Euro. Jnl of Applied Mathematics (2017), *vol.* 28, *pp.* 317–356.
doi:10.1017/S0956792516000346

A meeting point of entropy and bifurcations in cross-diffusion herding†

ANSGAR JÜNGEL[1], CHRISTIAN KUEHN[2] and LARA TRUSSARDI[1]

[1]*Institute for Analysis and Scientific Computing, Vienna University of Technology, Wiedner Hauptstraße 8–10, 1040 Wien, Austria*
emails: juengel@tuwien.ac.at, lara.trussardi@tuwien.ac.at
[2]*Faculty of Mathematics, Technical University of Munich, Boltzmannstraße 3, 85748 Garching, Germany*
email: ckuehn@ma.tum.de

(*Received 28 May 2015; revised 13 July 2016; accepted 14 July 2016; first published online 15 August 2016*)

A cross-diffusion system modelling the information herding of individuals is analysed in a bounded domain with no-flux boundary conditions. The variables are the species' density and an influence function which modifies the information state of the individuals. The cross-diffusion term may stabilize or destabilize the system. Furthermore, it allows for a formal gradient-flow or entropy structure. Exploiting this structure, the global-in-time existence of weak solutions and the exponential decay to the constant steady state is proved in certain parameter regimes. This approach does not extend to all parameters. We investigate local bifurcations from homogeneous steady states analytically to determine whether this defines the validity boundary. This analysis shows that generically there is a gap in the parameter regime between the entropy approach validity and the first local bifurcation. Next, we use numerical continuation methods to track the bifurcating non-homogeneous steady states globally and to determine non-trivial stationary solutions related to herding behaviour. In summary, we find that the main boundaries in the parameter regime are given by the first local bifurcation point, the degeneracy of the diffusion matrix and a certain entropy decay validity condition. We study several parameter limits analytically as well as numerically, with a focus on the role of changing a linear damping parameter as well as a parameter controlling the cross-diffusion. We suggest that our paradigm of comparing bifurcation-generated obstructions to the parameter validity of global-functional methods could also be of relevance for many other models beyond the one studied here.

Key words: Information herding, entropy method, large-time dynamics of solutions, numerical continuation, bifurcation

2000 Mathematics Subject Classification: 35K57, 35K20, 35B40, 35Q91

† AJ and LT acknowledge partial support from the European Union in the FP7-PEOPLE-2012-ITN Program under Grant Agreement Number 304617, the Austrian Science Fund (FWF), grants P22108, P24304, W1245, and the Austrian-French Program of the Austrian Exchange Service (ÖAD). CK acknowledges partial support by an APART fellowship of the Austrian Academy of Sciences (ÖAW) and by a Marie-Curie International Reintegration Grant by the EU/REA (IRG 271086).

1 Introduction

In this paper, we study the following cross-diffusion system:

$$\partial_t u_1 = \operatorname{div}(\nabla u_1 - g(u_1)\nabla u_2), \tag{1.1}$$

$$\partial_t u_2 = \operatorname{div}(\delta\nabla u_1 + \kappa\nabla u_2) + f(u_1) - \alpha u_2, \tag{1.2}$$

where $u_1 = u_1(t,x)$, $u_2 = u_2(t,x)$ for $(t,x) \in [0,T) \times \Omega$, $T > 0$ is the final time, $\Omega \subset \mathbb{R}^d$ ($d \geqslant 1$) is a bounded domain with sufficiently smooth boundary, ∇ denotes the gradient, $\operatorname{div} = \nabla\cdot$ is the divergence and $\partial_t = \frac{\partial}{\partial t}$ denotes the partial derivative with respect to time. The equations are supplemented by no-flux boundary conditions and suitable initial conditions:

$$\begin{aligned} (\nabla u_1 - g(u_1)\nabla u_2)\cdot\nu &= 0 \\ (\delta\nabla u_1 + \kappa\nabla u_2)\cdot\nu &= 0 \end{aligned} \quad \text{on } \partial\Omega,\ t > 0, \quad u_1(0,x) = u_1^0,\ u_2(0,x) = u_2^0 \quad \text{in } \Omega, \tag{1.3}$$

where ν denotes the outer unit normal vector to $\partial\Omega$. The function $u_1(x,t) \in [0,1]$ represents the density of individuals with information variable $x \in \Omega$ at time $t \geqslant 0$, and $u_2(x,t)$ is an influence function which modifies the information state of the individuals and possibly may lead to a herding (or aggregation) behaviour of individuals. The influence function acts through the term $g(u_1)\nabla u_2$ in (1.1). The non-negative bounded function $g(u_1)$ is assumed to vanish only at $u_1 = 0$ and $u_1 = 1$, which provides the bound $0 \leqslant u_1 \leqslant 1$ if $0 \leqslant u_1(0,x) \leqslant 1$. In particular, we assume that the influence becomes weak if the number of individuals at fixed $x \in \Omega$ is very low or close to the maximal value $u_1 = 1$, i.e., $g(0) = 0$ and $g(1) = 0$, which may enhance herding behaviour. The influence function is assumed to be modified by diffusive effects also due to the random behaviour of the density of the individuals with parameter $\delta > 0$, by the non-negative source term $f(u_1)$, relaxation with time with rate $\alpha > 0$, and diffusion with coefficient $\kappa > 0$.

If $\delta = 0$, equations (1.1), (1.2) can be interpreted as a non-linear variant of the chemotaxis Patlak–Keller–Segel model [33], where the function u_2 corresponds to the concentration of the chemoattractant. The model with non-linear mobility $g(u_1)$ was first analysed by Hillen and Painter [26], even for more general mobilities of the type $u_1\beta(u_1)\chi(u_2)$. Generally, the mobility $g(u_1) = u_1(1-u_1)$ models finite-size exclusion and prevents blow-up phenomena [44], which are known in the original Keller–Segel model. The convergence to equilibrium was shown in [29]. Such models were also employed to describe evolution of large human crowds driven by the dynamic field u_2 [7].

Systems (1.1) and (1.2) are the two possible models to describe the dynamics of information herding in a macroscopic setting. There exist other approaches to model herding behaviour, for instance, using kinetic equations [16] or agent-based models [37], but the focus in this paper is to understand the influence of the parameters δ and α on the solution from a mathematical viewpoint, i.e., to investigate the interplay between cross-diffusion and damping.

Equations (1.1) and (1.2) with $\delta > 0$ can be derived from an interacting "particle" system modelled by stochastic differential equations, at least in the case $g(u_1) = \text{const.}$ (see [22]). One expects that this derivation can be extended to the case of non-constant $g(u_1)$ but we do not discuss this derivation here. The above system with $g(u_1) = u_1$ was

https://doi.org/10.1017/S0956792516000346 Published online by Cambridge University Press

analysed in [24] in the Keller–Segel context. The additional cross-diffusion with $\delta > 0$ in (1.2) was motivated by the fact that it prevents the blow up of the solutions in two space dimensions, even for large initial densities and for arbitrarily small values of $\delta > 0$. The motivation to introduce this term in our model is different since the non-linear mobility $g(u_1)$ allows us to conclude that $u_1 \in [0, 1]$, thus preventing blow up without taking into account the cross-diffusion term $\delta \Delta u_1$. Our aim is to investigate the solutions to (1.1) and (1.2) for *all* values for δ, thus allowing for *destabilizing* cross-diffusion parameters $\delta < 0$.

One starting point to investigate the dynamics is to consider the functional structure of the equation. In this context, entropy methods are a possible tool [28]. The entropy structure can frequently be used to establish the existence of (weak) solutions. Furthermore, it is helpful for a quantitative analysis of the large-time dynamics of solutions for certain reaction–diffusion systems; see, e.g., [14]. The method quantifies the decay of a certain functional with respect to a steady state. An advantage is that the entropy approach can work globally, even for initial conditions far away from steady states. Moreover, the entropy structure may be formulated in the variational framework of gradient flows which allows one to analyse the geodesic convexity of their solutions [**?**, 36]. However, this global view indicates already that we may not expect that the approach is valid for all parameters in general non-linear systems. Indeed, in many situations, global methods only work for a certain range of parameters occurring in the system. The question is what happens for parameter values outside the admissible parameter range and near the validity boundary.

One natural conjecture is that upon variation of a single parameter, there exists a *single* critical parameter value associated to a first local bifurcation point δ_b beyond which a global functional approach does not extend. In particular, the homogeneous steady state upon which the entropy is built, could lose stability and new solution branches may appear in parameter space. Another possibility is that global bifurcation branches in parameter space are an obstruction. In our context, the generic situation is different from the two natural conjectures.

In the context of (1.1) and (1.2), the *main distinguished parameter* we are interested in is δ. Here, we shall state our results on an informal level. Carrying out the existence of weak solutions and the global decay to homogeneous steady states:

$$u^* = (u_1^*, u_2^*)$$

via an entropy approach, we find the following results:

(M1) Using the entropy approach, one may prove the existence of weak solutions to (1.1) and (1.2) in certain parameter regimes.

(M2) The global entropy decay to equilibrium does not extend to arbitrary negative δ. Suppose we fix all other parameters, then there exists a critical δ_e (to be defined below) such that global decay occurs only for $\delta > \delta_e$ ($\delta \neq 0$).

(M3) If we consider the limit $\alpha \to +\infty$, then we can extend the global decay up to

$$\delta^* := -\kappa/\gamma < 0, \qquad \text{where } \gamma := \max_{v \in [0,1]} g(v),$$

https://doi.org/10.1017/S0956792516000346 Published online by Cambridge University Press

i.e., global exponential decay to a steady state occurs for all $\delta > \delta^*(\delta \neq 0)$ if α is large enough.

(M4) In the limit $\alpha \to 0$, we find that $\delta_e \to +\infty$. In particular, the entropy method breaks down in this limiting regime in the formulation presented here.

