Available online at www.sciencedirect.com

ScienceDirect

Ann. I. H. Poincaré – AN 34 (2017) 1–29

ANNALES DE L'INSTITUT HENRI POINCARÉ
ANALYSE NON LINÉAIRE

www.elsevier.com/locate/anihpc

# Analysis of degenerate cross-diffusion population models with volume filling

Nicola Zamponi [1], Ansgar Jüngel [*,1]

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

Received 19 February 2015; received in revised form 11 August 2015; accepted 27 August 2015

Available online 11 September 2015

**Abstract**

A class of parabolic cross-diffusion systems modeling the interaction of an arbitrary number of population species is analyzed in a bounded domain with no-flux boundary conditions. The equations are formally derived from a random-walk lattice model in the diffusion limit. Compared to previous results in the literature, the novelty is the combination of general degenerate diffusion and volume-filling effects. Conditions on the nonlinear diffusion coefficients are identified, which yield a formal gradient-flow or entropy structure. This structure allows for the proof of global-in-time existence of bounded weak solutions and the exponential convergence of the solutions to the constant steady state. The existence proof is based on an approximation argument, the entropy inequality, and new nonlinear Aubin–Lions compactness lemmas. The proof of the large-time behavior employs the entropy estimate and convex Sobolev inequalities. Moreover, under simplifying assumptions on the nonlinearities, the uniqueness of weak solutions is shown by using the $H^{-1}$ method, the $E$-monotonicity technique of Gajewski, and the subadditivity of the Fisher information.



*Keywords:* Cross diffusion; Population dynamics; Gradient-flow structure; Entropy variables; Nonlinear Aubin–Lions lemmas

## 1. Introduction

In this paper, we analyze a class of multi-species population cross-diffusion systems with volume-filling effects. Such systems arise in various applications, like spatial segregation of interacting species [30], chemotactic cell migration in tissues [29], and ion transport through membranes [8]. Our model class can be derived from a system of random-walk master equations in the diffusion limit for a large class of transition rates (see Appendix A). The key novelty of our analysis is the identification of a new entropy or formal gradient-flow structure and the treatment of non-standard degeneracies in the diffusion coefficients, which significantly extends previous results in [21].

[*] Corresponding author.
*E-mail addresses:* nicola.zamponi@tuwien.ac.at (N. Zamponi), juengel@tuwien.ac.at (A. Jüngel).

[1] The authors acknowledge partial support from the Austrian Science Fund (FWF), grants P24304, P27352, and W1245. Furthermore, they thank Prof. Xiuqing Chen (Beijing) for an improvement of Lemma 9.

http://dx.doi.org/10.1016/j.anihpc.2015.08.003
0294-1449/

The diffusion systems have the form

$$\partial_t u - \operatorname{div}(A(u)\nabla u) = 0 \quad \text{in } \Omega,\ t > 0, \tag{1}$$

with boundary and initial conditions

$$(A(u)\nabla u)\cdot\nu = 0 \quad \text{on } \partial\Omega,\ t>0, \quad u(0) = u^0 \quad \text{in } \Omega. \tag{2}$$

Here, $\Omega \subset \mathbb{R}^d$ ($d \geq 1$) is a bounded domain, $A(u) = (A_{ij}(u)) \in \mathbb{R}^{n\times n}$ is a diffusion matrix, the function $u = (u_1,\ldots,u_n) : \Omega\times(0,\infty) \to \mathbb{R}^n$ is the vector of the proportions of the subpopulations, and $u_{n+1} = 1 - \sum_{i=1}^n u_i$ is the proportion of unoccupied space. In particular, $0 \leq u_i \leq 1$ for all $i = 1,\ldots,n+1$. The $i$th component of equations (1) and (2) has to be understood, respectively, as

$$\partial_t u_i - \sum_{j=1}^n \operatorname{div}(A_{ij}(u)\nabla u_j) = 0, \quad \sum_{j=1}^n A_{ij}(u)\nabla u_j\cdot\nu = 0.$$

The boundary condition in (2) means that the physical or biological system is isolated; the species cannot move through the boundary. For ease of presentation, we have neglected reaction and drift terms in the equations. We refer to Section 7 for a discussion of more general models.

The diffusion matrix in (1) is given by

$$A_{ij}(u) = \delta_{ij}p_i(u)q_i(u_{n+1}) + u_ip_i(u)q_i'(u_{n+1}) + u_iq_i(u_{n+1})\frac{\partial p_i}{\partial u_j}(u), \quad i,j = 1,\ldots,n, \tag{3}$$

where $\delta_{ij}$ is the Kronecker delta. The nonnegative functions $p_i$ and $q_i$ model the transition rates in the random-walk lattice model. The coefficients $A_{ij}$ are derived from this model in the diffusion limit (see Appendix A). The function $q_i$ vanishes when the cells are fully packed, i.e. if $\sum_{i=1}^n u_i = 1$, so $q_i(0) = 0$ and $q_i$ is nondecreasing. In the literature, several special models were considered and we review now some of them.

**Example 1.** *1. Population-dynamics models.* The case $n = 2$, $p_i(u) = a_{i0} + a_{i1}u_1 + a_{i2}u_2$ and $q_i(u_3) = 1$ for $i = 1,2$ was suggested by Shigesada, Kawasaki, and Teramoto [30] to describe the spatial segregation of interacting populations and to study the coexistence of two similar species. This model has attracted a lot of attention in the literature. One of the first existence results is due to Kim [22] who imposed some restrictions of the parameters $a_{ij}$. The tridiagonal case $a_{21} = 0$ was investigated, e.g., by Amann [1] and Le [23]. The first global existence result without any restriction on the diffusion coefficients (except positivity) was achieved in [19] in one space dimension and in [9,10] in several space dimensions. The case of concave functions $p_1$ and $p_2$ was analyzed by Desvillettes et al. [14], recently improved in [15]. The $n$-species case with superlinear functions $p_i(u)$ was investigated in [21]; also see [5] for a so-called relaxed system.

*2. Ion-transport models.* The case $p_i(u) = 1$ for $i = 1,\ldots,n$ and $q(u_{n+1}) = u_{n+1}$ was employed to describe the motility of biological cells [32] or the ion transport through nanopores [8]. The global existence of bounded weak solutions was proved in [7]. This result was generalized in [21] to a class of nondecreasing functions including all power functions $q(s) = s^\alpha$ with $\alpha \geq 1$. The models in [8,32] also include a drift term to account for electric effects, and we discuss these extensions in Section 7.

*3. Multi-species chemotaxis models.* A special case of the model in [29] is given by $p_i(u) = 1$ and $q(u_{n+1}) = u_{n+1}$, similar to the ion-transport model. In fact, the system in [29] contains additional terms which cannot be described by (3) since the transition rates assumed in [29] are not of the type $p_i(u)q_i(u_{n+1})$ (see (65) in Appendix A) but they equal $p_i(u) + q_i(u_{n+1})$. We refer to the discussion in Section 7. □

In the model classes (i) and (ii), either $p_i \equiv 1$ or $q_i \equiv 1$. In contrast, we investigate here a more general model class allowing for nonconstant functions $p_i$ *and* $q_i$. A guiding example is system (1) with diffusion coefficients (3) and $p_i(u) = u_1 + u_2$, $q_i(s) = s$ for $i = 1,2$, which models volume-filling effects in population systems. The diffusion matrix reads explicitly as

$$A(u) = \begin{pmatrix} u_1(1-u_1-u_2) + (u_1+u_2)(1-u_2) & u_1 \\ u_2 & u_2(1-u_1-u_2) + (u_1+u_2)(1-u_1) \end{pmatrix}. \tag{4}$$

We will show in Theorem 1 that (1) with this diffusion matrix possesses a global weak solution satisfying $0 \le u(t) \le 1$ for all $t > 0$. In fact, Theorem 1 is concerned with much more general models.

Let us mention some related results for cross-diffusion systems which became recently very popular in the mathematical literature. The variational structure of special classes of cross-diffusion systems, including geodesic convexity properties, was investigated in [33]; also see [24, Section 4.7]. Cross-diffusion systems like (1) with (3) and $q_i \equiv 1$ for $i = 1, \dots, n$ can be approximated by reaction–diffusion systems [26]. The nice feature is that the diffusion matrix of the reaction–diffusion system is diagonal; however, the number of variables doubles. This idea was exploited for the design of numerical schemes in [27].

The analysis of system (1) with diffusion matrix (3) faces a number of mathematical challenges. First, the equations are *strongly coupled* such that standard tools, like maximum principles and regularity theory, generally do not apply. Second, the diffusion matrix is generally *not positive definite* and thus, even the local-in-time existence of solutions is nontrivial. Third, since the variables $u_i$ are proportions, we need to prove *lower and upper bounds* for the solutions (here, $u_i \ge 0$ and $\sum_{i=1}^n u_i \le 1$), but maximum principle or invariant region methods seemingly do not apply. Fourth, the parabolic system may be *degenerate* (e.g. like in (4) for $u = (0, 1)$ or $u = (1, 0)$).

Some of these difficulties have been dealt with in, e.g., [21] under the assumption that the diffusion system has a formal entropy or gradient-flow structure, i.e., there exists a convex functional $h : \mathscr{D} \to \Omega$ (called entropy density), where $\mathscr{D} \subset \mathbb{R}^n$, such that the matrix $B = A(u)h''(u)^{-1}$ is positive semi-definite and (1) can be written as

$$\partial_t u - \operatorname{div}(B \nabla h'(u)) = 0, \tag{5}$$

where $h'(u)$ and $h''(u)$ are the Jacobian and Hessian of $h$, respectively. This formulation has two advantages: First, $H[u] = \int_\Omega h(u)dx$ is a Lyapunov functional along solutions $u(t)$ to (1)–(2),

$$\frac{dH}{dt}[u(t)] = \int_\Omega h'(u) \cdot \partial_t u dx = -\int_\Omega \nabla u : h''(u)A(u)\nabla u = -\int_\Omega \nabla w : B \nabla w dx \le 0, \tag{6}$$

where $w = h'(u)$ are called entropy variables. In particular, this yields a gradient-type estimate for $w$ or $u$. Second, if $h'$ is invertible on $\mathscr{D}$ (see Lemma 6), the original variable $u = (h')^{-1}(w)$ is an element of $\mathscr{D}$. Thus, if $\mathscr{D}$ is a bounded domain, we obtain lower and upper bounds for $u$ without the use of a maximum principle. In our situation, we define $\mathscr{D} = \{u \in \mathbb{R}^n : u_i > 0 \text{ for } i = 1, \dots, n, \sum_{j=1}^n u_j < 1\}$ such that $u_i$ is positive and bounded by one.

There remain still two issues for systems with diffusion coefficients (3). The first one is to identify a suitable entropy density $h$, the second one is the possible degeneracy. In the example given by (4), we choose

$$\begin{aligned} h(u) = &\sum_{i=1}^2 u_i(\log u_i - 1) + (1 - u_1 - u_2)(\log(1 - u_1 - u_2) - 1) \\ &+ (u_1 + u_2)(\log(u_1 + u_2) - 1) + 4, \end{aligned}$$

which yields the matrix

$$B = A(u)h''(u)^{-1} = \begin{pmatrix} u_1(u_1 + u_2)(1 - u_1 - u_2) & 0 \\ 0 & u_2(u_1 + u_2)(1 - u_1 - u_2) \end{pmatrix}.$$

At least one eigenvalue of $B$ vanishes if $u \in \partial\mathscr{D} = \{u_1 = 0,\ u_2 = 0,\ 1 - u_1 - u_2 = 0\}$. In this sense, system (1) is called to be of degenerate type. Generally, systems (1) are always of degenerate type since $q(0) = 0$. Here, we develop a technique to deal with such a degeneracy.

We overcome these issues by developing two main ideas. Our first key idea is the identification of a class of functions $p_i$ and $q_i$ for which we are able to define a novel entropy density. The second idea is the extension of the Aubin–Lions compactness lemma to non-standard degenerate cases. In the following, we detail these concepts.

We make the following structural hypotheses on the functions $p_i$ and $q_i$: There exist functions $q : [0, 1] \to \mathbb{R}$, $\chi : \overline{\mathscr{D}} \to \mathbb{R}$ and a number $\gamma > 0$ such that for all $i = 1, \dots, n$,

$$q(s) := q_i(s) > 0, \quad q'(s) \ge \gamma q(s) \text{ for } s \in (0, 1), \quad q(0) = 0, \quad q \in C^3([0, 1]), \tag{7}$$

$$p_i(u) = \exp\left(\frac{\partial \chi(u)}{\partial u_i}\right) \text{ for } u \in \mathscr{D}, \quad \chi \ge 0 \text{ is convex on } \overline{\mathscr{D}}, \quad \chi \in C^3(\overline{\mathscr{D}}). \tag{8}$$

Examples of functions $q$ and $p_i$ satisfying these conditions are given in Remark 2. We define the entropy density

$$h(u) = \sum_{i=1}^{n}(u_i \log u_i - u_i + 1) + \int_a^{u_{n+1}} \log q(s)ds + \chi(u), \quad u \in \mathscr{D}, \tag{9}$$

where $a \in (0, 1]$ is such that $\int_a^b \log q(s)ds \geq 0$ for all $b \in (0, 1)$, namely

$$a = \begin{cases} 1 & \text{if } q(1) \leq 1, \\ q^{-1}(1) & \text{if } q(1) > 1. \end{cases} \tag{10}$$

Notice that we require that all functions $q_i$ are the same and that $p_i$ possesses a particular structure. It seems to be difficult to treat more general cases, except imposing other conditions.

Surprisingly, system (1) with (3) partially decouples in the entropy variables. Indeed, we may write the formal gradient-flow formulation

$$\partial_t u_i - \operatorname{div}\left(q(u_{n+1})^2 \exp\frac{\partial h}{\partial u_i}(u)\nabla\frac{\partial h}{\partial u_i}\right) = 0, \quad i = 1, \ldots, n,$$

which makes the degenerate structure more apparent than (1). In particular, the transformed diffusion matrix $B$ in (5) is diagonal with elements $q(u_{n+1})^2 \exp(\partial h/\partial u_i)$, $i = 1, \ldots, n$. We also note that if $q \equiv 1$, we obtain $\partial_t u_i = \Delta(\exp(\partial h/\partial u_i)) = \Delta(u_i p_i(u))$. This structure was exploited in [14,15].

A computation, which is made rigorous below, shows that the following entropy inequality holds:

$$\frac{d}{dt}\int_\Omega h(u)dx + c\int_\Omega \left(q(u_{n+1})^2 \sum_{i=1}^{n} |\nabla u_i^{1/2}|^2 + |\nabla q(u_{n+1})^{1/2}|^2\right) dx \leq 0, \tag{11}$$

where $c > 0$ is some constant. We wish to deduce $L^2$ gradient estimates for $u_1, \ldots, u_n$, which are needed to apply the Aubin–Lions compactness lemma for a suitable approximated system. However, because of the degeneracy of $q$ (i.e. $q(0) = 0$), these estimates are nontrivial. We overcome this problem by proving two compactness results.

The first compactness result essentially states that if we have (i) uniform gradient estimates for the bounded sequences $(\xi_\varepsilon)$ and $(\xi_\varepsilon\eta_\varepsilon)$, (ii) a uniform estimate for the (discrete) time derivative of $\eta_\varepsilon$, and (iii) the strong convergence $\xi_\varepsilon \to \xi$ in $L^2$, then up to a subsequence, $\xi_\varepsilon f(\eta_\varepsilon) \to \xi f(\eta)$ in $L^2$ for any continuous function $f$ (Lemma 8). If $\xi_\varepsilon$ were strictly positive, the statement would be a consequence of the usual Aubin–Lions lemma [31]. Here, we are able to deal with functions $\xi_\varepsilon$ which may vanish locally. The case $f(s) = s$ was considered in [7,21].

The second compactness result is a generalization of the Aubin–Lions–Dubinskiĭ lemma; see, e.g., [11,25]. It states that if a bounded sequence $(u_\varepsilon)$ possesses a uniform estimate for the (discrete) time derivative and a uniform gradient estimate for $Q(u_\varepsilon)$ and $Q'(u_\varepsilon)$ for some nonnegative convex increasing function $Q$, then up to a subsequence, $u_\varepsilon \to u$ strongly in $L^2$ (Lemma 9). This result is complementary to the nonlinear Aubin–Lions lemma stated in [25] and generalizes the lemma in [11] stated for $Q(s) = s^\alpha$ with $\alpha > 1$.

Based on the above ideas, we prove three results. First, we show the global-in-time existence of bounded weak solutions to (1)–(3) satisfying the entropy inequality (11) (Theorem 1). Second, the entropy inequality and a convex Sobolev inequality allow us to show that $u_{n+1}(t)$ converges to the constant steady state in the $L^2$ sense. Moreover, if $q$ is strictly positive, this convergence also holds for $u_1(t), \ldots, u_n(t)$ (Theorem 4). Third, if $p_i \equiv 1$ for all $i = 1, \ldots, n$, there is a unique weak solution to (1)–(3). The proof combines the $H^{-1}$ method and the $E$-monotonicity technique of Gajewski [17].

