GLOBAL EXISTENCE ANALYSIS OF CROSS-DIFFUSION POPULATION SYSTEMS FOR MULTIPLE SPECIES

XIUQING CHEN, ESTHER S. DAUS, AND ANSGAR JÜNGEL

Abstract. The existence of global-in-time weak solutions to reaction-cross-diffusion systems for an arbitrary number of competing population species is proved. The equations can be derived from an on-lattice random-walk model with general transition rates. In the case of linear transition rates, it extends the two-species population model of Shigesada, Kawasaki, and Teramoto. The equations are considered in a bounded domain with homogeneous Neumann boundary conditions. The existence proof is based on a refined entropy method and a new approximation scheme. Global existence follows under a detailed balance or weak cross-diffusion condition. The detailed balance condition is related to the symmetry of the mobility matrix, which mirrors Onsager's principle in thermodynamics. Under detailed balance (and without reaction), the entropy is nonincreasing in time, but counter-examples show that the entropy may increase initially if detailed balance does not hold.

1. Introduction

Shigesada, Kawasaki, and Teramoto suggested in their seminal paper [24] a diffusive Lotka-Volterra system for two competing species, which is able to describe the segregation of the population and to show pattern formation when time increases. Starting from an on-lattice random-walk model, this system was extended to an arbitrary number of species in [30, Appendix]. While the existence analysis of global weak solutions to the two-species model is well understood by now [3, 4], only very few results for the n-species model under very restrictive conditions exist (see the discussion below). In this paper, we provide for the first time a global existence analysis for an arbitrary number of population species using the entropy method of [15], and we reveal an astonishing relation between the monotonicity of the entropy and the detailed balance condition of an associated Markov chain.

More specifically, we consider the reaction-cross-diffusion equations

$$(1)\qquad \partial_t u_i - \operatorname{div}\left(\sum_{j=1}^n A_{ij}(u)\nabla u_j\right) = f_i(u) \quad \text{in } \Omega,\ t>0, \quad i=1,\ldots,n,$$

Date: August 12, 2016.

2000 *Mathematics Subject Classification.* 35K51, 35Q92, 92D25, 60J10.

Key words and phrases. Population dynamics, Shigesada-Kawasaki-Teramoto system, competition model, detailed balance, entropy method, global existence of weak solutions, Onsager's principle.

The first author acknowledges support from the National Natural Science Foundation of China, grant 11471050. The last two authors acknowledge partial support from the Austrian Science Fund (FWF), grants P22108, P24304, and W1245.

with no-flux boundary and initial conditions

$$\sum_{j=1}^n A_{ij}(u)\nabla u_j\cdot\nu = 0 \quad \text{on } \partial\Omega,\ t>0, \quad u_i(\cdot,0)=u_i^0 \quad \text{in } \Omega. \tag{2}$$

Here, u_i models the density of the ith species, $u=(u_1,\ldots,u_n)$, $\Omega\subset\mathbb{R}^d$ ($d\geq 1$) is a bounded domain with Lipschitz boundary, and ν is the exterior unit normal vector to $\partial\Omega$. The diffusion coefficients are given by

$$A_{ij}(u)=\delta_{ij}p_i(u)+u_i\frac{\partial p_i}{\partial u_j}(u), \quad p_i(u)=a_{i0}+\sum_{k=1}^n a_{ik}u_k^s, \quad i,j=1,\ldots,n, \tag{3}$$

where a_{i0}, $a_{ij}\geq 0$ and $s>0$. The functions p_i are the transition rates of the underlying random-walk model [16, 30]. The source terms f_i are of Lotka-Volterra type,

$$f_i(u)=u_i\bigg(b_{i0}-\sum_{j=1}^n b_{ij}u_j\bigg), \quad i=1,\ldots,n, \tag{4}$$

and we suppose that b_{i0}, $b_{ij}\geq 0$ (competition case). Note that (1) can be written more compactly as

$$\partial_t u-\operatorname{div}(A(u)\nabla u)=f(u), \quad f(u)=(f_1(u),\ldots,f_n(u)).$$

State of the art. From a mathematical viewpoint, the analysis of (1)-(2) is highly non-trivial since the diffusion matrix $A(u)$ is neither symmetric nor generally positive definite. Although the maximum principle may be applied to prove the nonnegativity of the densities, it is generally not possible to show upper bounds. Moreover, there is no general regularity theory for diffusion systems, which makes the analysis very delicate. Equations (1) can be written in the form

$$\partial_t u_i-\Delta(u_ip_i(u))=f_i(u), \tag{5}$$

which allows for the proof of an L^{2+s} estimate by the duality method [8, 21], but we will not exploit this method in the paper.

The case of $n=2$ species and linear transition rates $s=1$ corresponds to the original population model of Shigesada, Kawasaki, and Teramota [24],

$$\begin{aligned}\partial_t u_1-\Delta\big(u_1(a_{10}+a_{11}u_1+a_{12}u_2)\big)&=f_1(u),\\ \partial_t u_2-\Delta\big(u_2(a_{20}+a_{21}u_1+a_{22}u_2)\big)&=f_2(u).\end{aligned} \tag{6}$$

The numbers a_{i0} are the diffusion coefficients, a_{ii} are the self-diffusion coefficients, and a_{ij} for $i\neq j$ are called the cross-diffusion coefficients. This model attracted a lot of attention in the mathematical literature. The first global existence result is due to Kim [17] who studied the equations in one space dimension, neglected self-diffusion, and assumed equal coefficients ($a_{ij}=1$). His result was extended to higher space dimensions in [11]. Most of the papers made restrictive structural assumptions, for instance supposing that the diffusion matrix is triangular ($a_{21}=0$), since this allows for the maximum principle in the second equation [1, 18, 20]. Another restriction is to suppose that the cross-diffusion coefficients are small, since in this situation the diffusion matrix becomes positive definite [11, 28].

Significant progress was made by Amann [1] who showed that a priori estimates in the $W^{1,p}$ norm with $p > d$ are sufficient for the solutions to general quasilinear parabolic systems to exist globally in time, and he applied his result to the triangular case. The first global existence result without any restriction on the diffusion coefficients (except positivity) was achieved in [14] in one space dimension and in [3, 4] in several space dimensions. The results were extended to the whole space in [12]. The existence of global classical solutions was proved in, e.g., [19], under suitable conditions on the coefficients.

Nonlinear transition rates, but still for two species, were analyzed by Desvillettes and co-workers, assuming sublinear ($0 < s < 1$) [9] or superlinear rates ($s > 1$) and the weak cross-diffusion condition $((s-1)/(s+1))^2 a_{12}a_{21} \le a_{11}a_{22}$ [10]. Similar results, but under a slightly stronger weak cross-diffusion hypothesis, were proved in [15].

As already mentioned, there are very few results for more than two species. The existence of positive stationary solutions and the stability of the constant equilibrium was investigated in [2, 23]. The existence of global weak solutions in one space dimension assuming a positive definite diffusion matrix was proved in [27], based on Amann's results. Using an entropy approach, the global existence of solutions was shown in [10] for three species under the condition $0 < s < 1/\sqrt{3}$ (which guarantees that $\det(A(u)) > 0$). To our knowledge, a global existence theorem under more general conditions seems to be not available in the literature. In this paper, we prove such a result and relate a structural condition on the coefficients a_{ij} with Onsager's principle of thermodynamics.

Key ideas. Before we state the main results, let us explain our strategy. The idea is to find a priori estimates by employing a Lyapunov functional approach with

$$\mathcal{H}[u] = \int_\Omega h(u)dx = \int_\Omega \sum_{i=1}^n \pi_i h_s(u_i)dx, \tag{7}$$

where $\pi_i > 0$ are some numbers and

$$h_s(z) = \begin{cases} z(\log z - 1) + 1 & \text{for } s = 1, \\ \dfrac{z^s - sz}{s-1} + 1 & \text{for } s \neq 1. \end{cases} \tag{8}$$

Because of the connection of our method to nonequilibrium thermodynamics [16, Section 4.3], we refer to $\mathcal{H}[u]$ as an entropy and to $h(u)$ as an entropy density. Introducing the so-called entropy variable $w = (w_1, \ldots, w_n)$ (called chemical potential in thermodynamics) by

$$w_i = \frac{\partial h}{\partial u_i}(u) = \begin{cases} \pi_i \log u_i & \text{for } s = 1, \\ \dfrac{s\pi_i}{s-1}(u_i^{s-1} - 1) & \text{for } s \neq 1, \end{cases}$$

equations (1) can be written as

$$\partial_t u(w) - \operatorname{div}(B(w)\nabla w) = f(u(w)), \quad B(w) = A(u)H(u)^{-1}, \tag{9}$$

where $u(w) := (h')^{-1}(w)$ is the inverse transformation and $H(u) = h''(u)$ is the Hessian of the entropy density. We claim that if $f = 0$ and $B(w)$ or, equivalently, $H(u)A(u)$ is positive semi-definite[1], $\mathcal{H}[u]$ is a Lyapunov functional along solutions to (1). Indeed, a (formal) computation shows that

$$\frac{d}{dt}\mathcal{H}[u] = -\int_\Omega \nabla w : B(w)\nabla w dx \le 0,$$

which implies that $t \mapsto \mathcal{H}[u(t)]$ is nonincreasing. The entropy method provides more than just the monotonicity of $\mathcal{H}[u]$. If, for instance, $z^\top H(u)A(u)z \ge \sum_{i=1}^n c_i u_i^{\alpha-2} z_i^2$ for some constants $\alpha > 0$, $c_i > 0$, it follows that

$$\frac{d}{dt}\mathcal{H}[u] + \frac{4}{\alpha^2}\int_\Omega \sum_{i=1}^n c_i |\nabla u_i^{\alpha/2}|^2 dx \le 0,$$

which yields gradient estimates for $u_i^{\alpha/2}$. This strategy was employed in many papers on cross-diffusion systems; see, e.g., [3, 4, 9, 12, 14, 15, 30]. In this paper, we introduce two new ideas which we explain for the case $s = 1$ ($s \ne 1$ is studied below).

It is known that the entropy (7) with $\pi_i = 1$ is a Lyapunov functional for the two-species model (6) with $f_1 = f_2 = 0$. This property is generally not satisfied for the corresponding n-species system. Our *first idea* is to introduce the numbers $\pi = (\pi_1, \ldots, \pi_n)$ in the entropy (7). It turns out that (7) is a Lyapunov functional and $H(u)A(u)$ is symmetric and positive definite if

$$\pi_i a_{ij} = \pi_j a_{ji} \quad \text{for all } i, j = 1, \ldots, n. \tag{10}$$

More precisely, this property is *equivalent* to the symmetry of $H(u)A(u)$ (see Proposition 19). We recognize (10) as the detailed balance condition for the Markov chain associated to (a_{ij}). The equivalence of the symmetry and the detailed balance condition is new but not surprising. In fact, the latter condition means that π is a reversible measure, and time-reversibility of a thermodynamic system is equivalent to the symmetry of the so-called Onsager matrix $B(w)$, so symmetry and reversibility are related both from a mathematical and physical viewpoint. We detail these relations in Section 5.1. In Section 2.1, we derive a refined estimate for $H(u)A(u)$ leading to

$$\frac{d}{dt}\mathcal{H}[u] + 4\int_\Omega \sum_{i=1}^n \pi_i a_{i0} |\nabla\sqrt{u_i}|^2 dx + 2\int_\Omega \sum_{i=1}^n \pi_i a_{ii} |\nabla u_i|^2 dx \le 0, \tag{11}$$

and thus giving an H^1 estimate for $\sqrt{u_i}$ (if $a_{i0} > 0$) and u_i (if $a_{ii} > 0$). This is the key estimate for the global existence result. (Below we also take into account the reaction terms (4).)

One may ask whether the detailed balance condition is necessary for the monotonicity of the entropy. It is not. We show that if self-diffusion dominates cross-diffusion in the

[1]We say that an arbitrary matrix $M \in \mathbb{R}^{n\times n}$ is positive (semi-) definite if $z^\top Mz > (\ge)\ 0$ for all $z \in \mathbb{R}^n$, $z \ne 0$.

sense

$$\eta_0 := \min_{i=1,\ldots,n}\left(a_{ii} - \frac{s}{2(s+1)}\sum_{j=1}^n\left(\sqrt{a_{ij}}-\sqrt{a_{ji}}\right)^2\right) > 0, \tag{12}$$

and detailed balance may be not satisfied, then the estimate leading to (11) still holds (with different constants), and global existence follows. (Throughout this paper, we set $\pi_i = 1$ when detailed balance does not hold.) However, if conditions (10) or (12) are both not satisfied, there exist coefficients a_{ij} and initial data u^0 such that $t\mapsto \mathcal{H}[u(t)]$ is increasing on $[0,t_0]$ for some $t_0>0$; see Section 5.3. Numerical experiments (not shown) indicate that after the initial increase, the entropy decays and, in fact, it stays bounded for all time. We conjecture that the entropy is bounded for all time for all nonnegative coefficients and nonnegative initial data and that global existence of weak solutions holds for any (positive) coefficients a_{ij}.

Our results can be extended to nonlinear transition rates of type (3). One may choose more general terms $a_{ij}u_j^{s_j}$ with different exponents s_j but the results are easier to formulate if all exponents are equal. Coefficients with exponents $s\neq 1$ were also considered in [9, 10, 15] but in the two-species case only. We generalize these results to the multi-species case for any $n\geq 2$. The entropy method has to be adapted since the inverse of $h_s'(z) = (s/(s-1))(z^{s-1}-1)$ cannot be defined on $\mathbb{R}$ and thus, $u(w) = (h')^{-1}(w)$ is not defined for all $w\in\mathbb{R}^n$. This issue can be overcome by regularization as in [9, 15]. In fact, we introduce

$$h_\varepsilon(u) = h(u) + \varepsilon\sum_{i=1}^n\big(u_i(\log u_i - 1)+1\big).$$

Then $h_\varepsilon' : (0,\infty)^n\to\mathbb{R}^n$ can be inverted and $(h_\varepsilon')^{-1}:\mathbb{R}^n\to(0,\infty)^n$ is defined on $\mathbb{R}^n$. As a consequence, $u_i = (h_\varepsilon')^{-1}(w)_i$ is positive for any $w\in\mathbb{R}^n$ and even strongly positive if w varies in a compact subset of $\mathbb{R}^n$.