We stress that the results for the global decay (M2)–(M4) may not be sharp, in the sense that one could potentially improve the validity boundary δ_e. Interestingly, we shall prove below that (M3) is indeed sharp for certain steady states, i.e., no improvement is possible in this limit. The proofs of (M1)–(M4) provide a number of technical challenges, which are discussed in more detail in Sections 2.1 and 3. We also note that the entropy method definitely does not extend to any negative δ. It is clear that a global decay to a homogeneous steady state for all initial conditions is impossible if bifurcating non-homogeneous steady state solutions exist as well. We use analytical local bifurcation theory for the stationary problem, based upon a modification of Crandall–Rabinowitz theory [31], to prove the following:

(M5) The bifurcation approach for homogeneous steady states can be carried out as long as

$$\delta \neq \delta_d := -\kappa/g(u_1^*).$$

On a generic open and connected domain, local bifurcations of simple eigenvalues occur for

$$\delta_b^n = \delta_d + \frac{1}{\mu_n}\left[f'(u_1^*) - \frac{\alpha}{g(u_1^*)}\right],$$

where μ_n are the eigenvalues of the negative Neumann Laplacian.

(M6) If $\alpha > 0$ is sufficiently *large* and fixed, $\delta_b^n < \delta_d < \delta^*$ and the bifurcation points accumulate at δ_d.

(M7) If $\alpha > 0$ is sufficiently *small* and fixed, $\delta_d < \delta_b^n$ and the bifurcation points again accumulate at δ_d.

Although these results are completely consistent with the global decay of the entropy functional, they do not yield global information about the bifurcation curves. In general, it is not possible to analytically characterize all global bifurcation for arbitrary non-linear systems. Therefore, we consider numerical continuation of the non-homogeneous steady-state solution branches (for spatial dimension $d = 1$). The continuation is carried out using `AUTO` [12]. Our numerical results show the following:

(M8) We regularize the numerical problem using a small parameter ρ to avoid higher dimensional bifurcation surfaces due to mass conservation.

(M9) The non-homogeneous steady-state bifurcation branches starting at the local bifurcation points extend in parameter space and contain multi-bump solutions, which deform into more localized (herding) states upon changing parameters.

(M10) A second continuation run considering $\rho \to 0$ yields non-trivial solutions for the original problem. In particular, solutions may have multiple transition layers (respectively concentration regions) and the ones with very few layers occupy the largest ranges in δ-parameter space.

https://doi.org/10.1017/S0956792516000346 Published online by Cambridge University Press

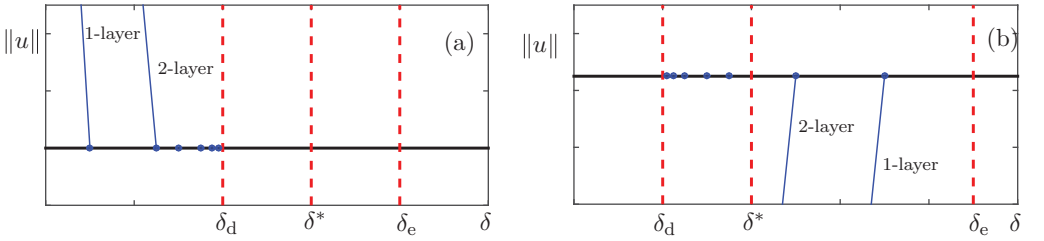

FIGURE 1. Sketch of the different bifurcation scenarios; for more detailed numerical calculations see Section 5. Only the main parameter δ is varied, a homogeneous branch is shown in black and bifurcation points and branches in blue (dots and curves). Only the first two non-trivial branches are sketched which contain solutions with one transition layer. (a) Case (C1) with $\alpha > 0$ sufficiently large; for a suitable choice of u^* and $\alpha \to +\infty$ all three vertical dashed red lines collapse onto one line. (b) Case (C2) with $\alpha > 0$ sufficiently small.

Combining all the results we conclude that we have the situations in Figure 1(a) and (b) for generic fixed parameter values and a generic fixed domain. These two main cases of interest are:

(C1) $\alpha > 0$ *sufficiently large*: In this limit, the entropy validity boundary, the analytical bifurcation approach and the numerical methods are organized around the singular limit at $\delta = \delta^*$. Indeed, note that

$$\delta^* = \delta_d, \quad \text{if } u = u_1^* \text{ maximizes } g(u) \text{ on } [0,1],$$

and we show below that $\delta_e \to \delta^*$ as $\alpha \to +\infty$. The generic picture for a homogeneous steady state so that u_1^* does not maximize g and α is moderate and fixed is given in Figure 1(a).

(C2) $\alpha > 0$ *sufficiently small*: In this case, the generic picture is shown in Figure 1(b). The entropy decay only occurs for very large values $\delta > \delta_e$. Interestingly, the approaches do not seem to collapse onto one singular limit in this case.

We remark that the condition $\kappa \neq -\delta g(u_1)$ does not only occur in the numerical continuation analysis. It occurs in the context of the entropy method as well as the analytical bifurcation calculation. It is precisely the condition for the vanishing of the determinant of the diffusion matrix that prevents pushing existence and decay techniques based upon global functionals further. The condition also prevents analytical bifurcation theory to work as the linearized problem does not yield a Fredholm operator. In some sense, this explains the singular limit as $\alpha \to +\infty$ from (C1). Although (C1) is quite satisfactory from a mathematical perspective, one drawback is that the forward problem may not be well-posed in a classical sense if $\delta < \delta_d$; of course, the stationary problem is s still well-defined.

For (C2), we cannot prove sharp global decay via an entropy functional. However, the first non-trivial branch of locally stable stationary herding solutions can be reached in forward time via a classical well-posed problem, and (C1) and (C2) always make sense for adiabatic parameter variation. Although we postpone the detailed mathematical study of

https://doi.org/10.1017/S0956792516000346 Published online by Cambridge University Press

the limit $\alpha \to 0$ to future work, the observations raise several interesting problems, which we discuss in the outlook at the end of this paper.

In summary, the main contribution of this work is to study the interplay between three different techniques available for reaction–diffusion systems with cross-diffusion: entropy methods, analytical local bifurcation and numerical global bifurcation theory. Furthermore, for each technique, we have to use, improve and apply the previously available methods to the herding model problem (1.1)–(1.3). Our results lead to clear insight on the subdivision of parameter space into regimes, where each method is particularly well-suited to describe the system dynamics. We identify two interesting singular limits and provide a detailed analysis for the limit of large damping. Furthermore, we compute via numerical continuation several solutions that are of interest for applications to herding behaviour using a two-parameter homotopy approach to desingularize the mass conservation. From an application perspective, we identify herding states with clustering of individuals in one, or just a few, distinct regions, as the ones occupying the largest parameter ranges. Hence, we expect applications to be governed by homogeneous stationary and relatively simple heterogeneous herding states.

There seem to be very few works [1,20] studying the parameter space interplay between global entropy-structure methods in comparison to local analytical and global numerical bifurcation approaches. Our work seems to be, to the best of our knowledge, the first analysis combining and comparing all three methods, and also the first to consider the global-functional and bifurcations interaction problem for cross-diffusion systems. In fact, our analysis suggests a general paradigm to improve our understanding of global methods for non-linear spatio-temporal systems, i.e., one major goal is to determine the parameter space *validity boundaries* between *different methods*.

The paper is organized as follows. In Section 2, we state our main results and provide an overview of the strategy for the proofs respectively the numerical methods employed. In particular, the entropy method results are considered in Section 2.1, the analytical local bifurcation in Section 2.2, and the numerical global bifurcation results in Section 2.3. The following sections contain the full details for the main results. The proofs using the entropy method are contained in Section 3, where the weak solution construction is carried out in Section 3.1 and the global decay is proved in Section 3.2. Section 4 proves the existence of local bifurcation points to non-trivial solutions upon decreasing δ. The details for the global numerical continuation results are reported in Section 5. We conclude in Section 6 with an outlook, where we discuss several open questions.

Notation: When operating with vectors we view them as column vectors and use $(\cdot)^\top$ to denote the transpose. We use the standard notation for L^p-spaces, $W^{k,p}$ for the Sobolev space with (weak) derivatives up to and including order k in L^p as well as the shorthand notation $W^{k,2} = H^k$; see [18] for details. Furthermore, $'$ denotes the associated dual space, when applied to a function space.

2 Main results

We describe the main results of this paper, obtained by either the entropy method or local analytical and global numerical bifurcation analysis.

https://doi.org/10.1017/S0956792516000346 Published online by Cambridge University Press

2.1 Entropy method

First, we show the global existence of weak solutions and their large-time decay to equilibrium. We observe that the diffusion matrix of systems (1.1) and (1.2) are neither symmetric nor positive definite which complicates the analysis. Local existence of (smooth) solutions follows from Amann's results [3] if the system is parabolic in the sense of Petrovskii, i.e., if the real parts of the eigenvalues of the diffusion matrix are positive. A sufficient condition for this statement is $\delta \geqslant \delta_{\mathrm{d}} = -\kappa/\gamma$. The challenge here is to prove the existence of *global* (weak) solutions.

