CORE Metadata, citation and similar papers at core.ac.uk
Provided by MUCC (Crossref)

© 2017 The Author(s).
This article is published with open access at Springerlink.com
0044-2275/17/010001-15
published online January 20, 2017
DOI 10.1007/s00033-017-0772-1

Zeitschrift für angewandte Mathematik und Physik ZAMP

A cross-diffusion system derived from a Fokker–Planck equation with partial averaging

Ansgar Jüngel and Nicola Zamponi

Abstract. A cross-diffusion system for two components with a Laplacian structure is analyzed on the multi-dimensional torus. This system, which was recently suggested by P.-L. Lions, is formally derived from a Fokker–Planck equation for the probability density associated with a multi-dimensional Itō process, assuming that the diffusion coefficients depend on partial averages of the probability density with exponential weights. A main feature is that the diffusion matrix of the limiting cross-diffusion system is generally neither symmetric nor positive definite, but its structure allows for the use of entropy methods. The global-in-time existence of positive weak solutions is proved and, under a simplifying assumption, the large-time asymptotics is investigated.

Mathematics Subject Classification. 35K45, 35K65, 35Q84.

Keywords. Cross-diffusion system, Fokker–Planck equation, Entropy methods, Global existence of weak solutions, Large-time asymptotics, Positivity of solutions.

1. Introduction

The aim of this paper is the analysis of the following cross-diffusion system

$$\partial_t u_i = \Delta\big(a(u_1/u_2)u_i\big) + \mu_i u_i, \quad t > 0, \quad u_i(0) = u_i^0 \geq 0 \quad \text{in } \mathbb{T}^d, \quad i = 1, 2, \tag{1}$$

where $\mathbb{T}^d$ is the d-dimensional torus with $d \geq 1, a : (0,\infty) \to (0,\infty)$ is a continuously differentiable function, and $\mu_i \in \mathbb{R}$. This system can be formally derived [7] from a $(d+1)$-dimensional Fokker–Planck equation for the probability density $f(x, y, t)$, where $x \in \mathbb{R}^d, y \in \mathbb{R}$. The function u_i is obtained from f by partial averaging,

$$u_i(x,t) = \int_{\mathbb{R}} f(x,y,t)e^{\lambda_i y}dy, \quad i = 1, 2,$$

μ_i is a function of λ_i, and $a(u_1/u_2)$ is related to the diffusion coefficients in the Fokker–Planck equation. Strictly speaking, Eq. (1) holds in $\mathbb{R}^d$ (or on some subset of $\mathbb{R}^d$), but we consider this equation on the torus for the sake of simplicity (and to avoid possible issues with boundary conditions). For details on the derivation, we refer to Sect. 2.

System (1) has been suggested by Lions [7], and the global-in-time existence of (weak) solutions has been identified as an open problem. In this paper, we solve this problem by applying the entropy method for diffusive equations.

The underlying Fokker–Planck equation for $f(x, y, t)$ models the time evolution of the value of a financial product in an idealized financial market, depending on various underlying assets or economic values. The function u_i is an average with respect to the variable y, which may be interpreted as the value

The authors acknowledge partial support from the Austrian Science Fund (FWF), Grants P22108, P24304, and W1245, and from the Austrian-French Project Amadeé of the Austrian Exchange Service (ÖAD), Grant FR 04/2016.

Birkhäuser

of an economic parameter, and the exponential weight emphasizes large positive or large negative values of y, depending on the sign of λ_i. We note that partial averaging is also employed to simplify chemical master equations [9]. Here, we are not interested in potential applications, but more in the refinement of mathematical tools to analyze (1).

We assume that there exist $a_0 > 0$ and $p \geq 0$ such that for all $r > 0$,

$$a(r) \geq r|a'(r)|, \quad a(r) \geq \frac{a_0}{r^p + r^{-p}}. \tag{2}$$

The first condition means that a grows at most linearly (see Lemma 6). The second condition is a technical assumption needed for the entropy method (see the proof of Lemma 5). Examples are $a(r) = 1$, which leads to uncoupled heat equations for u_1 and $u_2, a(r) = r^\alpha$ with $0 < \alpha \leq 1, a(r) = r^\beta/(1 + r^{\beta-1})$ with $\beta > 0$, and $a(r) = 1/r$. The last example gives the equations

$$\partial_t u_1 = \Delta u_2, \quad \partial_t u_2 = \Delta\left(\frac{u_2^2}{u_1}\right). \tag{3}$$

Surprisingly, this system corresponds (up to a factor) to an energy-transport model for semiconductors. Indeed, introducing the electron density $n := u_1$ and the electron temperature $\theta := u_2/u_1$, Eq. (3) can be written as

$$\partial_t n = \Delta(n\theta), \quad \partial_t(n\theta) = \Delta(n\theta^2).$$

A class of energy-transport models that includes the above example was analyzed in [13].

Another class of models which resembles (1) are the equations

$$\partial_t u_i = \Delta(p_i(u)u_i), \quad i = 1, \ldots, m, \tag{4}$$

modeling the time evolution of population densities u_i. These systems are analyzed in, e.g., [5,8], essentially for $m = 2$. In this application, p_i is often given by the sum $p_{i1}(u_1) + p_{i2}(u_2)$, and consequently, the results of [5,8] do not apply and we need to develop new ideas.

Our first main result is the global-in-time existence of weak solutions to (1).

Theorem 1. (Existence of weak solutions) *Let* (2) *hold and let* $T > 0, \alpha \geq p + 4$, $\mu_1, \mu_2 \in \mathbb{R}$, $0 \leq a \in C^1(0, \infty)$, $u^0 = (u_1^0, u_2^0) \in L^2(\mathbb{T}^d)^2$ *with* $u_1^0, u_2^0 \geq 0$ *in* $\mathbb{T}^d$ *and* $H[u^0] < \infty$. *Then, there exists a solution* $u = (u_1, u_2)$ *to* (1) *satisfying* $u_i > 0$ *in* $\mathbb{T}^d, t > 0, i = 1, 2$, *and*

$$\begin{aligned} &u_i,\ a(u_1/u_2)u_i \in L^\infty(0, T; L^2(\mathbb{T}^d)), \\ &\quad \nabla u_i,\ \nabla\big(a(u_1/u_2)u_i\big) \in L^2(0, T; L^2(\mathbb{T}^d)), \quad \partial_t u_i \in L^2(0, T; H^1(\mathbb{T}^d)'), \quad i = 1, 2. \end{aligned}$$

If additionally $\mu_i \leq 0$ *for* $i = 1, 2$, *we have the uniform bounds*

$$u_i,\ a(u_1/u_2)u_i \in L^\infty(0, \infty; L^2(\mathbb{T}^d)), \quad \nabla u_i,\ \nabla\big(a(u_1/u_2)u_i\big) \in L^2(0, \infty; L^2(\mathbb{T}^d)). \tag{5}$$

As mentioned above, the proof of this theorem is based on entropy methods. These methods have been originally developed to understand the large-time behavior of solutions; see, e.g., [2,12]. The "entropy" of system (1) is often understood as a convex Lyapunov functional which provides suitable nonlinear gradient estimates. In many situations, and also in the financial context presented here, the "entropy" has no physical counterpart. However, we claim that this notion is appropriate since it naturally generalizes physical situations. For details, we refer to [8].

Our key idea is to employ the functional

$$H[u] = \int_{\mathbb{T}^d} h(u)\mathrm{d}x, \quad h(u) = \left(\frac{u_1}{u_2}\right)^\alpha u_1^2 + \left(\frac{u_1}{u_2}\right)^{-\alpha} u_2^2 + u_1 - \log u_1 + u_2 - \log u_2, \tag{6}$$

where $\alpha \geq p+4$ and $u = (u_1, u_2) \in (0, \infty)^2$. We will show that

$$\frac{\mathrm{d}}{\mathrm{d}t}H[u] + \int_{\mathbb{T}^d} \left(\left(\frac{u_1}{u_2}\right)^{\alpha-p} + \left(\frac{u_1}{u_2}\right)^{p-\alpha} \right) \left(|\nabla u_1|^2 + |\nabla u_2|^2\right) \mathrm{d}x \leq CH[u] \tag{7}$$

for some constant $C > 0$ which vanishes if $\mu_1 = \mu_2 = 0$. In this situation, the mapping $t \mapsto H[u(t)]$ is nonincreasing; otherwise, for $\mu_i \neq 0, t \mapsto H[u(t)]$ is bounded on finite time intervals. We infer from the inequality $x + x^{-1} \geq 2$ for all $x > 0$ uniform bounds for $u_i(t)$ in $H^1(\mathbb{T}^d)$, which are needed for the compactness argument.