Unfortunately, the product $H_\varepsilon(u)A(u)$, where $H_\varepsilon(u) = h_\varepsilon''(u)$, is generally not positive definite and we need to approximate $A(u)$. In contrast to the approximations suggested in [9, 15], we employ a *non-diagonal* matrix; see (23) below. More specifically, we introduce $A_\varepsilon(u) = A(u) + \varepsilon A^0(u) + \varepsilon^\eta A^1(u)$ with non-diagonal $A^0(u)$, diagonal $A^1(u)$, and $\eta\leq 1/2$ such that

$$z^\top H_\varepsilon(u)A_\varepsilon(u)z \geq z^\top H(u)A(u)z \quad \text{for all } z\in\mathbb{R}^n.$$

The choice of the non-diagonal approximation satisfying this inequality is nontrivial, and this construction is our *second idea.*

Main results. First, we show that global existence of weak solutions holds for linear transition rates ($s=1$). In the following, we set $Q_T = \Omega\times(0,T)$.

Theorem 1 (Global existence for linear transition rates)**.** *Let $T>0$, $s=1$ and $u^0 = (u_1^0,\ldots,u_n^0)$ be such that $u_i^0\geq 0$ for $i=1,\ldots,n$ and $\int_\Omega h(u^0)dx<\infty$. Let either detailed balance and $a_{ii}>0$ for $i=1,\ldots,n$; or (12) hold. Then there exists a weak solution $u=(u_1,\ldots,u_n)$ to (1)-(2) satisfying $u_i\geq 0$ in Ω, $t>0$, and*

$$u_i\in L^2(0,T;H^1(\Omega)),\quad u_i\in L^\infty(0,T;L^1(\Omega)),$$

$$u_i \in L^{2+2/d}(Q_T), \quad \partial_t u_i \in L^{q'}(0,T;W^{1,q}(\Omega)'), \quad i=1,\ldots,n,$$

where $q = 2(d+1)$ *and* $q' = (2d+2)/(2d+1)$*. The solution* u *solves* (1) *in the weak sense*

$$\int_0^T \langle \partial_t u, \phi\rangle dt + \int_0^T \int_\Omega \nabla\phi : A(u)\nabla u dxdt = \int_0^T \int_\Omega f(u)\cdot\phi dxdt \tag{13}$$

for all test functions $\phi \in L^q(0,T;W^{1,q}(\Omega))$*, and the initial condition in* (2) *is satisfied in the sense of* $W^{1,q}(\Omega)'$.

The theorem can be generalized to the case of vanishing self-diffusion, i.e. $a_{ii} = 0$ if detailed balance, $a_{i0} > 0$, and $b_{ii} > 0$ hold; see Remark 12.

Our second result is concerned with nonlinear transition rates ($s \neq 1$). The entropy inequality yields the regularity $u_i \in L^{2s+2/d}(Q_T)$ which may not include L^2 for "small" exponents $s < 1$ and large dimensions d. For this reason, we need to suppose, in the sublinear case, the lower bound $s > 1 - 2/d$ and a weaker growth of the Lotka-Volterra terms:

$$f_i(u) = u_i\bigg(b_{i0} - \sum_{j=1}^n b_{ij}u_j^\sigma\bigg), \quad i=1,\ldots,n, \quad 0 \le \sigma < 2s-1+2/d. \tag{14}$$

The superlinear case ($s > 1$) is somehow easier than the sublinear one since the entropy inequality gives the higher regularity $u_i \in L^p(Q_T)$ with $p > 2$. On the other hand, we need a weak cross-diffusion constraint. More precisely, if detailed balance holds, we require that

$$\eta_1 := \min_{i=1,\ldots,n}\bigg(a_{ii} - \frac{s-1}{s+1}\sum_{j=1,\, j\neq i}^n a_{ij}\bigg) > 0, \tag{15}$$

and if detailed balance does not hold, we suppose that

$$\eta_2 := \min_{i=1,\ldots,n}\bigg(a_{ii} - \frac{1}{2(s+1)}\sum_{j=1,\, j\neq i}\big(s(a_{ij}+a_{ji}) - 2\sqrt{a_{ij}a_{ji}}\big)\bigg) > 0. \tag{16}$$

For $m \ge 2$ and $1 \le q \le \infty$ we introduce the space

$$W_\nu^{m,q}(\Omega) = \{\phi \in W^{m,q}(\Omega) : \nabla\phi\cdot\nu = 0 \text{ on } \partial\Omega\}. \tag{17}$$

Theorem 2 (Global existence for nonlinear transition rates)**.** *Let* $T > 0$*,* $s > \max\{0, 1 - 2/d\}$*, and* u^0 *be such that* $u_i^0 \ge 0$ *for* $i = 1,\ldots,n$ *and* $\int_\Omega h(u^0)dx < \infty$*. If* $s < 1$*, we suppose that* (14) *and either detailed balance and* $a_{ii} > 0$ *for* $i = 1,\ldots,n$*; or* (12) *hold. If* $s > 1$*, we suppose that* (4) *and either detailed balance and* (15) *or* (16) *hold. Then there exist a number* $2 \le q < \infty$ *and a weak solution* $u = (u_1,\ldots,u_n)$ *to* (1)-(2) *satisfying* $u_i \ge 0$ *in* Ω*,* $t > 0$*, and*

$$u_i^s \in L^2(0,T;H^1(\Omega)), \quad u_i \in L^\infty(0,T;L^{\max\{1,s\}}(\Omega)),$$
$$u_i \in L^{p(s)}(Q_T), \quad \partial_t u_i \in L^{q'}(0,T;W_\nu^{m,q}(\Omega)'), \quad i=1,\ldots,n,$$

where $p(s) = 2s + (2/d)\max\{1, s\}$, $1/q + 1/q' = 1$, *and* $m > \max\{1, d/2\}$. *The solution* u *solves* (1) *in the "very weak" sense*

$$\int_0^T \langle \partial_t u, \phi \rangle dt - \int_0^T \int_\Omega \sum_{i=1}^n u_i p_i(u) \Delta\phi_i dxdt = \int_0^T \int_\Omega f(u) \cdot \phi dxdt \tag{18}$$

for all $\phi = (\phi_1, \ldots, \phi_n) \in L^q(0, T; W_\nu^{m,q}(\Omega))$, *and the initial condition holds in the sense of* $W_\nu^{m,q}(\Omega)'$.

In the superlinear case, it can be shown that the solution satisfies (1) in the weak sense (13); see Remark 16. Moreover, for any $s > \max\{0, 1-2/d\}$, it is sufficient to consider test functions from $L^\beta(0, T; W_\nu^{2,\beta}(\Omega))$ with $1/\beta + 1/p(s) = 1$, and the initial condition holds in the sense of $W_\nu^{2,\beta}(\Omega)'$. We can generalize the theorem to the case of vanishing self-diffusion if either $s > \max\{1, d/2\}$; or $0 < s < 1$, $d = 1$, and $\sigma < s + 1$ hold; see Remark 17.

The lower bound $s > 1 - 2/d$ can be avoided if the regularity $u_i \in L^{2+s}(Q_T)$ holds, which is expected to follow from the duality method [8, 21]. Unfortunately, this method is not compatible with our approximation scheme (see (24) below). This issue can possibly be overcome by employing the scheme proposed in [10] which is specialized to diffusion systems like (5). In this paper, however, we prefer to employ scheme (24).

The paper is organized as follows. Section 2 is concerned with the positive definiteness of the matrices $H(u)A(u)$ and $H_\varepsilon(u)A_\varepsilon(u)$. The existence theorems are proved in Sections 3 and 4, respectively. In the final Section 5, we detail the connection between the detailed balance condition and the symmetry of $H(u)A(u)$, prove a nonlinear Aubin-Lions compactness lemma needed in the proof of Theorem 2, and show that the entropy may be increasing initially for special initial data.

2. Positive definiteness of the mobility matrix

We derive sufficient conditions for the positive definiteness of the matrix $H(u)A(u)$. Let $\mathbb{R}_+ = (0, \infty)$. Recall that

$$A_{ij}(u) = \delta_{ij}\bigg(a_{i0} + \sum_{k=1}^n a_{ik}u_k^s\bigg) + s a_{ij} u_i u_j^{s-1}, \quad H_{ij}(u) = \delta_{ij} s \pi_i u_i^{s-2}.$$

The following result is valid for any $s > 0$.

Lemma 3. *Let* $s > 0$. *Then, for any* $z \in \mathbb{R}^n$ *and* $u \in \mathbb{R}_+^n$,

$$\begin{aligned} z^\top H(u)A(u)z \geq{}& s\sum_{i=1}^n \pi_i a_{i0} u_i^{s-2} z_i^2 + s(1-s) \sum_{i,j=1,\, i\neq j}^n \pi_i a_{ij} u_j^s u_i^{s-2} z_i^2 \\ &+ s\sum_{i=1}^n \bigg((s+1)\pi_i a_{ii} - \frac{s}{2}\sum_{j=1}^n \big(\sqrt{\pi_i a_{ij}} - \sqrt{\pi_j a_{ji}}\big)^2 \bigg) u_i^{2(s-1)} z_i^2. \end{aligned} \tag{19}$$

Proof. The elements of the matrix $H(u)A(u)$ equal

$$
\begin{aligned}
(H(u)A(u))_{ij} &= \delta_{ij} s\pi_i \bigg(a_{i0} u_i^{s-2} + \sum_{k=1}^n a_{ik} u_k^s u_i^{s-2} \bigg) + s^2 a_{ij} (u_i u_j)^{s-1} \\
&= \delta_{ij} \big(s\pi_i a_{i0} u_i^{s-2} + s(s+1)\pi_i a_{ii} u_i^{2(s-1)} \big) \\
&\quad + \delta_{ij} s\pi_i \sum_{k=1,\, k\neq i}^n a_{ik} u_k^s u_i^{s-2} + (1-\delta_{ij}) s^2 \pi_i a_{ij} (u_i u_j)^{s-1}.
\end{aligned}
$$

Therefore, for $z \in \mathbb{R}^n$,

$$
\begin{aligned}
z^\top H(u)A(u)z &= s\sum_{i=1}^n \pi_i a_{i0} u_i^{s-2} z_i^2 + s(s+1) \sum_{i=1}^n \pi_i a_{ii} u_i^{2(s-1)} z_i^2 \\
&\quad + s \sum_{i,j=1,\, i\neq j}^n \pi_i a_{ij} u_j^s u_i^{s-2} z_i^2 + s^2 \sum_{i,j=1,\, i\neq j}^n \pi_i a_{ij} (u_i u_j)^{s-1} z_i z_j \qquad (20) \\
&=: I_1 + \cdots + I_4.
\end{aligned}
$$

The sum I_1 is the same as the first term on the right-hand side of (19), and I_2 equals the first part of the last term on this right-hand side. The remaining terms are written as

$$
\begin{aligned}
I_3 + I_4 = s^2 &\sum_{i,j=1,\, i\neq j}^n \pi_i a_{ij} u_j^s u_i^{s-2} z_i^2 + s(1-s) \sum_{i,j=1,\, i\neq j}^n \pi_i a_{ij} u_j^s u_i^{s-2} z_i^2 \\
&+ s^2 \sum_{i,j=1,\, i\neq j}^n \pi_i a_{ij} (u_i u_j)^{s-1} z_i z_j.
\end{aligned}
$$

The second term corresponds to the second term on the right-hand side of (19). Thus, it remains to prove that

$$
\begin{aligned}
J &:= s^2 \sum_{i,j=1,\, i\neq j}^n \pi_i a_{ij} u_j^s u_i^{s-2} z_i^2 + s^2 \sum_{i,j=1,\, i\neq j}^n \pi_i a_{ij} (u_i u_j)^{s-1} z_i z_j \\
&\geq -\frac{s^2}{2} \sum_{j=1}^n \big(\sqrt{\pi_i a_{ij}} - \sqrt{\pi_j a_{ji}} \big)^2 u_i^{2(s-1)} z_i^2.
\end{aligned}
$$

For this, we employ twice the inequality $b^2 + c^2 \geq 2bc$:

$$
\begin{aligned}
J = s^2 &\sum_{i,j=1,\, i<j}^n \pi_i a_{ij} u_j^s u_i^{s-2} z_i^2 + s^2 \sum_{i,j=1,\, i>j}^n \pi_i a_{ij} u_j^s u_i^{s-2} z_i^2 \\
&+ s^2 \sum_{i,j=1,\, i<j}^n \pi_i a_{ij} (u_i u_j)^{s-1} z_i z_j + s^2 \sum_{i,j=1,\, i>j}^n \pi_i a_{ij} (u_i u_j)^{s-1} z_i z_j
\end{aligned}
$$

$$
\begin{aligned}
&= s^2 \sum_{i,j=1,\, i<j}^n \Big(\pi_i a_{ij} u_j^s u_i^{s-2} z_i^2 + \pi_j a_{ji} u_i^s u_j^{s-2} z_j^2 + (\pi_i a_{ij} + \pi_j a_{ji})(u_i u_j)^{s-1} z_i z_j \Big) \\
&\geq s^2 \sum_{i,j=1,\, i<j}^n \Big(2\sqrt{\pi_i a_{ij} \pi_j a_{ji}} (u_i u_j)^{s-1} |z_i z_j| - (\pi_i a_{ij} + \pi_j a_{ji})(u_i u_j)^{s-1} |z_i z_j| \Big) \\
&= -s^2 \sum_{i,j=1,\, i<j}^n \big(\sqrt{\pi_i a_{ij}} - \sqrt{\pi_j a_{ji}} \big)^2 \big| (u_i^{s-1} z_i)(u_j^{s-1} z_j) \big| \\
&\geq -\frac{s^2}{2} \sum_{i,j=1,\, i<j}^n \big(\sqrt{\pi_i a_{ij}} - \sqrt{\pi_j a_{ji}} \big)^2 \big((u_i^{s-1} z_i)^2 + (u_j^{s-1} z_j)^2 \big) \\
&= -\frac{s^2}{2} \sum_{i,j=1,\, i\neq j}^n \big(\sqrt{\pi_i a_{ij}} - \sqrt{\pi_j a_{ji}} \big)^2 (u_i^{s-1} z_i)^2.
\end{aligned}
$$

This finishes the proof. □

2.1. **Sublinear and linear transition rates.**

For $s \leq 1$, Lemma 3 provides immediately the positive definiteness of $H(u)A(u)$ if detailed balance (10) holds. However, we can derive a sharper result.