The main challenge of (1.1) and (1.2) is that the diffusion matrix of the system is neither symmetric nor positive definite. The key idea of our analysis, similar as in [24], is to define a suitable entropy functional. The entropy is a special Lyapunov functional which provides suitable gradient estimates. Compared to Lyapunov functional techniques like in [25,43] (used for the case $\delta = 0$), the entropy method provides explicit decay rates and, in our case, L^∞ bounds without the use of a maximum principle. (Note that in the system at hand, the L^∞ bounds can be obtained by the standard maximum principle but there are systems where this can be achieved by using the entropy method only; see [28].) For this, we introduce the entropy density

$$h(u) = h_0(u_1) + \frac{u_2^2}{2\delta_0}, \quad u = (u_1, u_2)^\top \in [0,1] \times \mathbb{R},$$

where h_0 is defined as the second anti-derivative of $1/g$,

$$h_0(s) := \int_m^s \int_m^\sigma \frac{1}{g(t)}\,\mathrm{d}t\,\mathrm{d}\sigma, \quad s \in (0,1), \tag{2.1}$$

where $0 < m < 1$ is a fixed number, and

$$\delta_0 := \delta \quad \text{if } \delta > 0, \quad \delta_0 := \kappa/\gamma \quad \text{if } -\kappa/\gamma < \delta < 0.$$

It turns out that the so-called entropy variables $w = (w_1, w_2)^\top$ with $w_1 = h_0'(u_1)$ and $w_2 = u_2/\delta_0$ make the diffusion matrix positive semi-definite for all $\delta > \delta^* := -\kappa/\gamma$, $\delta \neq 0$. We remark that for $\delta = 0$, the method does not work and we do not cover this case. In the w-variables, we can formulate (1.1) and (1.2) equivalently as

$$\partial_t u = \mathrm{div}(B(w)\nabla w) + F(u),$$

where $u = u(w)$, $F(u) = (0, f(u_1) - \alpha u_2)^\top$ and

$$B(w) = \begin{pmatrix} g(u_1) & -\delta_0 g(u_1) \\ \delta g(u_1) & \delta_0 \kappa \end{pmatrix}. \tag{2.2}$$

The invertibility of the mapping $w \mapsto u(w)$ is guaranteed by Hypothesis (H3) below. We show in Lemma 4 below that $B(w)$ is positive semi-definite if $\delta > \delta^*$, $\delta \neq 0$. The global existence is based on the fact that the entropy

$$H(u(t)) = \int_\Omega \left(h_0(u_1(t)) + \frac{u_2(t)^2}{2\delta_0} \right) \mathrm{d}x \tag{2.3}$$

https://doi.org/10.1017/S0956792516000346 Published online by Cambridge University Press

is bounded on $[0,T]$ for any $T>0$; note that we write $u=u(t)$ here to emphasize the time dependence of H. A formal computation, which is made rigorous in Section 3.1, shows that

$$\frac{\mathrm{d}H}{\mathrm{d}t} = -\int_\Omega \left(\frac{|\nabla u_1|^2}{g(u_1)} + \left(\frac{\delta}{\delta_0}-1\right)\nabla u_1\cdot\nabla u_2 + \frac{\kappa}{\delta_0}|\nabla u_2|^2\right)\mathrm{d}x + \frac{1}{\delta_0}\int_\Omega (f(u_1)-\alpha u_2)u_2\,\mathrm{d}x. \tag{2.4}$$

The terms in the first bracket define a positive definite quadratic form if and only if $\delta>\delta^*$. The second integral is bounded since $f(u_1)$ is bounded. This shows that for some $\varepsilon_1(\delta)>0$,

$$\frac{\mathrm{d}H}{\mathrm{d}t} \leqslant -\varepsilon_1(\delta)\int_\Omega \left(\frac{|\nabla u_1|^2}{g(u_1)} + \frac{|\nabla u_2|^2}{\delta_0^2}\right)\mathrm{d}x + c, \tag{2.5}$$

where the constant $c>0$ depends on Ω, f and α. These gradient bounds are essential for the existence analysis.

Before we state the existence theorem, we make our assumptions precise:

(H1) $\Omega\subset\mathbb{R}^d$ with $\partial\Omega\in C^2$, $\alpha>0$, $\kappa>0$, $h(u^0)\in L^1(\Omega)$ with $u_1^0\in(0,1)$ a.e.

(H2) $f\in C^0([0,1])$ is non-negative.

(H3) $g\in C^2([0,1])$ is positive on $(0,1)$, $g(0)=g(1)=0$, $g(u)\leqslant\gamma$ for $u\in[0,1]$, where $\gamma>0$, and $\int_0^m \mathrm{d}s/g(s)=\int_m^1 \mathrm{d}s/g(s)=+\infty$ for some $0<m<1$.

The condition $g(u)\leqslant\gamma$ in $[0,1]$ in (H3) implies that $(u_1^0-m)^2/(2\gamma)\leqslant h_0(u_1^0)$ and hence, $h(u^0)\in L^1(\Omega)$ in (H1) yields $u_1^0\in L^2(\Omega)$ and $u_2^0\in L^2(\Omega)$. Hypothesis (H3) ensures that the function h_0 defined in (2.1) is well defined and of class C^4 (needed in Lemma 5). Its derivative h_0' is strictly increasing on $(0,1)$ with range $\mathbb{R}$, thus being invertible with inverse $(h_0')^{-1}:\mathbb{R}\to(0,1)$. For instance, the function $g(s)=s(1-s)$, $s\in[0,1]$, satisfies (H3) and $h_0(s)=s\log s+(1-s)\log(1-s)$, where log denotes the natural logarithm. A more general class of functions fulfilling (H3) is $g(s)=s^a(1-s)^b$ with $a,b\geqslant 1$.

Theorem 1 (Global existence) *Let assumptions (H1)–(H3) hold and let $\delta>-\kappa/\gamma$. Then, there exists a weak solution to* (1.1)–(1.3) *satisfying $0\leqslant u_1\leqslant 1$ in Ω, $t>0$ and*

$$u_1,\,u_2\in L^2_{\mathrm{loc}}(0,\infty;H^1(\Omega)),\quad \partial_t u_1,\,\partial_t u_2\in L^2_{\mathrm{loc}}(0,\infty;H^1(\Omega)').$$

The initial datum is satisfied in the sense of $H^1(\Omega;\mathbb{R}^2)'$.

We provide a brief overview of the proof. First, we discretize the equations in time using the implicit Euler scheme, which keeps the entropy structure. Since we are working in the entropy-variable formulation, we need to regularize the equations in order to be able to apply the Lax-Milgram lemma for the linearized problem. The existence of solutions to the non-linear problem then follows from the Leray–Schauder theorem, where the uniform estimate is a consequence of the entropy inequality (2.5). This estimate also provides bounds uniform in the approximation parameters. A discrete Aubin lemma in the

https://doi.org/10.1017/S0956792516000346 Published online by Cambridge University Press

version of [15] provides compactness, which allows us to perform the limit of vanishing approximation parameters.

Although the proof is similar to the existence proofs in [24, 28], the results of these papers are not directly applicable since our situation is more general than in [24, 28]. The main novelties of our existence analysis are the new entropy (2.3) and the treatment of destabilizing cross-diffusion ($\delta < 0$).

For the analysis of the large-time asymptotics, we introduce the constant steady state $u^* = (u_1^*, u_2^*)$, where

$$u_1^* = \overline{u}_1^0, \quad u_2^* = \frac{f(u_1^*)}{\alpha}, \qquad \text{with } \overline{u}_j^0 := \frac{1}{\mathrm{m}(\Omega)} \int_\Omega u_j^0(x)\,\mathrm{d}x,\ j \in \{1,2\},$$

and $\mathrm{m}(\Omega)$ denotes the Lebesgue measure of Ω. Furthermore, we define the relative entropy

$$H(u|u^*) = \int_\Omega h(u|u^*)\,\mathrm{d}x$$

with the entropy density

$$h(u|u^*) = h_0(u_1|u_1^*) + \frac{1}{2\delta_0}(u_2 - u_2^*)^2, \quad \text{where } h_0(u_1|u_1^*) = h_0(u_1) - h_0(u_1^*). \tag{2.6}$$

Note that u_1 conserves mass, i.e., $\overline{u}_1(t) := \mathrm{m}(\Omega)^{-1} \int_\Omega u_1(t)\,\mathrm{d}x$ is constant in time and $\overline{u}_1(t) = u_1^*$ for all $t > 0$. Thus, by Jensen's inequality, $h_0(u_1|u_1^*) \geqslant 0$.

Theorem 2 (Exponential decay) *Let assumptions (H1)–(H3) hold, let Ω be convex, let f be Lipschitz continuous with constant $c_L > 0$, and let*

$$\delta_0 \varepsilon_1(\delta) > \frac{\gamma}{\alpha} c_L^2 c_S, \tag{2.7}$$

where $\varepsilon_1(\delta) > 0$ and $c_S > 0$ are defined in Lemmas 4 and 5, respectively. Then, for $t > 0$,

$$H(u(t)|u^*) \leqslant \mathrm{e}^{-\chi(\delta)t} H(u^0|u^*), \quad \textit{where} \quad \chi(\delta) := \min\left\{ \frac{\varepsilon_1(\delta)}{c_S} - \frac{\gamma c_L^2}{\alpha \delta_0}, \alpha \right\} > 0. \tag{2.8}$$

Moreover, it holds for $t > 0$,

$$\|u_1(t) - u_1^*\|_{L^2(\Omega)} + \|u_2(t) - u_2^*\|_{L^2(\Omega)} \leqslant 2\sqrt{\max\{\gamma, \delta\} H(u^0|u^*)}\,\mathrm{e}^{-\chi(\delta)t/2}. \tag{2.9}$$

Recall that $\delta_0 = \kappa/\gamma$ if $\delta < 0$ and $\delta_0 = \delta$ if $\delta > 0$. The values for $\delta_0 \varepsilon_1(\delta)$ are illustrated in Figure 2. It turns out that (2.7) is fulfilled if either the additional diffusion $\delta > 0$ is sufficiently large or if γ/α is sufficiently small. The latter condition means that the influence of the drift term $g(u_1)\nabla u_2$ is "small" or that the relaxation $-\alpha u_2$ is "strong". The theorem states that in all these cases, the diffusion is sufficiently strong to lead to exponential decay to equilibrium. For all parameters fixed, except δ, we conclude from the condition (2.7) that there exists a δ_e such that exponential decay holds for $\delta > \delta_\mathrm{e}$

https://doi.org/10.1017/S0956792516000346 Published online by Cambridge University Press

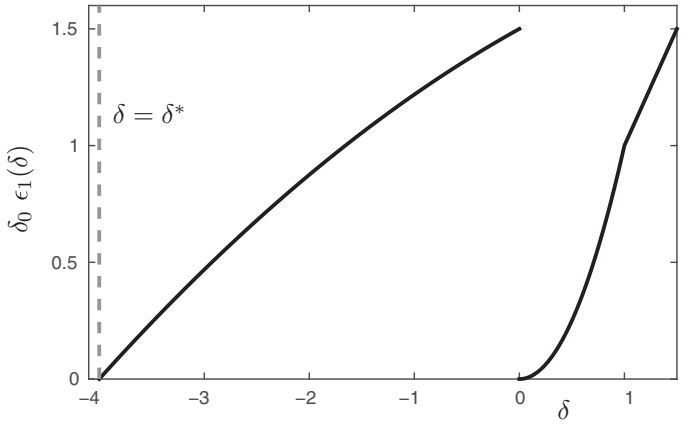

FIGURE 2. Illustration of $\delta_0\varepsilon_1(\delta)$ for $\kappa = 1$ and $\delta = \frac{1}{4}$ (black curves). The corresponding singular limit $\delta^* = -\kappa/\gamma = -4$ is also marked (grey-dashed vertical line).