The entropy method gives more than just the a priori estimate (7). Indeed, let us write (1) in divergence form:

$$\partial_t u - \operatorname{div}(A(u)\nabla u) = f(u), \quad t > 0, \quad u(0) = u^0 \quad \text{in } \mathbb{T}^d,$$

where the ith component of $\operatorname{div}(A(u)\nabla u)$ equals $\sum_{j=1}^d \sum_{k=1}^2 \partial_j(A_{ik}(u)\partial_j u_k), \partial_j = \partial/\partial x_j$, and $f(u) = (\mu_1 u_1, \mu_2 u_2)^\top$. The diffusion matrix

$$A(u) = \begin{pmatrix} a(u_1/u_2) + (u_1/u_2)a'(u_1/u_2) & -(u_1/u_2)^2 a'(u_1/u_2) \\ a'(u_1/u_2) & a(u_1/u_2) - (u_1/u_2)a'(u_1/u_2) \end{pmatrix} \tag{8}$$

is generally neither symmetric nor positive definite. Since the only eigenvalue of $A(u)$ is given by $\lambda = a(u_1/u_2) > 0$, the system is normally elliptic [1], and local-in-time existence of classical solutions can be expected. The difficulty is to prove the global-in-time existence. The entropy density $h(u)$ allows us to formulate (1) in new variables with a positive semidefinite diffusion matrix. Then, together with the a priori estimates from (7), global existence will be deduced. Indeed, defining the so-called entropy variable $w = (w_1, w_2)$ by $w_i = \partial h/\partial u_i$ $(i = 1, 2)$, Eq. (1) is equivalent to

$$\partial_t u - \operatorname{div}(B(w)\nabla w) = f(u), \quad t > 0, \quad u(0) = u^0 \quad \text{in } \mathbb{T}^d, \tag{9}$$

where $B(w) = A(u)h''(u)^{-1}$ is positive semidefinite (see Lemma 5) and $h''(u)$ is the Hessian matrix of $h(u)$. With this formulation, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}H[u] + \int_{\mathbb{T}^d} \nabla u : h''(u)A(u)\nabla u \mathrm{d}x = \int_{\mathbb{T}^d} f(u) \cdot w \mathrm{d}x,$$

where $A : B = \sum_{j=1}^d \sum_{k=1}^2 A_{kj}B_{kj}$ for two matrices $A = (A_{kj}), B = (B_{kj}) \in \mathbb{R}^{2\times d}$. The right-hand side can be bounded in terms of $H[u]$ [see (14)], and the integral on the left-hand side is related to the corresponding integral in (7).

The proof of Theorem 1 is based on a regularization of (9), the fixed-point theorem of Leray-Schauder, and the de-regularization limit. The compactness is obtained from the entropy estimate (7). This technique is similar to those employed in our works [8,13]. The novelty here is the (nontrivial) observation that the cross-diffusion system (1) possesses a convex Lyapunov functional, defined by (6). Moreover, compared to [8,13], we are facing additional technical difficulties due to the quotient u_1/u_2.

The second result concerns the large-time asymptotics in the case $\mu_i = 0$ for $i = 1, 2$.

Theorem 2. (Large-time asymptotics) *Let the assumptions of Theorem 1 hold and let $\mu_1 = \mu_2 = 0$. Then, the solution $u(t) = (u_1, u_2)(t)$ to (1) converges in $L^2(\mathbb{T}^d)$ to $\overline{u} = (\overline{u}_1, \overline{u}_2)$ as $t \to \infty$, where*

$$\overline{u}_i = \frac{1}{\operatorname{meas}(\mathbb{T}^d)} \int_{\mathbb{T}^d} u_i^0 \, dx, \quad i = 1, 2.$$

If $\mu_i < 0$ for $i = 1, 2$, we prove the exponential convergence of $u(t)$ to zero in $H^1(\mathbb{T}^d)'$, see Remark 9. For a discussion of the case $\mu_i > 0$, we refer to Remark 10.

The paper is organized as follows. In Sect. 2, we make precise the derivation of (1) from a Fokker–Planck equation. Some technical results are proved in Sect. 3. Section 4 is devoted to the proof of Theorem 1, and Theorem 2 is shown in Sect. 5.

2. Derivation of the cross-diffusion system (1)

We summarize the formal derivation of (1) from a Fokker–Planck equation as presented by Lions [7]. Consider the n-dimensional Itō process $X_t = (X_t^1, \ldots, X_t^n)$ on some probability space, driven by the n-dimensional Wiener process $W_t = (W_t^1, \ldots, W_t^n)$ with respect to some given filtration. We assume that X_t solves the stochastic differential equation

$$dX_t = \widetilde{\mu}_t(X_t)\mathrm{d}t + \sigma_t(X_t)dW_t, \quad t > 0,$$

where $\widetilde{\mu}_t = (\widetilde{\mu}_t^1, \ldots, \widetilde{\mu}_t^n)$, and $\sigma_t = (\sigma_t^{ij})_{i,j,=1,\ldots,n}$ is an $n \times n$ matrix. It is well known [10, Theorems 7.3.3, 8.2.1] that the probability density $f(x_1, \ldots, x_n, t)$ for X_t satisfies the Fokker–Planck (or forward Kolmogorov) equation

$$\partial_t f = \frac{1}{2}\sum_{i,j=1}^n \frac{\partial^2}{\partial x_i \partial x_j}(D_{ij}(\widehat{x})f) - \sum_{i=1}^n \frac{\partial}{\partial x_i}(\widetilde{\mu}^i(\widehat{x})f), \quad \widehat{x} \in \mathbb{R}^n, \quad t > 0,$$

where $D(\widehat{x}) = (D_{ij}(\widehat{x})) = \sigma(\widehat{x})\sigma(\widehat{x})^\top$ is the diffusion tensor and $\widehat{x} = (x_1, \ldots, x_n)$.

In the following, we set $\widetilde{\mu}_t = 0$ and $\sigma_t = \mathrm{diag}(\sigma_1, \ldots, \sigma_n)$. This means that we neglect correlations between the processes. Taking them into account will lead to first-order terms in the final equations; see Remark 3. Under the above simplifications, the Fokker–Planck equation becomes

$$\partial_t f = \frac{1}{2}\sum_{j=1}^n \frac{\partial^2}{\partial x_j^2}(\sigma_j^2 f), \quad \widehat{x} \in \mathbb{R}^n, \quad t > 0. \tag{10}$$

We assume that σ_j is a function of the partial averages

$$u_i(x,t) = \int_{\mathbb{R}} f(x, x_n, t)e^{\lambda_i x_n}\mathrm{d}x_n, \quad x = (x_1, \ldots, x_{n-1}), \quad i = 1, \ldots, m,$$

where λ_i are some given (pairwise different) parameters. Temporal averages appear, for instance, in the modeling of Asian options. Here, u_i may be interpreted as an average with respect to the ecocnomic parameter x_n. We may employ other weights than the exponential one, but this one is mathematically extremely convenient because of the property $\partial u_i/\partial x_n = \lambda_i u_i$ (see Remark 3). Multiplying (10) by $e^{\lambda_i x_n}$ and integrating with respect to $x_n \in \mathbb{R}$, a straightforward calculation shows that u_i solves

$$\partial_t u_i = \frac{1}{2}\sum_{j=1}^{n-1} \frac{\partial^2}{\partial x_j^2}(\sigma_j^2 u_i) + \frac{\lambda_i^2}{2}\sigma_n^2 u_i, \quad i = 1, \ldots, m. \tag{11}$$

We allow σ_j to depend on the partial averages, $\sigma_j = \sigma_j(u_1, \ldots, u_m)$.

