathbb{R}^d)} \left(\|f\|^2_{\mathrm{L}^2(A_m)} + \|\nabla f\|^2_{\mathrm{L}^2(A_m)}\right). \end{aligned}$$

As $p > 2$, we can interpolate $\|f\|_{\mathrm{L}^2(A_m)}$ between $\|f\|_{\mathrm{L}^1(A_m)}$ and $\|f\|_{\mathrm{L}^p(A_m)}$ and absorb the contribution of $\|f\|_{\mathrm{L}^p(A_m)}$ so that the claimed inequality follows. $\square$

The first lemma ensures the integrability and determines the sequence ϵ.

Lemma 22. *Assume the setup of Theorem 14. Then there exists a constant C_T and a decreasing sequence $(\epsilon_m)_m$ with $\epsilon_m \downarrow 0$ such that*

$$\|u_i^m\|_{\mathrm{L}^{\tilde{p}}([0,T)\times\Omega)} \le C_T$$

for $i = 1, \dots, n$ and $\tilde{p} = 2 + 1/d$.

Proof. For the regularisation kernel K^m and $\epsilon_m > 0$ we find by Theorem 13 a solution u_m which satisfies

(i) $\|u_i^m\|_{\mathrm{L}^\infty(0,T;\mathrm{L}^1(\Omega))} \le c$ for $i = 1, \dots, n$ (conservation of mass),

(ii) $\|\nabla u_i^m\|^2_{\mathrm{L}^2(0,T;\mathrm{L}^2(\Omega))} \le \epsilon_m^{-1} \exp(\epsilon_m^{-1}) c$ (ϵ-dependent estimate),

(iii) $\|\nabla u_i^m\|^2_{\mathrm{L}^2(0,T;\mathrm{L}^2(A_m))} \le c$ (dissipation estimate in the set A_m where the weights $w_i^m \equiv 1$)

for a constant c independent of m.

The parameters θ and p of the Gagliardo-Nirenberg in Lemma 21 are chosen such that $\theta p = 2$. Hence we find on A_m that for $i = 1, \dots, n$

$$\int_0^T \|u_i^m\|^p_{\mathrm{L}^p(A_m)} \mathrm{d}t \lesssim \int_0^T \left(\|\nabla u_i^m\|^2_{\mathrm{L}^2(A_m)} \, \|u_i^m\|^{(1-\theta)p}_{\mathrm{L}^1(A_m)} + \|u_i^m\|^p_{\mathrm{L}^1(A_m)} \right) \, \mathrm{d}t.$$

With the gradient control from the dissipation and the conservation of mass this shows

$$\int_{[0,T)\times A_m} |u_i^m|^p \, \mathrm{d}x \, \mathrm{d}t \le c_d$$

for a constant c_d independent of m.

As the domain Ω is assumed to have $\mathscr{C}^1$ boundary, we can also apply the argument of Lemma 21 to find over Ω that for $i = 1, \dots, n$

$$\int_0^T \|u_i^m\|^p_{\mathrm{L}^p(\Omega)} \mathrm{d}t \lesssim \int_0^T \left(\|\nabla u_i^m\|^2_{\mathrm{L}^2(\Omega)} \, \|u_i^m\|^{(1-\theta)p}_{\mathrm{L}^1(\Omega)} + \|u_i^m\|^p_{\mathrm{L}^1(\Omega)} \right) \, \mathrm{d}t.$$

Hence we find for a constant c_e independent of m that

$$\int_{[0,T)\times \Omega} |u_i^m|^p \, \mathrm{d}x \, \mathrm{d}t \le c_e \big(1 + \frac{c_e}{\epsilon_m} \exp(\epsilon_m^{-1})\big).$$

As $\tilde{p} < p$ we can find $q \in (0,1)$ so that the Hölder inequality implies

$$\|f\|_{\mathrm{L}^{\tilde{p}}([0,T)\times B)} \le (T\,|B|)^q \, \|f\|_{\mathrm{L}^p([0,T)\times B)}$$

for $B \subset \Omega$ and $f \in \mathrm{L}^p([0,T) \times B)$.