($\delta \neq 0$) and we see that

$$\lim_{\alpha\to+\infty} \delta_\mathrm{e} = \delta^* = -\kappa/\gamma$$

as a singular limit already discussed above. We remark that the exclusion of the decay for $\delta = 0$ seems to be purely technical and we conjecture that exponential decay also holds for $\delta = 0$. On the contrary, extensions to $\alpha \to 0$ are highly non-trivial and we can currently not cover this degenerate limiting case using entropy methods.

Theorem 2 is proved by differentiating the relative entropy $H(u|u^*)$ with respect to time, similar as in (2.4). We wish to estimate the gradient terms from below by a multiple of $H(u|u^*)$. The convex Sobolev inequality from Lemma 5 shows that the L^2-norm of $g(u_1)^{1/2}\nabla u_1$ is estimated from below by $\int_\Omega h_0(u_1|u_1^*)\,\mathrm{d}x$, up to a factor. The L^2-norm of ∇u_2 is estimated from below by a multiple of $\int_\Omega (u_2 - \overline{u}_2)^2\,\mathrm{d}x$, using the Poincaré inequality. However, the variable u_2 generally does not conserve mass and in particular, $\overline{u}_2 \neq u_2^*$. We exploit instead the relaxation term in (1.2) to achieve the estimate

$$H(u(t)|u^*) + \chi(\delta)\int_0^t H(u(s)|u^*)\,\mathrm{d}s \leqslant 0.$$

Then Gronwall's lemma gives the result. The difficulty is the estimate of the source term $f(u_1)$. This problem is overcome by controlling the expression involving $f(u_1)$ by taking into account the contribution coming from the convex Sobolev inequality. However, we need that δ is sufficiently large, i.e., cross-diffusion has to dominate reaction.

The above arguments hold on a formal level only. A second difficulty is to make these arguments rigorous since we need the test function $h_0'(u_1) - h_0'(u_1^*)$, which is undefined if $u_1 = 0$ or $u_1 = 1$ (since $h_0'(0) = -\infty$ and $h_0'(1) = +\infty$ by Hypothesis (H3)). The idea is to perform a transformation of variables in terms of so-called entropy variables which ensure that $0 < u_1 < 1$ in a time-discrete setting. Passing from the semi-discrete to the continuous case, the variable u_1 may satisfy $0 \leqslant u_1 \leqslant 1$ in the limit.

https://doi.org/10.1017/S0956792516000346 Published online by Cambridge University Press

2.2 Analytical bifurcation analysis

As outlined in the introduction, the first natural conjecture for the failure of the entropy method is to study bifurcations of the homogeneous steady states $u^* = (u_1^*, u_2^*)$, which solve

$$\begin{aligned} 0 &= \operatorname{div}(\nabla u_1 - g(u_1)\nabla u_2), \\ 0 &= \operatorname{div}(\delta \nabla u_1 + \kappa \nabla u_2) + f(u_1) - \alpha u_2, \end{aligned} \tag{2.10}$$

with the no-flux boundary conditions (1.3). To study the bifurcations of u^* under variation of δ, we use the right-hand side of (2.10) to define a bifurcation function and apply the theory of Crandall–Rabinowitz [10, 31]. The problem is that u^* is not an *isolated* bifurcation branch as a function of δ since fixing any initial mass yields a different one-dimensional family of homogeneous steady states with

$$u_1^* = \frac{1}{\mathrm{m}(\Omega)} \int_\Omega u_1(x)\, \mathrm{d}x \geqslant 0. \tag{2.11}$$

Hence, the standard approach has to be modified and we follow arguments that can be found in [9, 39, 45]. It is helpful to introduce some notations first. For $p > d$, let

$$\begin{aligned} \mathcal{X} &:= \{u \in W^{2,p}(\Omega) : \nabla u \cdot \nu = 0 \text{ on } \partial\Omega\}, \\ \mathcal{Y} &:= L^p(\Omega), \\ \mathcal{Y}_0 &:= \left\{u_1 \in L^p(\Omega) : \textstyle\int_\Omega u_1(x)\, \mathrm{d}x = 0\right\}, \end{aligned} \tag{2.12}$$

where the space $\mathcal{X}$ includes standard Neumann boundary conditions. Due to the Sobolev embedding theorem, we know that $W^{2,p}(\Omega)$ is continuously embedded in $C^{(1+\theta)}(\bar{\Omega})$ for some $\theta \in (0, 1)$. If Neumann boundary conditions hold, then our original boundary conditions (1.3) hold as well. However, the converse is only true if we can invert the diffusion matrix, i.e., as long as $\delta \neq \delta_d = -\frac{\kappa}{g(u_1)}$. In particular, we shall always assume for the local bifurcation analysis of homogeneous steady states that

$$\delta \neq \delta_{\mathrm{d}} = -\frac{\kappa}{g(u_1^*)}. \tag{2.13}$$

This implies that we may not find all possible bifurcations and the single point when the diffusion matrix vanishes has to be treated separately; we leave this as a goal for future work.

Next, we define the mapping $\mathcal{F} : \mathcal{X} \times \mathcal{X} \times \mathbb{R} \longrightarrow \mathcal{Y}_0 \times \mathcal{Y} \times \mathbb{R}$ by

$$\mathcal{F}(u_1, u_2, \delta) := \begin{pmatrix} \operatorname{div}(\nabla u_1 - g(u_1)\nabla u_2) \\ \delta \Delta u_1 + \kappa \Delta u_2 - \alpha u_2 + f(u_1) \\ \int_\Omega u_1(x)\, \mathrm{d}x - \mathrm{m}(\Omega) u_1^* \end{pmatrix}. \tag{2.14}$$

The first two terms are the usual bifurcation functions one would naturally define, the third term is used to isolate the bifurcation branch for the mapping $\mathcal{F}$, i.e., to avoid the problem with mass conservation, whilst the last two terms take into account the boundary conditions. We know that there exists a family of homogeneous steady state solutions:

$$\mathcal{F}(u_1^*, u_2^*, \delta) = 0$$

https://doi.org/10.1017/S0956792516000346 Published online by Cambridge University Press

for each $\delta \in \mathbb{R}$. The goal is to find the parameter values δ_b such that at $\delta = \delta_b$ a non-trivial (or non-homogeneous) branch of steady states is generated at the bifurcation point; see also Figure 1. We are going to check that $\mathcal{F}$ is C^1-smooth and the Fréchet derivative $D_u\mathcal{F}$ with respect to u at a point $\tilde{u} = (\tilde{u}_1, \tilde{u}_2)$ is given by

$$\mathcal{A}_\delta(\tilde{u})\begin{pmatrix} U_1 \\ U_2 \end{pmatrix} := D_u\mathcal{F}(\tilde{u}, \delta)\begin{pmatrix} U_1 \\ U_2 \end{pmatrix} = \begin{pmatrix} \Delta U_1 - \operatorname{div}[g'(\tilde{u}_1)(\nabla \tilde{u}_2)U_1 + g(\tilde{u}_1)\nabla U_2] \\ \delta \Delta U_1 + \kappa \Delta U_2 - \alpha U_2 + f'(\tilde{u}_1)U_1 \\ \int_\Omega U_1(x)\,dx \end{pmatrix}, \quad (2.15)$$

where $(U_1, U_2)^\top \in \mathcal{X} \times \mathcal{X}$ and $\mathcal{A}_\delta : \mathcal{X} \times \mathcal{X} \to \mathcal{Y}_0 \times \mathcal{Y} \times \mathbb{R}$. We already know from Theorem 2 that for all $\delta > \delta_e$ ($\delta \neq 0$), the homogeneous steady state u^* is globally stable. Clearly, this implies local stability as well and this fact can also be checked by studying the spectrum of $\mathcal{A}_\delta(u^*)$. From the structure of the cross-diffusion equations (1.1) and (1.2), one does expect destabilization of the homogeneous state upon decreasing δ.