We consider only the special case $m = 2, \sigma := \sigma_j$ for $j = 1, \ldots, n-1$, and σ_n is constant and positive. Setting $u = (u_1, u_2), \mu_i := \lambda_i^2\sigma_n/2$, we find that

$$\partial_t u_i = \frac{1}{2}\Delta(\sigma(u)^2 u_i) + \mu_i u_i, \quad x \in \mathbb{R}^{n-1}, \quad t > 0, \quad i = 1, 2. \tag{12}$$

In divergence form, this system is equivalent to

$$\partial_t u = \mathrm{div}(A(u)\nabla u), \quad \text{where } A(u) = \sigma\begin{pmatrix} \sigma + 2\partial_1\sigma u_1 & 2\partial_2\sigma u_1 \\ 2\partial_1\sigma u_2 & \sigma + 2\partial_2\sigma u_2 \end{pmatrix},$$

where $\partial_i\sigma = \partial\sigma/\partial u_i, i = 1, 2$. This system is of parabolic type in the sense of Petrovski if the real parts of the eigenvalues of A are nonnegative [1], i.e., if $\sigma + \partial_1\sigma u_1 + \partial_2\sigma u_2 \geq 0$ for all $u \in \mathbb{R}^2$. This requirement is fulfilled if, for instance, σ depends on the quotient u_1/u_2 only. Therefore, we set $\sigma(u)^2 = 2a(u_1/u_2)$. Then,

$$\partial_t u_i = \Delta(a(u_1/u_2)u_i) + \mu_i u_i, \quad x \in \mathbb{R}^{n-1}, \quad t > 0, \quad i = 1, 2,$$

is of parabolic type in the sense of Petrovski, and these equations correspond to (1).

Remark 3. (Generalizations) The general model for nonvanishing $\widetilde{\mu}_t^i$ and nondiagonal σ_t is derived as above, and the result reads as

$$\partial_t u_i = \frac{1}{2}\sum_{j,k=1}^{n-1} \frac{\partial^2}{\partial x_j \partial x_k}(D_{jk}u_i) - \frac{1}{2}\sum_{j=1}^{n-1} \frac{\partial}{\partial x_j}\big((2\widetilde{\mu}^j + \lambda_i D_{jn})u_i\big) + \frac{\lambda_i}{2}(2\widetilde{\mu}^n + \lambda_i D_{nn})u_i. \tag{13}$$

Compared to (11), this equation also contains first-order terms. If $\widetilde{\mu}_t^i = 0$ and σ_t is diagonal, we obtain m equations of the type (12). The analysis of cross-diffusion systems with more than two components is expected to be much more involved than for those with two components. For instance, the analysis of the cross-diffusion model (4) is rather well understood only in the case of $m = 2$ components, while the case of $m \geq 3$ equations requires additional properties [3].

Another generalization concerns nonexponential weights. For instance, we may define

$$u_i = \int_{\mathbb{R}} f(x,t)\sin(\lambda_i x_n)\mathrm{d}x_n, \quad i = 1, \ldots, m.$$

Choosing again $\widetilde{\mu}_t^i = 0$ and $\sigma_t = \mathrm{diag}(\sigma_1, \ldots, \sigma_n)$, we find that

$$\partial_t u_i = \frac{1}{2}\sum_{j=1}^{n-1} \frac{\partial^2}{\partial x_j^2}(\sigma_j(u)^2 u_i) - \frac{\lambda_i^2}{2}\sigma_n^2 u_i, \quad i = 1, \ldots, m.$$

This justifies the assumption $\mu_i \in \mathbb{R}$ in (1), but there seems to be no financial interpretation of the trigonometric weight functions. □

3. Some auxiliary lemmas

In this section, we prove some algebraic properties of the matrices $h''(u)$ and $A(u)$ and some estimates related to the entropy density $h(u)$ and the components of $A(u)$. Recall that $h(u)$ is defined in (6) and $A(u)$ in (8).

Lemma 4. *(Properties of h) Let $\alpha > 0$. The function $h : (0,\infty)^2 \to \mathbb{R}^2$, defined in (6), is convex, its derivative h' is invertible, and there exists $C_h > 0$ such that for all $u = (u_1, u_2) \in (0,\infty)^2$,*

$$h(u) \geq \frac{1}{2}(u_1^2 + u_2^2), \quad \sum_{i=1}^{2} \mu_i u_i \partial_i h(u) \leq C_h h(u), \tag{14}$$

where we recall that $\partial_i h = \partial h/\partial u_i$.

Proof. We proceed in several steps.

Step 1: h is convex We compute the first partial derivatives of h,

$$\partial_1 h(u) = (\alpha+2)u_1^{\alpha+1}u_2^{-\alpha} - \alpha u_1^{-\alpha-1}u_2^{\alpha+2} - u_1^{-1} + 1, \tag{15}$$

$$\partial_2 h(u) = (\alpha+2)u_1^{-\alpha}u_2^{\alpha+1} - \alpha u_1^{\alpha+2}u_2^{-\alpha-1} - u_2^{-1} + 1, \tag{16}$$

and the Hessian $h''(u) = H^{(1)} + H^{(2)} + H^{(3)}$, where

$$H^{(1)} = \begin{pmatrix} (\alpha+2)(\alpha+1)(u_1/u_2)^\alpha & -\alpha(\alpha+2)(u_1/u_2)^{\alpha+1} \\ -\alpha(\alpha+2)(u_1/u_2)^{\alpha+1} & \alpha(\alpha+1)(u_1/u_2)^{\alpha+2} \end{pmatrix},$$
$$H^{(2)} = \begin{pmatrix} \alpha(\alpha+1)(u_2/u_1)^{\alpha+2} & -\alpha(\alpha+2)(u_2/u_1)^{\alpha+1} \\ -\alpha(\alpha+2)(u_2/u_1)^{\alpha+1} & (\alpha+2)(\alpha+1)(u_2/u_1)^\alpha \end{pmatrix},$$
$$H^{(3)} = \begin{pmatrix} u_1^{-2} & 0 \\ 0 & u_2^{-2} \end{pmatrix}. \tag{17}$$

Since $\det H^{(1)} = \alpha(\alpha+2)(u_1/u_2)^{2(\alpha+1)} > 0, \det H^{(2)} = \alpha(\alpha+2)(u_2/u_1)^{2(\alpha+1)} > 0$, and the diagonal elements of $H^{(1)}, H^{(2)}$ are positive, the matrices $H^{(i)}, i = 1, 2, 3$, are positive definite and so does $h''(u)$. Thus, h is convex.

Step 2: h' is invertible Since the Hessian h'' is positive definite on $(0,\infty)^2$, h' is one-to-one and the image $R(h')$ is open. If $R(h')$ is also closed, it follows that $R(h') = \mathbb{R}^2$ which means that h' is surjective. For this, let $(w_n) \in R(h')$ for $n \in \mathbb{N}$ such that $w_n \to w$ as $n \to \infty$. We show that $w \in R(h')$. By definition, there exists $u_n > 0$ such that $w_n = h'(u_n)$ for $n \in \mathbb{N}$. The idea is to prove that $(u_n) = (u_{1,n}, u_{2,n})$ is a bounded and strictly positive sequence. This implies that, up to a subsequence, $u_n \to u \in (0,\infty)^2$ as $n \to \infty$. By continuity of h', we infer that $h'(u_n) \to h'(u)$ as $n \to \infty$. We already know that $h'(u_n) = w_n \to w$ which shows that $w = h'(u) \in R(h')$, and $R(h')$ is closed.

It remains to verify that there exist positive constants $m, M > 0$ such that $m \le u_{i,n} \le M$ for all $n \in \mathbb{N}, i = 1, 2$. We argue by contradiction. Let us assume that (up to a subsequence) $u_{1,n} \to 0$ as $n \to \infty$. Since $(w_{1,n}) = (\partial_1 h(u_n))$ is convergent, we deduce from (15) that $u_{2,n} \to 0$ as well. As a consequence,

$$\alpha u_{1,n} w_{1,n} + (\alpha+2) u_{2,n} w_{2,n} \to 0, \quad (\alpha+2) u_{1,n} w_{1,n} + \alpha u_{2,n} w_{2,n} \to 0.$$

Expanding these expressions yields

$$u_{1,n}^{-\alpha} u_{2,n}^{\alpha+2} \to \frac{1}{2}, \quad u_{1,n}^{\alpha+2} u_{2,n}^{-\alpha} \to \frac{1}{2},$$

and the product also converges, $u_{1,n}^2 u_{2,n}^2 \to 1/4$. This is absurd since (u_n) converges to zero. Therefore, $u_{1,n}$ is strictly positive. With an analogous argument, we conclude that $u_{2,n}$ is strictly positive too.

Let us assume that (up to a subsequence) $u_{1,n} \to \infty$ as $n \to \infty$. Again, the convergence of $(w_{1,n})$ and (15) imply that $u_{2,n} \to \infty$. Consequently,

$$\frac{\alpha}{u_{2,n}} w_{1,n} + \frac{\alpha+2}{u_{1,n}} w_{2,n} \to 0, \quad \frac{\alpha+2}{u_{2,n}} w_{1,n} + \frac{\alpha}{u_{1,n}} w_{2,n} \to 0,$$

from which we infer after expanding these expressions that $u_{2,n}/u_{1,n} \to 0$ and $u_{1,n}/u_{2,n} \to 0$, which is a contradiction. So, $(u_{1,n})$ is bounded, and the same conclusion holds for $(u_{2,n})$.