Lemma 4 (Detailed balance)**.** *Let $0 < s \leq 1$ and $\pi_i a_{ij} = \pi_j a_{ji}$ for all $i \neq j$. Then, for all $z \in \mathbb{R}^n$ and $u \in \mathbb{R}^n_+$,*

$$
\begin{aligned}
z^\top H(u)A(u)z &\geq s \sum_{i=1}^n \pi_i u_i^{s-2} \big(a_{i0} + (s+1) a_{ii} u_i^s \big) z_i^2 \\
&\quad + \frac{s^2}{2} \sum_{i,j=1,\, i\neq j}^n \pi_i a_{ij} (u_i u_j)^{s-1} \left(\sqrt{\frac{u_j}{u_i}} z_i + \sqrt{\frac{u_i}{u_j}} z_j \right)^2. \qquad (21)
\end{aligned}
$$

Proof. The sum of the terms I_1 and I_2 in (20) is exactly the first term on the right-hand side of (21). Using detailed balance, we find that

$$
\begin{aligned}
I_3 + I_4 &= \frac{s}{2} \sum_{i,j=1,\, i\neq j}^n \pi_i a_{ij} (u_i u_j)^{s-1} \frac{u_j}{u_i} z_i^2 + \frac{s}{2} \sum_{i,j=1,\, i\neq j}^n \pi_i a_{ij} (u_j u_i)^{s-1} \frac{u_i}{u_j} z_j^2 \\
&\quad + s^2 \sum_{i,j=1,\, i\neq j}^n \pi_i a_{ij} (u_i u_j)^{s-1} z_i z_j \\
&= \frac{s^2}{2} \sum_{i,j=1,\, i\neq j}^n \pi_i a_{ij} (u_i u_j)^{s-1} \frac{u_j}{u_i} z_i^2 + \frac{s^2}{2} \sum_{i,j=1,\, i\neq j}^n \pi_i a_{ij} (u_j u_i)^{s-1} \frac{u_i}{u_j} z_j^2 \\
&\quad + s^2 \sum_{i,j=1,\, i\neq j}^n \pi_i a_{ij} (u_i u_j)^{s-1} z_i z_j + \frac{s}{2}(1-s) \sum_{i,j=1,\, i\neq j}^n \pi_i a_{ij} (u_i u_j)^{s-1} \frac{u_j}{u_i} z_i^2
\end{aligned}
$$

$$+\frac{s}{2}(1-s)\sum_{i,j=1,\,i\neq j}^{n}\pi_i a_{ij}(u_ju_i)^{s-1}\frac{u_i}{u_j}z_j^2.$$

The sum of the first three terms equal the second term on the right-hand side of (21), and the remaining two terms are nonnegative since $s\le 1$. ☐

Remark 5. In the existence proof, we will choose $z_i=\nabla u_i$ (with a slight abuse of notation). Then the first term in (21) gives an estimate for $\nabla u_i^{s/2}$ in L^2 (if $a_{i0}>0$) and the better bound $\nabla u_i^s\in L^2$ (if $a_{ii}>0$). If $a_{ii}=0$, we lose the latter regularity. This loss can be compensated by the last term in (21) giving

$$(u_iu_j)^{s-1}\left|\sqrt{\frac{u_j}{u_i}}\nabla u_i+\sqrt{\frac{u_i}{u_j}}\nabla u_j\right|^2=\frac{4}{s^2}|\nabla(u_iu_j)^{s/2}|^2,\quad i\neq j,$$

and consequently a bound for $\nabla(u_iu_j)^{s/2}$ in L^2. This observation is used in Remark 12. ☐

Lemma 6 (Non detailed balance)**.** *Let $0<s\le 1$. If*

$$\eta_0:=\min_{i=1,\ldots,n}\left(a_{ii}-\frac{s}{2(s+1)}\sum_{j=1}^{n}\left(\sqrt{a_{ij}}-\sqrt{a_{ji}}\right)^2\right)\ge 0,$$

then $H(u)A(u)$ is positive definite. Under the slightly stronger condition $\eta_0>0$, it holds for all $z\in\mathbb{R}^n$ and $u\in\mathbb{R}^n_+$ that

$$z^\top H(u)A(u)z\ge s\sum_{i=1}^{n}a_{i0}u_i^{s-2}z_i^2+\eta_0 s(s+1)\sum_{i=1}^{n}u_i^{2(s-1)}z_i^2.$$

The lemma follows from Lemma 3 after choosing $\pi_i=1$ for $i=1,\ldots,n$. Observe that $\eta_0>0$ holds if $a_{ii}>0$ for all i and (a_{ij}) is symmetric.

It is possible to show the positive definiteness of $H(u)A(u)$ without any restriction on (a_{ij}) (except positivity) if we restrict the choice of the parameter s; see the following lemma.

Lemma 7. *Let $a_{ij}+a_{ji}>0$ for $i,j=1,\ldots,n$ and $0<s\le s_0$, where*

$$s_0:=\min_{i,j=1,\ldots,n}\frac{2\sqrt{a_{ij}a_{ji}}}{a_{ij}+a_{ji}}\le 1.$$

Then, for all $z\in\mathbb{R}^n$ and $u\in\mathbb{R}^n_+$,

$$z^\top H(u)A(u)z\ge s\sum_{i=1}^{n}a_{i0}u_i^{s-2}z_i^2+s(s+1)\sum_{i=1}^{n}a_{ii}u_i^{2(s-1)}z_i^2.$$

Proof. We choose $\pi_i=1$ for $i=1,\ldots,n$. With the notation of the proof of Lemma 3, we only need to show that $I_3+I_4\ge 0$. Employing the inequality $b^2+c^2\ge 2bc$, we find that

$$I_3+I_4=s\sum_{i,j=1,\,i<j}^{n}\Big(a_{ij}u_j^su_i^{s-2}z_i^2+a_{ji}u_i^su_j^{s-2}z_j^2+s(a_{ij}+a_{ji})(u_iu_j)^{s-1}z_iz_j\Big)$$

$$\geq s \sum_{i,j=1,\,i<j}^{n} \Big(2\sqrt{a_{ij}a_{ji}}(u_iu_j)^{s-1}|z_iz_j| - s(a_{ij}+a_{ji})(u_iu_j)^{s-1}|z_iz_j| \Big)$$
$$= s \sum_{i,j=1,\,i<j}^{n} (a_{ij}+a_{ji}) \bigg(\frac{2\sqrt{a_{ij}a_{ji}}}{a_{ij}+a_{ji}} - s \bigg) (u_iu_j)^{s-1}|z_iz_j|,$$

and this expression is nonnegative if $s \leq s_0$. ☐

2.2. **Superlinear transition rates.** Again, we assume first that detailed balance holds.

Lemma 8 (Detailed balance). *Let $s > 1$ and $\pi_i a_{ij} = \pi_j a_{ji}$ for all $i \neq j$. If*

$$\eta_1 := \min_{i=1,\ldots,n} \bigg(a_{ii} - \frac{s-1}{s+1} \sum_{j=1,\,j\neq i}^{n} a_{ij} \bigg) \geq 0,$$

then $H(u)A(u)$ is positive definite. Furthermore, if $\eta_1 > 0$, then, for all $z \in \mathbb{R}^n$ and $u \in \mathbb{R}^n_+$,

$$z^\top H(u)A(u)z \geq s \sum_{i=1}^{n} \pi_i a_{i0} u_i^{s-2} z_i^2 + \eta_1 s(s+1) \sum_{i=1}^{n} \pi_i u_i^{2(s-1)} z_i^2.$$

Proof. It is sufficient to estimate the sum $I_3 + I_4$, defined in the proof of Lemma 3:

$$I_3 + I_4 = s \sum_{i,j=1,\,i<j}^{n} \Big(\pi_i a_{ij} u_j^s u_i^{s-2} z_i^2 + \pi_j a_{ji} u_i^s u_j^{s-2} z_j^2 + s(\pi_i a_{ij} + \pi_j a_{ji})(u_iu_j)^{s-1} z_iz_j \Big)$$
$$\geq s \sum_{i,j=1,\,i<j}^{n} \Big(2\sqrt{\pi_i a_{ij}\pi_j a_{ji}}(u_iu_j)^{s-1}|z_iz_j| - s(\pi_i a_{ij} + \pi_j a_{ji})(u_iu_j)^{s-1}|z_iz_j| \Big)$$
$$= -s \sum_{i,j=1,\,i<j}^{n} \Big(s(\pi_i a_{ij} + \pi_j a_{ji}) - 2\sqrt{\pi_i a_{ij}\pi_j a_{ji}} \Big)(u_iu_j)^{s-1}|z_iz_j|$$
$$\geq -\frac{s}{2} \sum_{i,j=1,\,i<j}^{n} \Big(s(\pi_i a_{ij} + \pi_j a_{ji}) - 2\sqrt{\pi_i a_{ij}\pi_j a_{ji}} \Big)\Big((u_i^{s-1}z_i)^2 + (u_j^{s-1}z_j)^2 \Big)$$
$$= -\frac{s}{2} \sum_{i,j=1,\,i\neq j}^{n} \Big(s(\pi_i a_{ij} + \pi_j a_{ji}) - 2\sqrt{\pi_i a_{ij}\pi_j a_{ji}} \Big)(u_i^{s-1}z_i)^2.$$

This expression simplifies because of the detailed balance condition:

$$I_3 + I_4 \geq -s(s-1) \sum_{i,j=1,\,i\neq j}^{n} \pi_i a_{ij} (u_i^{s-1} z_i)^2,$$

and we end up with

$$z^\top H(u)A(u)z \geq s \sum_{i=1}^{n} \pi_i a_{i0} u_i^{s-2} z_i^2 + s(s+1) \sum_{i=1}^{n} \pi_i \bigg(a_{ii} - \frac{s-1}{s+1} \sum_{j=1,\,j\neq i}^{n} a_{ij} \bigg) u_i^{2(s-1)} z_i^2,$$

from which we conclude the result. □

Remark 9. Let $n = 2$. Then the condition $\eta_1 \geq 0$ on the coefficients (a_{ij}) becomes $a_{11} \geq a_{12}(s-1)/(s+1)$ and $a_{22} \geq a_{21}(s-1)/(s+1)$. The product

$$a_{11}a_{22} \geq \left(\frac{s-1}{s+1}\right)^2 a_{12}a_{21}$$

is the same as the condition imposed in [10, Section 5.1] but weaker than

$$a_{11}a_{22} \geq \left(\frac{s-1}{s}\right)^2 a_{12}a_{21},$$

which was needed in [15, Lemma 11]. Furthermore, under the slightly stronger condition $\eta_1 > 0$, that is

$$a_{11}a_{22} > \left(\frac{s-1}{s+1}\right)^2 a_{12}a_{21},$$

our weak solution satisfies the stronger estimate $u_i^s \in L^2(0,T;H^1(\Omega))$ than that in [10, Section 5.1]. □

Lemma 10 (Non detailed balance)**.** *Let $s > 1$ and let*

$$\eta_2 := \min_{i=1,\ldots,n}\left(a_{ii} - \frac{1}{2(s+1)}\sum_{j=1,\,j\neq i}\big(s(a_{ij}+a_{ji}) - 2\sqrt{a_{ij}a_{ji}}\big)\right) \geq 0.$$

Then $H(u)A(u)$ is positive definite. Moreover, if $\eta_2 > 0$, then, for all $z \in \mathbb{R}^n$ and $u \in \mathbb{R}_+^n$,

$$z^\top H(u)A(u)z \geq s\sum_{i=1}^n a_{i0}u_i^{s-2}z_i^2 + \eta_2 s(s+1)\sum_{i=1}^n u_i^{2(s-1)}z_i^2.$$

Proof. We choose $\pi_i = 1$ for $i = 1,\ldots,n$. Then, as in the previous proof,

$$I_3 + I_4 \geq -\frac{s}{2}\sum_{i,j=1,\,i\neq j}^n \big(s(a_{ij}+a_{ji}) - 2\sqrt{a_{ij}a_{ji}}\big)u_i^{2(s-1)}z_i^2$$

and

$$\begin{aligned} z^\top H(u)A(u)z &\geq s\sum_{i=1}^n a_{i0}u_i^{s-2}z_i^2 + s(s+1)\sum_{i=1}^n a_{ii}u_i^{2(s-1)}z_i^2 \\ &\quad - \frac{s}{2}\sum_{i,j=1,\,i\neq j}^n \big(s(a_{ij}+a_{ji}) - 2\sqrt{a_{ij}a_{ji}}\big)u_i^{2(s-1)}z_i^2 \\ &= s\sum_{i=1}^n a_{i0}u_i^{s-2}z_i^2 \\ &\quad + s(s+1)\sum_{i=1}^n\left(a_{ii} - \frac{1}{2(s+1)}\sum_{i,j=1,\,i\neq j}^n \big(s(a_{ij}+a_{ji}) - 2\sqrt{a_{ij}a_{ji}}\big)\right)u_i^{2(s-1)}z_i^2. \end{aligned}$$

By definition of η_2, the result follows. □

2.3. **Approximate matrices.** Our theory requires that the range of the derivative h' equals $\mathbb{R}^n$. Since this is not the case if $s \neq 1$, we need to approximate the entropy density and consequently also the diffusion matrix. The approximate entropy density

$$h_\varepsilon(u) = h(u) + \varepsilon \sum_{i=1}^n \big(u_i(\log u_i - 1) + 1\big) \tag{22}$$

possesses the property that the range of its derivative is $\mathbb{R}^n$. We set $H(u) = h''(u) = (\delta_{ij} s\pi_i u_i^{s-2})_{i,j=1,\ldots,n}$ for its Hessian and

$$\begin{aligned} &H_\varepsilon(u) = H(u) + \varepsilon H^0(u), \quad H^0_{ij}(u) = \delta_{ij} u_i^{-1}, \\ &A_\varepsilon(u) = A(u) + \varepsilon A^0(u) + \varepsilon^\eta A^1(u), \end{aligned} \tag{23}$$

where $\eta < 1/2$ and

$$\begin{aligned} A^0_{ij}(u) &= \delta_{ij}\frac{u_i}{\pi_i}\mu_i - (1-\delta_{ij})\frac{u_i}{\pi_i}a_{ji}, \quad A^1_{ij}(u) = \delta_{ij} u_i, \\ \mu_i &:= \frac{\pi_i}{2}\sum_{j=1,\, j\neq i}^n \left(\frac{a_{ji}}{\pi_i} + \frac{a_{ij}}{\pi_j}\right), \quad i = 1,\ldots,n. \end{aligned}$$

The approximation $\varepsilon^\eta A^1(u)$ is needed to achieve bounds for $\varepsilon^{(\eta+1)/2}\nabla u_i$ in L^2, which are necessary for the limit $\varepsilon \to 0$. The off-diagonal terms in $A^0(u)$ are needed to preserve the entropy structure in the sense that $H_\varepsilon(u)A_\varepsilon(u)$ is still positive definite. This is shown in the following lemma.