By splitting Ω into A_m and $\Omega \setminus A_m$ we therefore find (as $|A_m| \le |\Omega|$)

$$\begin{aligned} \|u_i^m\|_{\mathrm{L}^{\tilde{p}}([0,T)\times\Omega)} &\le \|u_i^m\|_{\mathrm{L}^{\tilde{p}}([0,T)\times A_m)} + \|u_i^m\|_{\mathrm{L}^{\tilde{p}}([0,T)\times(\Omega\setminus A_m))} \\ &\le T^q \, |\Omega|^q c_d^{1/p} + T^q \, |\Omega \setminus A_m|^q c_e^{1/p} \left(1 + \frac{\exp(\epsilon_m^{-1})}{\epsilon_m}\right)^{1/p}. \end{aligned}$$

As $|\Omega \setminus A_m| \to 0$ and $q \in (0,1)$, we can therefore find a sequence $\epsilon_m \downarrow 0$ such that $|\Omega \setminus A_m|^q (1 + \epsilon_m^{-1} \exp(\epsilon_m^{-1}))^{1/p}$ is bounded by a constant independent of m. The claim then follows directly from the given estimate. □

We can now proceed with the convergence result.

Proof of Theorem 14. Inside each good set $A_{\bar m}$, the dissipation and mass conservation give a uniform estimate for u_i^m in $\mathrm{L}^\infty(0,T;\mathrm{L}^1(A_m))$ and $\mathrm{L}^2(0,T;\mathrm{H}^1(A_m))$ for $i = 1,\dots,n$ and $m \ge \bar m$. By the equation this also gives a uniform estimate of the time-derivative in $\mathrm{L}^1(0,T;H^{-k}(A_m))$ for a large enough $k \in \mathbb{N}$ (depending only on dimension d). Hence on $A_{\bar m}$ we have compactness for u_i^m. As $A_m \uparrow \Omega$, a diagonal argument shows that along a subsequence (which we relabel with m) that for $i = 1,\dots,n$ there exist $u_i : [0,T) \times \Omega$ such that $u_i^m \to u_i$ a.e. Moreover, choosing ϵ_m as in Lemma 22 we find $u_i \in \mathrm{L}^{\tilde p}([0,T) \times \Omega)$.

By the dissipation inequality we find that

$$\|\sqrt{w_i}\nabla u_i^m\|_{\mathrm{L}^2([0,T)\times\Omega)}$$

is uniformly bounded. Hence along a subsequence $\sqrt{w_i^m}\nabla u_i^m$ converges weakly in L^2 to a limit ψ_i. As w_i is the constant 1 inside the set A_m and $A_m \uparrow \Omega$, it follows that $\sqrt{w_i^m}\nabla u_i^m \rightharpoonup \nabla u$ and $\nabla u \in \mathrm{L}^2$.

As u^m preserves the mass, is non-negative and satisfies the entropy-dissipation inequality, the same is true for the limit u by using the stated regularity. Moreover, the stated regularity gives the claimed convergence.

It thus remains to check that u is a weak solution. As u^m satisfies von Neumann boundary data and K^m vanishes at the boundary, the constructed solutions satisfy for all $\phi \in \mathscr{C}^\infty([0,T] \times \Omega$ with $\phi(T,\cdot) \equiv 0$ and $i = 1,\dots,n$ that

$$\begin{aligned}
&- \int_0^T \int_\Omega u_i^m \,\partial_t \phi + \epsilon_m \int_0^T \int_\Omega \nabla u_i^m \cdot \phi \\
&+ \int_0^T \int_\Omega \left(\prod_{k\neq i} \int_{x_k \in \Omega} \mathrm{d}x_k K^m(x_1,\dots,x_n) \sum_{j=1}^n a_{ij}(u_1^m(x_1),\dots,u_n^m(x_n)) \nabla u_j^m(x_j) \right) \cdot \nabla \phi \\
&= \int_\Omega u_i^{\mathrm{init}} \phi(0,\cdot).
\end{aligned}$$