Step 3: proof of (14) Observing that $x - \log x \ge 1$ for all $x > 0$, it follows that

$$h(u) \ge u_1^2 \left(\left(\frac{u_1}{u_2}\right)^\alpha + \left(\frac{u_2}{u_1}\right)^{\alpha+2} \right) = u_2^2 \left(\left(\frac{u_1}{u_2}\right)^{\alpha+2} + \left(\frac{u_2}{u_1}\right)^\alpha \right).$$

The elementary inequality $x^\alpha + (1/x)^{\alpha+2} \ge 1$ for $x > 0$ shows the first inequality in (14):

$$h(u) \ge \frac{u_1^2}{2} \left(\left(\frac{u_1}{u_2}\right)^\alpha + \left(\frac{u_2}{u_1}\right)^{\alpha+2} \right) + \frac{u_2^2}{2} \left(\left(\frac{u_1}{u_2}\right)^{\alpha+2} + \left(\frac{u_2}{u_1}\right)^\alpha \right) \ge \frac{1}{2}(u_1^2 + u_2^2).$$

For the second inequality in (14), we employ definition (6) of h and the elementary inequality $x-1\le 2(x-\log x)$ for $x>0$ to find that, if $C_h=2(\alpha+2)(|\mu_1|+|\mu_2|)$,

$$
\begin{aligned}
\sum_{i=1}^2 \mu_i u_i \partial_i h(u) &= \big(\mu_1(\alpha+2)-\mu_2\alpha\big)u_1^{\alpha+2}u_2^{-\alpha} + \big(\mu_2(\alpha+2)-\mu_1\alpha\big)u_1^{-\alpha}u_2^{\alpha+2} \\
&\quad + \mu_1(u_1-1)+\mu_2(u_2-1) \\
&\le C_h(u_1^{\alpha+2}u_2^{-\alpha}+u_1^{-\alpha}u_2^{\alpha+2}) + C_h(u_1-\log u_1+u_2-\log u_2).
\end{aligned}
$$

This finishes the proof. □

Next, we prove that $h''(u)A)(u)$ is positive semidefinite. Then, $B=A(u)h''(u)^{-1}$ in (9) is positive semidefinite too, since $z^\top A(u)h''(u)^{-1}z=(h''(u)^{-1}z)^\top h''(u)A(u)(h''(u)^{-1}z)\ge 0$ for $z\in\mathbb{R}^2$.

Lemma 5. (Positive semidefiniteness of $h''A$) *Let condition (2) hold. If $\alpha(\alpha+2)>1$, the matrix $h''(u)A(u)$ is positive semidefinite in $(0,\infty)^2$. Furthermore, if additionally $\alpha\ge p$, there exists a constant $\kappa=\kappa(\alpha)>0$ such that for all $u=(u_1,u_2)\in(0,\infty)^2$ and $z\in\mathbb{R}^2$,*

$$
z^\top h''(u)A(u)z \ge \kappa\bigg(\Big(\frac{u_1}{u_2}\Big)^{\alpha-p}+\Big(\frac{u_1}{u_2}\Big)^{p-\alpha}\bigg)|z|^2.
$$

Proof. Let $\alpha(\alpha+2)>1$ and let $M^{(i)}=(M^{(i)}_{jk}):=\frac12((H^{(i)}A)^\top+H^{(i)}A)$ be the symmetric part of $H^{(i)}A$, where $H^{(i)}$ with $i=1,2,3$ is defined in (17). A computation shows that

$$
\begin{aligned}
M^{(1)}_{11} &= (\alpha+2)\big((\alpha+1)a(u_1/u_2)+(u_1/u_2)a'(u_1/u_2)\big)(u_1/u_2)^\alpha, \\
\det M^{(1)} &= \big(\alpha(\alpha+2)a(u_1/u_2)^2-(u_1/u_2)^2a'(u_1/u_2)^2\big)(u_1/u_2)^{2\alpha+2}, \\
M^{(2)}_{11} &= \alpha\big((\alpha+1)a(u_1/u_2)-(u_1/u_2)a'(u_1/u_2)\big)(u_2/u_1)^{\alpha+2}, \\
\det M^{(2)} &= \big(\alpha(\alpha+2)a(u_1/u_2)^2-(u_1/u_2)^2a'(u_1/u_2)^2\big)(u_2/u_1)^{2\alpha+2}, \\
M^{(3)} &= \begin{pmatrix} (a(u_1/u_2)+(u/u_2)a'(u_1/u_2))u_1^{-2} & 0 \\ 0 & (a(u_1/u_2)-(u_1/u_2)a'(u_1/u_2))u_2^{-2} \end{pmatrix}.
\end{aligned}
$$

By the first condition in (2) and the positivity of α, we infer that $M^{(3)}$ is positive semidefinite and $M^{(1)}_{11}, M^{(2)}_{11}$ are positive for $u,v>0$. Moreover, since $\alpha(\alpha+2)>1$ by assumption, $\det(M^{(1)})>0$ and $\det(M^{(2)})>0$. Thus, by Sylvester's criterion, $(h''A)(u)$ is positive semidefinite for all $u\in(0,\infty)^2$.

Now let additionally $\alpha\ge p$. Then, the first condition in (2) shows that

$$
\begin{aligned}
\frac{\det M^{(1)}}{\operatorname{tr} M^{(1)}} &= \frac{\alpha(\alpha+2)a(u_1/u_2)^2-(u_1/u_2)^2a'(u_1/u_2)^2}{(\alpha(u_1/u_2)^2+\alpha+2)((\alpha+1)a(u_1/u_2)+(u_1/u_2)a'(u_1/u_2))}(u_1/u_2)^{\alpha+2} \\
&\ge \frac{(\alpha(\alpha+2)-1)a(u_1/u_2)^2}{(\alpha+2)((u_1/u_2)^2+1)(\alpha+2)a(u_1/u_2)}(u_1/u_2)^{\alpha+2} \\
&= k_1(\alpha)\frac{a(u_1/u_2)}{(u_1/u_2)^2+1}(u_1/u_2)^{\alpha+2},
\end{aligned}
$$

where $k(\alpha)=(\alpha(\alpha+2)-1)/(\alpha+2)^2$. In a similar way, we find that

$$
\begin{aligned}
\frac{\det M^{(2)}}{\operatorname{tr} M^{(2)}} &= \frac{\alpha(\alpha+2)a(u_1/u_2)^2-(u_1/u_2)^2a'(u_1/u_2)^2}{((\alpha+2)(u_1/u_2)^2+\alpha)((\alpha+1)a(u_1/u_2)-(u_1/u_2)a'(u_1/u_2))}(u_2/u_1)^{\alpha} \\
&\ge \frac{(\alpha(\alpha+2)-1)a(u_1/u_2)^2}{(\alpha+2)((u_1/u_2)^2+1)(\alpha+2)a(u_1/u_2)}(u_1/u_2)^{-\alpha} \\
&= k(\alpha)\frac{a(u_1/u_2)}{(u_1/u_2)^2+1}(u_1/u_2)^{-\alpha}.
\end{aligned}
$$

Since $\det M/\operatorname{tr} M$ is a lower bound for the eigenvalues of any symmetric positive definite matrix $M \in \mathbb{R}^{2\times 2}$ (and taking into account that $M^{(3)}$ is positive definite), we deduce that for $z \in \mathbb{R}^2$,

$$\begin{aligned} z^\top (h''A)(u)z &\geq k(\alpha)a(u_1/u_2)\frac{(u_1/u_2)^{\alpha+2} + (u_1/u_2)^{-\alpha}}{(u_1/u_2)^2+1}|z|^2 \\ &\geq \frac{1}{2}k(\alpha)a(u_1/u_2)((u_1/u_2)^\alpha + (u_1/u_2)^{-\alpha})|z|^2. \end{aligned}$$

In the last inequality, we have employed the elementary inequality $(x^{\alpha+2}+x^{-\alpha})/(x^2+1) \geq \frac{1}{2}(x^\alpha + x^{-\alpha})$ which is equivalent to $(x^2-1)(x^\alpha - x^{-\alpha}) \geq 0$, and this holds true for all $x > 0$. By the second condition in (2),

$$z^\top (h''A)(u)z \geq \frac{a_0}{2}k(\alpha)\frac{(u_1/u_2)^\alpha + (u_1/u_2)^{-\alpha}}{(u_1/u_2)^p + (u_1/u_2)^{-p}}|z|^2.$$

The inequality $(x^\alpha + x^{-\alpha})/(x^p + x^{-p}) \geq \frac{1}{2}(x^{\alpha-p} + x^{p-\alpha})$ is equivalent to $(x^{\alpha-p} - x^{p-\alpha})(x^p - x^{-p}) \geq 0$, which holds true for $x > 0$ since $\alpha - p \geq 0$ and $p \geq 0$. Therefore,

$$z^\top (h''A)(u)z \geq \frac{a_0}{4}k(\alpha)\big((u_1/u_2)^{\alpha-p} + (u_1/u_2)^{p-\alpha})\big)|z|^2,$$

which concludes the proof with $\kappa = a_0 k(\alpha)/4$. □

The following two lemmas concern elementary estimates for $a(r)$.