Theorem 3 *Let $u^* = (u_1^*, u_2^*)$ be a homogeneous steady state, consider the generic parameter case with $-\kappa \neq \delta g(u_1^*)$ and suppose all eigenvalues μ_n of the negative Neumann Laplacian on Ω are simple. Then, the following hold:*

(R1) *$D_u\mathcal{F}(\tilde{u}, \delta) : \mathcal{X} \times \mathcal{X} \to \mathcal{Y}_0 \times \mathcal{Y} \times \mathbb{R}$ is a Fredholm operator with index zero;*

(R2) *there exists a sequence of bifurcation points $\delta = \delta_b^n$ such that $\dim\left(\mathcal{N}[D_u\mathcal{F}(u^*, \delta_b^n)]\right) = 1$, where $\mathcal{N}[\cdot]$ denotes the nullspace;*

(R3) *there exist simple real eigenvalues $\lambda_n(\delta)$ of $\mathcal{A}_\delta(u^*)$, which satisfy $\lambda_n(\delta_b^n) = 0$. Furthermore, $\lambda_n(\delta)$ crosses the imaginary axis at δ_b^n with non-zero speed, i.e., $D_{\delta u}F(u^*, \delta_b^n)e_b^n \notin \mathcal{R}[\mathcal{A}_{\delta_b^n}]$, where $\mathcal{R}[\cdot]$ denotes the range and $\operatorname{span}[e_b^n] = \mathcal{N}[\mathcal{A}_{\delta_b^n}]$.*

The results from (R1)–(R3) hold quite generically (i.e., for $\delta \neq \delta_d$ and for generic domains [41]) and yield, upon applying a standard result by Crandall–Rabinowitz [10,11, 31], the existence of branches of non-trivial solutions:

$$(u_1[s], u_2[s], \delta[s]) \in \mathcal{X} \times \mathcal{X} \times \mathbb{R}, \qquad (u_1[0], u_2[0], \delta[0]) = (u_1^*, u_2^*, \delta_b^n),$$

where $s \in [-s_0, s_0]$ parametrizes the steady-state branch locally for some small $s_0 > 0$, and $(u_1[s], u_2[s], \delta[s]) \neq (u_1^*, u_2^*, \delta_b^n)$ for $s \in [-s_0, 0) \cup (0, s_0]$. Slightly more precise information about the branch can be obtained using the eigenfunction e_b and we refer to Section 4 for the details. The main conclusion of the bifurcation theorem is that we know that the entropy method cannot show the decay to steady state for all parameter regions. However, to track the non-trivial solution branches in parameter space, it is usually not possible to compute the global shape of all bifurcation branches analytically. In this case, numerical bifurcation analysis is extremely helpful.

2.3 Numerical bifurcation analysis

The results from Sections 2.1 and 2.2 do not provide a full exploration of the dynamical structure of the solutions for the parameter regime $\delta < \delta^*$. To understand this regime

https://doi.org/10.1017/S0956792516000346 Published online by Cambridge University Press

better, we study the bifurcations of (2.10) numerically for

$$f(s) = s(1-s), \qquad g(s) = s(1-s), \qquad s \in \Omega = [0,l] \subset \mathbb{R}, \tag{2.16}$$

for some interval length $l > 0$. Note that this yields a boundary-value problem involving two second-order ordinary differential equations (ODEs):

$$0 = \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\mathrm{d}u_1}{\mathrm{d}x} - g(u_1)\frac{\mathrm{d}u_2}{\mathrm{d}x}\right), \tag{2.17}$$

$$0 = \delta\frac{\mathrm{d}^2u_1}{\mathrm{d}x^2} + \kappa\frac{\mathrm{d}^2u_2}{\mathrm{d}x^2} - \alpha u_2 + f(u_1), \tag{2.18}$$

with boundary conditions:

$$0 = \frac{\mathrm{d}u_1}{\mathrm{d}x}(0) - g(u_1(0))\frac{\mathrm{d}u_2}{\mathrm{d}x}(0), \qquad 0 = \delta\frac{\mathrm{d}u_1}{\mathrm{d}x}(0) + \kappa\frac{\mathrm{d}u_2}{\mathrm{d}x}(0), \tag{2.19}$$

$$0 = \frac{\mathrm{d}u_1}{\mathrm{d}x}(1) - g(u_1(1))\frac{\mathrm{d}u_2}{\mathrm{d}x}(1), \qquad 0 = \delta\frac{\mathrm{d}u_1}{\mathrm{d}x}(1) + \kappa\frac{\mathrm{d}u_2}{\mathrm{d}x}(1). \tag{2.20}$$

An excellent available tool to study the problem (2.17)–(2.20) is the software `AUTO` [12] for numerical continuation of boundary-value problems; for other possible options and extensions, we refer to the discussion in Section 6. `AUTO` is precisely designed to deal with boundary-value problems for ODEs of the form

$$\frac{\mathrm{d}z}{\mathrm{d}x} = F(z;p), \qquad x \in [0,1], \quad G(w(0),w(1)) = 0, \tag{2.21}$$

where $F : \mathbb{R}^N \times \mathbb{R}^P \to \mathbb{R}^N$, $G : \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}^N$ and $p \in \mathbb{R}^P$ are parameters and $z = z(x) \in \mathbb{R}^N$ is the unknown vector. It is easy to re-write (2.17)–(2.20) as a system in the form (2.21) of four first-order ODEs, i.e., we get $N = 4$, consider the scaling $\tilde{x} = x/l$ to normalize the interval length to one, then drop the tilde for x again, and let

$$p_1 := \delta, \quad p_2 := \kappa, \quad p_3 := \alpha, \quad p_4 := l,$$

so $P = 4$ with primary bifurcation parameter δ. For more background on `AUTO` and on numerical continuation, we refer to [21, 30, 32]. In the setup (2.21), one can numerically continue the family of homogeneous solutions:

$$(u^*, \delta) = (u_1^*, u_2^*, \delta)$$

as a function of δ, i.e., to compute $u^* = u^*(\cdot;\delta)$ for δ in some specified parameter interval. Although this calculation yields bifurcation points for some δ values, it is not straightforward to use the formulation (2.17) and (2.18) to switch onto the non-homogeneous solution branches generated at the bifurcation point. The problem is due to the mass conservation since

$$\overline{u}_1 = \mathrm{m}(\Omega)^{-1}\int_\Omega u_1 \,\mathrm{d}x = u_1^*, \qquad u_2^* = \frac{f(u_1^*)}{\alpha}$$

https://doi.org/10.1017/S0956792516000346 Published online by Cambridge University Press

is a solution for every positive initial mass $\overline{u}_1^0$. In particular, the branch of solutions is not isolated and there exist parametric two-dimensional families of solutions. There are multiple ways to deal with this problem; see also Section 6. One possibility is to resolve the degeneracy of the problem via a small parameter $0 < \rho \ll 1$ and consider

$$0 = \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\mathrm{d}u_1}{\mathrm{d}x} - g(u_1)\frac{\mathrm{d}u_2}{\mathrm{d}x}\right) - \rho(u_1 - \overline{u}_1), \tag{2.22}$$

$$0 = \delta\frac{\mathrm{d}^2u_1}{\mathrm{d}x^2} + \kappa\frac{\mathrm{d}^2u_2}{\mathrm{d}x^2} - \alpha u_2 + f(u_1) \tag{2.23}$$

for a fixed positive parameter $\overline{u}_1 > 0$. In particular, upon setting

$$z_1 := u_1, \qquad z_2 := u_2, \qquad z_3 := \frac{\mathrm{d}u_1}{\mathrm{d}x}, \qquad z_4 := \frac{\mathrm{d}u_2}{\mathrm{d}x},$$

as well as

$$p_5 := \overline{u}_1, \qquad p_6 := \rho, \qquad P = 6,$$

we end up with a problem of the form (2.21) by transforming the two second-order ODEs to four first-order ODEs and re-labelling parameters. The vector field for the ODE–BVP we study numerically is then given by

$$F(z;p) = \begin{pmatrix} p_4 z_3 \\ p_4 z_4 \\ p_4[-g(z_1)f(z_1) + p_3 g(z_1)z_2 + p_2 g'(z_1)z_3 z_4 + p_2 p_6(z_1 - p_5)]/\mathcal{D}_g \\ p_4[-f(z_1) + p_3 z_2 - p_1 g'(z_1)z_3 z_4 - p_1(z_1 - p_5)p_6]/\mathcal{D}_g \end{pmatrix}, \tag{2.24}$$

where $\mathcal{D}_g := p_2 + p_1 g(z_1)$ and the detailed choices for the free parameters are discussed in Section 5. Observe that the system (2.24) becomes singular if $\mathcal{D}_g = 0$, which is precisely the condition $\delta \neq -\kappa/g(u_1)$ already discovered above. Therefore, we would need also for the numerical analysis a re-formulation (or de-singularization) of the problem to deal with this singularity and we postpone this problem to future work. As mentioned above, the primary bifurcation parameter we are going to be interested in is $\delta = p_1$. The main results of the numerical bifurcation analysis, which are presented in full detail in Section 5, are the following:

(B1) As predicted by the analytical results, we find the existence of local bifurcation points on the branch of homogeneous steady states in the parameter region with $\delta < \delta_\mathrm{d}$ for the case of sufficiently large α and for $\delta > \delta_\mathrm{d}$ for the case of sufficiently small α. At each bifurcation point on the homogeneous branch, a simple eigenvalue crosses the imaginary axis.

(B2) The non-trivial (i.e., non-homogeneous) solution branches consist of solutions of multiple "interfaces" or "layers"; branches originating further away from δ_d contain less layers. The branches can acquire sharper layers upon variation of further parameters which is important for information herding.

https://doi.org/10.1017/S0956792516000346 Published online by Cambridge University Press

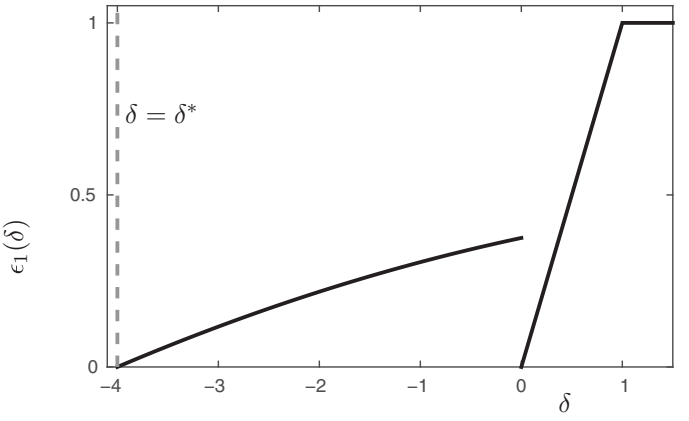

FIGURE 3. Illustration of $\varepsilon_1(\delta)$ for $\kappa = 1$ and $\delta = \frac{1}{4}$ (black curves). The corresponding singular limit $\delta^* = -\kappa/\gamma = -4$ is also marked (grey-dashed vertical line).