The paper is organized as follows. The main results are stated and commented in Section 2. Section 3 is devoted to the proof of some auxiliary results, like the positive semi-definiteness of the matrix $h''(u)A(u)$ and the Aubin–Lions compactness lemmas. The three main theorems are proved in Sections 4, 5, and 6, respectively. Extensions of our model are discussed in Section 7. Appendix A is concerned with the formal derivation of (1) from a random-walk lattice model.

## 2. Main results

We state our main theorems and detail the ideas of the proofs. The first theorem is concerned with the global existence of bounded weak solutions. Recall that

$$\mathscr{D} = \Big\{u = (u_1, \dots, u_n) \in \mathbb{R}^n : u_i > 0 \text{ for } i = 1, \dots, n,\ \sum_{j=1}^n u_j < 1\Big\}. \tag{12}$$

**Theorem 1** *(Global existence). Let $T > 0$, let $u^0 : \Omega \to \mathscr{D}$ be a measurable function such that $h(u^0) \in L^1(\Omega)$, and let $A(u)$ be given by (3). Assume that hypotheses (7) and (8) hold. Then:*

(i) *There exists a weak solution $u : \Omega \times (0,T) \to \overline{\mathscr{D}}$ to (1)–(2) satisfying $u_i \geq 0$, $u_{n+1} := 1 - \sum_{i=1}^n u_i \geq 0$, and*

$$q(u_{n+1})^{1/2},\ u_i^{1/2} q(u_{n+1})^{1/2},\ u_i p_i(u) q(u_{n+1})^{1/2} \in L^2(0,T; H^1(\Omega)), \tag{13}$$

$$u_i \in L^\infty(0,T; L^\infty(\Omega)), \quad \partial_t u_i \in L^2(0,T; H^1(\Omega)'), \quad i = 1, \dots, n. \tag{14}$$

*The function $u$ satisfies the weak formulation*

$$\sum_{i=1}^n \int_0^T \langle \partial_t u_i, \phi_i \rangle dt + \sum_{i=1}^n \int_0^T \int_\Omega \Big[ q(u_{n+1})^{1/2} \nabla\big(u_i p_i(u) q(u_{n+1})^{1/2}\big) - 3 u_i p_i(u) q(u_{n+1})^{1/2} \nabla q(u_{n+1})^{1/2} \Big] \cdot \nabla \phi_i \, dx dt = 0 \tag{15}$$

*for all $\phi_1, \dots, \phi_n \in L^2(0,T; H^1(\Omega))$, and $u(0) = u^0$ in the sense of $H^1(\Omega)'$. Here, $\langle \cdot, \cdot \rangle$ denotes the duality product of $H^1(\Omega)'$ and $H^1(\Omega)$.*

(ii) *The following entropy inequality holds:*

$$\int_\Omega h(u(t)) dx + c_0 \int_0^t \int_\Omega \left( \sum_{i=1}^n q(u_{n+1})^2 |\nabla u_i^{1/2}|^2 + |\nabla q(u_{n+1})^{1/2}|^2 \right) dx dt \leq \int_\Omega h(u^0) dx, \tag{16}$$

*where $c_0 = 4 p_0 \min\{1, \delta\} > 0$ with $p_0$ and $\delta$ being defined in (23) below.*

(iii) *If $\int_0^b |\log q(s)| ds = +\infty$ for all $0 < b < 1$ then $u_{n+1} > 0$ a.e. in $\Omega \times (0,T)$.*

**Remark 2.** We present examples of functions $q$ and $p_i$ satisfying (7) and (8), respectively. Hypothesis (7) is satisfied by $q(s) = s^\alpha$ for $s \in [0,1]$, where $\alpha \geq 1$. Indeed, the inequality $q'(s) \geq \gamma q(s)$ holds for all $s \in [0,1]$ with $\gamma := \alpha$. Another example class is given by $q(s) = \exp(f(s)) - 1$ with $f(0) = 0$ and $f'(s) \geq \gamma > 0$ for $s \in [0,1]$. A concrete example is $q(s) = \exp(s^\alpha) - 1$ with $0 < \alpha \leq 1$. A third example is $q(s) = \exp(-s^{-\alpha})$ with $\alpha > 0$ which satisfies the assumption stated in Theorem 1, part (iii).

Hypothesis (8) is satisfied by every function $p_i(u) = \widetilde{p}_i(u_i)$, where $\widetilde{p}_i \in C^1([0,1])$ is strictly positive and nondecreasing. Indeed, let us define

$$\chi_i(s) = \int_0^s \log \widetilde{p}_i(\sigma) d\sigma + k, \quad \chi(u) = \sum_{j=1}^n \chi_j(u_j)$$

for $s \in [0,1]$, $i = 1, \dots, n$, and $u = (u_1, \dots, u_n) \in \mathscr{D}$. Here, $k > 0$ is such that $\chi_i \geq 0$ in $[0,1]$. Since $\widetilde{p}_i$ is strictly positive and nondecreasing in $[0,1]$, it follows that $\chi''(u)$, given by

$$\frac{\partial^2 \chi}{\partial u_i \partial u_j}(u) = \delta_{ij} \frac{\widetilde{p}_i'(u_i)}{\widetilde{p}_i(u_i)}, \quad i, j = 1, \dots, n,$$

is positive semi-definite and $\chi : \overline{\mathscr{D}} \to [0, \infty)$ is convex. Furthermore, $\exp(\partial \chi / \partial u_i) = \widetilde{p}_i(u_i) = p_i(u)$ for $u \in \mathscr{D}$.

Another example is given by $p_i(u) = (\sum_{j=1}^n a_j u_j)^{a_i}$ with $a_i \geq 0$, $i = 1, \ldots, n$. Indeed, the function $\chi(u) = \sum_{j=1}^n a_j u_j (\log(\sum_{j=1}^n a_j u_j) - 1)$ is convex on $\overline{\mathscr{D}}$ and satisfies $\exp(\partial\chi/\partial u_i) = \exp(a_i \log(\sum_{j=1}^n a_j u_j)) = p_i(u)$. This example corresponds to the diffusion matrix (4) for $n = 2$ and $a_1 = a_2 = 1$. □

**Remark 3.** Let us give some concrete examples of diffusion matrices which satisfy hypotheses (7) and (8). In order to simplify the notation, we fix $n = 3$ but this is not essential. We choose $p_i \equiv 1$ and $q_i(s) = s^\alpha$ with $\alpha \geq 1$ for the first example:

$$A = u_4^{\alpha-1} \begin{pmatrix} u_4 + \alpha u_1 & \alpha u_1 & \alpha u_1 \\ \alpha u_2 & u_4 + \alpha u_2 & \alpha u_2 \\ \alpha u_3 & \alpha u_3 & u_4 + \alpha u_3 \end{pmatrix}.$$

The case $\alpha = 1$ was analyzed in [7], the case $\alpha > 1$ in [21]. The choice $\alpha = 1$ corresponds to the ion-transport model described in Example 1, no. 2 and no. 3. We may also choose $q_i(s) = \exp(s^\alpha) - 1$ with $0 < \alpha \leq 1$, which is new. Next, let $p_i \not\equiv 1$. We cannot choose $q_i \equiv 1$ in this situation which corresponds to the population-dynamics models described in Example 1, no. 1 but such models are analyzed in [9] for $n = 2$. Here, we may choose $p_i(u) = a_i u_i$ for $a_i > 0$ and $q_i(s) = s^\alpha$ for $\alpha \geq 1$ which gives the diffusion matrix

$$A = u_4^{\alpha-1} \begin{pmatrix} a_1u_1(2u_4 + \alpha u_1) & a_1u_1(u_4 + \alpha u_1) & a_1u_1(u_4 + \alpha u_1) \\ a_2u_2(u_4 + \alpha u_2) & a_2u_2(2u_4 + \alpha u_2) & a_2u_2(u_4 + \alpha u_2) \\ a_3u_3(u_4 + \alpha u_3) & a_3u_3(u_4 + \alpha u_3) & a_3u_3(2u_4 + \alpha u_3) \end{pmatrix}.$$

A final example (with $n = 2$) is given by $p_i(u) = u_1 + u_2$ and $q_i(s) = s$ for $i = 1, 2$; see (4). For all these examples, the existence result in Theorem 1 applies. For the examples with $p_i \equiv 1$, the weak solution is unique; see Theorem 5 below. □

The proof of Theorem 1 is based on an approximation and regularization of (1). More precisely, we consider the semi-discrete system

$$\frac{1}{\tau}(u(w^k) - u(w^{k-1})) = \operatorname{div}(B(w^k)\nabla w^k) + \tau^2(\Delta w^k + w^k) + \varepsilon\tau^2 \sum_{2\leq|\alpha|\leq m} (-1)^{|\alpha|-1} D^{2\alpha} w^k$$

with homogeneous Neumann boundary conditions, where $\tau > 0$, $\varepsilon > 0$, $m > d/2$, $u(w) = (h')^{-1}(w)$, $w^k$ approximates $w(k\tau)$, and $D^{2\alpha}$ is a partial derivative of order $2|\alpha|$, with $\alpha \in \mathbb{N}_0^d$ being a multiindex. Compared to [21], we need *two* regularization levels: the $H^1$ regularization given by $\Delta w^k + w^k$ and the $H^m$ regularization given by the sum over $\alpha$. The second regularization is needed to obtain approximate $L^\infty$ solutions (observe that $H^m(\Omega) \hookrightarrow L^\infty(\Omega)$), while the first one allows us to interpret the weak formulation in the larger space $H^{-1}$ instead of $H^{-m}$. This is needed to apply the generalized Aubin–Lions Lemmas 8 and 9, for which $H^{-1}$ is required.

The entropy inequality (11), adapted to the above problem, yields uniform $H^m$ estimates. Hence, applying the Leray–Schauder fixed-point theorem, we obtain the existence of semi-discrete $H^m$ solutions. The same entropy inequality provides a priori estimates uniform in $\tau$ and $\varepsilon$. First, we perform the limit $\varepsilon \to 0$, then the limit $\tau \to 0$. The latter limit is highly nontrivial since we have only an $L^2$ bound for $q(u_{n+1})\nabla u_i^{1/2}$, and $q(u_{n+1}) = 0$ at $u_{n+1} = 0$ is possible. This degeneracy will be overcome by the compactness result in Lemma 8.

The second result is about the large-time behavior of the solutions to the constant steady state given by

$$u_i^\infty = \frac{1}{|\Omega|}\int_\Omega u_i^0 dx, \; i = 1, \ldots, n, \quad u_{n+1}^\infty = 1 - \sum_{i=1}^n u_i^\infty.$$

We are able to prove exponential convergence of $u_{n+1}(t)$ and, under an additional assumption on $q$, also of $u_1(t), \ldots, u_n(t)$.

**Theorem 4** *(Convergence to steady state). Let $\Omega$ be convex, $u^0 \in L^1(\Omega; \mathscr{D})$, let $A(u)$ be given by (3), and assume that (7) and (8) hold. Furthermore, let $q \in C^3([0, 1])$ be such that $q'$ is strictly positive and $q/q'$ is concave on $(0, 1)$. Let $u : \Omega \times (0, T) \to \mathscr{D}$ be a weak solution to (1)–(2) in the sense of Theorem 1. Then*

$$\|u_{n+1}(t) - u_{n+1}^\infty\|_{L^2(\Omega)} \le C_1 e^{-\lambda_1 t}, \quad t \ge 0, \tag{17}$$

*where* $C_1 = (2/\gamma)^{1/2}\|h^*(u^0|u^\infty)\|_{L^1(\Omega)}^{1/2}$ *and* $\lambda_1 = c_0 q_1/(4c_S)$, $h^*$ *is the relative entropy density (see (19)),* $q_1 := \min_{s\in[0,1]} q'(s) > 0$, $c_0 > 0$ *is defined in Theorem 1, and* $c_S > 0$ *is the constant of the convex Sobolev inequality in Lemma 11. Moreover, if* $q_0 := \min_{s\in[0,1]} q(s) > 0$,

$$\|u_i(t) - u_i^\infty\|_{L^2(\Omega)} \le C_1 e^{-\lambda_2 t}, \quad t \ge 0, \quad i = 1, \dots, n, \tag{18}$$

*where* $\lambda_2 = c_0 q_0/c_L$ *and* $c_L > 0$ *is the constant in the logarithmic Sobolev inequality (see, e.g., [13, Lemma 1]).*

The convexity of $\Omega$ and the concavity of $q/q'$ is needed to apply the convex Sobolev inequality (see Lemma 11 below). For instance, $q/q'$ is concave for $q(s) = s^\alpha$ with $\alpha > 0$. The condition on the strict positivity of $q$ contradicts the assumption $q(0) = 0$ in hypothesis (7). However, Theorem 1 is also valid for functions $q(0) > 0$. In fact, the existence analysis is much easier in this case since the problem becomes nondegenerate.

The idea of the proof is to derive an inequality for the relative entropy

$$\int_\Omega h^*(u|u^\infty)dx = \int_\Omega \big(h(u) - h(u^\infty) - h'(u^\infty)\cdot(u - u^\infty)\big)dx. \tag{19}$$

A computation, which is made rigorous in Section 5, shows that

$$\frac{d}{dt}\int_\Omega \int_{u_{n+1}^0}^{u_{n+1}(t)} \log q(s)ds dx + c\int_\Omega |\nabla q(u_{n+1})^{1/2}|^2 dx \le 0$$

for some $c > 0$. The entropy dissipation can be bounded from below (up to a factor) by the relative entropy by means of the convex Sobolev inequality [2]. Together with the Gronwall lemma and the convexity of the relative entropy, this yields exponential convergence of $u_{n+1}(t)$ to $u_{n+1}^\infty$ in the $L^2$ norm. In a similar way, we obtain the entropy inequality

$$\frac{d}{dt}\int_\Omega \sum_{i=1}^n u_i(t)\log\frac{u_i(t)}{u_i^\infty}dx + c\int_\Omega q(u_{n+1})^2|\nabla u_i^{1/2}|^2 dx \le 0.$$

Here, the degeneracy of $q$ at $u_{n+1} = 0$ prevents the application of the logarithmic Sobolev inequality. For this reason, we assume that $q$ is strictly positive. Then, by Gronwall's lemma again, we deduce the exponential convergence of $u_i(t)$ to $u_i^\infty$ in the $L^2$ norm.

The idea of using the entropy functional and logarithmic Sobolev inequalities to prove the exponential decay of solutions was already employed for reaction–diffusion systems in [20] in a non-constructive way and more directly in, e.g., [12,13]. In fact, this idea goes back to Bakry and Emery [3], but they focused more on the derivation of convex Sobolev inequalities.

Our last theorem is a uniqueness result in the special case $p_i \equiv 1$. This includes the ion-transport model [8].

**Theorem 5** *(Uniqueness of solutions). Let the assumptions of Theorem 1 hold and let* $p_i \equiv 1$ *for* $i = 1, \dots, n$. *Then there exists a unique weak solution to (1)–(2) satisfying (13)–(14).*

The idea of the proof is to combine the $H^{-1}$ method and the $E$-monotonicity technique of Gajewski [17]. In fact, we exploit the special structure of (1) and (3) in the case $p_i \equiv 1$:

$$\partial_t u_i = \operatorname{div}\big(q(u_{n+1})\nabla u_i - u_i\nabla q(u_{n+1})\big), \quad i = 1, \dots, n.$$

Summing all these equations, we end up with a simple equation for $u_{n+1}$:

$$\partial_t u_{n+1} = \Delta Q(u_{n+1}), \quad Q'(s) = q(s) + (1-s)q'(s).$$

The uniqueness for $u_{n+1}$ is shown by the usual $H^{-1}$ method. The uniqueness for the remaining components $u_i$ is more difficult since we cannot easily treat the drift term. This is in contrast to the drift-diffusion equations for semiconductors, where a monotonicity property of the drift term can be exploited. Here, we employ the $E$-monotonicity method [17]. This method is based on the convexity of the logarithmic entropy. More precisely, define the distance

$$d(u,v)=\sum_{i=1}^n\int_\Omega\left(\xi(u_i)+\xi(v_i)-2\xi\left(\frac{u_i+v_i}{2}\right)\right)dx,$$
$$\xi(s)=s(\log s-1)+1,\quad s\geq 0.$$

A formal computation, which is made rigorous in Section 6, using the subadditivity of the Fisher information (see Lemma 10), shows that

$$\frac{d}{dt}d(u,v)\leq 0,\quad t>0,$$

and consequently, $d(u(t),v(t))\leq d(u(0),v(0))=0$ for $t>0$. Since $\xi$ is convex, we infer that $d(u(t),v(t))\geq 0$, which finally yields $u_i=v_i$ for $i=1,\dots,n$.

## 3. Auxiliary results

### *3.1. Invertibility of the entropy transformation*

We show that the transformation of variables $w=h'(u)$ can be inverted. Recall that the set $\mathscr{D}$ is defined in (12).