Lemma 6. *Given $r_0 > 0$ arbitrary, it holds that*

$$a(r) \leq \begin{cases} \frac{a(r_0)}{r_0} r & \text{for } r \geq r_0, \\ r_0 a(r_0)\frac{1}{r} & \text{for } r < r_0. \end{cases}$$

Proof. The first inequality in (2) implies that $r \mapsto a(r)/r$ is nonincreasing, while $r \mapsto a(r)r$ is nondecreasing. Writing these monotonicity properties in an explicit way gives the result. □

Lemma 7. *Let $\alpha \geq 2$. Then, for all $u_1, u_2 > 0$,*

$$\begin{aligned} a\left(\frac{u_1}{u_2}\right)^2 (u_1^2 + u_2^2) &\leq C_a\left(u_1^2 + u_2^2 + \frac{u_1^4}{u_2^2}\right) \\ &\leq \xi_\alpha C_a\left(\left(\frac{u_1}{u_2}\right)^\alpha u_1^2 + \left(\frac{u_1}{u_2}\right)^{-\alpha} u_2^2\right) \leq \xi_\alpha C_a h(u), \end{aligned}$$

where $C_a = a(1)^2$ and $\xi_\alpha > 0$ is a suitable constant which only depends on α.

Proof. The first inequality follows from an application of Lemma 6 with $r_0 = 1$. Indeed, if $u_1/u_2 \geq 1$, we obtain

$$a\left(\frac{u_1}{u_2}\right)^2 u_1^2 \leq a(1)^2\frac{u_1^2}{u_2^2}u_1^2, \quad a\left(\frac{u_1}{u_2}\right)^2 u_2^2 \leq a(1)^2 u_1^2,$$

while if $u_1/u_2 \leq 1$ we have

$$a\left(\frac{u_1}{u_2}\right)^2 u_1^2 \leq a(1)^2 u_2^2, \quad a\left(\frac{u_1}{u_2}\right)^2 u_2^2 \leq a(1)^2\frac{u_2^2}{u_1^2}u_2^2.$$

These inequalities show the claim with $C_a = a(1)^2$. The second inequality follows from

$$u_1^2 + u_2^2 + \frac{u_1^4}{u_2^2} + \frac{u_2^4}{u_1^2} = u_1u_2\left(\frac{u_1}{u_2} + \frac{u_2}{u_1} + \frac{u_1^3}{u_2^3} + \frac{u_2^3}{u_1^3}\right) \leq \xi_\alpha u_1u_2\left(\frac{u_1^{\alpha+1}}{u_2^{\alpha+1}} + \frac{u_2^{\alpha+1}}{u_1^{\alpha+1}}\right),$$

where $\xi_\alpha > 0$ is a suitable constant, which depends only on α. This finishes the proof. □

Lemma 8. *Recall that $A(u) = (A_{ij}(u))$ is given by (8). Then, there exists $C_A > 0$, only depending on $a(1)$, such that for all $u_1, u_2 > 0$,*

$$|A(u)| \leq C_A\left(1 + \left(\frac{u_1}{u_2}\right)^2 + \left(\frac{u_1}{u_2}\right)^{-2}\right).$$

Proof. We apply the first condition in (2) to find that

$$\sum_{i,j=1}^{2} |A_{ij}(u)| \leq a\left(\frac{u_1}{u_2}\right)\left(4 + \frac{u_1}{u_2} + \frac{u_2}{u_1}\right).$$

Then Lemma 6 with $r_0 = 1$ implies that

$$\sum_{i,j=1}^{2} |A_{ij}(u)| \leq a(1)\left(\frac{u_1}{u_2} + \frac{u_2}{u_1}\right)\left(4 + \frac{u_1}{u_2} + \frac{u_2}{u_1}\right).$$

This estimate and Young's inequality conclude the proof. □

4. Proof of Theorem 1

Let $T > 0, N \in \mathbb{N}, \tau = T/N$, and $m \in \mathbb{N}$ with $m > d/2$. Then, the embedding $H^m(\mathbb{T}^d) \hookrightarrow L^\infty(\mathbb{T}^d)$ is compact. Furthermore, let $w^{k-1} = (w_1^{k-1}, w_2^{k-1}) \in L^\infty(\mathbb{T}^d)^2$ be given and let $u^{k-1} = (h')^{-1}(w^{k-1})$. By Lemma 4, the pair $u^{k-1} = (u_1^{k-1}, u_2^{k-1})$ is well defined and we have $u^{k-1} \in L^\infty(\mathbb{T}^d)^2$. We wish to find $w_k = (w_1^k, w_2^k) \in H^m(\mathbb{T}^d)^2$ such that for all $\phi = (\phi_1, \phi_2) \in H^m(\mathbb{T}^d)^2$,

$$\begin{aligned} &\frac{1}{\tau}\int_{\mathbb{T}^d}(u^k - u^{k-1}) \cdot \phi \mathrm{d}x + \int_{\mathbb{T}^d} \nabla\phi : B(w^k)\nabla w^k \mathrm{d}x \\ &\quad + \tau\int_{\mathbb{T}^d}(D^m w^k \cdot D^m\phi + w^k \cdot \phi)\mathrm{d}x = \sum_{i=1}^{2}\mu_i\int_{\mathbb{T}^d} u_i^k \phi_i \mathrm{d}x, \end{aligned} \tag{18}$$

where $B(w^k) = A(u^k)h''(u^k)^{-1}$,

$$D^m w^k \cdot D^m\phi := \sum_{|\alpha|=m}\sum_{i=1}^{2} D^\alpha u_i^k D^\alpha \phi_i,$$

$\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{N}_0^d$ is a multiindex and $D^\alpha = \partial^{|\alpha|}/(\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d})$ a partial derivative of order $|\alpha|$.

Step 1: solution of (18) Let $\widehat{w} = (\widehat{w}_1, \widehat{w}_2) \in L^\infty(\mathbb{T}^d)^2$ and $\eta \in [0,1]$ be given. Set $\widehat{u} = (\widehat{u}_1, \widehat{u}_2) := (h')^{-1}(\widehat{w})$. We solve first the linear problem

$$a(w, \phi) = \eta F(\phi) \quad \text{for all } \phi \in H^m(\mathbb{T}^d)^2, \tag{19}$$

where

$$a(w, \phi) = \int_{\mathbb{T}^d}(D^m w^k \cdot D^m\phi + w^k \cdot \phi)\mathrm{d}x + \int_{\mathbb{T}^d}\nabla\phi : B(\widehat{w})\nabla w^k \mathrm{d}x,$$

$$F(\phi) = -\frac{1}{\tau}\int_{\mathbb{T}^d}(\widehat{u} - u^{k-1}) \cdot \phi \mathrm{d}x + \sum_{i=1}^{2}\mu_i\int_{\mathbb{T}^d}\widehat{u}_i^k \phi_i \mathrm{d}x.$$

Since $\widehat{w} \in L^\infty(\mathbb{T}^d)^2$ and h' is continuous in $(0,\infty)^2$, we have $\widehat{u} \in L^\infty(\mathbb{T}^d)^2$. This shows that F is continuous on $H^m(\mathbb{T}^d)$. The bilinear form a is continuous and coercive, by the generalized Poincaré inequality for H^m spaces [11, Chap. 2.1.4, Formula (1.39)] and the positive semidefiniteness of $B(\widehat{w})$ (see Lemma 5).