For the diffusion from d_{ij} with $i \neq j$ and $i,j = 1,\dots,n$ we must therefore show that for all test function ϕ

$$\begin{aligned}
&\int_0^T \int_{\Omega^n} K^m(x_1,\dots,x_n) u_i^m(t,x_i) \nabla u_j^m(t,x_j)\, \nabla\phi(t,x_i)\, \mathrm{d}x_1 \dots \mathrm{d}x_n \mathrm{d}t \\
&\to \int_0^T \int_\Omega u_i(t,x) \nabla u_j(t,x)\, \nabla \phi(t,x)\, \mathrm{d}x\, \mathrm{d}t.
\end{aligned}$$

We rewrite the nonlinear diffusion term as

$$\begin{aligned}
&\int_0^T \int_{\Omega^n} K^m(x_1,\dots,x_n) u_i^m(t,x_i) \nabla u_j^m(t,x_j)\, \nabla\phi(t,x_i)\, \mathrm{d}x_1 \dots \mathrm{d}x_n \mathrm{d}t \\
&= \int_0^T \int_\Omega \sqrt{w_j^m(x_j)} \nabla u_j^m(t,x_j) \psi^m(t,x_j)\, \mathrm{d}x_i\, \mathrm{d}t,
\end{aligned}$$

where

$$\psi^m(t, x_j) = \frac{1}{\sqrt{w_j^m(x_j)}} \prod_{k\neq j} \int_\Omega \mathrm{d}x_k\, K^m(x_1, \ldots, x_n) u_i^m(t, x_i)\, \nabla\phi(t, x_i).$$

By the definition of the weight w_j^m, we can apply Jensen's inequality for $\tilde{p} \geq 2$ to find that

$$\|\psi^m\|_{\mathrm{L}^{\tilde{p}}([0,T)\times\Omega)} \leq \|u_i^m \nabla\phi\|_{\mathrm{L}^{\tilde{p}}([0,T)\times\Omega)}.$$

By the proven convergence and regularity of u_i^m we find that $\psi^m \to u_i\nabla\phi$ a.e. in $[0, T) \times \Omega$. The previous inequality gives a uniform bound of ψ^m in $\mathrm{L}^{\tilde{p}}$ with $\tilde{p} > 2$ so that ψ^m converges strongly in L^2 to $u_i\nabla\phi$. As $\sqrt{w_j^m}\nabla u_j^m$ converges weakly in L^2 to ∇u_j, this proves the claimed convergence.

The other terms in the weak formulation converge more directly in the limit and we thus have found a weak solution. □

A. Microscopic reversibility

In the linear SKT model (5), the entropy was understood as reversiblity in a microscopic model in [6] and this gave us the intuition about the nonlocal entropy structure. In this appendix we discuss in the case of two species how the form in Remark 8 in the general regularisation on bounded domains by a kernel $K : \Omega^2 \to \mathbb{R}_{\geq 0}$ appears formally from the microscopic entropy structure.

In the microscopic picture of [6] we considered a spatial discretisation in the one-dimensional setting so that we have discrete positions $\{1, \ldots, N\}$. On this discrete setting we consider many particles of the two species 1 and 2 and we then obtain a reversible cross-diffusion behaviour if a pair consisting of a particle of species 1 at position i and a particle of species 2 at position j jumps together with rate $R_r(i, j)$ to the positions $i + 1$ and $j + 1$, respectively. Likewise the pair can jump with a rate $R_l(i, j)$ to $i - 1$ and $j - 1$, respectively. We then have the reversibility (and thus the entropy structure) if

$$R_r(i, j) = R_l(i + 1, j + 1). \tag{47}$$

In the formal mean-field limit we then find the evolution for the densities u_1 and u_2 the following nonlinear system

$$\begin{cases} \partial_t u_1(i) = \displaystyle\sum_{j=1}^{M} \Big\{ R_r(i-1, j) u_1(i-1) u_2(j) + R_l(i+1, j) u_1(i+1) u_2(j) \\ \qquad\qquad - (R_l(i, j) + R_r(i, j)) u_1(i) u_2(j) \Big\} \\ \partial_t u_2(j) = \displaystyle\sum_{i=1}^{M} \Big\{ R_r(i, j-1) u_1(i) u_2(j-1) + R_l(i, j+1) u_1(i) u_2(j+1) \\ \qquad\qquad - (R_l(i, j) + R_r(i, j)) u_1(i) u_2(j) \Big\} \end{cases}$$

for which we can indeed verify the entropy

$$H = \sum_{i=1}^{M} \Big[h(u_1(i)) + h(u_2(i)) \Big]$$

where $h'(x) = \log x$ as

$$\frac{\mathrm{d}}{\mathrm{d}t} H = -\sum_{i,j} R_r(i,j) \Big[\big(u_1(i{+}1) u_2(j{+}1) - u_1(i) u_2(j) \big) \big(\log(u_1(i{+}1) u_2(j{+}1)) - \log(u_1(i) u_2(j)) \big) \Big].$$