**Lemma 6.** *Let assumptions (7)–(8) hold. Then the function $h:\mathscr{D}\to\mathbb{R}$, defined in (9), is strictly convex, nonnegative, belongs to $C^2(\mathscr{D})$, and its gradient $h':\mathscr{D}\to\mathbb{R}^n$ is invertible. Moreover, the inverse of the Hessian $h'':\mathscr{D}\to\mathbb{R}^n$ is uniformly bounded.*

**Proof.** We first show that $h':\mathscr{D}\to\mathbb{R}^n$ is invertible. For this, we observe that

$$\frac{\partial h}{\partial u_i}=\log u_i-\log q\Big(1-\sum_{j=1}^n u_j\Big)+\frac{\partial\chi}{\partial u_i},\quad i=1,\dots,n.$$

The Jacobian of the function $g=(g_1,\dots,g_n):\mathscr{D}\to\mathbb{R}^n$, defined by $g_i(u)=\log u_i-\log q(1-\sum_{j=1}^n u_j)$, is positive definite since

$$\frac{\partial g_i}{\partial u_j}=\frac{\delta_{ij}}{u_i}+\frac{q'(u_{n+1})}{q(u_{n+1})}.$$

It is shown in Step 1 of the proof of Theorem 6 in [21] that $g:\mathscr{D}\to\mathbb{R}^n$ is invertible. Thus, we can define the function $f=h'\circ g^{-1}:\mathbb{R}^n\to\mathbb{R}^n$. Since $h''(u)$ and $g'(u)$ are nonsingular matrices for $u\in\mathscr{D}$, the Jacobian of $f$,

$$f'(y)=h''(g^{-1}(y))(g')^{-1}(g^{-1}(y)),$$

is nonsingular for $y\in\mathbb{R}^n$. Moreover, by the definitions of $f$ and $g$, we have

$$f(y)=y+\chi'(g^{-1}(y)),\quad y\in\mathbb{R}^n. \tag{20}$$

Hypothesis (8) states that $\chi'\in C^0(\overline{\mathscr{D}})\subset L^\infty(\mathscr{D})$, thus (20) implies that $|f(y)|\to\infty$ as $|y|\to\infty$. This property as well as the invertibility of the matrix $f'(u)$ allow us to apply Hadamard's global inverse theorem, showing that $f:\mathbb{R}^n\to\mathbb{R}^n$ is invertible. Consequently, also $h'=f\circ g:\mathscr{D}\to\mathbb{R}^n$ is invertible.

It remains to prove that the inverse of the Hessian of $h$ is bounded. Since $q'/q\geq 0$, $0<u_i<1$, and $\chi$ is convex in $\mathscr{D}$, the expression

$$\frac{\partial^2 h}{\partial u_i\partial u_j}=\frac{\delta_{ij}}{u_i}+\frac{q'(u_{n+1})}{q(u_{n+1})}+\frac{\partial^2\chi}{\partial u_i\partial u_j},\quad u\in\mathscr{D}, \tag{21}$$

shows that $v^\top h''(u)v\geq|v|^2$ for all $u\in\mathscr{D}$, $v\in\mathbb{R}^n$. We infer that all points in the spectrum of $h''$ are strictly positive in $\mathscr{D}$. In particular, $h$ is strictly convex. As $h''$ is symmetric, we conclude that the inverse of $h''$ is bounded in $\mathscr{D}$. □

*3.2. Positive definiteness of $HA$*

We show that the product $HA$ of the Hessian $H := h''(u)$ and the diffusion matrix $A = A(u)$ is positive definite. This result is needed to deduce gradient estimates for $u$; see (6).

**Lemma 7.** *Let assumptions (7)–(8) hold. Then the matrix $HA$ is symmetric and positive definite. More precisely, for all $u \in \mathscr{D}$ and $v \in \mathbb{R}^n$, we have*

$$v^\top (HA)v \geq p_0 q(u_{n+1}) \sum_{i=1}^n \frac{v_i^2}{u_i} + p_0\delta \frac{q'(u_{u_n+1})^2}{q(u_{n+1})} \left(\sum_{i=1}^n v_i\right)^2, \tag{22}$$

*where*

$$p_0 = \min_{1\leq i\leq n} \inf_{u\in\mathscr{D}} p_i(u) > 0, \quad \delta = \min\left\{\frac{1}{2}, \frac{2q(1/2)}{\sup_{1/2<s<1} q'(s)}\right\} > 0. \tag{23}$$

**Proof.** First, we verify the symmetry of $HA$. Using (21) and the definition of $A$, we find that

$$\begin{aligned}(HA)_{ij} &= \sum_{k=1}^n \left(\frac{\delta_{ik}}{u_i} + \frac{\partial^2\chi}{\partial u_i \partial u_k} + \frac{q'}{q}\right)\left(\delta_{kj} p_k q + u_k p_k q' + u_k q \frac{\partial p_k}{\partial u_j}\right) \\ &= \delta_{ij}\frac{p_i q}{u_i} + p_i q' + \frac{\partial p_i}{\partial u_j} q + \frac{\partial^2\chi}{\partial u_i \partial u_j} p_j q + \sum_{k=1}^n \frac{\partial^2\chi}{\partial u_i\partial u_k} u_k p_k q' \\ &\quad + \sum_{k=1}^n \frac{\partial^2\chi}{\partial u_i\partial u_k}\frac{\partial p_k}{\partial u_j} u_k q + p_j q' + \frac{(q')^2}{q}\sum_{k=1}^n p_k u_k + q'\sum_k u_k \frac{\partial p_k}{\partial u_j}.\end{aligned}$$

Dividing this equation by $q$, defining $\varphi = q'/q$, and taking into account that, by assumption (8),

$$\frac{\partial^2\chi}{\partial u_i\partial u_j} = \frac{1}{p_j}\frac{\partial p_j}{\partial u_i} = \frac{1}{p_i}\frac{\partial p_i}{\partial u_j} \quad \text{for } i, j = 1, \ldots, n,$$

we infer that

$$\begin{aligned}\frac{1}{q}(HA)_{ij} &= \delta_{ij}\frac{p_i}{u_i} + p_i\varphi + \frac{\partial p_i}{\partial u_j} + \frac{\partial p_j}{\partial u_i} + \sum_{k=1}^n \frac{\partial p_k}{\partial u_i} u_k\varphi \\ &\quad + \sum_{k=1}^n \frac{\partial p_k}{\partial u_i}\frac{\partial p_k}{\partial u_j}\frac{u_k}{p_k} + p_j\varphi + \varphi^2\sum_{k=1}^n p_k u_k + \varphi\sum_{k=1}^n u_k\frac{\partial p_k}{\partial u_j} \\ &= \delta_{ij}\frac{p_i}{u_i} + \frac{\partial p_i}{\partial u_j} + \frac{\partial p_j}{\partial u_i} + \sum_{k=1}^n \frac{u_k}{p_k}\frac{\partial p_k}{\partial u_i}\frac{\partial p_k}{\partial u_j} \\ &\quad + \varphi\left(p_i + p_j + \sum_{k=1}^n u_k\left(\frac{\partial p_k}{\partial u_i} + \frac{\partial p_k}{\partial u_j}\right)\right) + \varphi^2\sum_{k=1}^n p_k u_k,\end{aligned} \tag{24}$$

which proves the symmetry of $HA$.

Next, we show the lower bound (22). Since $p_i$ is strictly positive in $\mathscr{D}$, $p_i(u) = \lambda + \widehat{p}_i(u)$ for any $\lambda \in (0, p_0)$, where $p_0 > 0$ is defined in (23), and $\widehat{p}_i(u)$ is still strictly positive in $\mathscr{D}$. Then we can write (24) as $HA/q = M + \lambda N$ for two matrices $M = (M_{ij})$ and $N = (N_{ij})$, defined by

$$M_{ij} = \delta_{ij}\frac{\widehat{p}_i}{u_i} + \frac{\partial \widehat{p}_i}{\partial u_j} + \frac{\partial \widehat{p}_j}{\partial u_i} + \sum_{k=1}^n \frac{u_k}{\widehat{p}_k+\lambda}\frac{\partial \widehat{p}_k}{\partial u_i}\frac{\partial \widehat{p}_k}{\partial u_j}$$
$$+\varphi\left(\widehat{p}_i + \widehat{p}_j + \sum_{k=1}^n u_k\left(\frac{\partial \widehat{p}_k}{\partial u_i} + \frac{\partial \widehat{p}_k}{\partial u_j}\right)\right) + \varphi^2\sum_{k=1}^n \widehat{p}_k u_k,$$
$$N_{ij} = \frac{\delta_{ij}}{u_i} + 2\varphi + \varphi^2\sum_{k=1}^n u_k = \frac{\delta_{ij}}{u_i} + 2\varphi + \varphi^2(1-u_{n+1}).$$

Let $v \in \mathbb{R}^n$. Then $v^\top(HA/q)v = v^\top Mv + v^\top Nv$. We consider $v^\top Nv$ first:

$$v^\top Nv = \sum_{i=1}^n \frac{v_i^2}{u_i} + \varphi(2+\varphi(1-u_{n+1}))\left(\sum_{i=1}^n v_i\right)^2. \tag{25}$$

The inequalities

$$2q(s) + (1-s)q'(s) \geq (1-s)q'(s) \geq \frac{1}{2}q'(s) \qquad \text{for } 0 \leq s \leq \frac{1}{2},$$
$$2q(s) + (1-s)q'(s) \geq 2q(s) \geq \frac{2q(1/2)}{\sup_{1/2<\sigma<1} q'(\sigma)}q'(s) \qquad \text{for } \frac{1}{2} \leq s \leq 1,$$

imply that

$$2q(u_{n+1}) + (1-u_{n+1})q'(u_{n+1}) \geq \delta q'(u_{n+1}),$$

where $\delta > 0$ is defined in (23). Thus, (25) yields

$$v^\top Nv \geq \sum_{i=1}^n \frac{v_i^2}{u_i} + \delta\varphi^2\left(\sum_{i=1}^n v_i\right)^2.$$

Finally, we show that $v^\top Mv \geq 0$, which, together with the above estimate, proves the lemma. Using the definition of $M$, we compute

$$v^\top Mv = \sum_{i=1}^n \frac{\widehat{p}_i}{u_i}v_i^2 + \sum_{k=1}^n \frac{u_k}{\widehat{p}_k}\left(\sum_{i=1}^n v_i\frac{\partial \widehat{p}_k}{\partial u_i}\right)^2 + 2\sum_{i,j=1}^n \frac{\partial \widehat{p}_j}{\partial u_i}v_iv_j \tag{26}$$
$$+2\varphi\left(\sum_{j=1}^n v_j\right)\left(\sum_{i=1}^n \widehat{p}_iv_i + \sum_{k=1}^n u_k\sum_{i=1}^n v_i\frac{\partial \widehat{p}_k}{\partial u_i}\right) + \varphi^2\left(\sum_{k=1}^n u_k\widehat{p}_k\right)\left(\sum_{j=1}^n v_j\right)^2.$$

Let us consider the terms proportional to $\varphi$ and $\varphi^2$:

$$2\varphi\left(\sum_{j=1}^n v_j\right)\left(\sum_{i=1}^n \widehat{p}_iv_i + \sum_{k=1}^n u_k\sum_{i=1}^n v_i\frac{\partial \widehat{p}_k}{\partial u_i}\right) + \varphi^2\left(\sum_{k=1}^n u_k\widehat{p}_k\right)\left(\sum_{j=1}^n v_j\right)^2$$
$$= \left(\sum_{k=1}^n u_k\widehat{p}_k\right)\left[\varphi^2\left(\sum_{j=1}^n v_j\right)^2 + 2\varphi\left(\sum_{j=1}^n v_j\right)\frac{\sum_{i=1}^n \widehat{p}_iv_i + \sum_{k=1}^n u_k\sum_{i=1}^n v_i(\partial\widehat{p}_k/\partial u_i)}{\sum_{k=1}^n u_k\widehat{p}_k}\right]$$
$$= \left(\sum_{k=1}^n u_k\widehat{p}_k\right)\left[\varphi\sum_{j=1}^n v_j + \frac{\sum_{i=1}^n \widehat{p}_iv_i + \sum_{k=1}^n u_k\sum_{i=1}^n v_i(\partial\widehat{p}_k/\partial u_i)}{\sum_{k=1}^n u_k\widehat{p}_k}\right]^2$$
$$-\frac{\left(\sum_{i=1}^n \widehat{p}_iv_i + \sum_{k=1}^n u_k\sum_{i=1}^n v_i(\partial\widehat{p}_k/\partial u_i)\right)^2}{\sum_{k=1}^n u_k\widehat{p}_k}.$$

Inserting this expression into (26) yields

$$v^\top M v \geq \sum_{i=1}^n \frac{\widehat{p}_i}{u_i} v_i^2 + \sum_{k=1}^n \frac{u_k}{\widehat{p}_k}\left(\sum_{i=1}^n v_i \frac{\partial \widehat{p}_k}{\partial u_i}\right)^2 + 2\sum_{i,j=1}^n \frac{\partial \widehat{p}_j}{\partial u_i} v_i v_j - \frac{\left(\sum_{i=1}^n \widehat{p}_i v_i + \sum_{k=1}^n u_k \sum_{i=1}^n v_i (\partial \widehat{p}_k/\partial u_i)\right)^2}{\sum_{k=1}^n u_k \widehat{p}_k}.$$

We claim that the right-hand side can be written as a square. To see this, we introduce the vectors $y = (y_1, \ldots, y_n)$, $z = (z_1, \ldots, z_n) \in \mathbb{R}^n$ by

$$y_i = \sqrt{\frac{\widehat{p}_i}{u_i}} v_i + \sqrt{\frac{u_i}{\widehat{p}_i}} \sum_{k=1}^n v_k \frac{\partial \widehat{p}_i}{\partial u_k}, \quad z_i = \frac{\sqrt{u_i \widehat{p}_i}}{\sqrt{\sum_{k=1}^n u_k \widehat{p}_k}}, \quad i = 1, \ldots, n.$$

The properties

$$|z|^2 = 1, \quad |y|^2 = \sum_{i=1}^n \frac{\widehat{p}_i}{u_i} v_i^2 + \sum_{k=1}^n \frac{u_k}{\widehat{p}_k}\left(\sum_{i=1}^n v_i \frac{\partial \widehat{p}_k}{\partial u_i}\right)^2 + 2\sum_{i,j=1}^n \frac{\partial \widehat{p}_j}{\partial u_i} v_i v_j,$$

$$y \cdot z = \frac{\sum_{i=1}^n \widehat{p}_i v_i + \sum_{k=1}^n u_k \sum_{i=1}^n v_i (\partial \widehat{p}_k/\partial u_i)}{\sqrt{\sum_{k=1}^n u_k \widehat{p}_k}}$$

show that

$$v^\top M v \geq |y|^2 - (y \cdot z)^2 = |y - (y \cdot z) z|^2 \geq 0.$$

The lemma is proved. □

### *3.3. Generalized Aubin lemmas*

We prove two generalized Aubin lemmas for functions which are piecewise constant in time, extending results from [11,21].

**Lemma 8** *(Generalized Aubin lemma I). Let $(\xi^{(\tau)})$, $(\eta_1^{(\tau)}), \ldots, (\eta_n^{(\tau)})$ be sequences of functions which are piecewise constant in time with constant step size $\tau > 0$ and which are bounded in $L^\infty(0,T; L^\infty(\Omega))$. Furthermore, they satisfy the following properties:*

- *$\xi^{(\tau)} \to \xi$ strongly in $L^2(0,T; L^2(\Omega))$ as $\tau \to 0$.*
- *$\eta_i^{(\tau)} \rightharpoonup \eta_i$ weakly* in $L^\infty(0,T; L^\infty(\Omega))$ as $\tau \to 0$ for $i = 1, \ldots, n$.*
- *There exists $C > 0$ such that for all $\tau > 0$ and $i = 1, \ldots, n$,*

$$\|\xi^{(\tau)}\|_{L^2(0,T;H^1(\Omega))} + \|\xi^{(\tau)}\eta_i^{(\tau)}\|_{L^2(0,T;H^1(\Omega))} + \tau^{-1}\|\eta_i^{(\tau)} - \pi_\tau \eta_i^{(\tau)}\|_{L^2(\tau,T;H^1(\Omega)')} \leq C, \tag{27}$$

*where $\pi_\tau \eta_i^{(\tau)}(\cdot, t) = \eta_i^{(\tau)}(\cdot, t - \tau)$ for $\tau \leq t \leq T$ is a shift operator. Let $D \subset \mathbb{R}^n$ be a compact domain such that $\eta^{(\tau)}(x,t) = (\eta_1^{(\tau)}, \ldots, \eta_n^{(\tau)})(x,t) \in D$ for a.e. $(x,t) \in \Omega \times (0,T)$. Then, for all $f \in C^0(D; \mathbb{R}^n)$, up to a subsequence, as $\tau \to 0$,*

$$\xi^{(\tau)} f(\eta^{(\tau)}) \to \xi f(\eta) \quad \text{strongly in } L^2(0,T; L^2(\Omega)).$$

Since $(\xi^{(\tau)})$ and $(\eta_i^{(\tau)})$ are assumed to be bounded in $L^\infty(0,T; L^\infty(\Omega))$, the strong convergence also holds in $L^p(0,T; L^p(\Omega))$ for all $p < \infty$. This theorem extends [21, Lemma 13], proved for $f(s) = s$, to arbitrary continuous functions $f$.