Hence, the Lax-Milgram lemma provides a unique solution $w = (w_1, w_2) \in H^m(\mathbb{T}^d)^2 \hookrightarrow L^\infty(\mathbb{T}^d)^2$ to (19). This defines the fixed-point operator $S : L^\infty(\mathbb{T}^d)^2 \times [0,1] \to L^\infty(\mathbb{T}^d)^2, S(\widehat{w}, \eta) = w$, where w solves (19).

It holds clearly $S(w, 0) = 0$. Standard arguments show that S is continuous (see, e.g., the proof of Lemma 5 in [8]). Because of the compact embedding $H^m(\mathbb{T}^d) \hookrightarrow L^\infty(\mathbb{T}^d)$, the mapping S is even compact. In order to apply the Leray-Schauder fixed-point theorem, it remains to prove a uniform bound for all fixed points of $S(\cdot, \eta)$ in $L^\infty(\mathbb{T}^d)^2$.

Let $w \in L^\infty(\mathbb{T}^d)^2$ be such a fixed point, i.e., a solution to (19) with $\widehat{u}$ replaced by $u := (h')^{-1}(w)$. The uniform bound will be a consequence of the entropy inequality. For this, we employ the test function w in (19):

$$\begin{aligned}\frac{\eta}{\tau}\int_{\mathbb{T}^d}(u-u^{k-1})\cdot w\mathrm{d}x + \int_{\mathbb{T}^d}\nabla w : B(w)\nabla w\mathrm{d}x + \tau\int_{\mathbb{T}^d}\big(|D^m w|^2 + |w|^2\big)\mathrm{d}x \\ = \eta\sum_{i=1}^2\mu_i\int_{\mathbb{T}^d}u_i w_i\mathrm{d}x. \qquad (20)\end{aligned}$$

By the convexity of h, it follows that

$$h(u) - h(u^{k-1}) \le h'(u)\cdot(u-u^{k-1}) = (u-u^{k-1})\cdot w.$$

Moreover, by (9) and Lemma 5, we have

$$\begin{aligned}\int_{\mathbb{T}^d}\nabla w : B(w)\nabla w\mathrm{d}x &= \int_{\mathbb{T}^d}\nabla u : (h''A)(u)\nabla u\mathrm{d}x \\ &\ge \kappa\int_{\mathbb{T}^d}\left(\left(\frac{u_1}{u_2}\right)^{\alpha-p} + \left(\frac{u_1}{u_2}\right)^{p-\alpha}\right)|\nabla u|^2\mathrm{d}x.\end{aligned}$$

Taking into account the second estimate in (14), we infer that

$$\eta\sum_{i=1}^2\mu_i\int_{\mathbb{T}^d}u_i w_i\mathrm{d}x = \eta\sum_{i=1}^2\mu_i\int_{\mathbb{T}^d}u_i\partial_i h(u)\mathrm{d}x \le C_h\int_{\mathbb{T}^d}h(u)\mathrm{d}x.$$

Therefore, (20) becomes

$$\begin{aligned}\frac{\eta}{\tau}\int_{\mathbb{T}^d}h(u)\mathrm{d}x + \kappa\int_{\mathbb{T}^d}\left(\left(\frac{u_1}{u_2}\right)^{\alpha-p} + \left(\frac{u_1}{u_2}\right)^{p-\alpha}\right)|\nabla u|^2\mathrm{d}x \\ + \tau\int_{\mathbb{T}^d}\big(|D^m w|^2 + |w|^2\big)\mathrm{d}x \le \frac{\eta}{\tau}\int_{\mathbb{T}^d}h(u^{k-1})\mathrm{d}x + C_h\int_{\mathbb{T}^d}h(u)\mathrm{d}x. \qquad (21)\end{aligned}$$

Choosing $\tau < 1/C_h$, this shows that w is uniformly bounded in $H^m(\mathbb{T}^d)$. Thus, we can apply the fixed-point theorem of Leray-Schauder to conclude the existence of a weak solution $w^k := w$ with $u^k = h'(w^k)$ to (18) with $\eta = 1$.

Step 2: a priori estimates Inequality (21) shows, for $w = w^k, u = u^k$, and $\eta = 1$, that

$$\begin{aligned}(1 - C_h\tau)\int_{\mathbb{T}^d}h(u^k)\mathrm{d}x + \kappa\tau\int_{\mathbb{T}^d}\left(\left(\frac{u_1^k}{u_2^k}\right)^{\alpha-p} + \left(\frac{u_1^k}{u_2^k}\right)^{p-\alpha}\right)|\nabla u^k|^2\mathrm{d}x \\ + \tau^2\int_{\mathbb{T}^d}\big(|D^m w^k|^2 + |w^k|^2\big)\mathrm{d}x \le \int_{\mathbb{T}^d}h(u^{k-1})\mathrm{d}x.\end{aligned}$$

We sum (21) for $k=1,\dots,j$ and divide the resulting inequality by $1-C_h\tau$ (recall that we have chosen $\tau<1/C_h$):

$$\int_{\mathbb{T}^d} h(u^j)\mathrm{d}x + \frac{\kappa\tau}{1-C_h\tau}\sum_{k=1}^{j}\int_{\mathbb{T}^d}\bigg(\Big(\frac{u_1^k}{u_2^k}\Big)^{\alpha-p}+\Big(\frac{u_1^k}{u_2^k}\Big)^{p-\alpha}\bigg)|\nabla u^k|^2\mathrm{d}x$$
$$+\frac{\tau^2}{1-C_h\tau}\sum_{k=1}^{j}\int_{\mathbb{T}^d}\big(|D^m w^k|^2+|w^k|^2\big)\mathrm{d}x$$
$$\le \frac{1}{1-C_h\tau}\int_{\mathbb{T}^d} h(u^0)\mathrm{d}x + \frac{C_h\tau}{1-C_h\tau}\sum_{k=1}^{j-1}\int_{\mathbb{T}^d} h(u^k)\mathrm{d}x.$$

We apply the discrete Gronwall inequality [4] to obtain for $j\tau\le T$,

$$\int_{\mathbb{T}^d} h(u^j)\mathrm{d}x + \tau\sum_{k=1}^{j}\int_{\mathbb{T}^d}\bigg(\Big(\frac{u_1^k}{u_2^k}\Big)^{\alpha-p}+\Big(\frac{u_1^k}{u_2^k}\Big)^{p-\alpha}\bigg)|\nabla u^k|^2\mathrm{d}x$$
$$+\tau^2\sum_{k=1}^{j}\int_{\mathbb{T}^d}\big(|D^m w^k|^2+|w^k|^2\big)\mathrm{d}x \le C, \tag{22}$$

where $C>0$ denotes a constant which is independent of τ (and independent of T if $\mu_i\le 0$) but dependent on the initial entropy $H[u^0]$.

We define the piecewise constant functions in time $w^{(\tau)}(x,t)=w^k(x)$ and $u^{(\tau)}(x,t)=u^k(x)$ for $x\in\mathbb{T}^d$ and $t\in((k-1)\tau,k\tau]$, $k=1,\dots,j$. Furthermore, we introduce the shift operator $\sigma_\tau u^{(\tau)}(x,t)=u^{k-1}(x)$ for $x\in\mathbb{T}^d, t\in((k-1)\tau,k\tau]$. With this notation, we can rewrite (20) (with $\eta=1$) as

$$\frac{1}{\tau}\int_0^T\int_{\mathbb{T}^d}(u^{(\tau)}-\sigma_\tau u^{(\tau)})\cdot\phi\,\mathrm{d}x\mathrm{d}t + \int_0^T\int_{\mathbb{T}^d}\nabla\phi : (h''A)(u^{(\tau)})\nabla u^{(\tau)}\mathrm{d}x\mathrm{d}t$$
$$+\tau\int_0^T\int_{\mathbb{T}^d}\big(D^m w^{(\tau)}\cdot D^m\phi + w^{(\tau)}\cdot\phi^{(\tau)}\big)\mathrm{d}x\mathrm{d}t + \sum_{i=1}^{2}\mu_i\int_0^T\int_{\mathbb{T}^d} u_i^{(\tau)}\phi_i\mathrm{d}x\mathrm{d}t \tag{23}$$

and (22) as

$$\int_{\mathbb{T}^d} h(u^{(\tau)}(t))\mathrm{d}x + \int_0^t\int_{\mathbb{T}^d}\bigg(\Big(\frac{u_1^{(\tau)}}{u_2^{(\tau)}}\Big)^{\alpha-p}+\Big(\frac{u_1^{(\tau)}}{u_2^{(\tau)}}\Big)^{p-\alpha}\bigg)|\nabla u^{(\tau)}|^2\mathrm{d}x\mathrm{d}s$$
$$+\tau\int_0^t\int_{\mathbb{T}^d}\big(|D^m w^{(\tau)}|^2+|w^{(\tau)}|^2\big)\mathrm{d}x\mathrm{d}s\le C, \tag{24}$$