Lemma 11. *Let $s > 0$. Then, for all $z \in \mathbb{R}^n$ and $u \in \mathbb{R}^n_+$,*

$$z^\top H_\varepsilon(u)A_\varepsilon(u)z \geq z^\top H(u)A(u)z + \varepsilon^\eta s\sum_{i=1}^n \pi_i u_i^{s-1} z_i^2 + \varepsilon^{\eta+1}\sum_{i=1}^n z_i^2.$$

Proof. We decompose the product $H_\varepsilon(u)A_\varepsilon(u)$ as

$$\begin{aligned} H_\varepsilon(u)A_\varepsilon(u) = H(u)A(u) &+ \varepsilon^\eta H_\varepsilon(u)A^1(u) + \varepsilon\big(H^0(u)A(u) + H(u)A^0(u)\big) \\ &+ \varepsilon^2 H^0(u)A^0(u). \end{aligned}$$

The ε^2-term becomes

$$\begin{aligned} (H^0(u)A^0(u))_{ij} &= \sum_{k=1}^n \delta_{ik} u_k^{-1}\left(\delta_{kj}\frac{u_k}{\pi_k}\mu_k - (1-\delta_{kj})\frac{u_k}{\pi_k}a_{jk}\right) \\ &= \delta_{ij}\frac{\mu_i}{\pi_i} - (1-\delta_{ij})\frac{a_{ji}}{\pi_i}. \end{aligned}$$

We obtain for $z \in \mathbb{R}^n$:

$$z^\top H^0(u)A^0(u)z = \sum_{i=1}^n \frac{\mu_i}{\pi_i}z_i^2 - \sum_{i,j=1,\, i\neq j}^n \frac{a_{ji}}{\pi_i}z_i z_j$$

$$\begin{aligned}&\geq \sum_{i=1}^n \frac{\mu_i}{\pi_i}z_i^2 - \frac12\sum_{i,j=1,\,i\neq j}^n \frac{a_{ji}}{\pi_i}(z_i^2+z_j^2) \\ &= \sum_{i=1}^n \frac{\mu_i}{\pi_i}z_i^2 - \frac12\sum_{i=1}^n\bigg(\sum_{j=1,\,j\neq i}^n\frac{a_{ji}}{\pi_i}\bigg)z_i^2 - \frac12\sum_{i=1}^n\bigg(\sum_{j=1,\,j\neq i}^n\frac{a_{ij}}{\pi_j}\bigg)z_i^2 \\ &= 0.\end{aligned}$$

Next, we consider the ε-terms:

$$\begin{aligned}(H^0(u)A(u))_{ij} &= \sum_{k=1}^n \delta_{ik}u_i^{-1}\bigg(\delta_{kj}\bigg(a_{k0}+\sum_{\ell=1}^n a_{k\ell}u_\ell^s + sa_{kk}u_k^s\bigg) + (1-\delta_{kj})sa_{kj}u_j^{s-1}u_k\bigg) \\ &= \delta_{ij}\bigg(a_{i0}u_i^{-1}+\sum_{\ell=1}^n a_{i\ell}u_\ell^s u_i^{-1} + sa_{ii}u_i^{s-1}\bigg) + (1-\delta_{ij})sa_{ij}u_j^{s-1}, \\ (H(u)A^0(u))_{ij} &= \sum_{k=1}^n \delta_{ik}s\pi_i u_i^{s-2}\bigg(\delta_{kj}\frac{u_k}{\pi_k}\mu_k - (1-\delta_{kj})\frac{u_k}{\pi_k}a_{jk}\bigg) \\ &= \delta_{ij}su_i^{s-1}\mu_i - (1-\delta_{ij})sa_{ji}u_i^{s-1}.\end{aligned}$$

Summing these expressions and neglecting some positive contributions, we find that

$$\begin{aligned}z^\top\big(H^0(u)A(u)+H(u)A^0(u)\big)z &\geq \sum_{i=1}^n (a_{i0}u_i^{-1}+sa_{ii}u_i^{s-1})z_i^2 \\ &\quad + s\sum_{i,j=1}^n(1-\delta_{ij})a_{ij}u_j^{s-1}z_iz_j - s\sum_{i,j=1}^n(1-\delta_{ij})a_{ji}u_i^{s-1}z_iz_j \\ &= \sum_{i=1}^n \big(a_{i0}u_i^{-1}+sa_{ii}u_i^{s-1}\big)z_i^2 \geq s\sum_{i=1}^n a_{ii}u_i^{s-1}z_i^2.\end{aligned}$$

Here we see how we constructed $A^0_{ij}(u)$: The off-diagonal coefficients are chosen in such a way that the mixed terms in z_iz_j cancel, and the diagonal elements (namely μ_i) are sufficiently large to obtain positive definiteness of $H^0(u)A^0(u)$. Finally, we have $(H_\varepsilon(u)A^1(u))_{ij}$ $= \delta_{ij}(s\pi_i u_i^{s-1}+\varepsilon)$ and

$$z^\top H_\varepsilon(u)A^1(u)z = \sum_{i=1}^n(s\pi_iu_i^{s-1}+\varepsilon)z_i^2,$$

which proves the lemma. □

3. Linear transition rates: proof of Theorem 1

In this section, we prove Theorem 1. Let $T>0$, $N\in\mathbb{N}$, $\tau=T/N$, $\varepsilon>0$, and $m\in\mathbb{N}$ with $m>d/2$. This ensures that the embedding $H^m(\Omega)\hookrightarrow L^\infty(\Omega)$ is compact. We assume that $u_i^0(x)\in[a,b]$ for $x\in\Omega$, $i=1,\ldots,n$, where $0<a<b<\infty$. Then, clearly, $w^0=$

$h'(u^0) \in L^\infty(\Omega;\mathbb{R}^n)$. For general $u_i^0 \geq 0$, we may first consider $u_\varepsilon^0 = (Q_\varepsilon(u_1^0),\ldots,Q_\varepsilon(u_n^0))$, where $0<\varepsilon<1$ and Q_ε is the cut-off function

$$Q_\varepsilon(z) = \begin{cases} \varepsilon & \text{for } 0 \leq z < \varepsilon, \\ z & \text{for } \varepsilon \leq z < \varepsilon^{-1/2}, \\ \varepsilon^{-1/2} & \text{for } z \geq \varepsilon^{-1/2}, \end{cases}$$

and then pass to the limit $\varepsilon \to 0$. We leave the details to the reader.

Step 1: solution of an approximated problem. Given $w^{k-1} \in L^\infty(\Omega;\mathbb{R}^n)$ for $k\in\mathbb{N}$, we wish to find $w^k \in H^m(\Omega;\mathbb{R}^n)$ such that

$$\begin{aligned} &\frac{1}{\tau}\int_\Omega \big(u(w^k) - u(w^{k-1})\big)\cdot \phi dx + \int_\Omega \nabla\phi : B(w^k)\nabla w^k dx \\ &\quad + \varepsilon \int_\Omega \bigg(\sum_{|\alpha|=m} D^\alpha w^k \cdot D^\alpha \phi + w^k\cdot\phi\bigg)dx = \int_\Omega f(u(w^k))\cdot\phi dx \end{aligned} \tag{24}$$

for all $\phi \in H^m(\Omega;\mathbb{R}^n)$. Here, $u(w^k) = (h')^{-1}(w^k)$, $B(w^k) = A(u(w^k))H(u(w^k))^{-1}$, $\alpha = (\alpha_1,\ldots,\alpha_d) \in \mathbb{N}_0^n$ with $|\alpha| = \alpha_1 + \ldots + \alpha_d = m$ is a multiindex, and $D^\alpha = \partial^{|\alpha|}/(\partial x_1^{\alpha_1}\cdots\partial x_d^{\alpha_d})$ is a partial derivative of order m. If $k=1$, we define $w^0 = h'(u^0)$. Equation (24) is an implicit Euler discretization of (1) including an H^m regularization term.

We recall that the entropy is given by

$$\mathcal{H}[u] = \int_\Omega h(u)dx = \int_\Omega \sum_{i=1}^n \pi_i h_1(u_i)dx, \quad h_1(u_i) = u_i(\log u_i - 1) + 1.$$

Then the entropy variables equal $w_i = \partial h/\partial u_i = \pi_i \log u_i$. In particular, $h' : \mathbb{R}_+^n \to \mathbb{R}^n$ is invertible on $\mathbb{R}^n$, i.e., Hypothesis (H1) in [15] is satisfied. By Lemmas 4 and 6, $H(u)A(u)$ is positive definite, i.e., Hypothesis (H2) in [15] holds as well. (At this step, we only need that $H(u)A(u)$ is positive semi-definite.) Furthermore, f_i grows at most linearly which implies that

$$\sum_{i=1}^n f_i(u)\pi_i \log u_i \leq C_f(1+h(u)),$$

where $C_f>0$ depends only on (b_{ij}) and π. This means that Hypothesis (H3) in [15] is also satisfied. Thus, we can apply Lemma 5 in [15] giving a weak solution $w^k \in H^m(\Omega;\mathbb{R}^n)$ to (24) satisfying the discrete entropy inequality

$$\begin{aligned} &(1-C_f\tau)\int_\Omega h(u(w^k))dx + \tau\int_\Omega \nabla w^k : B(w^k)\nabla w^k dx \\ &\quad + \varepsilon\tau \int_\Omega \bigg(\sum_{|\alpha|=m} |D^\alpha w^k|^2 + |w^k|^2\bigg)dx \leq \int_\Omega h(u(w^{k-1}))dx + C_f\tau \mathrm{meas}(\Omega). \end{aligned} \tag{25}$$

Step 2: uniform estimates. We set $u^k = u(w^k)$ and introduce the piecewise in time constant functions $w^{(\tau)}(x,t) = w^k(x)$ and $u^{(\tau)}(x,t) = u^k(x)$ for $x\in\Omega$, $t\in((k-1)\tau,k\tau]$. At time $t=0$, we set $w^{(\tau)}(\cdot,0) = h'(u^0) = w^0$ and $u^{(\tau)}(\cdot,0) = u^0$. Let $u^{(\tau)} = (u_1^{(\tau)},\ldots,u_n^{(\tau)})$.

We define the backward shift operator $(\sigma_\tau u^{(\tau)})(x,t) = u(w^{k-1}(x))$ for $x \in \Omega$, $t \in ((k-1)\tau, k\tau]$. Then $u^{(\tau)}$ solves

$$\begin{aligned}\frac{1}{\tau}\int_0^T\int_\Omega(u^{(\tau)} - \sigma_\tau u^{(\tau)})\cdot\phi dxdt + \int_0^T\int_\Omega \nabla\phi : B(w^{(\tau)})\nabla w^{(\tau)}dxdt \\ + \varepsilon\int_0^T\int_\Omega\bigg(\sum_{|\alpha|=m} D^\alpha w^{(\tau)}\cdot D^\alpha\phi + w^{(\tau)}\cdot\phi\bigg)dxdt = \int_0^T\int_\Omega f(u^{(\tau)})\cdot\phi dxdt\end{aligned} \tag{26}$$

for piecewise constant functions $\phi : (0,T) \to H^m(\Omega;\mathbb{R}^n)$. By a density argument, this equation also holds for all $\phi \in L^2(0,T;H^m(\Omega;\mathbb{R}^n))$ [22, Prop. 1.36].

By Lemmas 4 and 6, we have

$$\nabla w^k : B(w^k)\nabla w^k = \nabla u^k : H(u^k)A(u^k)\nabla u^k \geq 2\eta_0\sum_{i=1}^n |\nabla u_i^k|^2,$$

where $\eta_0 = \min_{i=1,\ldots,n}\pi_i a_{ii} > 0$ if detailed balance holds, and $\eta_0 > 0$ is given by (12) otherwise. By the generalized Poincaré inequality [26, Chapter 2, Section 1.4], it holds that

$$\int_\Omega\bigg(\sum_{|\alpha|=m}|D^\alpha w^k|^2 + |w^k|^2\bigg)dx \geq C_P\|w^k\|^2_{H^m(\Omega)},$$

where $C_P > 0$ is the Poincaré constant. Then the discrete entropy inequality (25) gives

$$\begin{aligned}(1 - C_f\tau)\int_\Omega h(u^k)dx + 2\eta_0\tau\int_\Omega|\nabla u^k|^2dx + \varepsilon C_P\tau\|w^k\|^2_{H^m(\Omega)} \\ \leq \int_\Omega h(u^{k-1})dx + C_f\tau\mathrm{meas}(\Omega).\end{aligned}$$

Summing these inequalities over $k = 1,\ldots,j$, it follows that

$$\begin{aligned}(1 - C_f\tau)\int_\Omega h(u^j)dx + 2\eta_0\tau\sum_{j=1}^k\int_\Omega|\nabla u^k|^2dx + \varepsilon C_P\tau\sum_{j=1}^k\|w^k\|^2_{H^m(\Omega)} \\ \leq \int_\Omega h(u^0)dx + C_f\tau\sum_{k=1}^{j-1}\int_\Omega h(u^k)dx + C_fT\mathrm{meas}(\Omega).\end{aligned}$$

By the discrete Gronwall inequality [6], if $\tau < 1/C_f$,

$$\int_\Omega h(u^j)dx + \tau\sum_{j=1}^k\int_\Omega|\nabla u^k|^2dx + \varepsilon\tau\sum_{j=1}^k\|w^k\|^2_{H^m(\Omega)} \leq C,$$

where here and in the following, $C > 0$ denotes a generic constant independent of τ and ε. Then, observing that the entropy density dominates the L^1 norm and consequently, $u^{(\tau)}$ is uniformly bounded in $L^\infty(0,T;L^1(\Omega;\mathbb{R}^n))$, we obtain

$$\|u^{(\tau)}\|_{L^\infty(0,T;L^1(\Omega))} + \|u^{(\tau)}\|_{L^2(0,T;H^1(\Omega))} + \varepsilon^{1/2}\|w^{(\tau)}\|_{L^2(0,T;H^m(\Omega))} \leq C. \tag{27}$$

We wish to derive more a priori estimates. Set $Q_T = \Omega\times(0,T)$. The Gagliardo-Nirenberg inequality with $p = 2 + 2/d$ and $\theta = 2d(p-1)/(dp+2p) \in [0,1]$ (such that $\theta p = 2$) yields for $i = 1,\ldots,n$,

$$\begin{aligned} \|u_i^{(\tau)}\|_{L^p(Q_T)}^p &= \int_0^T \|u_i^{(\tau)}\|_{L^p(\Omega)}^p dt \le C\int_0^T \|u_i^{(\tau)}\|_{H^1(\Omega)}^{\theta p}\|u_i^{(\tau)}\|_{L^1(\Omega)}^{(1-\theta)p}dt \\ &\le C\|u_i^{(\tau)}\|_{L^\infty(0,T;L^1(\Omega))}^{(1-\theta)p}\|u_i^{(\tau)}\|_{L^2(0,T;H^1(\Omega))}^{\theta p} \le C. \end{aligned} \tag{28}$$