(B3) At the local bifurcation points, we observe the emergence of two symmetric branches of solutions for the case when the non-linearities are identical quadratic non-linearities of the form $s \mapsto s(1-s)$.

(B4) We also construct non-homogeneous solutions for $\rho = 0$ by a homotopy continuation step first continuing onto the non-trivial branches in δ and then decreasing ρ to zero in a second continuation step.

(B5) Furthermore, we also study the shape deformation of non-trivial solutions upon variation of κ and the domain length l. The numerical results show that the main interesting structures of the problem have already been obtained by just varying δ and α.

3 Entropy method – proofs

3.1 Proof of Theorem 1

First, we prove that the new diffusion matrix $B(w)$, defined in (2.2), is positive semi-definite if δ is not too negative.

Lemma 4 *Assume (H3) and $\delta > -\kappa/\gamma$, $\delta \neq 0$. Then, the matrix $B(w)$ is positive semi-definite, and there exists $\varepsilon_1(\delta) > 0$ such that for all $z = (z_1, z_2)^\top \in \mathbb{R}^2$, $w \in \mathbb{R}^2$:*

$$z^\top B(w) z \geqslant \varepsilon_1(\delta)(g(u_1)z_1^2 + z_2^2).$$

It holds $\varepsilon_1(\delta) \to 0$ as $\delta \searrow 0$ and $\delta \searrow -\kappa/\gamma$ (see Figure 3).

For later use, we note that the lemma implies that

$$\nabla w : B(w)\nabla w \geqslant \varepsilon_1(\delta)\left(\frac{|\nabla u_1|^2}{g(u_1)} + \frac{|\nabla u_2|^2}{\delta_0^2}\right), \tag{3.1}$$

https://doi.org/10.1017/S0956792516000346 Published online by Cambridge University Press

where $w = (w_1, w_2) = (h_0'(u_1), u_2/\delta_0)$ are the entropy variables introduced in the introduction and $A : B = \sum_{i,j} A_{ij}B_{ij}$ for two matrices $A = (A_{ij})$, $B = (B_{ij})$.

Proof Let $z = (z_1, z_2)^\top \in \mathbb{R}^2$. Then,

$$z^\top B(w)z = g(u_1)z_1^2 - (\delta_0 - \delta)g(u_1)z_1z_2 + \delta_0\kappa z_2^2.$$

If $\delta > 0$, then $\delta_0 = \delta$ and the mixed term vanishes, showing the claim for $\varepsilon_1(\delta) = \min\{1, \delta\kappa\}$. If $-\kappa/\gamma < \delta < 0$, we have $\delta_0 = \kappa/\gamma$. We make the (non-optimal) choice

$$\varepsilon_0 = \varepsilon_0(\delta) = \frac{1}{2}\left(1 - \frac{1}{4}\left(1 - \frac{\gamma\delta}{\kappa}\right)^2\right) > 0.$$

Then, $\varepsilon_0 < 1 - (1 - \gamma\delta/\kappa)^2/4$, which is equivalent to $(\kappa - \gamma\delta)^2 < 4(1 - \varepsilon_0)\kappa^2$. Thus, using $g(u_1) \leqslant \gamma$ (see assumption (H3)),

$$\begin{aligned} z^\top B(w)z &= g(u_1)z_1^2 - \left(\frac{\kappa}{\gamma} - \delta\right)g(u_1)z_1z_2 + \frac{\kappa^2}{\gamma}z_2^2 \\ &= \varepsilon_0 g(u_1)z_1^2 + (1 - \varepsilon_0)g(u_1)\left(z_1 - \frac{(\kappa - \gamma\delta)z_2}{2\gamma(1 - \varepsilon_0)}\right)^2 \\ &\quad + \frac{1}{\gamma}\left(\kappa^2 - \frac{(\kappa - \gamma\delta)^2}{4\gamma(1 - \varepsilon_0)}g(u_1)\right)z_2^2 \\ &\geqslant \varepsilon_0 g(u_1)z_1^2 + \frac{1}{\gamma}\left(\kappa^2 - \frac{(\kappa - \gamma\delta)^2}{4(1 - \varepsilon_0)}\right)z_2^2. \end{aligned}$$

In view of the choice of ε_0, the bracket on the right-hand side is positive, and the claim follows after choosing $\varepsilon_1(\delta) = \min\{\varepsilon_0(\delta), [\kappa^2 - (\kappa - \gamma\delta)^2/(4(1 - \varepsilon_0(\delta)))]/\gamma\} > 0$ for $-\kappa/\gamma < \delta < 0$. □

The proof of Theorem 1 is based on the solution of a time-discrete and regularized problem.

Step 1 (Solution of an approximate problem) Let $T > 0$, $N \in \mathbb{N}$, $\tau = T/N$, $\varepsilon > 0$ and $n \in \mathbb{N}$ such that $n > d/2$. Then, $H^n(\Omega; \mathbb{R}^2) \hookrightarrow L^\infty(\Omega; \mathbb{R}^2)$. Let $w^{k-1} \in L^\infty(\Omega; \mathbb{R}^2)$ be given. If $k = 1$, we define $w^0 = h'(u^0)$. We wish to find $w^k \in H^n(\Omega; \mathbb{R}^2)$ such that

$$\begin{aligned} &\frac{1}{\tau}\int_\Omega (u(w^k) - u(w^{k-1}))\cdot\phi\,\mathrm{d}x + \int_\Omega \nabla\phi : B(w^k)\nabla w^k\,\mathrm{d}x \\ &\quad + \varepsilon\int_\Omega\left(\sum_{|\beta|=n} D^\beta w^k \cdot D^\beta\phi + w^k\cdot\phi\right)\mathrm{d}x = \int_\Omega F(u(w^k))\cdot\phi\,\mathrm{d}x \end{aligned} \tag{3.2}$$

for all $\phi \in H^n(\Omega; \mathbb{R}^2)$, where $\beta \in \mathbb{N}_0^n$ is a multi-index, D^β is the corresponding partial derivative, $u(w) = (h')^{-1}(w)$ for $w \in \mathbb{R}$, and we recall that $F(u) = (0, f(u_1) - \alpha u_2)^\top$. By definition of h_0, we find that $u_1(w) \in (0, 1)$, thus avoiding any degeneracy at $u_1 = 0$ or $u_1 = 1$.

The existence of a solution to (3.2) will be shown by a fixed-point argument. In order to define the fixed-point operator, let $y \in L^\infty(\Omega; \mathbb{R}^2)$ and $\eta \in [0, 1]$ be given. We solve

https://doi.org/10.1017/S0956792516000346 Published online by Cambridge University Press

the linear problem

$$a(w,\phi) = G(\phi) \quad \text{for all } \phi \in H^n(\Omega;\mathbb{R}^2), \tag{3.3}$$

where

$$a(w,\phi) = \int_\Omega \nabla\phi : B(y)\nabla w \,\mathrm{d}x + \varepsilon \int_\Omega \left(\sum_{|\beta|=n} D^\beta w \cdot D^\beta \phi + w \cdot \phi \right) \mathrm{d}x,$$
$$G(\phi) = -\frac{\eta}{\tau} \int_\Omega \big(u(y) - u(w^{k-1})\big) \,\mathrm{d}x + \eta \int_\Omega F(u(y)) \cdot \phi \,\mathrm{d}x.$$

The forms a and G are bounded on $H^n(\Omega;\mathbb{R}^2)$. Moreover, in view of the positive semi-definiteness of $B(y)$ and the generalized Poincaré inequality (see Chap. II.1.4 in [40]), the bilinear form a is coercive:

$$a(w,w) \geqslant \varepsilon \int_\Omega \left(\sum_{|\beta|=n} |D^\beta w|^2 + |w|^2 \right) \mathrm{d}x \geqslant \varepsilon c \|w\|_{H^n(\Omega)} \quad \text{for } w \in H^n(\Omega;\mathbb{R}^2).$$

By the Lax–Milgram lemma, there exists a unique solution $w \in H^n(\Omega;\mathbb{R}^2) \hookrightarrow L^\infty(\Omega;\mathbb{R}^2)$ to (3.3). This defines the fixed-point operator $S : L^\infty(\Omega;\mathbb{R}^2) \times [0,1] \to L^\infty(\Omega;\mathbb{R}^2)$, $S(y,\eta) = w$.