**Proof of Lemma 8.** The proof is based on the compactness result in [21, Lemma 13], whose proof goes back to [7], and an induction and approximation argument. We perform the proof in two steps. In the first step $f$ is assumed to be a monomial, in the second step we approximate an arbitrary continuous function by a polynomial and apply the Stone–Weierstrass theorem. We set $Q_T = \Omega \times (0, T)$.

*Step 1.* Let $f(\eta) = \eta^\alpha := \eta_1^{\alpha_1} \cdots \eta_n^{\alpha_n}$, where $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}_0^n$ is a multiindex. The proof is an induction argument on the rank $|\alpha| = \sum_{i=1}^n \alpha_i \geq 0$ of the multiindex. If $|\alpha| = 0$, the statement is trivially true. Let us assume that $\xi^{(\tau)}(\eta^{(\tau)})^\alpha \to \xi\eta^\alpha$ strongly in $L^2(Q_T)$ as $\tau \to 0$ for all $\alpha \in \mathbb{N}_0^n$ with $|\alpha| \leq k$, $k \geq 0$. Let $\alpha \in N_0^n$ be a multiindex such that $|\alpha| = k + 1 \geq 1$. Then there exists an index $i_0 \in \{1, \ldots, n\}$ such that $\alpha_{i_0} \geq 1$. Hence, we can define the multiindex $\beta$ such that $\beta_j = \alpha_j - \delta_{i_0,j}$ for $j = 1, \ldots, n$ and $|\beta| = k$.

Introduce $y^{(\tau)} = \xi^{(\tau)}(\eta^{(\tau)})^\beta$ and $y = \xi\eta^\beta$. Clearly, $(y^{(\tau)})$ is bounded in $L^\infty(0, T; L^\infty(\Omega))$. Since the multiindex $\beta$ has rank $k$ and thus satisfies the induction assumption, $y^{(\tau)} \to y$ strongly in $L^2(Q_T)$. We claim that $(y^{(\tau)})$ and $(y^{(\tau)}\eta_{i_0}^{(\tau)})$ are bounded in $L^2(0, T; H^1(\Omega))$. Indeed, it follows from (27) that $\xi^{(\tau)}\nabla\eta_i^{(\tau)} = \nabla(\xi^{(\tau)}\eta_i^{(\tau)}) - \eta_i^{(\tau)}\nabla\xi^{(\tau)}$ is uniformly bounded in $L^2(Q_T)$. As a consequence,

$$\begin{aligned}\nabla y^{(\tau)} &= (\eta^{(\tau)})^\beta\nabla\xi^{(\tau)} + \xi^{(\tau)}\nabla(\eta^\beta)\\ &= (\eta^{(\tau)})^\beta\nabla\xi^{(\tau)} + \sum_{k:\beta_k>0} \beta_k(\eta_k^{(\tau)})^{\beta_k-1}\left(\prod_{i\neq k}(\eta_i^{(\tau)})^{\beta_i}\right)\xi^{(\tau)}\nabla\eta_k^{(\tau)}\end{aligned}$$

is uniformly bounded in $L^2(Q_T)$, and $(y^{(\tau)})$ is bounded in $L^2(0, T; H^1(\Omega))$. In a similar way, we can show that $(y^{(\tau)}\eta_{i_0}^{(\tau)})$ is bounded in $L^2(0, T; H^1(\Omega))$. Applying [21, Lemma 13] to the sequences $(y^{(\tau)})$ and $(\eta_{i_0}^{(\tau)})$, we infer that there exists a subsequence, which is not relabeled, such that $y^{(\tau)}\eta_{i_0}^{(\tau)} \to y\eta_{i_0}$ strongly in $L^2(Q_T)$, which means, by definition of $y^{(\tau)}$ and $\beta$, that $\xi^{(\tau)}(\eta^{(\tau)})^\beta \to \xi\eta^\beta$ strongly in $L^2(Q_T)$.

*Step 2.* It follows from the previous step that the statement of the lemma is true if $f$ is a multivariate polynomial. Let $f \in C^0(D; \mathbb{R}^n)$ be given. Since $D$ is compact, we may apply the Stone–Weierstrass approximation theorem to obtain, for any $\varepsilon > 0$, a multivariate polynomial $P : D \to \mathbb{R}^n$ such that $|f(\eta) - P(\eta)| < \varepsilon$ for $\eta \in D$. Since $(\xi^{(\tau)})$ and $\xi$ are bounded in $L^\infty$, we have for some $C > 0$, which does not depend on $\varepsilon$,

$$\|\xi^{(\tau)}f(\eta^{(\tau)}) - \xi^{(\tau)}P(\eta^{(\tau)})\|_{L^2(Q_T)} \leq C\varepsilon, \quad \|\xi P(\eta) - \xi f(\eta)\|_{L^2(Q_T)} \leq C\varepsilon.$$

Thus,

$$\begin{aligned}\|\xi^{(\tau)}f(\eta^{(\tau)}) - \xi f(\eta)\|_{L^2(Q_T)} &\leq \|\xi^{(\tau)}f(\eta^{(\tau)}) - \xi^{(\tau)}P(\eta^{(\tau)})\|_{L^2(Q_T)}\\ &\quad + \|\xi^{(\tau)}P(\eta^{(\tau)}) - \xi P(\eta)\|_{L^2(Q_T)} + \|\xi P(\eta) - \xi f(\eta)\|_{L^2(Q_T)}\\ &\leq 2C\varepsilon + \|\xi^{(\tau)}P(\eta^{(\tau)}) - \xi P(\eta)\|_{L^2(Q_T)}.\end{aligned}$$

Since $P$ is a polynomial, the first step of the proof applies and the last term on the right-hand side converges to zero as $\tau \to 0$ (at least for a subsequence), resulting in

$$\limsup_{\tau\to 0} \|\xi^{(\tau)}f(\eta^{(\tau)}) - \xi f(\eta)\|_{L^2(Q_T)} \leq 2C\varepsilon.$$

Since $\varepsilon > 0$ is arbitrary and the left-hand side does not depend on $\varepsilon$, it must vanish, finishing the proof. □

**Lemma 9** *(Generalized Aubin lemma II). Let $(u^{(\tau)})$ be a sequence of functions which are piecewise constant in time with constant step size $\tau > 0$. Let $Q \in C^0(\mathbb{R})$ be strictly monotone such that both $Q$ and $Q^{-1}$ are Lipschitz continuous, and assume that there exist $C_1, C_2 > 0$ such that for all $\tau > 0$,*

$$\begin{aligned}\|Q(u^{(\tau)})\|_{L^2(0,T;H^1(\Omega))} &\leq C_1,\\ \tau^{-1}\|u^{(\tau)} - \pi_\tau u^{(\tau)}\|_{L^2(\tau,T;H^1(\Omega)')} &\leq C_2.\end{aligned}$$

*Then there exists $u \in L^2(0, T; H^1(\Omega))$ such that, up to a subsequence,*

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

*Moreover, if* $(u^{(\tau)})$ *is uniformly bounded in* $L^\infty$ *but only* $Q$ *is Lipschitz continuous, the convergence holds in* $L^p$ *for all* $1 \le p < \infty$.

This result generalizes Theorem 3a in [11], stated for $Q(s) = s^m$ with $m > 0$. A related result has been proved in [25, Theorem 1]. Instead of the global Lipschitz bound on $Q$ it is assumed that the function $|Q'|$ is bounded from below by a positive value near $\pm\infty$ and that the set $\{x : Q'(x) = 0\}$ is finite. Thus, our result is related to that one in [25].

**Proof of Lemma 9.** Denoting the Lipschitz bound of $Q$ by $L > 0$, we can estimate as follows:

$$\begin{aligned}
&\|Q(u^{(\tau)}) - \pi_\tau Q(u^{(\tau)})\|^2_{L^2(\tau,T;L^2(\Omega))} \\
&= \iint\limits_{\{u^{(\tau)} \neq \pi_\tau u^{(\tau)}\}} \frac{Q(u^{(\tau)}) - \pi_\tau Q(u^{(\tau)})}{u^{(\tau)} - \pi_\tau u^{(\tau)}} \big(u^{(\tau)} - \pi_\tau u^{(\tau)}\big)\big(Q(u^{(\tau)}) - \pi_\tau Q(u^{(\tau)})\big)dxdt \\
&\le L \int_\tau^T \int_\Omega \big(u^{(\tau)} - \pi_\tau u^{(\tau)}\big)\big(Q(u^{(\tau)}) - \pi_\tau Q(u^{(\tau)})\big)dxdt \\
&\le L\|u^{(\tau)} - \pi_\tau u^{(\tau)}\|_{L^2(\tau,T;H^1(\Omega)')}\|Q(u^{(\tau)}) - \pi_\tau Q(u^{(\tau)})\|_{L^2(\tau,T;H^1(\Omega))} \\
&\le 2LC_1C_2\tau.
\end{aligned}$$

Observe that the product $(u^{(\tau)} - \pi_\tau u^{(\tau)})(Q(u^{(\tau)}) - \pi_\tau Q(u^{(\tau)}))$ is nonnegative since $Q$ is monotone. Applying Theorem 1 of [16] and the compactness of the embedding $H^1(\Omega) \hookrightarrow L^2(\Omega)$, there exists a subsequence which is not relabeled such that, as $\tau \to 0$, $Q(u^{(\tau)}) \to z$ in $L^2(0,T;L^2(\Omega))$ for some function $z$. In particular, we may assume that the convergence also holds pointwise a.e. Furthermore, $u^{(\tau)} \to u := Q^{-1}(z)$ a.e. The facts that $Q^{-1}$ is Lipschitz continuous and $(Q(u^{(\tau)}))$ is bounded in $L^2$ imply that also $(u^{(\tau)})$ is bounded in $L^2$. By the dominated convergence theorem, $u^{(\tau)} \to u$ strongly in $L^2(0,T;L^2(\Omega))$. Finally, if $(u^{(\tau)})$ is uniformly bounded, we may apply the dominated convergence theorem to $((u^{(\tau)})^p)$ for some $p < \infty$ and conclude the strong convergence in $L^p$. For this result, we do not need the assumption that $Q^{-1}$ is Lipschitz continuous. □

### *3.4. Further results*

We show that the Fisher information $\int_\Omega |\nabla\sqrt{u}|^2 d\mu$ is subadditive, and we recall a convex Sobolev inequality.

**Lemma 10.** *Let* $\mu$ *be an absolutely continuous measure with respect to the Lebesgue measure, and let* $f, g : \Omega \to [0,\infty)$ *be measurable, bounded, positive functions such that* $\sqrt{f}, \sqrt{g} \in H^1(\Omega, d\mu)$. *Then*

$$\int_\Omega |\nabla\sqrt{f+g}|^2 d\mu \le \int_\Omega |\nabla\sqrt{f}|^2 d\mu + \int_\Omega |\nabla\sqrt{g}|^2 d\mu.$$

This result was proven in [28, Section 3.6] in a slightly different context. For the convenience of the reader, we present the (short) proof.

**Proof of Lemma 10.** We define the function $F : [0,1] \to \mathbb{R}$ by

$$F(s) = \int_\Omega |\nabla\sqrt{f}|^2 d\mu + \int_\Omega |\nabla\sqrt{sg}|^2 d\mu - \int_\Omega |\nabla\sqrt{f+sg}|^2 d\mu, \quad s \in [0,1].$$

Then $F(0) = 0$ and $F'(s) \ge 0$ for all $s \in [0,1]$ since

$$\begin{aligned}
F'(s) &= \int_\Omega |\nabla\sqrt{g}|^2 d\mu - \int_\Omega \nabla\sqrt{f+sg}\cdot\nabla\left(\frac{g}{\sqrt{f+sg}}\right)d\mu \\
&= \int_\Omega |\nabla\sqrt{g}|^2 d\mu - \int_\Omega \nabla\sqrt{f+sg}\cdot\left(\frac{2\sqrt{g}\nabla\sqrt{g}}{\sqrt{f+sg}} - \frac{g}{f+sg}\nabla\sqrt{f+sg}\right)d\mu \\
&= \int_\Omega |\nabla\sqrt{g}|^2 d\mu + \int_\Omega \frac{g}{f+sg}|\nabla\sqrt{f+sg}|^2 d\mu - 2\int_\Omega \frac{\sqrt{g}}{\sqrt{f+sg}}\nabla\sqrt{g}\cdot\nabla\sqrt{f+sg}d\mu \\
&= \int_\Omega \left|\nabla\sqrt{g} - \frac{\sqrt{g}}{\sqrt{f+sg}}\nabla\sqrt{f+sg}\right|^2 d\mu \ge 0.
\end{aligned}$$

We conclude that $F(1) \ge 0$ which shows the lemma. □

**Lemma 11.** *Let $\Omega \subset \mathbb{R}^d$ ($d \ge 1$) be a convex domain and let $g \in C^4$ be a convex function such that $1/g''$ is concave. Then there exists $c_S > 0$ such that for all integrable functions $u$ with integrable $g(u)$ and $g''(u)|\nabla u|^2$,*

$$\frac{1}{|\Omega|}\int_\Omega g(u)dx - g\left(\frac{1}{|\Omega|}\int_\Omega udx\right) \le \frac{c_S}{|\Omega|}\int_\Omega g''(u)|\nabla u|^2 dx,$$

*where $|\Omega|$ denotes the measure of $\Omega$.*

A proof can be found in [4, Prop. 7.6.1] or [2, Remark 3.8].

## 4. Proof of Theorem 1

We divide the proof into several steps.

### 4.1. Time discretization and regularization of system (1)

We recall the definition of the entropy variable $w = h'(u)$ for $u \in \mathcal{D}$, where $h$ is defined in (9). Lemma 6 shows that $h'$ is invertible, thus we may define $u = (h')^{-1}(w)$ for $w \in \mathbb{R}^n$ and we may set $u(w) = u$. By Lemma 7, the matrix $B(w) = A(u)(h'')^{-1}(u)$ is positive definite for all $w \in \mathbb{R}$ and $u = u(w)$. We introduce a time discretization for (1). Let $T > 0$, $N \in \mathbb{N}$, and let $\tau = T/N$ be the time step size. Furthermore, let $0 < \varepsilon < 1$ be a regularization parameter and let $m \in \mathbb{N}$ be such that $H^m(\Omega) \hookrightarrow L^\infty(\Omega)$ compactly (i.e. choose $m > d/2$). Given $w^{k-1} \in H^m(\Omega;\mathbb{R}^n)$, we wish to find $w^k \in H^m(\Omega;\mathbb{R}^n)$ which solves the discretized and regularized problem

$$\frac{1}{\tau}\int_\Omega (u(w^k) - u(w^{k-1}))\cdot\phi dx + \int_\Omega \nabla\phi : B(w^k)\nabla w^k dx + \tau^2 b_\varepsilon(\phi, w^k) = 0 \tag{28}$$

for $\phi \in H^m(\Omega;\mathbb{R}^n)$, where

$$b_\varepsilon(\phi, w^k) = \int_\Omega (\phi\cdot w^k + \nabla\phi : \nabla w^k)dx + \varepsilon\sum_{2\le|\alpha|\le m} D^\alpha\phi\cdot D^\alpha w^k dx, \tag{29}$$

and $D^\alpha$ is a partial derivative of order $|\alpha|$. We prove the existence of weak solutions to (28).

**Lemma 12.** *Let (7)–(8) hold and let $u^0 : \Omega \to \mathcal{D}$ be measurable such that $h(u^0) \in L^1(\Omega)$. Then there exists a sequence of solutions $w^k \in H^m(\Omega;\mathbb{R}^n)$ to (28) satisfying the discrete entropy inequality*

$$\int_\Omega h(u(w^k))dx + \tau\int_\Omega \nabla w^k : B(w^k)\nabla w^k dx + \tau^3 b_\varepsilon(w^k, w^k) \le \int_\Omega h(u(w^{k-1}))dx. \tag{30}$$

**Proof.** The idea is to apply the Leray–Schauder fixed-point theorem. Let $y \in L^\infty(\Omega; \mathbb{R}^n)$ and $\eta \in [0, 1]$ be given. We first solve the linear problem

$$a(w, \phi) = F(\phi) \quad \text{for all } \phi \in H^m(\Omega; \mathbb{R}^n), \tag{31}$$

where

$$a(w, \phi) = \int_\Omega \nabla\phi : B(y)\nabla w dx + \tau^2 b_\varepsilon(w, \phi),$$

$$F(\phi) = -\frac{\eta}{\tau}\int_\Omega (u(y) - u(w^{k-1})) \cdot \phi dx.$$

The forms $a$ and $F$ are bounded on $H^m(\Omega; \mathbb{R}^n)$. The matrix $B(y) = A(u(y))h''(u(y))^{-1}$ is positive semi-definite,

$$v^\top B(y)v = [h''(u(y))^{-1}v]^\top h''(u(y))A(u(y))[h''(u(y))^{-1}v] \geq 0$$

for all $v \in \mathbb{R}^n$, thanks to (22). Hence, the bilinear form $a$ is coercive:

$$a(w, w) \geq \varepsilon\tau^2\|w\|^2_{H^m(\Omega)} \quad \text{for } w \in H^m(\Omega; \mathbb{R}^n).$$

Therefore, we can apply the Lax–Milgram lemma to infer the existence of a unique solution $w \in H^m(\Omega; \mathbb{R}^n) \hookrightarrow L^\infty(\Omega; \mathbb{R}^n)$ to (31). This defines the fixed-point operator $S : L^\infty(\Omega; \mathbb{R}^n) \times [0, 1] \to L^\infty(\Omega; \mathbb{R}^n)$, $S(y, \eta) = w$, where $w$ solves (31).