For the formal limit of the discrete system to a PDE, we denote the centred discrete Laplacian

$$(\Delta_d f)(i) = f(i+1) + f(i-1) - 2f(i).$$

We can then rewrite the evolution as

$$\partial_t u_1(i) = \Delta_d \left(\sum_j \frac{R_l(i,j) + R_r(i,j)}{2} u_1(i) u_2(j) \right)$$
$$+ \frac{1}{2} \sum_j \Big\{ u_1(i{+}1) u_2(j) \Big[R_l(i{+}1,j) - R_r(i{+}1,j) \Big] + u_1(i{-}1) u_2(j) \Big[R_r(i{-}1,j) - R_r(i{+}1,j) \Big] \Big\}$$

and likewise for u_2. By the microscopic reversibility (47) we note that this is exactly the discrete form of the regularisation found in (17).

References

[1] Herbert Amann. "Erratum: "Dynamic theory of quasilinear parabolic systems. III. Global existence" [Math. Z. 202 (1989), no. 2, 219–250; MR1013086 (90i:35125)]". In: *Math. Z.* 205.2 (1990), p. 331. ISSN: 0025-5874. DOI: 10.1007/BF02571246. URL: https://doi.org/10.1007/BF02571246.

[2] Andrea Braides. "Local Minimization, Variational Evolution and Γ-Convergence". In: *Lecture Notes in Mathematics* (2014). ISSN: 1617-9692. DOI: 10.1007/978-3-319-01982-6. URL: http://dx.doi.org/10.1007/978-3-319-01982-6.

[3] L. Chen and A. Jüngel. "Analysis of a parabolic cross-diffusion population model without self-diffusion". In: *J. Differential Equations* 224.1 (2006), pp. 39–59. ISSN: 0022-0396. DOI: 10.1016/j.jde.2005.08.002. URL: http://dx.doi.org/10.1016/j.jde.2005.08.

[4] Li Chen, Esther S. Daus, A. Holzinger, and A. Jüngel. *Rigorous derivation of population cross-diffusion systems from moderately interacting particle systems.* 2020. arXiv: 2010.12389v1 [math.AP].

[5] Xiuqing Chen, Esther S. Daus, and Ansgar Jüngel. "Global existence analysis of cross-diffusion population systems for multiple species". In: *Arch. Ration. Mech. Anal.* 227.2 (2018), pp. 715–747. ISSN: 0003-9527. DOI: 10.1007/s00205-017-1172-6. URL: https://doi.org/10.1007/s00205-017-1172-6.

[6] Esther S. Daus, Laurent Desvillettes, and Helge Dietert. "About the entropic structure of detailed balanced multi-species cross-diffusion equations". In: *Journal of Differential Equations* 266.7 (2019), 3861–3882. ISSN: 0022-0396. DOI: 10.1016/j.jde.2018.09.020. URL: http://dx.doi.org/10.1016/j.jde.2018.09.020.

[7] L. Desvillettes, T. Lepoutre, A. Moussa, and A. Trescases. "On the entropic structure of reaction-cross diffusion systems". In: *Comm. Partial Differential Equations* 40.9 (2015), pp. 1705–1747. ISSN: 0360-5302. DOI: 10.1080/03605302.2014.998837. URL: https://doi.org/10.1080/03605302.2014.998837.

[8] Lawrence Evans. *Partial Differential Equations*. American Mathematical Society, 2010. ISBN: 9781470411442. DOI: 10.1090/gsm/019. URL: http://dx.doi.org/10.1090/gsm/019.