By construction, $S(y,0) = 0$ for all $y \in L^\infty(\Omega;\mathbb{R}^2)$, and standard arguments show that S is continuous and compact, observing that the embedding $H^n(\Omega;\mathbb{R}^2) \hookrightarrow L^\infty(\Omega;\mathbb{R}^2)$ is compact. It remains to prove a uniform bound for all fixed points of $S(\cdot,\eta)$. Let $w \in L^\infty(\Omega;\mathbb{R}^2)$ be such a fixed point. Then, w solves (3.3) with y replaced by w. With the test function $\phi = w$, we find that

$$\frac{\eta}{\tau} \int_\Omega (u(w) - u(w^{k-1})) \cdot w \,\mathrm{d}x + \int_\Omega \nabla w : B(w)\nabla w \,\mathrm{d}x + \varepsilon \int_\Omega \left(\sum_{|\beta|=n} |D^\beta w|^2 + |w|^2 \right) \mathrm{d}x = \eta \int_\Omega F(u(w)) \cdot w \,\mathrm{d}x. \tag{3.4}$$

Since $h_0'' = 1/g > 0$ on $(0,1)$, h_0 is convex. Consequently, $h_0(x) - h_0(y) \leqslant h_0'(x)(x-y)$ for all $x, y \in [0,1]$. Choosing $x = u(w)$ and $y = u(w^{k-1})$ and using $h_0'(u(w)) = w$, this gives

$$\frac{\eta}{\tau} \int_\Omega (u(w) - u(w^{k-1})) \cdot w \,\mathrm{d}x \geqslant \frac{\eta}{\tau} \int_\Omega \big(h(u(w)) - h(u(w^{k-1}))\big) \,\mathrm{d}x.$$

Since $u_1 = u_1(w) \in (0,1)$ and f is continuous, there exists $f_M = \max_{s\in[0,1]} f(s)$ and thus,

$$\int_\Omega F(u(w)) \cdot w \,\mathrm{d}x \leqslant \int_\Omega (f_M - \alpha u_2) u_2 \,\mathrm{d}x \leqslant c_f,$$

https://doi.org/10.1017/S0956792516000346 Published online by Cambridge University Press

where $c_f > 0$ only depends on f_M and α. Hence, (3.4) can be estimated as follows:

$$\eta\int_\Omega h(u(w))\,\mathrm{d}x + \tau\int_\Omega \nabla w : B(w)\nabla w\,\mathrm{d}x + \varepsilon\tau\int_\Omega\left(\sum_{|\beta|=n}|D^\beta w|^2 + |w|^2\right)\mathrm{d}x \tag{3.5}$$
$$\leqslant \eta\tau c_f + \eta\int_\Omega h(u(w^{k-1}))\,\mathrm{d}x.$$

This yields an H^n bound for w uniform in η (but not uniform in τ or ε). The Leray–Schauder fixed-point theorem shows the existence of a solution $w \in H^n(\Omega;\mathbb{R}^2)$ to (3.3) with y replaced by w and with $\eta = 1$, which is a solution to (3.2).

Step 2 (Uniform bounds) Let w^k be a solution to (3.2). Set $w^{(\tau)}(x,t) = w^k(x)$ and $u^{(\tau)}(x,t) = u(w^k(x))$ for $x \in \Omega$ and $t \in ((k-1)\tau, k\tau]$, $k = 1,\ldots,N$. At time $t = 0$, we set $w^{(\tau)}(\cdot,0) = h_0'(u^0)$ and $u^{(\tau)}(0) = u^0$. We introduce the shift operator $(\sigma_\tau u^{(\tau)})(t) = u(w^{k-1})$ for $t \in ((k-1)\tau, k\tau]$, $k = 1,\ldots,N$. Then, $u^{(\tau)}$ solves

$$\frac{1}{\tau}\int_0^T\int_\Omega (u^{(\tau)} - \sigma_\tau u^{(\tau)})\cdot\phi\,\mathrm{d}x\,\mathrm{d}t + \int_0^T\int_\Omega \nabla\phi : B(w^{(\tau)})\nabla w^{(\tau)}\,\mathrm{d}x\,\mathrm{d}t \tag{3.6}$$
$$+\varepsilon\int_0^T\int_\Omega\left(\sum_{|\beta|=n} D^\beta w^{(\tau)}\cdot D^\beta\phi + w^{(\tau)}\cdot\phi\right)\mathrm{d}x\,\mathrm{d}t = \int_0^T\int_\Omega F(u^{(\tau)})\cdot\phi\,\mathrm{d}x\,\mathrm{d}t$$

for piecewise constant functions $\phi : (0,T) \to H^n(\Omega;\mathbb{R}^2)$. By density, the weak formulation also holds for all $L^2(0,T;H^n(\Omega;\mathbb{R}^2))$.

We have shown in Step 1 that the solution $w = w^k$ satisfies the entropy estimate (3.5). By (3.1), we obtain the gradient estimate

$$\int_\Omega \nabla w^k : B(w^k)\nabla w^k\,\mathrm{d}x \geqslant \varepsilon_1(\delta)\min\{\gamma^{-1},\delta_0^{-2}\}\int_\Omega(|\nabla u_1^k|^2 + |\nabla u_2^k|^2)\,\mathrm{d}x,$$

since $g(u_1^k) \leqslant \gamma$. Thus, we obtain from (3.5) the following entropy inequality:

$$\int_\Omega h(u^k)\,\mathrm{d}x + c_0\tau\int_\Omega(|\nabla u_1^k|^2 + |\nabla u_2^k|^2)\,\mathrm{d}x \tag{3.7}$$
$$+\varepsilon\tau\int_\Omega\left(\sum_{|\beta|=n}|D^\beta w^k|^2 + |w^k|^2\right)\mathrm{d}x \leqslant c_f\tau + \int_\Omega h(u^{k-1})\,\mathrm{d}x,$$

where $c_0 = \varepsilon_1(\delta)\min\{\gamma^{-1},\delta_0^{-2}\}$. Adding these inequalities leads to

$$\int_\Omega h(u^k)\,\mathrm{d}x + c_0\tau\sum_{j=1}^k\int_\Omega(|\nabla u_1^j|^2 + |\nabla u_2^j|^2)\,\mathrm{d}x$$
$$+\varepsilon\tau\sum_{j=1}^k\int_\Omega\left(\sum_{|\beta|=n}|D^\beta w^k|^2 + |w^k|^2\right)\mathrm{d}x \leqslant c_f k\tau + \int_\Omega h(u^0)\,\mathrm{d}x.$$

Since

$$\int_\Omega h(u^k)\,\mathrm{d}x = \int_\Omega\left(h_0(u_1^k) + \frac{(u_2^k)^2}{2\delta_0}\right)\mathrm{d}x \geqslant \frac{1}{2\delta_0}\int_\Omega (u_2^k)^2\,\mathrm{d}x,$$

https://doi.org/10.1017/S0956792516000346 Published online by Cambridge University Press

the above estimate shows the following uniform bounds:

$$\|u_1^{(\tau)}\|_{L^\infty(0,T;L^\infty(\Omega))} + \|u_2^{(\tau)}\|_{L^\infty(0,T;L^2(\Omega))} \leqslant c, \tag{3.8}$$

$$\|u_1^{(\tau)}\|_{L^2(0,T;H^1(\Omega))} + \|u_2^{(\tau)}\|_{L^2(0,T;H^1(\Omega))} \leqslant c, \tag{3.9}$$

$$\sqrt{\varepsilon}\|w^{(\tau)}\|_{L^2(0,T;H^n(\Omega))} \leqslant c, \tag{3.10}$$

where $c > 0$ denotes here and in the following a constant which is independent of ε or τ (but possibly depending on T).

In order to derive a uniform estimate for the discrete time derivative, let $\phi \in L^2(0,T;H^n(\Omega))$. Then, setting $Q_T = \Omega \times (0,T)$,

$$\begin{aligned}
\frac{1}{\tau}\left|\int_\tau^T \int_\Omega (u_1^{(\tau)} - \sigma_\tau u_1^{(\tau)})\phi \,\mathrm{d}x\,\mathrm{d}t\right| &\leqslant \big(\|\nabla u_1^{(\tau)}\|_{L^2(Q_T)} + \|g(u_1^{(\tau)})\|_{L^\infty(Q_T)}\|\nabla u_2^{(\tau)}\|_{L^2(Q_T)}\big) \\
&\quad \times \|\nabla\phi\|_{L^2(Q_T)} + \varepsilon\|w_1^{(\tau)}\|_{L^2(0,T;H^n(\Omega))}\|\phi\|_{L^2(0,T;H^n(\Omega))} \\
&\leqslant c\sqrt{\varepsilon}\|\phi\|_{L^2(0,T;H^n(\Omega))} + c\|\phi\|_{L^2(0,T;H^1(\Omega))}, \\
\frac{1}{\tau}\left|\int_\tau^T \int_\Omega (u_2^{(\tau)} - \sigma_\tau u_2^{(\tau)})\phi \,\mathrm{d}x\,\mathrm{d}t\right| &\leqslant \big(\delta\|\nabla u_1^{(\tau)}\|_{L^2(Q_T)} + \kappa\|\nabla u_2^{(\tau)}\|_{L^2(Q_T)}\big)\|\nabla\phi\|_{L^2(Q_T)} \\
&\quad + \varepsilon\|w_1^{(\tau)}\|_{L^2(0,T;H^n(\Omega))}\|\phi\|_{L^2(0,T;H^n(\Omega))} + \big(\|f(u_1^{(\tau)})\|_{L^2(Q_T)} + \alpha\|u_2^{(\tau)}\|_{L^2(Q_T)}\big)\|\phi\|_{L^2(Q_T)} \\
&\leqslant c\sqrt{\varepsilon}\|\phi\|_{L^2(0,T;H^n(\Omega))} + c\|\phi\|_{L^2(0,T;H^1(\Omega))},
\end{aligned} \tag{3.11}$$

which shows that

$$\tau^{-1}\|u^{(\tau)} - \sigma_\tau u^{(\tau)}\|_{L^2(0,T;(H^n(\Omega))')} \leqslant c. \tag{3.12}$$

Step 3 (The limit $(\varepsilon,\tau) \to 0$) The uniform estimates (3.9) and (3.12) allow us to apply the discrete Aubin lemma in the version of [15], showing that, up to a subsequence which is not relabelled, as $(\varepsilon,\tau) \to 0$,

$$\begin{aligned}
u^{(\tau)} &\to u \quad \text{strongly in } L^2(0,T;L^2(\Omega)) \text{ and a.e. in } Q_T, \\
u^{(\tau)} &\rightharpoonup u \quad \text{weakly in } L^2(0,T;H^1(\Omega)), \\
\tau^{-1}(u^{(\tau)} - \sigma_\tau u^{(\tau)}) &\rightharpoonup \partial_t u \quad \text{weakly in } L^2(0,T;(H^n(\Omega))'), \\
\varepsilon w^{(\tau)} &\to 0 \quad \text{strongly in } L^2(0,T;H^n(\Omega)).
\end{aligned} \tag{3.13}$$