It holds that $S(y, 0) = 0$ for all $y \in L^\infty(\Omega; \mathbb{R}^n)$. Furthermore, standard arguments show that $S$ is continuous (see e.g. the proof of Lemma 5 in [21]). It remains to prove a uniform bound for all fixed points $S(\cdot, \eta)$ in $L^\infty(\Omega; \mathbb{R}^n)$. Let $w \in L^\infty(\Omega; \mathbb{R}^n)$ be such a fixed point. Then $w$ solves (31) with $y$ replaced by $w$. With the test function $\phi = w$, we find that

$$\frac{\eta}{\tau}\int_\Omega (u(w) - u(w^{k-1})) \cdot w dx + \int_\Omega \nabla w : B(w)\nabla w dx + \tau^2 b_\varepsilon(w, w) = 0. \tag{32}$$

The convexity of $h$ implies that $h(x) - h(y) \leq h'(u) \cdot (x - y)$ for all $x, y \in \mathcal{D}$. Choosing $x = u(w)$ and $y = u(w^{k-1})$ and employing $h'(u(w)) = w$, this gives

$$\frac{\eta}{\tau}\int_\Omega (u(w) - u(w^{k-1})) \cdot w dx \geq \frac{\eta}{\tau}\int_\Omega \big(h(u(w)) - h(u(w^{k-1}))\big)dx.$$

Taking into account the positive semi-definiteness of $B(w)$, we infer from (32) that

$$\eta\int_\Omega h(u(w))dx + \varepsilon\tau^3\|w\|^2_{H^m(\Omega)} \leq \eta\int_\Omega h(u(w^{k-1}))dx.$$

This yields an $H^m$ bound for $w$ uniform in $\eta$ (but not uniform in $\varepsilon$ and $\tau$). By the Leray–Schauder fixed-point theorem, we conclude the existence of a solution $w \in H^m(\Omega; \mathbb{R}^n)$ to (31) with $y$ replaced by $w$ and $\eta = 1$. □

We derive some a priori estimates uniform in $\varepsilon$ and $\tau$. In the following, we set $u^k = u(w^k)$ for $k \geq 1$, where $(w^k)$ solves (28).

**Lemma 13.** *Under the assumptions of Lemma 12, there exists a constant $C > 0$ such that for all $\varepsilon, \tau > 0$,*

$$\begin{aligned}\int_\Omega h(u^k)dx &+ 4\tau p_0\sum_{j=1}^k\int_\Omega q(u^j_{n+1})\sum_{i=1}^n|\nabla(u^j_i)^{1/2}|^2dx \\ &+ 4\tau p_0\delta\sum_{j=1}^k\int_\Omega|\nabla q(u^j_{n+1})^{1/2}|^2dx + \tau^3\sum_{j=1}^k b_\varepsilon(w^j, w^j) \leq \int_\Omega h(u^0)dx,\end{aligned} \tag{33}$$

*where $p_0$ and $\delta$ are defined in (23).*

**Proof.** By Lemma 12, the sequence $(w^k)$ satisfies (28). Then, taking into account the identity $\nabla w^k : B(w^k)\nabla w^k = \nabla u^k : h''(u^k)A(u^k)\nabla u^k$, we deduce that

$$\int_\Omega h(u^k)dx + \tau\int_\Omega \nabla u^k : h''(u^k)A(u^k)\nabla u^k dx + \tau^3 b_\varepsilon(w^k, w^k) \le \int_\Omega h(u^{k-1})dx.$$

Resolving this recursion yields

$$\int_\Omega h(u^k)dx + \sum_{j=1}^k \tau\int_\Omega \nabla u^j : h''(u^j)A(u^j)\nabla u^j dx + \tau^3\sum_{j=1}^k b_\varepsilon(w^j, w^j) \le \int_\Omega h(u^0)dx.$$

Then the conclusion follows from Lemma 7 and $|\sum_{i=1}^n \nabla u_i^j|^2 = |\nabla u_{n+1}^j|^2$. □

*4.2. The limit $\varepsilon \to 0$*

Let $(w^k)$ be a sequence of solutions to (28). We fix $k \in \{1, \dots, n\}$ and set $u_i^{(\varepsilon)} = u_i^k$ $(i = 1, \dots, n+1)$ and $w_i^{(\varepsilon)} = w_i^k$ $(i = 1, \dots, n)$. The identity

$$\begin{aligned}(B(w^k)\nabla w^k)_i &= (A(u^k)\nabla u^k)_i \\ &= q(u_{n+1}^k)^{1/2}\nabla\big(u_i^k p_i(u^k)q(u_{n+1}^k)^{1/2}\big) - 3u_i^k p_i(u^k)q(u_{n+1}^k)^{1/2}\nabla q(u_{n+1}^k)^{1/2}\end{aligned}$$

shows that $u^k$ solves

$$\begin{aligned}&\frac{1}{\tau}\int_\Omega (u^k - u^{k-1})\cdot\phi dx + \sum_{i=1}^n\int_\Omega \big[q(u_{n+1}^j)^{1/2}\nabla\big(u_i^j p_i(u^j)q(u_{n+1}^j)^{1/2}\big) \\ &\quad - 3u_i^j p_i(u^j)q(u_{n+1}^j)^{1/2}\nabla q(u_{n+1}^j)^{1/2}\big]\cdot\nabla\phi_i dx + \tau^2 b_\varepsilon(w^j, \phi) = 0\end{aligned} \tag{34}$$

for all $\phi = (\phi_1, \dots, \phi_n) \in H^m(\Omega; \mathbb{R}^n)$. We wish to pass to the limit $\varepsilon \to 0$ in (34).

By Lemma 13 and definition (29) of $b_\varepsilon$, we have

$$\varepsilon\tau^3\sum_{j=1}^k \|w^j\|_{H^m(\Omega)}^2 + \tau^3\sum_{j=1}^k \|w^j\|_{H^1(\Omega)}^2 \le C, \tag{35}$$

where here and in the following, $C > 0$ denotes a generic constant independent of $\varepsilon$ and $\tau$. Thus, because of the boundedness of $(h'')^{-1}$ (see Lemma 6),

$$\|\nabla u^{(\varepsilon)}\|_{L^2(\Omega)} = \|(h''(u^{(\varepsilon)}))^{-1}\nabla w^{(\varepsilon)}\|_{L^2(\Omega)} \le C\|\nabla w^{(\varepsilon)}\|_{L^2(\Omega)} \le C\tau^{-3/2}.$$

Together with the $L^\infty$ bound for $(u^{(\varepsilon)})$, this implies that

$$\|u^{(\varepsilon)}\|_{H^1(\Omega)} \le C\tau^{-3/2}.$$

Therefore, up to subsequences, as $\varepsilon \to 0$,

$$u^{(\varepsilon)} \rightharpoonup u \quad \text{weakly in } H^1(\Omega), \quad u^{(\varepsilon)} \to u \quad \text{strongly in } L^2(\Omega) \text{ and a.e. in } \Omega,$$

since $H^1(\Omega)$ embeds compactly into $L^2(\Omega)$. We infer that $u_{n+1}^{(\varepsilon)} = 1 - \sum_{i=1}^n u_i^{(\varepsilon)} \to u_{n+1} := 1 - \sum_{i=1}^n u_i$ strongly in $L^2(\Omega)$ and a.e. in $\Omega$. The $L^\infty$ and $H^1$ bounds for $(u^{(\varepsilon)})$ as well as the $L^2$ bound for $\nabla q(u_{n+1}^{(\varepsilon)})^{1/2}$ in (33) show that

$$\begin{aligned}&\nabla\big(u_i^{(\varepsilon)} p_i(u^{(\varepsilon)})q(u_{n+1}^{(\varepsilon)})^{1/2}\big) \\ &\quad = u_i^{(\varepsilon)} p_i(u^{(\varepsilon)})\nabla q(u_{n+1}^{(\varepsilon)})^{1/2} + q(u_{n+1}^{(\varepsilon)})^{1/2}\sum_{j=1}^n\left(\delta_{ij}p_i(u^{(\varepsilon)}) + u_i^{(\varepsilon)}\frac{\partial p_i}{\partial u_j}(u^{(\varepsilon)})\right)\nabla u_j^{(\varepsilon)}\end{aligned}$$

is uniformly bounded in $L^2(\Omega)$ and hence,

$$\|u_i^{(\varepsilon)}p_i(u^{(\varepsilon)})q(u_{n+1}^{(\varepsilon)})^{1/2}\|_{H^1(\Omega)} \le C\tau^{-1/2}.$$

We employ the a.e. convergence of $(u^{(\varepsilon)})$ and $(u_{n+1}^{(\varepsilon)})$ and the continuity of $p_i$ and $q$ to obtain

$$u_i^{(\varepsilon)}p_i(u^{(\varepsilon)})q(u_{n+1}^{(\varepsilon)})^{1/2}) \to u_i p_i(u)q(u_{n+1})^{1/2} \quad \text{a.e. in } \Omega,$$

and, by the dominated convergence theorem, strongly in $L^2(\Omega)$. Thus, using the $H^1$ bound,

$$u_i^{(\varepsilon)}p_i(u^{(\varepsilon)})q(u_{n+1}^{(\varepsilon)})^{1/2} \rightharpoonup u_i p_i(u)q(u_{n+1})^{1/2} \quad \text{weakly in } H^1(\Omega).$$

Similar arguments, using the uniform estimates coming from (33), show that

$$q(u_{n+1}^{(\varepsilon)})^{1/2} \to q(u_{n+1})^{1/2} \quad \text{strongly in } L^2(\Omega) \text{ and weakly in } H^1(\Omega), \tag{36}$$

$$q(u_{n+1}^{(\varepsilon)})^{1/2}(u_i^{(\varepsilon)})^{1/2} \rightharpoonup q(u_{n+1})^{1/2}u_i^{1/2} \quad \text{weakly in } H^1(\Omega). \tag{37}$$

It follows from the bound (35) that, up to subsequences,

$$\varepsilon w^{(\varepsilon)} \to 0 \quad \text{strongly in } H^m(\Omega), \quad w^{(\varepsilon)} \rightharpoonup w \quad \text{weakly in } H^1(\Omega).$$

We set $u^k := u$. The above convergences hold for all $k = 1, \ldots, N$, where $T = N\tau$. Thus, we obtain a sequence of limit functions $(u^j)$. The above convergence results are sufficient to pass to the limit $\varepsilon \to 0$ in (34), resulting in

$$\frac{1}{\tau}\int_\Omega (u^k - u^{k-1})\cdot\phi dx + \sum_{i=1}^n \int_\Omega \big[q(u_{n+1}^j)^{1/2}\nabla\big(u_i^j p_i(u^j)q(u_{n+1}^j)^{1/2}\big) \\ - 3u_i^j p_i(u)q(u_{n+1}^j)^{1/2}\nabla q(u_{n+1}^j)^{1/2}\big]\cdot\nabla\phi_i dx + \tau^2\int_\Omega (w\cdot\phi + \nabla w : \nabla\phi)dx = 0 \tag{38}$$

for $\phi \in H^m(\Omega;\mathbb{R}^n)$. By density, this relation also holds for all $\phi \in H^1(\Omega;\mathbb{R}^n)$. Note that generally we cannot identify $w$ with $(h')^{-1}(u)$ anymore but this is not needed in the remaining proof.

Finally, we wish to pass to the limit $\varepsilon \to 0$ in (33), where $u^k$ has to be replaced by $u^{(\varepsilon)}$. Since

$$q(u_{n+1}^{(\varepsilon)})^{1/2}\nabla(u_i^{(\varepsilon)})^{1/2} = \nabla\big(q(u_{n+1}^{(\varepsilon)})^{1/2}(u_i^{(\varepsilon)})^{1/2}\big) - (u_i^{(\varepsilon)})^{1/2}\nabla q(u_{n+1}^{(\varepsilon)})^{1/2}, \tag{39}$$

the strong convergence $(u_i^{(\varepsilon)})^{1/2} \to u_i^{1/2}$ in $L^4(\Omega)$ and the weak convergences (36) and (37) imply that

$$\begin{aligned} q(u_{n+1}^{(\varepsilon)})^{1/2}\nabla(u_i^{(\varepsilon)})^{1/2} &\rightharpoonup \nabla\big(q(u_{n+1})^{1/2}u_i^{1/2}\big) - u_i^{1/2}\nabla q(u_{n+1})^{1/2} \\ &= q(u_{n+1})^{1/2}\nabla u_i^{1/2} \quad \text{weakly in } L^1(\Omega). \end{aligned}$$

In fact, since by (33),

$$\|q(u_{n+1}^{(\varepsilon)})^{1/2}\nabla(u_i^{(\varepsilon)})^{1/2}\|_{L^2(\Omega)} \le C\tau^{-1/2},$$

the above weak convergence also holds in $L^2(\Omega)$. In particular, by the weak lower semi-continuity of the $L^2$ norm,

$$\begin{aligned} \liminf_{\varepsilon\to 0}\int_\Omega q(u_{n+1}^{(\varepsilon)})|\nabla(u_i^{(\varepsilon)})^{1/2}|^2 dx &\ge \int_\Omega q(u_{n+1})|\nabla u_i^{1/2}|^2 dx, \\ \liminf_{\varepsilon\to 0}\int_\Omega |\nabla q(u_{n+1}^{(\varepsilon)})^{1/2}|^2 dx &\ge \int_\Omega |\nabla q(u_{n+1})|^2 dx, \\ \liminf_{\varepsilon\to 0}\|w^{(\varepsilon)}\|_{H^1(\Omega)}^2 &\ge \|w\|_{H^1(\Omega)}^2. \end{aligned}$$

Recall that $u^k = u$ and $w^k = w$. Passing to the limit inferior $\varepsilon \to 0$ in (33) and observing that $b_\varepsilon(w^{(\varepsilon)}, w^{(\varepsilon)}) \ge \|w^{(\varepsilon)}\|_{H^1(\Omega)}^2$, we infer that

$$\int_\Omega h(u^k)dx + 4\tau p_0 \sum_{j=1}^k \int_\Omega q(u_{n+1}^j) \sum_{i=1}^n |\nabla (u_i^j)^{1/2}|^2 dx \\ + 4\tau p_0 \delta \sum_{j=1}^k \int_\Omega |\nabla q(u_{n+1})^{1/2}|^2 dx + \tau^3 \sum_{j=1}^k \|w^j\|_{H^1(\Omega)}^2 \le \int_\Omega h(u^0)dx. \tag{40}$$

### 4.3. The limit $\tau \to 0$

We set $u^{(\tau)}(x,t) = u^k(x)$ and $w^{(\tau)}(x,t) = w^k(x)$ for $x \in \Omega$, $t \in ((k-1)\tau, k\tau]$. Equation (38) can be formulated as

$$\frac{1}{\tau}\int_\tau^T \int_\Omega (u^{(\tau)} - \pi_\tau u^{(\tau)}) \cdot \phi dxdt + \sum_{i=1}^n \int_\tau^T \int_\Omega \Big[ q(u_{n+1}^{(\tau)})^{1/2} \nabla \big( u_i^{(\tau)} p_i(u^{(\tau)}) q(u_{n+1}^{(\tau)})^{1/2} \big) \\ - 3u_i^{(\tau)} p_i(u^{(\tau)}) q(u_{n+1}^{(\tau)})^{1/2} \nabla q(u_{n+1})^{1/2} \Big] \cdot \nabla \phi_i dxdt \\ + \tau^2 \int_\tau^T \int_\Omega (w^{(\tau)} \cdot \phi + \nabla w^{(\tau)} : \nabla \phi) dxdt = 0 \tag{41}$$

for all $\phi(t) \in H^1(\Omega; \mathbb{R}^n)$ being piecewise constant in time and, by density, for all $\phi \in L^2(0,T; H^1(\Omega))$. Inequality (40) becomes

$$\int_\Omega h(u^{(\tau)}(T))dx + 4p_0 \int_0^T \int_\Omega q(u_{n+1}^{(\tau)}) \sum_{i=1}^n |\nabla (u_i^{(\tau)})^{1/2}|^2 dxdt \\ + 4p_0\delta \int_0^T \int_\Omega |\nabla q(u_{n+1}^{(\tau)})^{1/2}|^2 dxdt + \tau^2 \int_0^T \|w^{(\tau)}\|_{H^1(\Omega)}^2 dt \le \int_\Omega h(u^0)dx.$$