where $t\in((j-1)\tau,j\tau]$. It follows that

$$\|w^{(\tau)}\|_{L^2(0,T;H^m(\mathbb{T}^d))}\le C\tau^{-1/2}. \tag{25}$$

By Lemmas 7, 4, and estimate (24), we find that

$$\int_{\mathbb{T}^d}\bigg(\bigg|a\bigg(\frac{u_1^{(\tau)}}{u_2^{(\tau)}}\bigg)u_1^{(\tau)}\bigg|^2+\bigg|a\bigg(\frac{u_1^{(\tau)}}{u_2^{(\tau)}}\bigg)u_2^{(\tau)}\bigg|^2\bigg)\mathrm{d}x \le 3C_a\int_{\mathbb{T}^d}h(u^{(\tau)})\mathrm{d}x\le C, \tag{26}$$

$$\int_{\mathbb{T}^d}\big((u_1^{(\tau)})^2+(u_2^{(\tau)})^2\big)\mathrm{d}x\le\int_{\mathbb{T}^d}h(u^{(\tau)})\mathrm{d}x\le C. \tag{27}$$

Moreover, using Lemmas 8 and (24),

$$\begin{aligned}\int_0^T\int_{\mathbb{T}^d}&\Big(\big|\nabla\big(a(u_1^{(\tau)}/u_2^{(\tau)})u_1^{(\tau)}\big)\big|^2+\big|\nabla\big(a(u_1^{(\tau)}/u_2^{(\tau)})u_2^{(\tau)}\big)\big|^2\Big)\mathrm{d}x\mathrm{d}t\\
&=\int_{\mathbb{T}^d}|A(u^{(\tau)})\nabla u^{(\tau)}|^2\mathrm{d}x\le\int_0^T\int_{\mathbb{T}^d}|A(u^{(\tau)})|^2|\nabla u^{(\tau)}|^2\mathrm{d}x\mathrm{d}t\\
&\le C_A\int_0^T\int_{\mathbb{T}^d}\bigg(1+\bigg(\frac{u_1^{(\tau)}}{u_2^{(\tau)}}\bigg)^4+\bigg(\frac{u_2^{(\tau)}}{u_1^{(\tau)}}\bigg)^4\bigg)|\nabla u^{(\tau)}|^2\mathrm{d}x\mathrm{d}t\\
&\le C\int_0^T\int_{\mathbb{T}^d}\bigg(\bigg(\frac{u_1^{(\tau)}}{u_2^{(\tau)}}\bigg)^{\alpha-p}+\bigg(\frac{u_1^{(\tau)}}{u_2^{(\tau)}}\bigg)^{p-\alpha}\bigg)|\nabla u^{(\tau)}|^2\mathrm{d}x\mathrm{d}t\le C.\end{aligned} \tag{28}$$

The last but one inequality follows from the elementary estimate $1+y^4\le y^{\alpha-p}+y^{p-\alpha}$ for $y>0$ which holds because of the assumption $\alpha-p\ge 4$. Estimates (26)–(28) yield for $i=1,2$,

$$\|u_i^{(\tau)}\|_{L^\infty(0,T;L^2(\mathbb{T}^d))}+\|\nabla u_i^{(\tau)}\|_{L^2(0,T;L^2(\mathbb{T}^d))}\le C, \tag{29}$$

$$\|a(u_1^{(\tau)}/u_2^{(\tau)})u_i\|_{L^\infty(0,T;L^2(\mathbb{T}^d))}+\|\nabla(a(u_1^{(\tau)}/u_2^{(\tau)})u_i^{(\tau)})\|_{L^2(0,T;L^2(\mathbb{T}^d))}\le C. \tag{30}$$

These estimates are uniform in $T>0$ if $\mu_i\le 0$.

Next, we derive a uniform estimate for the discrete time derivative $(u^{(\tau)}-\sigma_\tau u^{(\tau)})/\tau$. For $\phi\in L^2(0,T;H^m(\mathbb{T}^d))$, we estimate

$$\begin{aligned}\frac{1}{\tau}\bigg|\int_0^T(u^{(\tau)}-\sigma_\tau u^{(\tau)})\cdot\phi\mathrm{d}x\mathrm{d}t\bigg|&\le\|A(u^{(\tau)})\nabla u^{(\tau)}\|_{L^2(0,T;L^2(\mathbb{T}^d))}\|\nabla\phi\|_{L^2(0,T;L^2(\mathbb{T}^d))}\\
&\quad+\tau\|w^{(\tau)}\|_{L^2(0,T;H^m(\mathbb{T}^d))}\|\phi\|_{L^2(0,T;H^m(\mathbb{T}^d))}\\
&\quad+\max\{\mu_1,\mu_2\}\|u^{(\tau)}\|_{L^2(0,T;L^2(\mathbb{T}^d))}\|\phi\|_{L^2(0,T;L^2(\mathbb{T}^d))}\\
&\le C\|\phi\|_{L^2(0,T;H^m(\mathbb{T}^d))},\end{aligned}$$

taking into account the bounds (25), (28), and (29). Therefore,

$$\tau^{-1}\|u^{(\tau)}-\sigma_\tau u^{(\tau)}\|_{L^2(0,T;H^m(\mathbb{T}^d)')}\le C. \tag{31}$$

Step 3: limit $\tau\to 0$ Estimates (29) and (31) allow us to apply the Aubin–Lions lemma in the discrete version of [6] to obtain the existence of a subsequence, which is not relabeled, such that, as $\tau\to 0$,

$$u_i^{(\tau)}\to u_i\quad\text{strongly in } L^2(0,T;L^2(\mathbb{T}^d))\text{ and a.e.},\quad i=1,2.$$

Moreover, by (25), (29), and (31), for the same subsequence and $i = 1, 2$,

$$\begin{aligned}\tau w_i^{(\tau)} &\to 0 \quad \text{strongly in } L^2(0,T;H^m(\mathbb{T}^d)),\\ \nabla u_i^{(\tau)} &\rightharpoonup \nabla u_i \quad \text{weakly in } L^2(0,T;L^2(\mathbb{T}^d)),\\ \tau^{-1}(u_i^{(\tau)} - \sigma_\tau u_i^{(\tau)}) &\rightharpoonup \partial_t u_i \quad \text{weakly in } L^2(0,T;H^m(\mathbb{T}^d)').\end{aligned}$$

The pointwise convergence of $(u_i^{(\tau)})$, Fatou's lemma, and estimate (24) imply that, for a.e. $t \in (0,T)$,

$$\begin{aligned}\sum_{i=1}^2 \int_{\mathbb{T}^d} (u_i(t) - \log u_i(t))\mathrm{d}x &\le \liminf_{\tau\to 0} \sum_{i=1}^2 \int_{\mathbb{T}^d} \big(u_i^{(\tau)}(t) - \log u_i^{(\tau)}(t)\big)\mathrm{d}x\\ &\le \liminf_{\tau\to 0} \int_{\mathbb{T}^d} h(u^{(\tau)}(t))\mathrm{d}x \le C.\end{aligned}$$

This means that $u_i > 0$ a.e. in $\mathbb{T}^d \times (0,T)$.

Estimate (26) and (28) show that, up to a subsequence,

$$a(u_1^{(\tau)}/u_2^{(\tau)})u_i^{(\tau)} \rightharpoonup q_i \quad \text{weakly in } L^2(0,T;H^1(\mathbb{T}^d)),\ i = 1, 2,$$

where $q_i \in L^2(0,T;H^1(\mathbb{T}^d))$. We wish to identify q_i. To this end, let us define $\chi_\varepsilon^{(\tau)} = \mathbf{1}_{\{u_1^{(\tau)}\ge\varepsilon,\, u_2^{(\tau)}\ge\varepsilon\}}$ and $\chi_\varepsilon = \mathbf{1}_{\{u_1\ge\varepsilon,\, u_2\ge\varepsilon\}}$, where $\mathbf{1}_A$ denotes the characteristic function on the set A. Clearly, $\chi_\varepsilon^{(\tau)} \to \chi_\varepsilon$ strongly in $L^s(0,T;L^s(\mathbb{T}^d))$ for all $1 \le s < \infty$. We infer that

$$\chi_\varepsilon^{(\tau)} a(u_1^{(\tau)}/u_2^{(\tau)})u_i^{(\tau)} \rightharpoonup \chi_\varepsilon a(u_1/u_2)u_i \quad \text{weakly in } L^s(0,T;L^s(\mathbb{T}^d)), \quad 1 \le s < 2.$$

We deduce that $q_i = a(u_1/u_2)u_i$ on the set $\{u_1 \ge \varepsilon,\ u_2 \ge \varepsilon\}$. Since $\varepsilon > 0$ is arbitrary and $u_i > 0$ a.e. in $\mathbb{T}^d \times (0,T)$, this identification holds, in fact, a.e. in $\mathbb{T}^d \times (0,T)$.