In order to apply a compactness result, we need a uniform estimate for the discrete time derivative of $u^{(\tau)}$. Let $q = 2(d+1)$ and $\phi \in L^q(0,T;W^{m,q}(\Omega;\mathbb{R}^n))$. Then $1/p+1/q+1/2 = 1$ and, by Hölder's inequality,

$$\begin{aligned} \frac{1}{\tau}\left|\int_0^T\int_\Omega (u^{(\tau)} - \sigma_\tau u^{(\tau)})\cdot\phi dxdt\right| &\le \sum_{i,j=1}^n \|A_{ij}(u^{(\tau)})\|_{L^p(Q_T)}\|\nabla u_j^{(\tau)}\|_{L^2(Q_T)}\|\nabla\phi_i\|_{L^q(Q_T)} \\ &\quad + \varepsilon\|w^{(\tau)}\|_{L^2(0,T;H^m(\Omega))}\|\phi\|_{L^2(0,T;H^m(\Omega))} \\ &\quad + \|f(u^{(\tau)})\|_{L^{q'}(Q_T)}\|\phi\|_{L^q(Q_T)}, \end{aligned}$$

where $q' = (2d+2)/(2d+1)$. Estimate (28) and the linear growth of $A_{ij}(u^{(\tau)})$ with respect to $u^{(\tau)}$ show that the first term on the right-hand side is bounded. The second term is bounded because of (27). Finally, $|f_i(u^{(\tau)})|$ is growing at most like $(u_i^{(\tau)})^2$ such that

$$\|f(u^{(\tau)})\|_{L^{q'}(Q_T)} \le C\big(1 + \|u^{(\tau)}\|_{L^{2q'}(Q_T)}^2\big) \le C,$$

since $2q' \le p$. We conclude that

$$\tau^{-1}\|u^{(\tau)} - \sigma_\tau u^{(\tau)}\|_{L^{q'}(0,T;W^{m,q}(\Omega)')} \le C. \tag{29}$$

Step 3: the limit $(\varepsilon,\tau)\to 0$. In view of (27) and (29), we can apply the Aubin-Lions lemma in the version of [13], which yields the existence of a subsequence, which is not relabeled, such that, as $(\tau,\varepsilon)\to 0$,

$$u^{(\tau)} \to u \quad \text{strongly in } L^2(Q_T) \text{ and a.e.,} \tag{30}$$

$$u^{(\tau)} \rightharpoonup u \quad \text{weakly in } L^2(0,T;H^1(\Omega)), \tag{31}$$

$$\varepsilon w^{(\tau)} \to 0 \quad \text{strongly in } L^2(0,T;H^m(\Omega)), \tag{32}$$

$$\tau^{-1}(u^{(\tau)} - \sigma_\tau u^{(\tau)}) \rightharpoonup \partial_t u \quad \text{weakly in } L^{q'}(0,T;W^{m,q}(\Omega)'), \tag{33}$$

where $u = (u_1,\ldots,u_n)$. In view of the a.e. convergence (30) and the uniform bound (28), we have

$$u^{(\tau)} \to u \quad \text{strongly in } L^\gamma(Q_T) \text{ for all } \gamma < 2 + 2/d. \tag{34}$$

Then, together with (31),

$$u_i^{(\tau)}\nabla u_j^{(\tau)} \rightharpoonup u_i\nabla u_j \quad \text{weakly in } L^1(Q_T).$$

We deduce from the $L^{q'}(Q_T)$ bound for $A(u^{(\tau)})\nabla u^{(\tau)}$ that

$$B(w^{(\tau)})\nabla w^{(\tau)} = A(u^{(\tau)})\nabla u^{(\tau)} \rightharpoonup A(u)\nabla u \quad \text{weakly in } L^{q'}(Q_T).$$

Furthermore, taking into account (34) and the uniform bound for $f_i(u^{(\tau)})$ in $L^{q'}(Q_T)$,

$$f_i(u^{(\tau)}) \rightharpoonup f_i(u) \quad \text{weakly in } L^{q'}(Q_T).$$

Then (32) and (33) allow us to perform the limit $(\varepsilon,\tau)\to 0$ in (26) with $\phi \in L^q(0,T; W^{m,q}(\Omega))$, which directly yields (13). Since $\partial_t u = \operatorname{div}(A(u)\nabla u)+f(u) \in L^{q'}(0,T;W^{1,q}(\Omega)')$, a density argument shows that the weak formulation holds for all $\phi \in L^q(0,T;W^{1,q}(\Omega))$. Moreover, $u_i \in W^{1,q'}(0,T;W^{1,q}(\Omega)') \hookrightarrow C^0([0,T];W^{1,q}(\Omega)')$, which shows that the initial condition is satisfied in $W^{1,q}(\Omega)'$. This ends the proof.

Remark 12 (Detailed balance and vanishing self-diffusion)**.** In the detailed balance case, we may allow for vanishing self-diffusion. If $a_{ii}=0$ but $a_{i0}>0$, Lemma 4 implies that only $\nabla(u_i^{(\tau)})^{1/2}$ is bounded in $L^2(Q_T)$. This situation was considered in [4] for the two-species case, and we sketch the generalization to the n-species case.

Applying the Gagliardo-Nirenberg inequality similarly as in Step 2 of the previous proof, we conclude that $(u_i^{(\tau)})^{1/2} \in L^{\widetilde{p}}(Q_T)$ with $\widetilde{p}=2+4/d$. Then

$$\|\nabla u_i^{(\tau)}\|_{L^{\widetilde{q}}(Q_T)} = 2\|(u_i^{(\tau)})^{1/2}\|_{L^{\widetilde{p}}(Q_T)}\,\|\nabla(u_i^{(\tau)})^{1/2}\|_{L^2(Q_T)} \le C, \quad \widetilde{q}=\frac{d+2}{d+1},$$

and thus, $(u_i^{(\tau)})$ is bounded in $L^{\widetilde{q}}(0,T;W^{1,\widetilde{q}}(\Omega))$ instead of $L^2(0,T;H^1(\Omega))$. This loss of regularity is problematic for the estimate of the discrete time derivative of $u_i^{(\tau)}$. In order to compensate this, we need the last sum in (21). Indeed, Remark 5 shows that for any $i\neq j$, $(u_i^{(\tau)}u_j^{(\tau)})^{1/2}$ is bounded in $L^2(0,T;H^1(\Omega))$. Moreover, $(u_i^{(\tau)}u_j^{(\tau)})^{1/2}$ is bounded in $L^\infty(0,T;L^1(\Omega))$. We infer from the Gagliardo-Nirenberg inequality that $(u_i^{(\tau)}u_j^{(\tau)})^{1/2}$ is bounded in $L^p(Q_T)$ with $p=2+2/d$.

Next we exploit the structure of the equations,

$$\sum_{j=1}^n A_{ij}(u^{(\tau)})\nabla u_j^{(\tau)} = \nabla(u_i^{(\tau)}p_i(u^{(\tau)})), \quad p_i(u^{(\tau)}) = a_{i0}+\sum_{j=1}^n a_{ij}u_j^{(\tau)}.$$

Thus, to show that $A_{ij}(u^{(\tau)})\nabla u_j^{(\tau)}$ is bounded, we only need to verify that $\nabla(u_i^{(\tau)}u_j^{(\tau)})$ is bounded:

$$\|\nabla(u_i^{(\tau)}u_j^{(\tau)})\|_{L^{q'}(Q_T)} \le 2\|(u_i^{(\tau)}u_j^{(\tau)})^{1/2}\|_{L^p(Q_T)}\|\nabla(u_i^{(\tau)}u_j^{(\tau)})^{1/2}\|_{L^2(Q_T)} \le C,$$

where $q'=(2d+2)/(2d+1)$. The estimate for the Lotka-Volterra term is more delicate since we have only the regularity $u_i^{(\tau)} \in L^{1+1/d}(Q_T)$. Here, we need to suppose that $b_{ii}>0$, since this assumption provides an estimate for $(u_i^{(\tau)})^2\log u_i^{(\tau)}$ in $L^1(Q_T)$. Then the discrete time derivative of $u_i^{(\tau)}$ is bounded in $L^1(0,T;W^{m,q}(\Omega)')$ – but not in $L^{q'}(0,T;W^{m,q}(\Omega)')$. By the Aubin-Lions lemma, there exists a subsequence (not relabeled) such that, as $(\varepsilon,\tau)\to 0$,

$$u_i^{(\tau)} \to u_i \quad \text{strongly in } L^{q'}(Q_T).$$

The problem now is to show that (a subsequence of) the discrete time derivative of $u_i^{(\tau)}$ converges to $\partial_t u_i$ since $L^1(0,T;W^{m,q}(\Omega)')$ is not reflexive. The idea is to apply a result from [29] which provides a criterium for weak compactness in $L^1(0,T;X)$, where X is a reflexive Banach space. For details, we refer to [4]. □

4. Nonlinear transition rates: proof of Theorem 2

The strategy of the proof is similar to the proof of Theorem 1 but the nonlinear transition rates complicate the proof significantly. As outlined in Section 2.3, we approximate the entropy density by (22) and the diffusion matrix by (23). Again, we assume without loss of generality that $u_i^0(x) \in [a,b]$ for $x \in \Omega$, $i = 1,\ldots,n$, where $0 < a < b < \infty$.

Step 1: solution of an approximated problem. We employ the transformation $w_i = \partial h_\varepsilon/\partial u_i$ and define $B_\varepsilon(w) = A_\varepsilon(u(w))H_\varepsilon(u(w))^{-1}$. Given $w^{k-1} \in L^\infty(\Omega;\mathbb{R}^n)$, we wish to find $w^k \in H^m(\Omega;\mathbb{R}^n)$ solving

$$\begin{aligned}\frac{1}{\tau}\int_\Omega \big(u(w^k) - u(w^{k-1})\big)\cdot\phi dx + \int_\Omega \nabla\phi : B_\varepsilon(w^k)\nabla w^k dx \\ + \varepsilon\int_\Omega\bigg(\sum_{|\alpha|=m} D^\alpha w^k\cdot D^\alpha\phi + w^k\cdot\phi\bigg)dx = \int_\Omega f(u(w^k))\cdot\phi dx \end{aligned}\tag{35}$$

for all $\phi \in H^m(\Omega;\mathbb{R}^n)$. If $k = 1$, we define $w^0 = h_\varepsilon'(u^0)$ such that $u(w^0) = u^0$.

The construction of h_ε ensures that Hypothesis (H1) of [15] is satisfied. By Lemma 11, Hypothesis (H2) holds as well. Also Hypothesis (H3) holds true since, for some $C_f > 0$,

$$f(u)\cdot w = \sum_{I=1}^n\bigg(b_{i0} - \sum_{j=1}^n b_{ij}u_j^\sigma\bigg)(su_i^s + \varepsilon u_i\log u_i) \le C_f(1 + h_\varepsilon(u)),$$

where $\sigma = 1$ if $s > 1$ and $0 \le \sigma \le \max\{0, 2s-1+2/d\}$ if $s < 1$. We apply Lemma 5 in [15] to deduce the existence of a weak solution $w^k \in H^m(\Omega;\mathbb{R}^n)$ to the above problem, which satisfies the discrete entropy inequality

$$\begin{aligned}(1 - C_f\tau)\int_\Omega h_\varepsilon(u(w^k))dx + \tau\int_\Omega \nabla w^k : B_\varepsilon(w^k)\nabla w^k dx \\ + \varepsilon\tau\int_\Omega\bigg(\sum_{|\alpha|=m}|D^\alpha w^k|^2 + |w^k|^2\bigg)dx \le \int_\Omega h_\varepsilon(u(w^{k-1}))dx + C_f\tau\mathrm{meas}(\Omega).\end{aligned}\tag{36}$$

Setting $u^k := u(w^k)$ and employing Lemma 11, the second integral can be estimated as follows:

$$\begin{aligned}\int_\Omega \nabla w^k : B_\varepsilon(w^k)\nabla w^k dx &= \int_\Omega u^k : H_\varepsilon(u^k)A_\varepsilon(u^k)\nabla u^k dx \\ &\ge s(s+1)\int_\Omega\sum_{i=1}^n \min\{a_{ii}\pi_i, \eta_0, \eta_1\pi_i, \eta_2\}(u_i^k)^{2(s-1)}|\nabla u_i^k|^2 dx \\ &\quad + \varepsilon^\eta s\int_\Omega\sum_{i=1}^n \pi_i(u_i^k)^{s-1}|\nabla u_i^k|^2 dx + \varepsilon^{\eta+1}\int_\Omega\sum_{i=1}^n|\nabla u_i^k|^2 dx\end{aligned}\tag{37}$$

$$\geq C_s \int_\Omega \sum_{i=1}^n |\nabla (u_i^k)^s|^2 dx$$
$$+ \frac{4\varepsilon^\eta s}{(s+1)^2} \int_\Omega \sum_{i=1}^n \pi_i |\nabla (u_i^k)^{(s+1)/2}|^2 dx + \varepsilon^{\eta+1} \int_\Omega \sum_{i=1}^n |\nabla u_i^k|^2 dx,$$

where $C_s = s^{-1}(s+1)\min\{a_{11}\pi_1, \ldots, a_{nn}\pi_n, \eta_0, \eta_1\pi_1, \ldots, \eta_1\pi_n, \eta_2\}$.