This gives the following uniform estimates:

$$\|q(u_{n+1}^{(\tau)})^{1/2} \nabla (u_i^{(\tau)})^{1/2}\|_{L^2(0,T;L^2(\Omega))} + \|q(u_{n+1}^{(\tau)})^{1/2}\|_{L^2(0,T;H^1(\Omega))} \le C, \tag{42}$$

$$\tau \|w^{(\tau)}\|_{L^2(0,T;H^1(\Omega))} \le C. \tag{43}$$

These bounds as well as the $L^\infty$ bound for $(u_i^{(\tau)})$ show that

$$\nabla \big( u_i^{(\tau)} p_i(u^{(\tau)}) q(u_{n+1}^{(\tau)})^{1/2} \big) = u_i^{(\tau)} p_i(u^{(\tau)}) \nabla q(u_{n+1}^{(\tau)})^{1/2} \\ + q(u_{n+1}^{(\tau)})^{1/2} \sum_{j=1}^n \left( \delta_{ij} p_i(u^{(\tau)}) + u_i^{(\tau)} \frac{\partial p_i}{\partial u_j}(u^{(\tau)}) \right) \nabla u_j^{(\tau)} \\ = u_i^{(\tau)} p_i(u^{(\tau)}) \nabla q(u_{n+1}^{(\tau)})^{1/2} \\ + 2\sum_{j=1}^n (u_j^{(\tau)})^{1/2} \left( \delta_{ij} p_i(u^{(\tau)}) + u_i^{(\tau)} \frac{\partial p_i}{\partial u_j}(u^{(\tau)}) \right) q(u_{n+1}^{(\tau)})^{1/2} \nabla (u_j^{(\tau)})^{1/2}$$

is uniformly bounded in $L^2(0,T; L^2(\Omega))$ and consequently,

$$\|u_i^{(\tau)} p_i(u^{(\tau)}) q(u_{n+1}^{(\tau)})^{1/2}\|_{L^2(0,T;H^1(\Omega))} \le C. \tag{44}$$

Similarly, (42) yields the estimate

$$\|(u_i^{(\tau)})^{1/2} q(u_{n+1}^{(\tau)})^{1/2}\|_{L^2(0,T;H^1(\Omega))} \le C. \tag{45}$$

Thus, the $L^\infty$ bound on $(u_i^{(\tau)})$ and estimates (42) and (43) give

$$\begin{aligned} &\tau^{-1}\|u^{(\tau)} - \pi_\tau u^{(\tau)}\|_{L^2(\tau,T;H^1(\Omega)')} \\ &\quad \le \sum_{i=1}^n \|q(u_{n+1}^{(\tau)})^{1/2}\|_{L^\infty(\tau,T;L^\infty(\Omega))} \big\|\nabla\big(u_i^{(\tau)} p_i(u^{(\tau)}) q(u_{n+1}^{(\tau)})^{1/2}\big)\big\|_{L^2(\tau,T;L^2(\Omega))} \\ &\qquad + 3\sum_{i=1}^n \|u_i^{(\tau)} p_i(u^{(\tau)}) q(u_{n+1})^{1/2}\|_{L^\infty(\tau,T;L^\infty(\Omega))} \|\nabla q(u_{n+1})^{1/2}\|_{L^2(0,T;L^2(\Omega))} \\ &\qquad + \tau^2 \|w^{(\tau)}\|^2_{L^2(\tau,T;H^1(\Omega))} \le C. \end{aligned} \tag{46}$$

Now, we define the function $Q(s) = \int_0^s q(\sigma)^{1/2} d\sigma$ for $s \in [0,1]$. Then $Q \in C^1([0,1])$ is strictly increasing. By assumption (7), $q(u_{n+1}^{(\tau)})/q'(u_{n+1}^{(\tau)})$ is uniformly bounded a.e. and thus,

$$\nabla Q(u_{n+1}^{(\tau)}) = \frac{Q'(u_{n+1}^{(\tau)})}{Q''(u_{n+1}^{(\tau)})} \nabla Q'(u_{n+1}^{(\tau)}) = \frac{2q(u_{n+1}^{(\tau)})}{q'(u_{n+1}^{(\tau)})} \nabla Q'(u_{n+1}^{(\tau)})$$

is uniformly bounded in $L^2(0,T;L^2(\Omega))$. We conclude that

$$\|Q(u_{n+1}^{(\tau)})\|_{L^2(0,T;H^1(\Omega))} \le C. \tag{47}$$

Estimates (46)–(47) show that the assumptions of Lemma 9 are fulfilled, and we infer the existence of a subsequence, which is not relabeled, such that, as $\tau \to 0$,

$$u_{n+1}^{(\tau)} \to u_{n+1} \quad \text{strongly in } L^r(0,T;L^r(\Omega)), \quad r < \infty. \tag{48}$$

This result, the bound (42), and the continuity of $q$ imply that

$$q(u_{n+1}^{(\tau)})^{1/2} \to q(u_{n+1})^{1/2} \quad \text{strongly in } L^r(0,T;L^r(\Omega)), \quad r < \infty, \tag{49}$$

$$q(u_{n+1}^{(\tau)})^{1/2} \rightharpoonup q(u_{n+1})^{1/2} \quad \text{weakly in } L^2(0,T;H^1(\Omega)). \tag{50}$$

Using the $L^\infty$ bound for $(u_i^{(\tau)})$, we have, up to a subsequence, $u_i^{(\tau)} \rightharpoonup^* u_i$ weakly* in $L^\infty(0,T;L^\infty(\Omega))$ as $\tau \to 0$. This convergence also holds in $L^2$. Thus, (48) implies that the relation $u_{n+1}^{(\tau)} = 1 - \sum_{i=1}^n u_i^{(\tau)}$ is satisfied by the limit function, $u_{n+1} = 1 - \sum_{i=1}^n u_i$. The set $\{v \in L^2(0,T;L^2(\Omega)) : v \ge 0 \text{ a.e. in } \Omega \times (0,T)\}$ is (strongly) closed and convex. Hence, it is also weakly closed, and the property $u_i^{(\tau)} \ge 0$ holds in the limit, i.e. $u_i \ge 0$ a.e. in $\Omega \times (0,T)$.

We turn to the convergence properties of the sequences $(u_i^{(\tau)})$ for $i = 1, \ldots, n$. We cannot expect strong convergence of $(u_i^{(\tau)})$, but the generalized Aubin–Lions Lemma 8 shows that the product $f(u^{(\tau)})q(u_{n+1}^{(\tau)})^{1/2}$ converges strongly, where $f$ is any continuous function. To make this precise, we verify the assumptions of Lemma 8. Set $\xi^{(\tau)} := q(u_{n+1}^{(\tau)})^{1/2}$ and $\eta_i^{(\tau)} := u_i^{(\tau)}$. Because of the $L^\infty$ bounds for $(u_i^{(\tau)})$, up to a subsequence,

$$\eta_i^{(\tau)} \rightharpoonup^* \eta_i = u_i \quad \text{weakly* in } L^\infty(0,T;L^\infty(\Omega)).$$

Furthermore, by (49), $\xi^{(\tau)} \to \xi = q(u_{n+1})^{1/2}$ strongly in $L^2(0,T;L^2(\Omega))$. Estimates (42), (45), and (46) show that the assumptions of Lemma 8 are satisfied, and we conclude the existence of a subsequence (not relabeled) such that

$$f(u^{(\tau)})q(u_{n+1}^{(\tau)})^{1/2} = f(\eta_i^{(\tau)})\xi^{(\tau)} \to f(\eta)\xi = f(u_i)q(u_{n+1})^{1/2} \quad \text{strongly in } L^2(0,T;L^2(\Omega))$$

for any function $f \in C^0(\overline{\mathscr{D}};\mathbb{R}^n)$. We choose $f(s) = s_i^{1/2}$ and $f(s) = s_i p_i(s)$ for $s = (s_i) \in \overline{\mathscr{D}}$. Then

$$\begin{aligned} (u_i^{(\tau)})^{1/2} q(u_{n+1}^{(\tau)})^{1/2} &\to u_i^{1/2} q(u_{n+1})^{1/2} \quad \text{strongly in } L^2(0,T;L^2(\Omega)), \\ u_i^{(\tau)} p_i(u^{(\tau)}) q(u_{n+1}^{(\tau)})^{1/2} &\to u_i p_i(u) q(u_{n+1})^{1/2} \quad \text{strongly in } L^2(0,T;L^2(\Omega)). \end{aligned} \tag{51}$$

We conclude from the bounds (44) and (45) that the above sequences converge weakly in $L^2(0,T;H^1(\Omega))$ and the limit functions can be identified:

$$(u_i^{(\tau)})^{1/2}q(u_{n+1}^{(\tau)})^{1/2} \rightharpoonup u_i^{1/2}q(u_{n+1})^{1/2} \quad \text{weakly in } L^2(0,T;H^1(\Omega)), \tag{52}$$

$$u_i^{(\tau)}p_i(u^{(\tau)})q(u_{n+1}^{(\tau)})^{1/2} \rightharpoonup u_ip_i(u)q(u_{n+1})^{1/2} \quad \text{weakly in } L^2(0,T;H^1(\Omega)). \tag{53}$$

We infer from estimate (46) that

$$\tau^{-1}(u_i^{(\tau)} - \pi_\tau u_i^{(\tau)}) \rightharpoonup \partial_t u_i \quad \text{weakly in } L^2(0,T;H^1(\Omega)'), \quad i = 1,\dots,n.$$

Moreover, taking into account (43),

$$\tau^2 w^{(\tau)} \to 0 \quad \text{strongly in } L^2(0,T;H^1(\Omega)).$$

These convergence results as well as the convergences (49)–(51) and (53) allow us to perform the limit $\tau \to 0$ in (41), which yields the weak formulation (15).

### 4.4. *Entropy inequality and positivity*

It remains to verify the entropy inequality (16) and the (conditional) positivity of $u_{n+1}$. Since the entropy density $h$ is convex and continuous, it is weakly lower semi-continuous [6, Corollary 3.9]. Thus, by the weak convergence of $(u_i^{(\tau)}(t))$,

$$\int_\Omega h(u(t))dx \le \liminf_{\tau\to 0} \int_\Omega h(u^{(\tau)}(t))dx \quad \text{for a.e. } t > 0.$$

Employing the convergences (48), (49), and (52), it follows that

$$q(u_{n+1}^{(\tau)})\nabla(u_i^{(\tau)})^{1/2} = q(u_{n+1}^{(\tau)})^{1/2}\nabla\big(q(u_{n+1}^{(\tau)})^{1/2}(u_i^{(\tau)})^{1/2}\big) - q(u_{n+1}^{(\tau)})^{1/2}(u_i^{(\tau)})^{1/2}\nabla q(u_{n+1}^{(\tau)})^{1/2}$$

converges weakly in $L^1$, but because of the $L^2$ bound (42) this convergence also holds in $L^2$:

$$q(u_{n+1}^{(\tau)})\nabla(u_i^{(\tau)})^{1/2} \rightharpoonup q(u_{n+1})\nabla u_i^{1/2} \quad \text{weakly in } L^2(0,T;L^2(\Omega)).$$

These results, together with (50), allow us to pass to the limit inferior $\tau \to 0$ in (40), yielding (16).

Finally, assume that

$$\int_0^b |\log q(s)|ds = +\infty \quad \text{for all } 0 < b < 1. \tag{54}$$

We deduce from the discrete entropy inequality (40) and definition (9) of $h$ that

$$\int_\Omega \int_a^{u_{n+1}^{(\tau)}(x,t)} \log q(s)dsdx \le \int_\Omega h(u^{(\tau)}(x,t))dx \le \int_\Omega h(u^0)dx \quad \text{for a.e. } t > 0.$$

Then, by the strong convergence (48) of $(u_{n+1}^{(\tau)})$ and the nonnegativity of $\int_a^b \log q(s)ds \ge 0$, we can apply Fatou's lemma yielding

$$\int_\Omega \int_a^{u_{n+1}(x,t)} \log q(s)dsdx \le \int_\Omega h(u^0)dx.$$

In particular, $\int_a^{u_{n+1}(x,t)} \log q(s)ds < \infty$ for a.e. $x \in \Omega$. We conclude from this fact and assumption (54) that $u_{n+1}(x,t) > 0$ for a.e. $x \in \Omega$ and $t \in (0,T)$, which ends the proof.

## 5. Proof of Theorem 4

We define the relative entropy density

$$h^*(u|u^\infty) = h(u) - h(u^\infty) - h'(u^\infty)\cdot(u-u^\infty) \quad \text{for } u\in\mathbb{R}^n. \tag{55}$$

We split $h^*$ in several parts, $h^* = h_1^* + h_2^* + h_3^*$, each of which is nonnegative, where

$$h_1^*(u|u^\infty) = \sum_{i=1}^n \left( u_i \log\frac{u_i}{u_i^\infty} - u_i + u_i^\infty \right),$$

$$h_2^*(u_{n+1}|u^\infty) = \int_{u_{n+1}^\infty}^{u_{n+1}} \log\frac{q(s)}{q(u_{n+1}^\infty)} ds = \int_1^{u_{n+1}/u_{n+1}^\infty} \log\frac{q(\sigma u_{n+1}^\infty)}{q(u_{n+1}^\infty)} u_{n+1}^\infty d\sigma,$$

$$h_3^*(u|u^\infty) = \chi(u) - \chi(u^\infty) - \sum_{i=1}^n (u_i - u_i^\infty)\log p_i(u^\infty),$$

where $\chi$ is defined in (8). The entropy inequality (16) and the $L^1$ conservation of $u(t)$ give

$$\begin{aligned}\int_\Omega h^*(u(t)|u^\infty)dx + c_0\int_0^t\int_\Omega \Big( q(u_{n+1})^2 \sum_{i=1}^n |\nabla u_i^{1/2}|^2 + |\nabla q(u_{n+1})^{1/2}|^2 \Big) dx ds \\ \le \int_\Omega h^*(u^0|u^\infty)dx, \quad t>0.\end{aligned} \tag{56}$$

We prove now that the above entropy inequality, reduced to an inequality for $h_2^*$, and the convex Sobolev inequality in Lemma 11 yield exponential convergence of $u_{n+1}(t)$, while the entropy estimate for $h_1^*$ and the logarithmic Sobolev inequality allow us to conclude the convergence of $u_i(t)$ for $i=1,\ldots,n$.

*Step 1: Exponential convergence of* $u_{n+1}(t)$. Let $g(s) = \int_1^s \log q(\sigma u_{n+1}^\infty) d\sigma$ for $s\in[0,1]$. This function is convex since $g''(s) = u_{n+1}^\infty q'(s u_{n+1}^\infty)/q(s u_{n+1}^\infty) > 0$ by assumption. Again by assumption, $1/g'' = (u_{n+1}^\infty)^{-1} q/q'$ is concave. Choosing $\phi_i = 1$ in the weak formulation (15) and summing the equations from $i=1,\ldots,n$, it follows that $\int_\Omega u_{n+1}(t)/u_{n+1}^\infty dx = \int_\Omega u_{n+1}^0/u_{n+1}^\infty dx = |\Omega|$ for $t>0$, and in particular,

$$g\left( \frac{1}{|\Omega|}\int_\Omega \frac{u_{n+1}}{u_{n+1}^\infty} dx \right) = g(1) = 0.$$

Thus, we may apply the convex Sobolev inequality in the version of Lemma 11:

$$\begin{aligned}\frac{1}{|\Omega|}\int_\Omega h_2^*(u_{n+1}|u^\infty)dx &= \frac{u_{n+1}^\infty}{|\Omega|}\int_\Omega g\left(\frac{u_{n+1}}{u_{n+1}^\infty}\right)dx \le \frac{c_S u_{n+1}^\infty}{|\Omega|}\int_\Omega g''\left(\frac{u_{n+1}}{u_{n+1}^\infty}\right)\left|\nabla\frac{u_{n+1}}{u_{n+1}^\infty}\right|^2 dx \\ &= \frac{c_S}{|\Omega|}\int_\Omega \frac{q'(u_{n+1})}{q(u_{n+1})}|\nabla u_{n+1}|^2 dx.\end{aligned}$$

By assumption, $q'$ is strictly positive on $[0,1]$, i.e. $0 < q_1 \le q'(s)$ for $s\in[0,1]$, so

$$\begin{aligned}\frac{1}{|\Omega|}\int_\Omega h_2^*(u_{n+1}|u^\infty)dx &\le \frac{c_S}{q_1|\Omega|}\int_\Omega \frac{q'(u_{n+1})^2}{q(u_{n+1})}|\nabla u_{n+1}|^2 dx \\ &= \frac{4c_S}{q_1|\Omega|}\int_\Omega |\nabla q(u_{n+1})^{1/2}|^2 dx.\end{aligned}$$

Therefore, (56) yields

$$\int_\Omega h_2^*(u_{n+1}(t)|u_{n+1}^\infty)dx + \frac{c_0 q_1}{4c_S}\int_0^t\int_\Omega h_2^*(u_{n+1}(t)|u_{n+1}^\infty)dxds \le \int_\Omega h^*(u^0|u^\infty)dx,$$

and Gronwall's lemma gives

$$\int_\Omega h_2^*(u_{n+1}(t)|u_{n+1}^\infty)dx \le e^{-c_0q_1t/(4c_S)}\int_\Omega h^*(u^0|u^\infty)dx. \tag{57}$$

The strict positivity of $q'$ implies that the function $s \mapsto h_2^*(s|u_{n+1}^\infty)$ is strictly convex. Moreover, $h_2^*(u_{n+1}^\infty|u_{n+1}^\infty) = 0$ and $(h_2^*)'(u_{n+1}^\infty|u_{n+1}^\infty) = 0$. Therefore, by a Taylor expansion, $h_2^*(u_{n+1}|u_{n+1}^\infty) \ge (\gamma/2)(u_{n+1} - u_{n+1}^\infty)^2$. Inserting this inequality in (57) gives (17).