Consequently, we may perform the limit $\tau \to 0$ in (23) to deduce that u is a weak solution to (1) with test functions $L^2(0,T;H^m(\mathbb{T}^d)')$. However, since $a(u_1/u_2)u_i \in L^2(0,T;H^1(\mathbb{T}^d))$, we employ a density argument to infer that (1) also holds for $L^2(0,T;H^1(\mathbb{T}^d)')$. Since $u_i \in L^2(0,T;H^1(\mathbb{T}^d))$ and $\partial_t u_i \in L^2(0,T;H^1(\mathbb{T}^d)')$, it follows that $u_i \in C^0([0,T];L^2(\mathbb{T}^d))$, so the initial datum is satisfied in $L^2(\mathbb{T}^d)$. Finally, since the bounds are uniform in T if $\mu_i \le 0$, the statement (5) follows.

5. Proof of Theorem 2

Theorem 1 allows us to employ the test equations $u_1 - \bar{u}_1, u_2 - \bar{u}_2$ in (1), respectively:

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_{\mathbb{T}^d} \sum_{i=1}^2 (u_i - \bar{u}_i)^2 \mathrm{d}x = -\int_{\mathbb{T}^d} \sum_{i=1}^2 \nabla u_i \cdot \nabla(a(u)u_i)\mathrm{d}x,$$

which, together with Theorem 1, implies that $(\mathrm{d}/\mathrm{d}t)\int \sum_{i=1}^2 (u_i - \bar{u}_i)^2 \mathrm{d}x \in L^1(0,\infty)$. Consequently, the limit

$$\lim_{t\to\infty} \int_{\mathbb{T}^d} \sum_{i=1}^2 (u_i(t) - \bar{u}_i)^2 \mathrm{d}x = \int_{\mathbb{T}^d} \sum_{i=1}^2 (u_i^0 - \bar{u}_i)^2 \mathrm{d}x + \int_0^\infty \frac{\mathrm{d}}{\mathrm{d}t} \int_{\mathbb{T}^d} \sum_{i=1}^2 (u_i - \bar{u}_i)^2 \mathrm{d}x\, \mathrm{d}t$$

exists and is finite. Poincaré's inequality and Theorem 1 imply that

$$\int_{\mathbb{T}^d} \sum_{i=1}^2 (u_i - \bar{u}_i)^2 \mathrm{d}x \le C_P \int_{\mathbb{T}^d} \sum_{i=1}^2 |\nabla u_i|^2 \mathrm{d}x \in L^1(0,\infty),$$

which means that $\lim_{t\to\infty}\int_{\mathbb{T}^d}\sum_{i=1}^2(u_i(t)-\bar u_i)^2\mathrm{d}x=0$. This finishes the proof.

Remark 9. If $\mu_i<0$ for $i=1,2$, we can prove the exponential convergence of the solution $u(t)$ to (1) in $H^1(\mathbb{T}^d)'$ by using the dual method. Indeed, let $\phi_i\in L^2(0,T;H^1(\mathbb{T}^d))$ be the unique solution to $-\Delta\phi_i=u_i(t)$ in $\mathbb{T}^d$ and $\int_{\mathbb{T}^d}\phi_i\mathrm{d}x=0, i=1,2$. Employing $\phi=(\phi_1,\phi_2)$ as a test function in (1), we find after a straightforward computation that

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{T}^d}\big(|\nabla\phi_1|^2+|\nabla\phi_2|^2\big)\mathrm{d}x+\int_{\mathbb{T}^d}a(u_1/u_2)(u_1^2+u_2^2)\mathrm{d}x=\int_{\mathbb{T}^d}\big(\mu_1|\nabla\phi_1|^2+\mu_2|\nabla\phi_2|^2\big)\mathrm{d}x.$$

Then Gronwall's lemma implies that

$$\int_{\mathbb{T}^d}|\nabla\phi(t)|^2\mathrm{d}x\le e^{\max\{\mu_1,\mu_2\}t}\int_{\mathbb{T}^d}|\nabla\phi(0)|^2\mathrm{d}x,\quad t>0.$$

Since $\|u_i\|_{H^1(\mathbb{T}^d)'}=\|\phi_i\|_{H^1(\mathbb{T}^d)}$, we conclude that $\|u_i(t)\|_{H^1(\mathbb{T}^d)'}\le C\exp(-\kappa t)$ for $t>0$, where $\kappa=-\max\{\mu_1,\mu_2\}>0$ and $C>0$ depends on u^0. $\square$

Remark 10. In the case $\mu_i>0$ for $i=1,2$, we cannot expect equilibration rates, since the solution grows in the L^2 norm as $t\to\infty$. This growth can be made precise if $\mu:=\mu_1=\mu_1>0$. Indeed, $u_i^*=e^{-\mu t}u_i$ solves

$$\partial_t u_i^*=\Delta\big(a(u_1^*/u_2^*)u_i^*\big),\quad t>0,\quad u_i^*(0)=u_i^0\quad\text{in }\mathbb{T}^d,\quad i=1,2,$$

and Theorem 2 shows that $u_i^*(t)\to\overline{u}_i$ in $L^2(\mathbb{T}^d)$ as $t\to\infty$, which translates to $\|e^{-\mu t}u_i(t)-\overline{u}_i\|_{L^2(\mathbb{T}^d)}\to 0$. $\square$

Acknowledgements

Open access funding provided by Austrian Science Fund (FWF).

Open Access. This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

[1] Amann, H.: Nonhomogeneous linear and quasilinear elliptic and parabolic boundary value problems. In: Schmeisser, H.J., Triebel, H. (eds.) Function Spaces, Differential Operators and Nonlinear Analysis, p. 9126. Teubner, Stuttgart (1993)

[2] Arnold, A., Markowich, P., Toscani, G., Unterreiter, A.: On convex Sobolev inequalities and the rate of convergence to equilibrium for Fokker–Planck type equations. Commun. Part. Differ. Eqs. **26**, 43–100 (2001)

[3] Chen, X., Daus, E., Jüngel, A.: Global existence analysis of cross-diffusion population systems for multiple species. Submitted for publication (2016). arXiv:1608.03696

[4] Clark, D.: Short proof of a discrete Gronwall inequality. Discrete Appl. Math. **16**, 279–281 (1987)

[5] Desvillettes, L., Lepoutre, T., Moussa, A., Trescases, A.: On the entropic structure of reaction-cross diffusion systems. Commun. Part. Differ. Eqs. **40**, 1705–1747 (2015)

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

[7] Lions, P.-L.: Some new classes of nonlinear Kolmogorov equations. Talk at the 16th Pauli Colloquium, Wolfgang-Pauli Institute, Vienna, November 18 (2015)

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

[9] Menz, S., Latorre, J., Schütte, C., Huisinga, W.: Hybrid stochastic-deterministic solution of the chemical master equation. SIAM Multiscale Model. Simul. **10**, 1232–1262 (2012)
[10] Øksendal, B.: Stochastic Differential Equations. Springer, Berlin (2003)
[11] Temam, R.: Infinite-Dimensional Dynamical Systems in Mechanics and Physics, 2nd edn. Springer, New York (1997)
[12] Villani, C.: Optimal Transport. Old and New. Springer, Berlin (2009)
[13] Zamponi, N., Jüngel, A.: Global existence analysis for degenerate energy-transport models for semiconductors. J. Differ. Eqs. **258**, 2339–2363 (2015)

Ansgar Jüngel and Nicola Zamponi
Institute for Analysis and Scientific Computing
Vienna University of Technology
Wiedner Hauptstraße 8-10
1040 Vienna
Austria
e-mail: juengel@tuwien.ac.at

Nicola Zamponi
e-mail: nicola.zamponi@tuwien.ac.at

(Received: January 19, 2016; revised: November 8, 2016)