To finish this step, we wish to write the "very weak" formulation for the solution $u^{(\tau)}$, which is defined from u^k as in the previous section. First, we observe that

$$(B_\varepsilon(w^k)\nabla w^k)_i = (A_\varepsilon(u^k)\nabla u^k)_i = \varepsilon(A^0(u^k)\nabla u^k)_i + \varepsilon^\eta (A^1(u^k)\nabla u^k)_i + \nabla(u_i^k p_i(u^k))$$
$$= \varepsilon(A^0(u^k)\nabla u^k)_i + \frac{\varepsilon^\eta}{2}\nabla(u_i^k)^2 + \nabla(u_i^k p_i(u^k)).$$

Next, we choose a test function $\phi = (\phi_1, \ldots, \phi_n) \in L^q(0,T;\, W_\nu^{m,q}(\Omega))$, where $m > \max\{1, d/2\}$ and $q \geq 2$ will be determined below. Recall that $W_\nu^{m,q}(\Omega)$ is defined in (17). Integrating by parts in (35), $u^{(\tau)}$ solves

$$\frac{1}{\tau}\int_0^T \int_\Omega \big(u^{(\tau)} - \sigma_\tau u^{(\tau)}\big)\cdot \phi dxdt - \int_0^T \int_\Omega \sum_{i=1}^n u_i^{(\tau)} p_i(u^{(\tau)}) \Delta\phi_i dxdt$$
$$+ \varepsilon \int_0^T \int_\Omega \nabla\phi : A^0(u^{(\tau)})\nabla u^{(\tau)} dxdt - \frac{\varepsilon^\eta}{2}\int_0^T \int_\Omega \sum_{i=1}^n (u_i^{(\tau)})^2 \Delta\phi_i dxdt \tag{38}$$
$$+ \varepsilon \int_0^T \int_\Omega \bigg(\sum_{|\alpha|=m} D^\alpha w^{(\tau)} \cdot D^\alpha \phi + w^{(\tau)}\cdot\phi\bigg) dxdt = \int_0^T \int_\Omega f(u^{(\tau)})\cdot \phi dxdt.$$

Step 2: uniform estimates. Arguing as in Step 2 of the proof of Theorem 1, we obtain from (36) and (37) for suffiently small $\tau > 0$ the following uniform estimates.

Lemma 13. *It holds for $i = 1, \ldots, n$ that*

$$\|u_i^{(\tau)}\|_{L^\infty(0,T;L^{\max\{1,s\}}(\Omega))} + \|(u_i^{(\tau)})^s\|_{L^2(0,T;H^1(\Omega))} \leq C, \tag{39}$$
$$\varepsilon^{\eta/2}\|(u_i^{(\tau)})^{(s+1)/2}\|_{L^2(0,T;H^1(\Omega))} + \varepsilon^{(\eta+1)/2}\|u_i^{(\tau)}\|_{L^2(0,T;H^1(\Omega))} \leq C, \tag{40}$$
$$\varepsilon^{1/2}\|w_i^{(\tau)}\|_{L^2(0,T;H^m(\Omega))} \leq C. \tag{41}$$

Here, we used the fact that $\int_\Omega h_\varepsilon(u^0)dx$ is uniformly bounded and that $s < 1$ implies that $u^{(\tau)} \leq C(1 + h(u^{(\tau)}))$ for some $C > 0$, from which we deduce that $(u_i^{(\tau)})$ is bounded in $L^\infty(0,T;L^1(\Omega))$. We need more a priori estimates.

Lemma 14. *Let $s > \max\{0, 1-2/d\}$. It holds that*

$$\|u^{(\tau)}\|_{L^{p(s)}(Q_T)} + \varepsilon^{\eta/r(s)}\|u^{(\tau)}\|_{L^{r(s)}(Q_T)} \leq C, \tag{42}$$

where $p(s) = 2s + (2/d)\max\{1,s\}$ and $r(s) = s + 1 + (2/d)\max\{1,s\} > 2$.

Proof. The estimates are consequences of Lemma 13 and the Gagliardo-Nirenberg inequality. First, let $s < 1$. We employ the Gagliardo-Nirenberg inequality, with $\theta = ds/(ds+1) \in (0,1)$:

$$\begin{aligned}\|u_i^{(\tau)}\|_{L^{p(s)}(Q_T)}^{p(s)} &= \int_0^T \|(u_i^{(\tau)})^s\|_{L^{p(s)/s}(\Omega)}^{p(s)/s} dt \le C\int_0^T \|(u_i^{(\tau)})^s\|_{H^1(\Omega)}^{\theta p(s)/s} \|(u_i^{(\tau)})^s\|_{L^{1/s}(\Omega)}^{(1-\theta)p(s)/s} dt \\ &\le C\|u_i^{(\tau)}\|_{L^\infty(0,T;L^1(\Omega))}^{(1-\theta)p(s)} \int_0^T \|(u_i^{(\tau)})^s\|_{H^1(\Omega)}^{\theta p(s)/s} dt, \quad i = 1,\ldots,n.\end{aligned}$$

It holds that $\theta p(s)/s = 2$. By (39), $\|u^{(\tau)}\|_{L^{p(s)}(Q_T)} \le C$.

Next, let $s > 1$. Then, with $\theta = d/(d+1) \in (0,1)$,

$$\begin{aligned}\|(u_i^{(\tau)})^s\|_{L^{2+2/d}(Q_T)}^{2+2/d} &\le C\int_0^T \|(u_i^{(\tau)})^s\|_{H^1(\Omega)}^{\theta(2+2/d)} \|(u_i^{(\tau)})^s\|_{L^1(\Omega)}^{(1-\theta)(2+2/d)} dt \\ &\le C\|(u_i^{(\tau)})^s\|_{L^2(0,T;H^1(\Omega)}^2 \|u_i^{(\tau)}\|_{L^\infty(0,T;L^s(\Omega))}^{s(1-\theta)(2+2/d)} \le C,\end{aligned}$$

again taking into account estimate (39). This shows that $(u^{(\tau)})$ is bounded in $L^{p(s)}(Q_T)$.

Finally, let $\max\{0, 1-2/d\} < s < 1$. Then $r(s) = s+1+2/d$. We apply the Gagliardo-Nirenberg inequality with $\theta = d(s+1)/(2+d(s+1)) \in (0,1)$ such that $\theta \cdot 2r(s)/(s+1) = 2$,

$$\begin{aligned}\varepsilon^\eta \|u_i^{(\tau)}\|_{L^{r(s)}(Q_T)}^{r(s)} &= \varepsilon^\eta \|(u_i^{(\tau)})^{(s+1)/2}\|_{L^{2r(s)/(s+1)}(Q_T)}^{2r(s)/(s+1)} \\ &\le \varepsilon^\eta C\int_0^T \|(u_i^{(\tau)})^{(s+1)/2}\|_{H^1(\Omega)}^{2r(s)\theta/(s+1)} \|(u_i^{(\tau)})^{(s+1)/2}\|_{L^{2/(s+1)}(\Omega)}^{2r(s)(1-\theta)/(s+1)} dt \\ &\le C\varepsilon^\eta \|(u_i^{(\tau)})^{(s+1)/2}\|_{L^2(0,T;H^1(\Omega))}^2 \|u_i^{(\tau)}\|_{L^\infty(0,T;L^1(\Omega))}^{(1-\theta)r(s)} \le C,\end{aligned}$$

using (39) and (40). If $s > 1$, we have $r(s) = s+1+2s/d$, and applying the Gagliardo-Nirenberg inequality with $\theta = d(s+1)/(2s+d(s+1)) \in (0,1)$, we obtain in a similar way as above

$$\begin{aligned}\varepsilon^\eta \|u_i^{(\tau)}\|_{L^{r(s)}(Q_T)}^{r(s)} &= \varepsilon^\eta \|(u_i^{(\tau)})^{(s+1)/2}\|_{L^{2r(s)/(s+1)}(Q_T)}^{2r(s)/(s+1)} \\ &\le \varepsilon^\eta C\int_0^T \|(u_i^{(\tau)})^{(s+1)/2}\|_{H^1(\Omega)}^{2r(s)\theta/(s+1)} \|(u_i^{(\tau)})^{(s+1)/2}\|_{L^{2s/(s+1)}(\Omega)}^{2r(s)(1-\theta)/(s+1)} dt \\ &\le C\varepsilon^\eta \|(u_i^{(\tau)})^{(s+1)/2}\|_{L^2(0,T;H^1(\Omega))}^2 \|u_i^{(\tau)}\|_{L^\infty(0,T;L^s(\Omega))}^{(1-\theta)r(s)} \le C.\end{aligned}$$

This shows the lemma. ☐

Lemma 15. *Let $s > \max\{0, 1-2/d\}$ and $m > \max\{1, d/2\}$. Then there exist $2 \le q < \infty$ and $C > 0$ such that*

$$\tau^{-1}\|u^{(\tau)} - \sigma_\tau u^{(\tau)}\|_{L^{q'}(0,T;W^{m,q}(\Omega)')} \le C, \tag{43}$$

and $1/q + 1/q' = 1$.

Proof. Let $\phi \in L^q(0,T;W_\nu^{m,q}(\Omega))$, where $q \geq 2$ has to be determined. Recall that $W_\nu^{m,q}(\Omega)$ is defined in (17) and that $m > \max\{1,d/2\}$. Then, by (38),

$$\begin{aligned}
\tau^{-1}\bigg|\int_\Omega (u^{(\tau)} - \sigma_\tau u^{(\tau)})\cdot\phi dx\bigg| &\leq \varepsilon\|w^{(\tau)}\|_{L^2(0,T;H^m(\Omega))}\|\phi\|_{L^2(0,T;H^m(\Omega))} \\
&\quad + \sum_{i=1}^n \|u_i^{(\tau)}p_i(u^{(\tau)})\|_{L^{q'}(Q_T)}\|\Delta\phi_i\|_{L^q(Q_T)} + \varepsilon\sum_{i,j=1}^n \|A_{ij}^0(u^{(\tau)})\nabla u_j^{(\tau)}\|_{L^{q'}(Q_T)}\|\nabla\phi_j\|_{L^q(Q_T)} \\
&\quad + \frac{\varepsilon^\eta}{2}\sum_{i=1}^n \|(u_i^{(\tau)})^2\|_{L^{q'}(Q_T)}\|\Delta\phi_i\|_{L^q(Q_T)} + \|f(u^{(\tau)})\|_{L^{q'}(Q_T)}\|\phi\|_{L^q(Q_T)} \\
&=: I_1 + \cdots + I_5,
\end{aligned} \tag{44}$$

where $1/q + 1/q' = 1$.

By (41), I_1 is bounded. We deduce from (42) that $u_i^{(\tau)}(u_j^{(\tau)})^s$ is uniformly bounded in $L^{p(s)/(s+1)}(Q_T)$, and so does $u_i^{(\tau)}p_i(u^{(\tau)})$. As $s > 1 - 2/d$, we have $q_1 := p(s)/(s+1) > 1$. We conclude that I_2 is bounded with $q' \leq \min\{2, q_1\}$.

Since $A_{ij}^0(u^{(\tau)})$ depends linearly on $u^{(\tau)}$, it is sufficient to prove that $\varepsilon u_i^{(\tau)}\nabla u_j^{(\tau)}$ is uniformly bounded in some $L^{q_2}(Q_T)$ for all i,j. Let $q_2 = 2r(s)/(r(s)+2)$, where $r(s) = s + 1 + 2/d$ is defined in Lemma 14. As $r(s) > 2$, it holds that $q_2 > 1$. Then, by Hölder's inequality, (40), and (42),

$$\varepsilon^{\eta/r(s)+(\eta+1)/2}\|u_i^{(\tau)}\nabla u_j^{(\tau)}\|_{L^{q_2}(Q_T)} \leq \varepsilon^{\eta/r(s)}\|u_i^{(\tau)}\|_{L^{r(s)}(Q_T)}\cdot\varepsilon^{(\eta+1)/2}\|\nabla u_j^{(\tau)}\|_{L^2(Q_T)} \leq C.$$

The property $r(s) > 2$ also implies that $\eta/r(s) + (\eta+1)/2 < 1$. This shows the bound on I_3 with $q' \leq \min\{2, q_2\}$.

Set $q_3 = r(s)/2 > 1$. Using the second estimate in (42) and $1 - 2/r(s) > 0$, we find that

$$\varepsilon^\eta\|(u_i^{(\tau)})^2\|_{L^{q_3}(Q_T)} = \varepsilon^{(1-2/r(s))\eta}\big(\varepsilon^{\eta/r(s)}\|u_i^{(\tau)}\|_{L^{r(s)}(Q_T)}\big)^2 \leq C,$$

proving that I_4 is bounded with $q' \leq \min\{2, q_3\}$.

Finally, in view of (14), $|f_i(u^{(\tau)})|$ grows at most like $(u_i^{(\tau)})^{1+\sigma}$, where $\sigma = 1$ if $s > 1$ and $\sigma < 2s - 1 + 2/d$ if $s < 1$. Therefore, we have $q_4 := p(s)/(1+\sigma) > 1$ and

$$\|f(u^{(\tau)})\|_{L^{q_4}(Q_T)} \leq C\big(1 + \|u^{(\tau)}\|^{1+\sigma}_{L^{(1+\sigma)q_4}(Q_T)}\big) = C\big(1 + \|u^{(\tau)}\|^{1+\sigma}_{L^{p(s)}(Q_T)}\big) \leq C.$$

Hence, I_5 is bounded with $q' \leq \min\{2, q_4\}$. We conclude that the lemma follows with $q' := \min\{2, q_1, q_2, q_3, q_4\} > 1$ and $q = q'/(q'-1)$. □

Step 3: the limit $(\varepsilon,\tau) \to 0$. Estimates (39) and (43) allow us to apply the nonlinear Aubin-Lions lemma (Theorem 21 if $s \geq 1/2$ or Theorem 22 if $s < 1/2$) to obtain the existence of a subsequence which is not relabeled such that, as $(\varepsilon,\tau) \to 0$,

$$u^{(\tau)} \to u \quad \text{strongly in } L^\gamma(Q_T) \quad \text{for all } 1 \leq \gamma < p(s).$$

In particular, $u^{(\tau)} \to u$ a.e. in Q_T. By estimates (39), (41), and (43), we have, up to subsequences,

$$\begin{aligned} (u_i^{(\tau)})^s \rightharpoonup u_i^s &\quad \text{weakly in } L^2(0,T;H^1(\Omega)), \\ \varepsilon w^{(\tau)} \to 0 &\quad \text{strongly in } L^2(0,T;H^m(\Omega)), \\ \tau^{-1}(u^{(\tau)} - \sigma_\tau u^{(\tau)}) \rightharpoonup \partial_t u &\quad \text{weakly in } L^{q'}(0,T;W^{m,q}(\Omega)'). \end{aligned}$$

We have shown in the proof of Lemma 15 that $(u_i^{(\tau)} p_i(u^{(\tau)}))$ is bounded in $L^{p(s)/(s+1)}(Q_T)$. Taking into account the a.e. convergence $u_i^{(\tau)} p_i(u^{(\tau)}) \to u_i p_i(u)$ in Q_T, we infer that

$$u_i^{(\tau)} p_i(u^{(\tau)}) \to u_i p_i(u) \quad \text{strongly in } L^1(Q_T).$$

Furthermore, we proved that $(\varepsilon^{\eta/r(s)+(\eta+1)/2} A_{ij}^0(u^{(\tau)}) \nabla u_j^{(\tau)})$ is bounded in $L^{q_2}(Q_T)$ with $q_2 = 2r(s)/(r(s)+2)$ such that

$$\begin{aligned} \varepsilon A_{ij}^0(u^{(\tau)}) \nabla u_j^{(\tau)} &= \varepsilon^{1-\eta/r(s)-(\eta+1)/2} \cdot \varepsilon^{\eta/r(s)+(\eta+1)/2} A_{ij}^0(u^{(\tau)}) \nabla u_j^{(\tau)} \\ &\to 0 \quad \text{strongly in } L^1(Q_T). \end{aligned}$$

Here, we used the fact that $\eta/r(s) + (\eta+1)/2 < 1$ such that $\varepsilon^{1-\eta/r(s)-(\eta+1)/2} \to 0$ as $\varepsilon \to 0$. We know from (42) that $(\varepsilon^{\eta/r(s)} u_i^{(\tau)})$ is bounded in $L^2(Q_T)$. Consequently,

$$\varepsilon^\eta (u_i^{(\tau)})^2 = \varepsilon^{\eta(1-2/r(s))} \big(\varepsilon^{\eta/r(s)} u_i^{(\tau)} \big)^2 \to 0 \quad \text{strongly in } L^1(Q_T),$$

since $\varepsilon^{\eta(1-2/r(s))} \to 0$ as $\varepsilon \to 0$ because of $r(s) > 2$. Finally, $f_i(u^{(\tau)}) \to f_i(u)$ a.e. and the uniform bound $\|f_i(u^{(\tau)}\|_{L^{q_4}(Q_T)} \le C$ with $q_4 = p(s)/(1+\sigma) > 1$ imply that

$$f_i(u^{(\tau)}) \to f_i(u) \quad \text{strongly in } L^1(Q_T).$$

Then, performing the limit $(\varepsilon,\tau) \to 0$ in (38) with $\phi \in L^\infty(0,T;W_\nu^{m,\infty}(\Omega))$, it follows that u solves (18) for such test functions. A density argument shows that, in fact, u solves (18) for $\phi \in L^q(0,T;W_\nu^{m,q}(\Omega))$, finishing the proof.