*Step 2: Convergence for* $(u_i(t))$. We assume that $q(s) \ge q_0 > 0$ for $s \in [0,1]$. It follows from the entropy inequality (56) that

$$\int_\Omega h_1^*(u(t)|u^\infty)dx + c_0q_0\int_0^t\int_{\Omega_\varepsilon}\sum_{i=1}^n|\nabla u_i^{1/2}|^2dxds \le \int_\Omega h^*(u^0|u^\infty)dx, \quad t>0.$$

We apply the logarithmic Sobolev inequality on bounded domains with constant $c_L > 0$ [13, Lemma 1],

$$\int_{\Omega_\varepsilon} h_1^*(u(t)|u^\infty)dx = \sum_{i=1}^n\int_\Omega u_i\log\frac{u_i}{u_i^\infty}dx \le c_L\sum_{i=1}^n\int_\Omega|\nabla u_i^{1/2}|^2dx.$$

Inserting this inequality into the entropy estimate gives

$$\int_\Omega h_1^*(u(t)|u^\infty)dx + \frac{c_0q_0}{c_L}\int_{\Omega_\varepsilon} h_1^*(u(t)|u^\infty)dx \le \int_\Omega h^*(u^0|u^\infty)dx, \quad t>0,$$

and then, Gronwall's lemma shows that

$$\int_\Omega h_1^*(u(t)|u^\infty)dx \le e^{-c_0q_0t/c_L}\int_\Omega h^*(u^0|u^\infty)dx, \quad t>0.$$

Finally, since $h_1^\infty(u^\infty|u^\infty) = |(h_1^*)'(u^\infty,u^\infty)| = 0$, and $\partial^2h_1^*/\partial u_i\partial u_j = \delta_{ij}/u_i \ge \delta_{ij}$ for $u \in \overline{\mathcal{D}}$, we obtain $h_1^*(u|u^\infty) \ge |u-u^\infty|^2$, which proves estimate (18) and finishes the proof.

## 6. Proof of Theorem 5

Let $u = (u_1,\dots,u_n)$ and $v = (v_1,\dots,v_n)$ be two bounded weak solutions to (1)–(2). Since $p_i \equiv 1$ for all $i = 1,\dots,n$ by assumption, (1) becomes

$$\partial_t u_i = \operatorname{div}\big(q(u_{n+1})\nabla u_i - u_i\nabla q(u_{n+1})\big), \quad i=1,\dots,n. \tag{58}$$

Summing these equations from $i=1,\dots,n$, the equation for $u_{n+1} = 1-\sum_{i=1}^n u_i$ reads as

$$\partial_t u_{n+1} = \operatorname{div}\big(q(u_{n+1})\nabla u_{n+1} + (1-u_{n+1})\nabla q(u_{n+1})\big) = \Delta Q(u_{n+1}), \tag{59}$$

where $Q(s) = \int_0^s(q(\sigma)+(1-\sigma)q'(\sigma))d\sigma$ for $0 \le s \le 1$. Furthermore, $\nabla Q(u_{n+1})\cdot\nu = 0$ on $\partial\Omega$, $t>0$ and $u_{n+1}(0) = u_{n+1}^0 := 1-\sum_{i=1}^n u_i^0$, and similar equations holds for $v_{n+1}$. Since $Q$ is a nondecreasing function, we can apply first the $H^{-1}$ method to (59) to show uniqueness for the $(n+1)$th component, i.e. $u_{n+1} = v_{n+1}$. Second, we employ the convexity of the entropy to prove that $u_i = v_i$ for $i=1,\dots,n$.

*Step 1: Uniqueness for* $u_{n+1}$. Let $t>0$ and let $\zeta(t) \in H^1(\Omega)$ be the unique solution to

$$-\Delta\zeta(t) = (u_{n+1}-v_{n+1})(t) \quad \text{in } \Omega, \quad \nabla\zeta\cdot\nu = 0 \quad \text{on } \Omega.$$

We know that $u_{n+1} - v_{n+1} \in L^2(0,T;L^2(\Omega))$. Thus, $t \mapsto \zeta(t)$ is Bochner integrable and $\zeta \in L^2(0,T;H^1(\Omega))$. As $\partial_t(u_{n+1} - v_{n+1}) \in L^2(0,T;H^1(\Omega)')$, we have even the regularity $\Delta\partial_t\zeta \in L^2(0,T;H^1(\Omega)')$. Therefore, using (59), we obtain for a.e. $t > 0$,

$$\frac{1}{2}\frac{d}{dt}\int_\Omega |\nabla\zeta|^2 dx = \langle -\Delta\partial_t\zeta, \zeta\rangle = \langle \partial_t(u_{n+1} - v_{n+1}), \zeta\rangle$$
$$= -\int_\Omega \nabla\big(Q(u_{n+1}) - Q(v_{n+1})\big)\cdot\nabla\zeta dx$$
$$= -\int_\Omega \big(Q(u_{n+1}) - Q(v_{n+1})\big)(u_{n+1} - v_{n+1})dx.$$

Here, $\langle\cdot,\cdot\rangle$ again denotes the duality pairing of $H^1(\Omega)'$ and $H^1(\Omega)$. The right-hand side is nonpositive since $Q$ is nondecreasing. This implies that

$$\int_\Omega |\nabla\zeta(t)|^2 dx \le \int_\Omega |\nabla\zeta(0)|^2 dx, \quad t > 0.$$

At time $t = 0$, $-\Delta\zeta(0) = (u_{n+1} - v_{n+1})(0) = 0$ in $\Omega$, thus $\nabla\zeta(0) = 0$. Hence, $|\nabla\zeta(t)| = 0$ a.e. in $\Omega$, which gives $(u_{n+1} - v_{n+1})(t) = -\Delta\zeta(t) = 0$ in $\Omega$.

*Step 2: Uniqueness for* $(u_1,\ldots,u_n)$. Let $0 < \varepsilon < 1$. Similarly as in [17], we introduce the distance

$$d_\varepsilon(u.v) = \sum_{i=1}^n \int_\Omega \left(\xi_\varepsilon(u_i) + \xi_\varepsilon(v_i) - 2\xi_\varepsilon\left(\frac{u_i + v_i}{2}\right)\right)dx,$$

where $\xi_\varepsilon(s) = (s+\varepsilon)(\log(s+\varepsilon) - 1) + 1,\ s \ge 0$.

As $\xi_\varepsilon$ is convex, we have $\xi_\varepsilon(u_i) + \xi_\varepsilon(v_i) - 2\xi_\varepsilon((u_i + v_i)/2) \ge 0$ in $\Omega$ and hence, $d_\varepsilon(u_i, v_i) \ge 0$. We need the regularization $\varepsilon > 0$ since $u_i$ and $v_i$ are only nonnegative and thus, expressions like $\log((u_i + v_i)/2)$ may be undefined. Since $u_{n+1} = v_{n+1}$ by Step 1, we may abbreviate $q := q(u_{n+1}) = q(v_{n+1})$. Then, using (58), we compute

$$\frac{d}{dt}d_\varepsilon(u,v) = \sum_{i=1}^n \bigg(\langle\partial_t u_i, \log(u_i+\varepsilon)\rangle + \langle\partial_t v_i, \log(v_i+\varepsilon)\rangle$$
$$- \left\langle \partial_t(u_i + v_i), \log\left(\frac{u_i + v_i}{2} + \varepsilon\right)\right\rangle\bigg)$$
$$= -\sum_{i=1}^n \int_\Omega \bigg((q\nabla u_i - u_i\nabla q)\cdot\frac{\nabla u_i}{u_i + \varepsilon} + (q\nabla v_i - v_i\nabla q)\cdot\frac{\nabla v_i}{v_i+\varepsilon}$$
$$- \big(q\nabla(u_i+v_i) - (u_i+v_i)\nabla q\big)\cdot\frac{\nabla(u_i+v_i)}{u_i+v_i+2\varepsilon}\bigg)dx.$$

Rearranging the terms, we arrive at

$$\frac{d}{dt}d_\varepsilon(u,v) = -\sum_{i=1}^n\int_\Omega\left(\frac{|\nabla u_i|^2}{u_i+\varepsilon} + \frac{|\nabla v_i|^2}{v_i+\varepsilon} - \frac{|\nabla(u_i+v_i)|^2}{u_i+v_i+2\varepsilon}\right)q dx \tag{60}$$
$$+ \sum_{i=1}^n\int_\Omega\left(\frac{u_i}{u_i+\varepsilon} - \frac{u_i+v_i}{u_i+v_i+2\varepsilon}\right)\nabla q\cdot\nabla u_i dx$$
$$+ \sum_{i=1}^n\int_\Omega\left(\frac{v_i}{v_i+\varepsilon} - \frac{u_i+v_i}{u_i+v_i+2\varepsilon}\right)\nabla q\cdot\nabla v_i dx$$

$$
\begin{aligned}
&= -4\sum_{i=1}^n \int_\Omega \Big(|\nabla\sqrt{u_i+\varepsilon}|^2 + |\nabla\sqrt{v_i+\varepsilon}|^2 - |\nabla\sqrt{u_i+v_i+2\varepsilon}|^2\Big) q dx \\
&\quad + 2\sum_{i=1}^n \int_\Omega \left(\frac{u_i}{u_i+\varepsilon} - \frac{u_i+v_i}{u_i+v_i+2\varepsilon}\right)\sqrt{q}\nabla\sqrt{q}\cdot\nabla u_i dx \\
&\quad + 2\sum_{i=1}^n \int_\Omega \left(\frac{v_i}{v_i+\varepsilon} - \frac{u_i+v_i}{u_i+v_i+2\varepsilon}\right)\sqrt{q}\nabla\sqrt{q}\cdot\nabla v_i dx.
\end{aligned}
$$

Now, we apply Lemma 10 with $d\mu = qdx$ and $f = u_i + \varepsilon$, $g = v_i + \varepsilon$, showing that the first integral on the right-hand side is nonnegative. We observe that $d_\varepsilon(u(0), v(0)) = 0$ as $u$ and $v$ have the same initial data. Thus, integrating (62) in time, we obtain

$$
\begin{aligned}
d_\varepsilon(u(t), v(t)) \le 2\sum_{i=1}^n \int_0^t \int_\Omega \left(\frac{u_i}{u_i+\varepsilon} - \frac{u_i+v_i}{u_i+v_i+2\varepsilon}\right)\sqrt{q}\nabla\sqrt{q}\cdot\nabla u_i dx \\
+ 2\sum_{i=1}^n \int_0^t \int_\Omega \left(\frac{v_i}{v_i+\varepsilon} - \frac{u_i+v_i}{u_i+v_i+2\varepsilon}\right)\sqrt{q}\nabla\sqrt{q}\cdot\nabla v_i dx.
\end{aligned} \tag{61}
$$

Since $\nabla\sqrt{q}, \sqrt{q}\nabla u_i, \sqrt{q}\nabla v_i \in L^2(0,T;H^1(\Omega))$ and

$$
\left|\frac{u_i}{u_i+\varepsilon}\right| \le 1, \quad \left|\frac{v_i}{v_i+\varepsilon}\right| \le 1, \quad \left|\frac{u_i+v_i}{u_i+v_i+2\varepsilon}\right| \le 1,
$$

the dominated convergence implies that the right-hand side of (61) tends to zero as $\varepsilon \to 0$. From the nonnegativity of $d_\varepsilon$ we deduce that $d_\varepsilon(u(t), v(t)) \to 0$ as $\varepsilon \to 0$, which means that

$$
\xi_\varepsilon(u_i) + \xi_\varepsilon(v_i) - 2\xi_\varepsilon\left(\frac{u_i+v_i}{2}\right) \to 0 \quad \text{as } \varepsilon \to 0 \quad \text{a.e. in } \Omega\times(0,\infty). \tag{62}
$$

According to Taylor's formula, there are functions $\theta_\varepsilon, \eta_\varepsilon : \Omega\times(0,\infty)$ such that

$$
\begin{aligned}
&\xi_\varepsilon(u_i) = \xi_\varepsilon\left(\frac{u_i+v_i}{2} + \frac{u_i-v_i}{2}\right) \\
&= \xi_\varepsilon\left(\frac{u_i+v_i}{2}\right) + \xi_\varepsilon'\left(\frac{u_i+v_i}{2}\right)\frac{u_i-v_i}{2} + \frac{1}{2}\xi_\varepsilon''\left(\theta_\varepsilon\frac{u_i+v_i}{2} + (1-\theta_\varepsilon)u_i\right)\left(\frac{u_i-v_i}{2}\right)^2, \\
&\xi_\varepsilon(v_i) = \xi_\varepsilon\left(\frac{u_i+v_i}{2} - \frac{u_i-v_i}{2}\right) \\
&= \xi_\varepsilon\left(\frac{u_i+v_i}{2}\right) - \xi_\varepsilon'\left(\frac{u_i+v_i}{2}\right)\frac{u_i-v_i}{2} + \frac{1}{2}\xi_\varepsilon''\left(\eta_\varepsilon\frac{u_i+v_i}{2} + (1-\eta_\varepsilon)v_i\right)\left(\frac{u_i-v_i}{2}\right)^2.
\end{aligned}
$$

Adding these identities and employing the estimate $\xi_\varepsilon''(s) = (s+\varepsilon)^{-1} \ge 1/2$ for $0 \le s \le 1$, we infer that

$$
\xi_\varepsilon(u_i) + \xi_\varepsilon(v_i) - 2\xi_\varepsilon\left(\frac{u_i+v_i}{2}\right) \ge \frac{1}{8}(u_i - v_i)^2.
$$

This estimate and (62) prove that $u_i = v_i$ in $\Omega\times(0,\infty)$ for $i = 1, \dots, n$.

**Remark 14.** The uniqueness proof provides the continuous dependence on the initial data with respect to the distance $d$ under the condition that $u_{n+1} = v_{n+1}$, which means that the proportion of the unoccupied space is fixed. Indeed, integrating (60) in time, we obtain (compare to (61))

$$d_\varepsilon(u(t),v(t)) \le d_\varepsilon(u(0),v(0)) + 2\sum_{i=1}^n \int_0^t \int_\Omega \left(\frac{u_i}{u_i+\varepsilon} - \frac{u_i+v_i}{u_i+v_i+2\varepsilon}\right)\sqrt{q}\nabla\sqrt{q}\cdot\nabla u_i dx$$
$$+ 2\sum_{i=1}^n \int_0^t \int_\Omega \left(\frac{v_i}{v_i+\varepsilon} - \frac{u_i+v_i}{u_i+v_i+2\varepsilon}\right)\sqrt{q}\nabla\sqrt{q}\cdot\nabla v_i dx.$$

With the same arguments as in the uniqueness proof, we may pass to the limit $\varepsilon \to 0$, which yields

$$d(u(t),v(t)) \le d(u(0),v(0)), \quad t \ge 0.$$

The distance can be bounded from below by the $L^2$ norm but not from above. This is only possible when positive lower bounds for $u_i$ and $v_i$ are available. □

## 7. Extensions

In this section, we discuss some extensions of the diffusion system (1).

*Reaction terms.* Cross-diffusion systems with reaction terms,

$$\partial_t u - \operatorname{div}(A(u)\nabla u) = f(u) \quad \text{in } \Omega,\ t > 0, \tag{63}$$

can be treated similarly as in [21]. More precisely, if there is a constant $c_f > 0$ such that $f(u)\cdot h'(u) \le c_f(1+h(u))$ for all $u \in \mathscr{D}$, then there exists a global weak solution to (2) and (63). The proof proceeds as for Theorem 1, where the right-hand side of the entropy inequality (33) has to be replaced by

$$\int_\Omega h(u^0)dx + \tau\int_\Omega f(u)\cdot h'(u)dx \le \int_\Omega h(u^0)dx + \tau c_f \int_\Omega (1+h(u))dx.$$

Then, for sufficiently small $\tau > 0$, the integral $\tau c_f \int_\Omega h(u)dx$ can be absorbed by the left-hand side of (33). For instance, reaction terms of Lotka–Volterra type

$$f_i(u) = u_i\left(1 - \sum_{j=1}^n s_{ij}u_j\right), \quad i = 1,\dots,n,\ s_{ij} \ge 0,$$

are admissible. The large-time behavior result is valid only under an additional condition on $f(u)$, namely $f(u)\cdot h'(u) \le 0$ for $u \in \mathscr{D}$. If we suppose conservation of the "total mass", i.e. $\sum_{i=1}^n f_i(u) = 0$, the $H^{-1}$ method allows us to prove uniqueness for $u_{n+1}$. Uniqueness for the remaining components $u_i$ follows if there exists $C > 0$ such that for all $u_i$ and $v_i$,

$$\sum_{i=1}^n \left(f_i(u)\log\frac{2u_i}{u_i+v_i} + f_i(v)\log\frac{2v_i}{u_i+v_i}\right) \le C\sum_{i=1}^n \left(\xi(u_i)+\xi(v_i) - 2\xi\left(\frac{u_i+v_i}{2}\right)\right),$$

where $\xi(s) = s(\log s - 1) + 1$ (see the proof of Theorem 5). More general conditions on $f(u)$ can be found in [18].