[9] Joaquin Fontbona and Sylvie Méléard. "Non local Lotka-Volterra system with cross-diffusion in an heterogeneous medium". In: *Journal of Mathematical Biology* 70.4 (2014), 829–854. ISSN: 1432-1416. DOI: 10.1007/s00285-014-0781-z. URL: http://dx.doi.org/10.1007/s00285-014-0781-z.

[10] Luan T. Hoang, Truyen V. Nguyen, and Tuoc V. Phan. "Gradient Estimates and Global Existence of Smooth Solutions to a Cross-Diffusion System". In: *SIAM Journal on Mathematical Analysis* 47.3 (2015), pp. 2122–2177. DOI: 10.1137/140981447. eprint: http://dx.doi.org/10.1137/140981447. URL: http://dx.doi.org/10.1137/140981447.

[11] Ansgar Jüngel. "The boundedness-by-entropy method for cross-diffusion systems". In: *Nonlinearity* 28.6 (2015), 1963–2001. ISSN: 1361-6544. DOI: 10.1088/0951-7715/28/6/1963. URL: http://dx.doi.org/10.1088/0951-7715/28/6/1963.

[12] Thomas Lepoutre and Ayman Moussa. "Entropic structure and duality for multiple species cross-diffusion systems". In: *Nonlinear Anal.* 159 (2017), pp. 298–315. ISSN: 0362-546X. DOI: 10.1016/j.na.2017.02.008. URL: https://doi.org/10.1016/j.na.2017.02.008.

[13] Thomas Lepoutre, Michel Pierre, and Guillaume Rolland. "Global well-posedness of a conservative relaxed cross diffusion system". In: *SIAM J. Math. Anal.* 44.3 (2012), pp. 1674–1693. ISSN: 0036-1410. DOI: 10.1137/110848839. URL: https://doi.org/10.1137/110848839.

[14] Alexander Mielke. "On Evolutionary Γ-Convergence for Gradient Systems". In: *Lecture Notes in Applied Mathematics and Mechanics* (2016), 187–249. ISSN: 2197-6732. DOI: 10.1007/978-3-319-26883-5_3. URL: http://dx.doi.org/10.1007/978-3-319-26883-

[15] A. Moussa. "Some variants of the classical Aubin-Lions lemma". In: *J. Evol. Equ.* 16.1 (2016), pp. 65–93. ISSN: 1424-3199. DOI: 10.1007/s00028-015-0293-3. URL: http://dx.doi.org/10.1007/s00028-015-0293-3.

[16] Ayman Moussa. "From Nonlocal to Classical Shigesada–Kawasaki–Teramoto Systems: Triangular Case with Bounded Coefficients". In: *SIAM Journal on Mathematical Analysis* 52.1 (2020), 42–64. ISSN: 1095-7154. DOI: 10.1137/19m1243294. URL: http://dx.doi.org/10.1137/19m1243294.

[17] Sylvia Serfaty. "Gamma-convergence of gradient flows on Hilbert and metric spaces and applications". In: *Discrete & Continuous Dynamical Systems - A* 31.4 (2011), 1427–1451. ISSN: 1553-5231. DOI: 10.3934/dcds.2011.31.1427. URL: http://dx.doi.org/10.3934/dcds.2011.31.1427.

[18] Nanako Shigesada, Kohkichi Kawasaki, and Ei Teramoto. "Spatial segregation of interacting species". In: *Journal of Theoretical Biology* 79.1 (1979), 83–99. ISSN: 0022-5193. DOI: 10.1016/0022-5193(79)90258-3. URL: http://dx.doi.org/10.1016/0022-5193(79)90

[19] Elias M. Stein. *Singular Integrals and Differentiability Properties of Functions (PMS-30)*. Princeton University Press, 1970. ISBN: 9781400883882. DOI: 10.1515/9781400883882. URL: http://dx.doi.org/10.1515/9781400883882.

[20] Jonathan Zinsl and Daniel Matthes. "Transport distances and geodesic convexity for systems of degenerate diffusion equations". In: *Calculus of Variations and Partial Differential Equations* 54.4 (2015), 3397–3438. ISSN: 1432-0835. DOI: 10.1007/s00526-015-0909-z. URL: http://dx.doi.org/10.1007/s00526-015-0909-z.