Remark 16 (Weak formulation)**.** In the superlinear case $s > 1$, the solution constructed in the previous proof satisfies (1) even in the weak sense (13) with test functions $\phi \in L^q(0,T;W^{1,q}(\Omega))$. In order to see this, it is sufficient to show that

$$A_{ij}(u^{(\tau)}) \nabla u_j^{(\tau)} \rightharpoonup A_{ij}(u) \nabla u_j \quad \text{weakly in } L^{q'}(Q_T)$$

for some $1 < q' \le 2$. Because of the structure of A_{ij}, we only need to verify that

$$\begin{aligned} u_i^{(\tau)} (u_j^{(\tau)})^{s-1} \nabla u_j^{(\tau)} \rightharpoonup u_i u_j^{s-1} \nabla u_j &\quad \text{weakly in } L^{q'}(Q_T), \\ (u_i^{(\tau)})^s \nabla u_j^{(\tau)} \rightharpoonup u_i^s \nabla u_j &\quad \text{weakly in } L^{q'}(Q_T). \end{aligned}$$

Indeed, we have the convergences $u_i^{(\tau)} \to u_i$ strongly in $L^\gamma(Q_T)$ for any $2 < \gamma < p(s)$ and $(u_i^{(\tau)})^s \rightharpoonup u_i^s$ weakly in $L^2(0,T;H^1(\Omega))$ and hence,

$$u_i^{(\tau)} (u_j^{(\tau)})^{s-1} \nabla u_j^{(\tau)} = s^{-1} u_i^{(\tau)} \nabla (u_j^{(\tau)})^s \rightharpoonup s^{-1} u_i \nabla u_j^s = u_i u_j^{s-1} \nabla u_j \quad \text{weakly in } L^{q'}(Q_T),$$

choosing $q' = 2\gamma/(\gamma+2) > 1$. For the remaining convergence, we need to integrate by parts. It holds for $\phi_i \in L^q(0,T;W^{2,q}_\nu(\Omega))$ that

$$\begin{aligned}\int_0^T\int_\Omega (u_i^{(\tau)})^s\nabla u_j^{(\tau)}\cdot\nabla\phi_i dxdt \\ &= -\int_0^T\int_\Omega u_j^{(\tau)}\nabla(u_i^{(\tau)})^s\cdot\nabla\phi_i dxdt - \int_0^T\int_\Omega (u_i^{(\tau)})^s u_j^{(\tau)}\Delta\phi_i dxdt \\ &\to -\int_0^T u_j\nabla u_i^s\cdot\nabla\phi_i dxdt - \int_0^T\int_\Omega u_i^s u_j\Delta\phi_i dxdt = \int_0^t\int_\Omega u_i^s\nabla u_j\cdot\nabla\phi_i dxdt.\end{aligned}$$

A density argument shows that the weak formulation also holds for $\phi_i \in L^q(0,T;W^{1,q}(\Omega))$. ☐

Remark 17 (Vanishing self-diffusion)**.** Assume that $a_{i0} > 0$ and $a_{ii} = 0$. The difficulty is that we obtain a uniform bound only for $\nabla(u_i^{(\tau)})^{s/2}$ instead for $\nabla(u_i^{(\tau)})^s$ in $L^2(Q_T)$. In order to compensate this loss of regularity, we need additional assumptions, namely either $s > \max\{1, d/2\}$ (superlinear rates); or $0 < s < 1$, $d = 1$, and $\sigma < s+1$ (sublinear rates). Under these conditions, the statement of Theorem 2 holds true.

For the proof, we remark that the regularity for $(u_i^{(\tau)})^s$ in $L^2(0,T;H^1(\Omega))$ is employed in the estimate of $u_i^{(\tau)}$ in $L^{p(s)}(Q_T)$. If only $(u_i^{(\tau)})^{s/2}$ is bounded in $L^2(0,T;H^1(\Omega))$, the Gagliardo-Nirenberg inequality gives a weaker result: for $0 < s < 1$ with $\theta = ds/(ds+2)$ and $\rho = s + 2/d$,

$$\begin{aligned}\|u_i^{(\tau)}\|^\rho_{L^\rho(Q_T)} &= \|(u_i^{(\tau)})^{s/2}\|^{2\rho/s}_{L^{2\rho/s}(Q_T)} \le C\int_0^T \|(u_i^{(\tau)})^{s/2}\|^{2\theta\rho/s}_{H^1(\Omega)}\|(u_i^{(\tau)})^{s/2}\|^{2(1-\theta)\rho/s}_{L^{2/s}(\Omega)}dt \\ &\le C\|(u_i^{(\tau)})^{s/2}\|^2_{L^2(0,T;H^1(\Omega))}\|u_i^{(\tau)}\|^{(1-\theta)\rho}_{L^\infty(0,T;L^1(\Omega))} \le C,\end{aligned}$$

since $2\theta\rho/s = 2$; and for $s > 1$ with $\theta = d/(d+2)$ and $\rho = s + 2s/d$,

$$\|u_i^{(\tau)}\|^\rho_{L^\rho(Q_T)} \le C\|(u_i^{(\tau)})^{s/2}\|^2_{L^2(0,T;H^1(\Omega))}\|u_i^{(\tau)}\|^{(1-\theta)\rho}_{L^\infty(0,T;L^s(\Omega))} \le C,$$

since $2\theta\rho/s = 2$. Consequently, $(u_i^{(\tau)})$ is bounded in $L^\rho(Q_T)$ with $\rho = s + (2/d)\max\{1,s\}$.

We claim that this estimate is sufficient to derive a bound for the discrete time derivative. Since the ε-terms in (44) do not need the estimate for $u_i^{(\tau)}$ in $L^\rho(Q_T)$, it is sufficient to bound $u_i^{(\tau)}p_i(u^{(\tau)})$ and $(u_i^{(\tau)})^{\sigma+1}$ in some $L^{q'}(Q_T)$ with $q' > 1$. This is possible as long as $\rho > s+1$ and $\rho > \sigma+1$, respectively. If $0 < s < 1$, these two inequalities are equivalent to $d = 1$ and $\sigma < s - 1 + d/2 = s+1$. If $s > 1$ (in this case $\sigma = 1$), they give the restriction $s > d/2$, thus $s > \max\{1, d/2\}$. This shows the claim. ☐

5. Additional and auxiliary results

5.1. **Detailed balance condition.** We wish to interpret the detailed balance condition (10) and to explain how the numbers π_i can be computed from the coefficients (a_{ij}). We assume that the coefficients are normalized in the sense that $a_{ij} \ge 0$ and $\sum_{k=1,\,k\neq j} a_{kj} \le 1$ for all i, j. The idea is to use a probabilistic approach, interpreting the coefficients a_{ij} as the transition rates between two discrete states i and j of the state space $S := \{1, \dots, n\}$. Then

$$a_{ij} = \mathrm{P}(X_k = j | X_{k-1} = i)$$

is the conditional probability for a random variable $X : \mathbb{N} \to S$. This variable represents the Markov chain associated to the stochastic matrix $Q = (Q_{ij})_{i,j} \in \mathbb{R}^{n\times n}$, defined by $Q_{ij} = a_{ij}$ for $i \neq j$ and $Q_{ii} = 1 - \sum_{i=1,\, i\neq j} a_{ij}$ for $i = 1, \dots, n$. A Markov chain is called reversible if there exists a probability distribution $\pi = (\pi_1, \dots, \pi_n)$ on S (called an invariant measure) such that

$$\pi_i a_{ij} = \pi_j a_{ji}, \quad i, j = 1, \dots, n. \tag{45}$$

The Markov chain can be interpreted as a directed graph, where the states $i \in S$ are the nodes and the edges are labeled by the probabilities a_{ij} going from state i to state j.

The state space S can be partitioned into so-called communicating classes. We write $i \to j$ if there exist $i_0, i_1, \dots, i_{n+1} \in S$ such that $a_{i_0,i_1} a_{i_1,i_2} \cdots a_{i_n,i_{n+1}} > 0$ for $i_0 = i$ and $i_{n+1} = j$. We say that i communicates with j if both $i \to j$ and $j \to i$. A set of states $\sigma \subset S$ is a communicating class if every pair in σ communicates with each other. This defines an equivalence relation, and communicating classes are the equivalence classes.

Consider the following properties:

(A1) For all $i, j \in S$, it holds that either $a_{ij} = a_{ji} = 0$ or $a_{ij} a_{ji} > 0$.

(A2) For any periodic cycle $i_0, i_1, \dots, i_{m+1} = i_0$,

$$\prod_{k=0}^{m} a_{i_k,i_{k+1}} = \prod_{k=0}^{m} a_{i_{k+1},i_k}.$$

The detailed balance condition (45) implies (A1) and (A2). It is shown in [25] that the converse is true and that the invariant measure π can be constructed explicitly.

Proposition 18. *Let (A1)-(A2) hold. Then there exists an invariant measure $\pi = (\pi_1, \dots, \pi_n)$ such that the detailed balance condition* (45) *is satisfied. Moreover, π can be computed explicitly by choosing an i_0 in each communicating class and defining π_j for i_0 and j belonging in the same class by*

$$\pi_j := \prod_{k=1}^{n-1} \frac{a_{i_k,i_{k+1}}}{a_{i_{k+1},i_k}}$$

depending only on i_0 and j, where $i_1, i_2, \dots, i_n = j$ are such that $a_{i_k,i_{k+1}} > 0$ for $k = 0, \dots, n-1$.

For instance, if $n = 3$, we need to suppose (according to (A2)) that

$$a_{12} a_{23} a_{31} = a_{13} a_{32} a_{21}, \tag{46}$$

and the invariant measure is given by $\pi = c(1, a_{12}/a_{21}, a_{13}/a_{31})$, where $c = (1 + a_{12}/a_{21} + a_{13}/a_{31})^{-1}$.

The following result relates the detailed balance condition and the symmetry of the matrix $H(u)A(u)$.

Proposition 19. *The following three properties are equivalent:*

(i) *Graph-theoretical condition: (A1) and (A2) hold.*
(ii) *Detailed balance condition:* $\pi_i a_{ij} = \pi_j a_{ji}$ *for* $i \neq j$.
(iii) *Symmetry: The matrix* $H(u)A(u)$ *is symmetric.*

Proof. The implication (i) $\Leftarrow$ (ii) is shown in Proposition 18. The converse can be proved directly using the detailed balance condition. Finally, the equivalence (ii) $\Leftrightarrow$ (iii) follows from an explicit calculation of $H(u)A(u)$. ☐

Remark 20. The equivalence of the symmetry of $H(u)A(u)$ and the detailed balance condition is related to the Onsager principle of thermodynamics. Indeed, the diffusion matrix $B = A(u)H(u)^{-1}$ in

$$\partial_t u - \operatorname{div}(B\nabla w) = f(u),$$

where $w = h'(u)$ is the vector of entropy variables, is the Onsager matrix which is symmetric, according to Onsager, if and only if the thermodynamic system is time-reversible. Time-reversibility means that the Markov chain associated to the matrix (a_{ij}) is reversible, and the symmetry of B is equivalent to the symmetry of $H(u)A(u)$. Thus, the equivalence (ii) $\Leftrightarrow$ (iii) corresponds to the equivalence of the symmetry of B and the time-reversibility. For details on the detailed balance principle in thermodynamics, we refer to [7]. ☐

5.2. **Nonlinear Aubin-Lions lemmas.** Let $\Omega \subset \mathbb{R}^d$ ($d \geq 1$) be a bounded domain with Lipschitz boundary. Let $(u^{(\tau)})$ be a family of nonnegative functions which are piecewise constant in time with uniform time step size $\tau > 0$. We introduce the time shift operator $(\sigma_\tau u^{(\tau)})(t) = u^{(\tau)}(t-\tau)$ for $t \geq \tau$.

If there exist uniform estimates for the gradient $(\nabla u^{(\tau)})$ and the discrete time derivative $\tau^{-1}(u^{(\tau)} - \sigma_\tau u^{(\tau)})$, then, by the Aubin-Lions theorem and under suitable conditions on the spaces, $(u^{(\tau)})$ is relatively compact in some L^q space. In the case of nonlinear transition rates, we obtain uniform estimates only for $(\nabla (u^{(\tau)})^s)$, where $s > 0$. Then relative compactness follows from a nonlinear version of the Aubin-Lions theorem [5]. We recall a special case of this result.