*Drift terms.* In the presence of environmental or electric potentials or of chemotactic signal concentrations, the diffusion system contains additional drift terms,

$$\partial_t u - \operatorname{div}(A(u)\nabla u + D(u)\nabla\phi) = 0 \quad \text{in } \Omega,\ t > 0, \tag{64}$$

where $D(u) = (D_{ij}(u))$ is an $n\times n$ matrix and the $i$th component of $D(u)\nabla\phi$ is given by $\sum_{j=1}^n D_{ij}(u)\nabla\phi_j$, where $\phi_j = \phi_j(x)$ is some potential. Assume that $h$ is such that $\nabla u : h''(u)A(u)\nabla u \ge \sum_{i=1}^n g_i(u)|\nabla u_i|^2$ for some nonnegative functions $g_i(u)$. Then, using the test function $h'(u)$ in the weak formulation of (64), we compute

$$\frac{d}{dt}\int_\Omega h(u)dx = -\int_\Omega \nabla u : h''(u)A(u)\nabla u dx - \int_\Omega \nabla u : h''(u)D(u)\nabla\phi dx$$
$$\le -\frac{1}{2}\sum_{i=1}^n \int_\Omega g_i(u)|\nabla u_i|^2 dx + \frac{1}{2}\sum_{k=1}^n \int_\Omega G_k(u)|\nabla\phi_k|^2 dx,$$

where we employed the Cauchy–Schwarz inequality and have set $G_k(u) = \sum_{i,j=1}^n g_i(u)^{-1}(H_{ij})^2(D_{jk}(u))^2$ with $H = h''(u)$. Thus, if $\nabla\phi_i$ is bounded in $L^2$ and $G_k(u)$ in $L^\infty$, we achieve some gradient estimates, which are the basis for the existence analysis. An example is the ion-transport model [8]

$$A_{ij}(u) = u_i \quad \text{for } i \neq j, \quad A_{ii}(u) = u_i + u_{n+1}, \quad D_{ij}(u) = u_i u_{n+1}\delta_{ij}$$

for $i, j = 1, \ldots, n$. The entropy density can be defined by

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

Then the Hessian $h''(u)$ does not depend on $\phi_i$. A formal computation, using the Cauchy–Schwarz inequality and the identity $\sum_{i=1}^n \nabla u_i = -\nabla u_{n+1}$, gives

$$\begin{aligned}
\frac{d}{dt}\int_\Omega h(u)dx &= -\sum_{i=1}^n \int_\Omega u_i u_{n+1}\left|\nabla\left(\log\frac{u_i}{u_{n+1}} + \phi_i\right)\right|^2 dx \\
&\le -\sum_{i=1}^n \int_\Omega u_i u_{n+1}\left(\frac{1}{2}\left|\nabla\log\frac{u_i}{u_{n+1}}\right|^2 - |\nabla\phi_i|^2\right)dx \\
&= -\sum_{i=1}^n \int_\Omega \big(2u_{n+1}|\nabla u_i^{1/2}|^2 + |\nabla u_{n+1}|^2 + 2|\nabla u_{n+1}^{1/2}|^2\big)dx + \sum_{i=1}^n \int_\Omega u_i u_{n+1}|\nabla\phi_i|^2 dx.
\end{aligned}$$

As $q(s) = s$ in this model, we find the same estimates as in the proof of Theorem 1 (also see [7, Section 3.2]). This shows that our strategy can be adapted to cross-diffusion systems with drift.

*Other diffusion coefficients.* Our main assumption on the transition rates is that they are given by the product of $p_i(u)$ and $q_i(u_{n+1})$ (see Appendix A). Also other choices are possible. An example is the diffusion system of [29], which is derived from a stochastic lattice model by assuming that the transition rates are given by $p_i(u) + q_i(u_{n+1})$ for some special functions $p_i$ and $q_i$. The diffusion matrix has the structure

$$A(u) = \begin{pmatrix} \alpha_1(1-u_2) + u_2 & (\alpha_1 - 1)u_1 \\ (\alpha_2 - 1)u_2 & \alpha_2(1-u_1) + u_1 \end{pmatrix},$$

where $\alpha_1, \alpha_2 > 0$. The corresponding diffusion system possesses the entropy density

$$h(u) = \sum_{i=1}^2 u_i(\log u_i - 1) + (1 - u_1 - u_2)(\log(1 - u_1 - u_2) - 1), \quad u = (u_1, u_2) \in \mathscr{D},$$

and the new diffusion matrix $B = h''(u)^{-1}A(u)$, given by

$$B = \begin{pmatrix} (\alpha_1(1 - u_1 - u_2) + u_2)u_1 & -u_1u_2 \\ -u_1u_2 & (\alpha_2(1 - u_1 - u_2) + u_1)u_2 \end{pmatrix},$$

is symmetric and positive semi-definite on $\overline{\mathscr{D}}$. For our analysis, we need bounds from $h''(u)A(u)$ (see Lemma 7), which are less obvious since

$$\begin{aligned}
\nabla u_1^\top h''(u)A(u)\nabla u_2 = {}& \frac{a_1|(1-u_2)\nabla u_1 + u_1\nabla u_2|^2}{u_1(1-u_1-u_2)} + \frac{a_2|u_2\nabla u_1 + (1-u_1)\nabla u_2|^2}{u_2(1-u_1-u_2)} \\
&+ 4|\nabla\sqrt{u_1u_2}|^2,
\end{aligned}$$

only yielding an $L^2$ bound for $\nabla\sqrt{u_1u_2}$ in $L^2$.

## Conflict of interest statement

There is no conflict of interest.

## Appendix A. Formal derivation of the $n$-species population model

We derive formally the cross-diffusion system (1) from a master equation for a discrete-space random walk in the diffusion limit. We consider random walks on a one-dimensional lattice only, since the derivation can be extended in a straightforward manner to the higher-dimensional situation. The lattice is given by cells $x_j$ ($j \in \mathbb{Z}$) with the uniform cell distance $h = x_j - x_{j-1} > 0$. The proportions of the $i$th population in the $j$th cell at time $t > 0$ is denoted by $u_i(x_j) = u_i(x_j, t)$. The species move from the $j$th cell into the neighboring cells $j \pm 1$ with the transition rates $T_i^{j,\pm}$. The master equations are given by

$$\partial_t u_i(x_j) = T_i^{j-1,+} u_i(x_{j-1}) + T_i^{j+1,-} u_i(x_{j+1}) - (T_i^{j,+} + T_j^{i,-}) u_i(x_j), \quad i = 1, \ldots, n,$$

and the transition rates are defined as

$$T_i^{j,\pm} = \sigma_0 p_i(u(x_j)) q_i(u_{n+1}(x_j)), \quad u_{n+1}(x_j) = 1 - \sum_{k=1}^{n} u_k(x_j), \tag{65}$$

where $u = (u_1, \ldots, u_n)$. The quantities $p_i(u(x_j))$ and $q_i(u_{n+1}(x_{j\pm1}))$ measure the tendency of the species $i$ to leave the $j$th cell or to move into the $j$th cell from one of the neighboring cells, respectively. More precisely, $u_i(x_j)$ denotes a volume fraction of occupancy and $u_{n+1}$ the volume fraction not occupied by the species. Our assumption is that the transition rates, measuring the occupancy and the non-occupancy, separate, resulting in the product of $p_i$ and $q_i$. Other choices are possible (see [29] for an example), but the analytical treatment of the corresponding diffusion systems is not obvious.

For the derivation of the diffusion model, it is convenient to introduce the following abbreviations:

$$p_i^j = p_i(u_1(x_j), \ldots, u_n(x_j)), \quad q_i^j = q_i(u_{n+1}(x_j)),$$
$$\partial_k p_i^j = \frac{\partial p_i}{\partial u_k}(u_1(x_j), \ldots, u_n(x_j)), \quad \partial q_i^j = q_i'(u_{n+1}(x_j)).$$

Thus, we can rewrite the master equation as

$$\sigma_0^{-1} \partial_t u_i^j = q_i^j (p_i^{j-1} u_i^{j-1} + p_i^{j+1} u_i^{j+1}) - p_i^j u_i^j (q_i^{j+1} + q_i^{j-1}). \tag{66}$$

Set $D = \partial_x$. We compute the Taylor expansions of $p_i$ and $q_i$ ($i = 1, \ldots, n$) and replace $u_k^{j\pm1} - u_k^j$ by the Taylor expansion $\pm h D u_k^j + \frac{1}{2} h^2 D^2 u_i^j + O(h^3)$. Then, collecting all terms up to order $O(h^2)$, we arrive at

$$p_i^{j\pm1} = p_i^j + h \sum_{k=1}^{n} \partial_k p_i^j D u_k^j + \frac{h^2}{2} \left( \sum_{k=1}^{n} \partial_k p_i^j D^2 u_k^j + \sum_{k,\ell=1}^{n} \partial_{k\ell}^2 p_i^j D u_k^j D u_\ell^j \right) + O(h^3),$$

$$\begin{aligned} q_i^{j\pm1} &= q_i^j \pm h \partial p_i^j D u_{n+1}^j + \frac{h^2}{2} \big( \partial q_i^j D^2 u_{n+1}^j + \partial^2 q_i^j (D u_{n+1}^j)^2 \big) + O(h^3) \\ &= q_i^j \mp h \partial q_i^j \sum_{k=1}^{n} D u_k^j + \frac{h^2}{2} \left( -\partial q_i^j \sum_{k=1}^{n} D^2 u_k^j + \partial^2 q_i^j \sum_{k,\ell=1}^{n} D u_k^j D u_\ell^j \right) + O(h^3). \end{aligned}$$

In the last step, we have used $u_{n+1} = 1 - \sum_{k=1}^{n} u_k$. We insert these expressions into (66) and rearrange the terms. It turns out that the terms of order $O(1)$ and $O(h)$ cancel, and we end up with

$$\begin{aligned} \sigma_0^{-1} h^{-2} \partial_t u_i^j &= \sum_{k=1}^{n} D^2 u_k^j (q_i^j p_i^j \delta_{ik} + q_i^j u_i^j \partial_k p_i^j + p_i^j u_i^j \partial q_i^j) \\ &\quad + \sum_{k,\ell=1}^{n} D u_k^j D u_\ell^j (2 q_i^j \partial_k p_i^j \delta_{i\ell} + q_i^j u_i^j \partial_{k\ell}^2 p_i^j - p_i^j u_i^j \partial^2 q_i^j). \end{aligned}$$

We choose $\sigma_0 = h^{-2}$ and pass to the limit $h \to 0$:

$$\partial_t u_i = \sum_{k=1}^n D^2 u_k \left( q_i p_i \delta_{ik} + q_i u_i \frac{\partial p_i}{\partial u_k} + p_i u_i q_i' \right)$$
$$+ \sum_{k,\ell=1}^n D u_k D u_\ell \left( 2q_i \frac{\partial p_i}{\partial u_k} \delta_{i\ell} + q_i u_i \frac{\partial^2 p_i}{\partial u_k \partial u_\ell} - p_i u_i q_i'' \right).$$

A lengthy but straightforward computation shows that the last sum equals

$$\sum_{k=1}^n D u_k D \left( q_i p_i \delta_{ik} + q_i u_i \frac{\partial p_i}{\partial u_k} + p_i u_i q_i' \right),$$

and we end up with

$$\partial_t u_i = D \sum_{k=1}^n D u_k \left( q_i p_i \delta_{ik} + q_i u_i \frac{\partial p_i}{\partial u_k} + p_i u_i q_i' \right),$$

which is the one-dimensional version of (1).

## References


[1] H. Amann, Dynamic theory of quasilinear parabolic systems. III. Global existence, Math. Z. 202 (1989) 219–250.
[2] A. Arnold, P. Markowich, G. Toscani, A. Unterreiter, On convex Sobolev inequalities and the rate of convergence to equilibrium for Fokker–Planck type equations, Commun. Partial Differ. Equ. 26 (2001) 43–100.
[3] D. Bakry, M. Emery, Diffusions hypercontractives, in: Séminaire de probabilités XIX, 1983/1984, in: Lect. Notes Math., vol. 1123, Springer, Berlin, 1985, pp. 177–206.
[4] D. Bakry, I. Gentil, M. Ledoux, Analysis and Geometry of Markov Diffusion Operators, Springer, Cham, 2014.
[5] M. Bendahmane, T. Lepoutre, A. Marrocco, B. Perthame, Conservative cross diffusions and pattern formation through relaxation, J. Math. Pures Appl. 92 (2009) 651–667.
[6] H. Brézis, Functional Analysis, Sobolev Spaces and Partial Differential Equations, Springer, New York, 2011.
[7] M. Burger, M. Di Francesco, J.-F. Pietschmann, B. Schlake, Nonlinear cross-diffusion with size exclusion, SIAM J. Math. Anal. 42 (2010) 2842–2871.
[8] M. Burger, B. Schlake, W.-T. Wolfram, Nonlinear Poisson–Nernst–Planck equations for ion flux through confined geometries, Nonlinearity 25 (2012) 961–990.
[9] L. Chen, A. Jüngel, Analysis of a multi-dimensional parabolic population model with strong cross-diffusion, SIAM J. Math. Anal. 36 (2004) 301–322.
[10] L. Chen, A. Jüngel, Analysis of a parabolic cross-diffusion population model without self-diffusion, J. Differ. Equ. 224 (2006) 39–59.
[11] X. Chen, A. Jüngel, J.-G. Liu, A note on Aubin–Lions–Dubinskii lemmas, Acta Appl. Math. 133 (2014) 33–43.
[12] L. Desvillettes, K. Fellner, Exponential decay toward equilibrium via entropy methods for reaction–diffusion equations, J. Math. Anal. Appl. 319 (2006) 157–176.
[13] L. Desvillettes, K. Fellner, Exponential convergence to equilibrium for a nonlinear reaction–diffusion systems arising in reversible chemistry, in: System Modelling and Optimization, Proceedings of the IFIP TC 7 Conference 2013, in: Adv. Inform. Commun. Techn., vol. 443, 2014, pp. 96–104.
[14] L. Desvillettes, T. Lepoutre, A. Moussa, Entropy, duality and cross diffusion, SIAM J. Math. Anal. 46 (2014) 820–853.
[15] L. Desvillettes, T. Lepoutre, A. Moussa, A. Trescases, On the entropic structure of reaction–cross diffusion systems, Preprint, arXiv:1410.7377, 2014.
[16] M. Dreher, A. Jüngel, Compact families of piecewise constant functions in $L^p(0, T; B)$, Nonlinear Anal. 75 (2012) 3072–3077.
[17] H. Gajewski, On a variant of monotonicity and its application to differential equations, Nonlinear Anal. 22 (1994) 73–80.
[18] H. Gajewski, I. Skrypnik, On the uniqueness problem for nonlinear parabolic equations, Discrete Contin. Dyn. Syst. 10 (2004) 315–336.
[19] G. Galiano, M. Garzón, A. Jüngel, Semi-discretization in time and numerical convergence of solutions of a nonlinear cross-diffusion population model, Numer. Math. 93 (2003) 655–673.
[20] A. Glitzky, K. Gröger, R. Hünlich, Free energy and dissipation rate for reaction–diffusion processes of electrically charged species, Appl. Anal. 60 (1996) 201–217.
[21] A. Jüngel, The boundedness-by-entropy principle for cross-diffusion systems, Nonlinearity 28 (2015) 1963–2001.
[22] J. Kim, Smooth solutions to a quasi-linear system of diffusion equations for a certain population model, Nonlinear Anal. 8 (1984) 1121–1144.
[23] D. Le, Cross diffusion systems in $n$ spatial dimensional domains, Indiana Univ. Math. J. 51 (2002) 625–643.
[24] M. Liero, A. Mielke, Gradient structures and geodesic convexity for reaction–diffusion systems, Philos. Trans. R. Soc. A 371 (2005) 20120346 (28 pages) 2013.
[25] A. Moussa, Some variants of the classical Aubin–Lions lemma, Preprint, arXiv:1401.7231v3, 2014.
[26] H. Murakawa, A relation between cross-diffusion and reaction–diffusion, Discrete Contin. Dyn. Syst., Ser. S 5 (2012) 147–158.

[27] H. Murakawa, Error estimates for discrete-time approximations of nonlinear cross-diffusion systems, SIAM J. Numer. Anal. 52 (2014) 955–974.

[28] F. Pacard, A. Unterreiter, A variational analysis of the thermal equilibrium state of charged quantum fluids, Commun. Partial Differ. Equ. 20 (1995) 885–900.

[29] K. Painter, Continuous models for cell migration in tissues and applications to cell sorting via differential chemotaxis, Bull. Math. Biol. 71 (2009) 1117–1147.

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

[31] J. Simon, Compact sets in the space $L^p(0, T; B)$, Ann. Math. Pura Appl. 146 (1987) 65–96.

[32] M. Simpson, K. Landman, B. Hughes, Multi-species simple exclusion processes, Physica A 388 (2009) 399–406.

[33] J. Zinsl, D. Matthes, Transport distances and geodesic convexity for systems of degenerate diffusion equations, to appear in Calc. Var. Partial Differ. Equ. (2015).