Theorem 21 (Nonlinear Aubin-Lions lemma for $s \geq 1/2$)**.** *Let* $s \geq 1/2$, $m \geq 0$, $1 \leq q < \infty$, *and there exists* $C > 0$ *such that for all* $\tau > 0$,

$$\|(u^{(\tau)})^s\|_{L^2(0,T;W^{1,q}(\Omega))} + \tau^{-1}\|u^{(\tau)} - \sigma_\tau u^{(\tau)}\|_{L^1(\tau,T;H^m(\Omega)')} \leq C.$$

Then there exists a subsequence of $(u^{(\tau)})$, *which is not relabeled, such that, as* $\tau \to 0$,

$$u^{(\tau)} \to u \quad \text{strongly in } L^{2s}(0,T;L^{ps}(\Omega)),$$

where $p \geq \max\{1, 1/s\}$ *is such that the embedding* $W^{1,q}(\Omega) \hookrightarrow L^p(\Omega)$ *is compact.*

Theorem 21 can be extended to the case $s < 1/2$ if $(u^{(\tau)})$ is additionally bounded in $L^\infty(0,T;L^1(\Omega))$ which generally follows from the entropy inequality. This result is new.

Theorem 22 (Nonlinear Aubin-Lions lemma for $s < 1/2$)**.** *Let* $\max\{0, 1/2 - 1/d\} < s < 1/2$, $m \geq 0$, *and there exists* $C > 0$ *such that for all* $\tau > 0$,

$$\|u^{(\tau)}\|_{L^\infty(0,T;L^1(\Omega))} + \|(u^{(\tau)})^s\|_{L^2(0,T;H^1(\Omega))} + \tau^{-1}\|u^{(\tau)} - \sigma_\tau u^{(\tau)}\|_{L^1(\tau,T;H^m(\Omega)')} \leq C.$$

Then there exists a subsequence of $(u^{(\tau)})$*, which is not relabeled, such that, as* $\tau \to 0$*,*

$$u^{(\tau)} \to u \quad \textit{strongly in } L^1(0,T;L^1(\Omega)).$$

Proof. The result follows from Theorem 21 and the Hölder inequality. Indeed, we have

$$\begin{aligned}\|\nabla(u^{(\tau)})^{1/2}\|_{L^2(0,T;L^{1/(1-s)}(\Omega))} &= (2s)^{-1}\|(u^{(\tau)})^{1/2-s}\nabla(u^{(\tau)})^s\|_{L^2(0,T;L^{1/(1-s)}(\Omega))} \\ &\le (2s)^{-1}\|(u^{(\tau)})^{(1-2s)/2}\|_{L^\infty(0,T;L^{2/(1-2s)}(\Omega))}\|\nabla(u^{(\tau)})^s\|_{L^2(0,T;L^2(\Omega))} \\ &= \|u^{(\tau)}\|^{(1-2s)/2}_{L^\infty(0,T;L^1(\Omega))}\|\nabla(u^{(\tau)})^s\|_{L^2(0,T;L^2(\Omega))} \le C.\end{aligned}$$

Therefore, $(u^{(\tau)})^{1/2}$ is uniformly bounded in $L^2(0,T;W^{1,1/(1-s)}(\Omega))$. By Rellich-Kondrachov's theorem, the embedding $W^{1,1/(1-s)}(\Omega) \hookrightarrow L^2(\Omega)$ is compact for $s>0$ if $d\le 2$ and $s > 1/2-1/d$ if $d\ge 3$. Applying Theorem 21 with $s=1/2$, $q=1/(1-s)$, and $p=2$, we infer that $(u^{(\tau)})$ is relatively compact in $L^1(0,T;L^1(\Omega))$. □

5.3. **Increasing entropies.** If detailed balance or a weak cross-diffusion condition hold, we have shown that the entropy is nonincreasing in time along solutions to (1)-(2). In this section, we show that the entropy may be increasing for small times if these conditions do not hold. To simplify the presentation, we restrict ourselves to the case $n=3$ (three species), $s=1$ (linear transition rates), and $\Omega=(0,1)$.

Lemma 23 (Vanishing diffusion coefficients a_{i0})**.** *Let* $a_{13}=a_{32}=a_{21}=1$ *and* $a_{ij}=0$ *else. For any* $\varepsilon>0$*, there exist initial data* u^0 *such that*

$$\frac{d\mathcal{H}}{dt}[u^0] \ge \frac{1}{\varepsilon}.$$

In particular, if $t\mapsto \mathcal{H}[u(t)]$ *is continuous, there exists* $t_0>0$ *such that* $t\mapsto\mathcal{H}[u(t)]$ *is increasing on* $[0,t_0]$*.*

Proof. Observe that (46) is not satisfied, and hence detailed balance does not hold. Furthermore, we have

$$H(u)A(u) = \begin{pmatrix} 1/u_1 & 0 & 0 \\ 0 & 1/u_2 & 0 \\ 0 & 0 & 1/u_3 \end{pmatrix}\begin{pmatrix} u_3 & 0 & u_1 \\ u_2 & u_1 & 0 \\ 0 & u_3 & u_2 \end{pmatrix} = \begin{pmatrix} u_3/u_1 & 0 & 1 \\ 1 & u_1/u_2 & 0 \\ 0 & 1 & u_2/u_3 \end{pmatrix}.$$

Let $0<\varepsilon<0.5$ and define $u^0=(u_1^0,u_2^0,u_3^0)$ by $u_1^0(x)=1$ for $x\in(0,1)$ and

$$u_2^0(x) = \begin{cases} 3 & \text{for } 0<x<0.5, \\ 3-\varepsilon^{-1}(x-0.5) & \text{for } 0.5<x<0.5+\varepsilon, \\ 2 & \text{for } 0.5+\varepsilon<x<1, \end{cases}$$

$$u_3^0(x) = \begin{cases} 9 & \text{for } 0<x<0.5, \\ 9+\varepsilon^{-1}(x-0.5) & \text{for } 0.5<x<0.5+\varepsilon, \\ 10 & \text{for } 0.5+\varepsilon<x<1, \end{cases}$$

Then

$$\int_0^1 (\partial_x u^0)^\top H(u^0)A(u^0)\partial_x u^0 dx = \frac{1}{\varepsilon^2}\int_{0.5}^{0.5+\varepsilon}\left(\frac{1}{u_2^0(x)} - 1 + \frac{u_2^0(x)}{u_3^0(x)}\right)dx$$

$$\leq \frac{1}{\varepsilon}\left(\frac{1}{2} - 1 + \frac{3}{9}\right) = -\frac{1}{6\varepsilon},$$

which implies that $(d\mathcal{H}/dt)[u^0] \geq 1/(6\varepsilon)$. □

One may ask if a similar result as above holds if the diffusion coefficients a_{i0} do not vanish, since they give positive contributions to the entropy production. The next lemma shows that the entropy may be increasing even if $a_{i0} > 0$ is chosen arbitrarily.

Lemma 24 (Positive diffusion coefficients a_{i0})**.** *Let $a_{13} = a_{32} = a_{21} = 1$, $a_{i0} > 0$ for $i = 1, 2, 3$, and $a_{ij} = 0$ else. For any $\varepsilon > 0$, there exist initial data u^0 such that*

$$\frac{d\mathcal{H}}{dt}[u^0] \geq \frac{1}{\varepsilon}.$$

In particular, if $t \mapsto \mathcal{H}[u(t)]$ is continuous, there exists $t_0 > 0$ such that $t \mapsto \mathcal{H}[u(t)]$ is increasing on $[0, t_0]$.

Proof. We choose the initial datum

$$u_1^0(x) = \frac{a_{20}(2a_{20} + a_{30})}{8a_{20} + 4a_{30}},$$
$$u_2^0(x) = \begin{cases} 4a_{20} & \text{for } 0 < x < 0.5, \\ a_{20}(4 - \varepsilon^{-1}(x - 0.5)) & \text{for } 0.5 < x < 0.5 + \varepsilon, \\ 3a_{20} & \text{for } 0.5 + \varepsilon < x < 1, \end{cases}$$
$$u_3^0(x) = \begin{cases} 8a_{20} + 4a_{30} & \text{for } 0 < x < 0.5, \\ a_{20}(8 - \varepsilon^{-1}(x - 0.5)) + 4a_{30} & \text{for } 0.5 < x < 0.5 + \varepsilon, \\ 9a_{20} + 4a_{30} & \text{for } 0.5 + \varepsilon < x < 1, \end{cases}$$

Then

$$\int_0^1 (\partial_x u^0)^\top H(u^0)A(u^0)\partial_x u^0 dx$$
$$= \int_{0.5}^{0.5+\varepsilon} \left(\frac{u_1^0}{u_2^0}(\partial_x u_2^0)^2 + \frac{a_{20}}{u_2^0}(\partial_x u_2^0)^2 + \partial_x u_2^0 \partial_x u_3^0 + \frac{u_2^0 + a_{30}}{u_3^0}(\partial_x u_3^0)^2\right) dx$$
$$\leq \frac{a_{20}^2}{\varepsilon^2} \int_{0.5}^{0.5+\varepsilon} \left(\frac{2a_{20} + a_{30}}{3(8a_{20} + 4a_{30})}\frac{a_{20}}{3a_{20}} - 1 + \frac{4a_{20} + a_{30}}{8a_{20} + 4a_{30}}\right) dx$$
$$\leq \frac{a_{20}^2}{\varepsilon^2}\left(\frac{1}{12} - \frac{1}{3} - 1 + \frac{1}{2}\right) = -\frac{a_{20}^2}{12\varepsilon},$$

which proves the result. □

References

[1] H. Amann. Dynamic theory of quasilinear parabolic systems. III. Global existence. *Math. Z.* 202 (1989), 219-250.

[2] R. Anguelov and H.-M. Tenkam. Lyapunov functional for a class of multi-species models with cross diffusion. *Biomath. Commun.* 1 (2014).

[3] L. Chen and A. Jüngel. Analysis of a multi-dimensional parabolic population model with strong cross-diffusion. *SIAM J. Math. Anal.* 36 (2004), 301-322.
[4] L. Chen and A. Jüngel. Analysis of a parabolic cross-diffusion population model without self-diffusion. *J. Diff. Eqs.* 224 (2006), 39-59.
[5] X. Chen, A. Jüngel, and J.-G. Liu. A note on Aubin-Lions-Dubinskiĭ lemmas. *Acta Appl. Math.* 133 (2014), 33-43.
[6] D. Clark. Short proof of a discrete Gronwall inequality. *Discrete Appl. Math.* 16 (1987), 279-281.
[7] S. de Groot and P. Mazur. *Nonequilibrium Thermodynamics.* North Holland, Ansterdam, 1962.
[8] L. Desvillettes and K. Fellner. Duality and entropy methods for reversible reaction-diffusion equations with degenerate diffusion. *Math. Meth. Appl. Sci.* 38 (2015), 3432-3443.
[9] L. Desvillettes, T. Lepoutre, and A. Moussa. Entropy, duality, and cross diffusion. *SIAM J. Math. Anal.* 46 (2014), 820-853.
[10] L. Desvillettes, T. Lepoutre, A. Moussa, and A. Trescases. On the entropic structure of reaction-cross diffusion systems. *Commun. Partial Diff. Eqs.* 40 (2015), 1705-1747.
[11] P. Deuring. An initial-boundary value problem for a certain density-dependent diffusion system. *Math. Z.* 194 (1987), 375-396.
[12] M. Dreher. Analysis of a population model with strong cross-diffusion in unbounded domains. *Proc. Roy. Soc. Edinburgh Sec. A* 138 (2008), 769-786.
[13] M. Dreher and A. Jüngel. Compact families of piecewise constant functions in $L^p(0,T;B)$. *Nonlin. Anal.* 75 (2012), 3072-3077.
[14] G. Galiano, M. Garzón, and A. Jüngel. Semi-discretization in time and numerical convergence of solutions of a nonlinear cross-diffusion population model. *Numer. Math.* 93 (2003), 655-673.
[15] A. Jüngel. The boundedness-by-entropy method for cross-diffusion systems. *Nonlinearity* 28 (2015), 1963-2001.
[16] A. Jüngel. *Entropy Methods for Diffusive Partial Differential Equations.* Springer Briefs in Mathematics, Springer, 2016.
[17] J. Kim. Smooth solutions to a quasi-linear system of diffusion equations for a certain population model. *Nonlin. Anal.* 8 (1984), 1121-1144.
[18] D. Le. Cross diffusion systems in n spatial dimensional domains. *Indiana Univ. Math. J.* 51 (2002), 625-643.
[19] D. Le. Global existence for a class of strongly coupled parabolic systems. *Annali Matem.* 185 (2006), 133-154.
[20] Y. Lou, W.-M. Ni, and Y. Wu. On the global existence of a cross-diffusion system. *Discrete Contin. Dyn. Sys.* 4 (1998), 193-203.
[21] M. Pierre. Global existence in reaction-diffusion systems with control of mass: a survey. *Milan J. Math.* 78 (2010), 417-455.
[22] T. Roubíček. *Nonlinear Partial Differential Equations with Applications.* Birkhäuser, Basel, 2005.
[23] K. Ryu and I. Ahn. Coexistence states of certain population models with nonlinear diffusions among multi-species. *Dyn. Contin. Discrete Impulsive Sys. A* 12 (2005), 235-246.
[24] N. Shigesada, K. Kawasaki, and E. Teramoto. Spatial segregation of interacting species. *J. Theor. Biol.* 79 (1979), 83-99.
[25] P. Suomela. Invariant measures of time-reversible Markov chains. *J. Appl. Prop.* 16 (1979), 226-229.
[26] R. Temam. *Infinite-dimensional Dynamical Systems in Mechanics and Physics.* 2nd edition. Springer, New York, 1997.
[27] Z. Wen and S. Fu. Global solutions to a class of multi-species reaction-diffusion systems with cross-diffusions arising in population dynamics. *J. Comput. Appl. Math.* 230 (2009), 34-43.
[28] A. Yagi. Global solution to some quasilinear parabolic systems in population dynamics. *Nonlin. Anal.* 21 (1993), 603-630.
[29] K. Yosida. *Functional Analysis.* 4th edition. Springer, Berlin, 1974.

[30] N. Zamponi and A. Jüngel. Analysis of degenerate cross-diffusion population models with volume filling. To appear in *Ann. H. Poincaré, Anal. Non Lin.*, 2016.

School of Sciences, Beijing University of Posts and Telecommunications, Beijing 100876, China

E-mail address: buptxchen@yahoo.com

Institute for Analysis and Scientific Computing, Vienna University of Technology, Wiedner Hauptstrasse 8–10, 1040 Wien, Austria

E-mail address: esther.daus@tuwien.ac.at

Institute for Analysis and Scientific Computing, Vienna University of Technology, Wiedner Hauptstrasse 8–10, 1040 Wien, Austria

E-mail address: juengel@tuwien.ac.at