Because of the L^∞ bound (3.8) for $(u_1^{(\tau)})$, we have

$$g(u_1^{(\tau)}) \rightharpoonup^* g(u_1), \quad f(u_1^{(\tau)}) \rightharpoonup^* f(u_1) \quad \text{weakly* in } L^\infty(0,T;L^\infty(\Omega))$$

(and even strongly in $L^p(Q_T)$ for any $p < \infty$). Thus, we can pass to the limit $(\varepsilon,\tau) \to 0$ in (3.6) to obtain a solution to

$$\begin{aligned}
&\int_0^T \langle \partial_t u_1, \phi\rangle \,\mathrm{d}t + \int_0^T \int_\Omega (\nabla u_1 - g(u_1)\nabla u_2)\phi \,\mathrm{d}x\,\mathrm{d}t = 0, \\
&\int_0^T \langle \partial_t u_2, \phi\rangle \,\mathrm{d}t + \int_0^T \int_\Omega (\delta\nabla u_1 + \kappa\nabla u_2)\phi \,\mathrm{d}x\,\mathrm{d}t = \int_0^T \int_\Omega (f(u_1) - \alpha u_2)\phi \,\mathrm{d}x\,\mathrm{d}t
\end{aligned}$$

https://doi.org/10.1017/S0956792516000346 Published online by Cambridge University Press

for all $\phi \in L^2(0,T;H^n(\Omega))$. In fact, performing the limit $\varepsilon \to 0$ and then $\tau \to 0$, we see from (3.11) that $\partial_t u \in L^2(0,T;(H^1(\Omega))')$ and hence, the weak formulation also holds for all $\phi \in L^2(0,T;H^1(\Omega))$. It contains the no-flux boundary conditions (1.3). Moreover, the initial conditions are satisfied in the sense of $(H^1(\Omega;\mathbb{R}^2))'$; see Step 3 of the proof of Theorem 2 in [28]. This finishes the proof.

3.2 Proof of Theorem 2

We recall first the following convex Sobolev inequality which is used to estimate the gradient terms in the entropy inequality.

Lemma 5 *Let $\Omega \subset \mathbb{R}^d$ ($d \geqslant 1$) be a convex domain and let $\phi \in C^4$ be a convex function such that $1/\phi''$ is concave. Then, there exists $c_S > 0$ such that for all integrable functions u with integrable $\phi(u)$ and $\phi''(u)|\nabla u|^2$,*

$$\frac{1}{\mathrm{m}(\Omega)}\int_\Omega \phi(u)\,\mathrm{d}x - \phi\left(\frac{1}{\mathrm{m}(\Omega)}\int_\Omega u\,\mathrm{d}x\right) \leqslant \frac{c_S}{\mathrm{m}(\Omega)}\int_\Omega \phi''(u)|\nabla u|^2\,\mathrm{d}x,$$

where $\mathrm{m}(\Omega)$ denotes the measure of Ω.

Proof The lemma is a consequence of Proposition 7.6.1 in [6] after choosing the probability measure $d\mu = \mathrm{d}x/\mathrm{m}(\Omega)$ on Ω and the differential operator $L = \Delta - x\cdot\nabla$, which satisfies the curvature condition $CD(1,\infty)$ since $\Gamma_2(u) = \frac{1}{2}(|\nabla^2 u|^2 + |\nabla u|^2) \geqslant \frac{1}{2}|\nabla u|^2 = \Gamma(u)$. Another proof can be found in [4, Section 3.4]. □

Step 1 (Uniform bound for the L^1 norm of u_1^k) The L^1 norm of u_1^k is not conserved but we are able to control its L^1 norm. For this, let $w^k \in H^n(\Omega;\mathbb{R}^2)$ be a solution to (3.2) and set $u_1^k = u_1(w^k)$. We introduce the notation $\overline{v} = \mathrm{m}(\Omega)^{-1}\int_\Omega v(x)\,\mathrm{d}x$ for any integrable function v. This implies that $u_1^* = \overline{u}_1^0$. Employing the test function $\phi = (1,0)$ in (3.2), we find that $\overline{u}_1^k = \overline{u}_1^{k-1} - \varepsilon\tau\overline{w}_1^k$. Solving the recursion gives

$$\overline{u}_1^k = \overline{u}_1^0 - \varepsilon\tau\sum_{j=1}^k \overline{w}_1^j = u_1^* - \varepsilon\tau\sum_{j=1}^k \overline{w}_1^j,$$

and by (3.10), we conclude that

$$|\overline{u}_1^{(\tau)}(t) - u_1^*| \leqslant \varepsilon\|w_1^{(\tau)}\|_{L^1(0,t;L^1(\Omega))} \leqslant \sqrt{\varepsilon}c,$$

where $\overline{u}_1^{(\tau)}(t) = \overline{u}_1^k$ for $t \in ((k-1)\tau, k\tau]$. Consequently, as $(\varepsilon,\tau) \to 0$, the convergence (3.13) shows that $\overline{u}_1(t) = u_1^*$ for $t > 0$.

Step 2 (Estimate of the relative entropy) We employ the test function:

$$\phi = (h_0'(u_1^k) - h_0'(u_1^*), (u_2^k - u_2^*)/\delta_0) = (w_1^k - h_0'(u_1^*), w_2^k - u_2^*/\delta_0)$$

https://doi.org/10.1017/S0956792516000346 Published online by Cambridge University Press

in (3.2) to obtain

$$\begin{aligned}0 &= \frac{1}{\tau}\int_\Omega \Big((u_1^k-u_1^{k-1})(h_0'(u_1^k)-h_0'(u_1^*))+\frac{1}{\delta_0}(u_2^k-u_2^{k-1})(u_2^k-u_2^*)\Big)\,\mathrm{d}x \\ &\quad + \int_\Omega \nabla w^k : B(w^k)\nabla w^k\,\mathrm{d}x + \varepsilon\int_\Omega \Big(\sum_{|\beta|=n}|D^\beta w^k|^2 + w_1^k(w_1^k-h_0'(u_1^*)) \\ &\quad + w_2^k(w_2^k-u_2^*)/\delta_0)\Big)\,\mathrm{d}x - \frac{1}{\delta_0}\int_\Omega (f(u_1^k)-\alpha u_2^k)(u_2^k-u_2^*)\,\mathrm{d}x \\ &=: I_1+\cdots+I_4.\end{aligned} \tag{3.14}$$

For the first integral, we employ the convexity of h_0:

$$\begin{aligned}(u_1^k-u_1^{k-1})(h_0'(u_1^k)-h_0'(u_1^*)) &\geqslant (h_0(u_1^k)-h_0(u_1^{k-1}))-h_0'(u_1^*)(u_1^k-u_1^{k-1}),\\ (u_2^k-u_2^{k-1})(u_2^k-u_2^*) &\geqslant \frac{1}{2}\big((u_2^k-u_2^*)^2-(u_2^{k-1}-u_2^*)^2\big),\end{aligned}$$

which yields

$$\begin{aligned}I_1 &\geqslant \frac{1}{\tau}\int_\Omega (h_0(u_1^k)-h_0(u_1^{k-1}))\,\mathrm{d}x - \frac{h_0'(u_1^*)}{\tau}\int_\Omega (u_1^k-u_1^{k-1})\,\mathrm{d}x \\ &\quad + \frac{1}{2\delta_0\tau}\int_\Omega \big((u_2^k-u_2^*)^2-(u_2^{k-1}-u_2^*)^2\big)\,\mathrm{d}x.\end{aligned}$$

By (3.1), it follows that

$$I_2 \geqslant \varepsilon_1(\delta)\int_\Omega\Big(\frac{|\nabla u_1^k|^2}{g(u_1^k)}+\frac{|\nabla u_2^k|^2}{\delta_0^2}\Big)\,\mathrm{d}x = \varepsilon_1(\delta)\int_\Omega\Big(h_0''(u_1^k)|\nabla u_1^k|^2+\frac{|\nabla u_2^k|^2}{\delta_0^2}\Big)\,\mathrm{d}x.$$

Lemma 5 then shows that

$$I_2 \geqslant \frac{\varepsilon_1(\delta)}{c_S}\int_\Omega (h_0(u_1^k)-h_0(\overline{u}_1^k))\,\mathrm{d}x + \frac{\varepsilon_1(\delta)}{\delta_0^2}\int_\Omega |\nabla u_2^k|^2\,\mathrm{d}x.$$

The third integral in (3.14) is estimated by using Young's inequality:

$$I_3 \geqslant \frac{\varepsilon}{2}\int_\Omega \big((w_1^k)^2+(w_2^k)^2-h_0'(u_1^*)^2-\delta_0^{-2}(u_2^*)^2\big)\,\mathrm{d}x \geqslant -\frac{\varepsilon}{2}\int_\Omega \big(h_0'(u_1^*)^2+\delta_0^{-2}(u_2^*)^2\big)\,\mathrm{d}x.$$

Summarizing these estimates, we infer from (3.14) that

$$\begin{aligned}&\int_\Omega (h_0(u_1^k)-h_0(u_1^{k-1}))\,\mathrm{d}x - h_0'(u_1^*)\int_\Omega (u_1^k-u_1^{k-1})\,\mathrm{d}x \\ &\quad + \frac{1}{2\delta_0}\int_\Omega \big((u_2^k-u_2^*)^2-(u_2^{k-1}-u_2^*)^2\big)\,\mathrm{d}x \\ &\quad + \frac{\varepsilon_1(\delta)\tau}{c_S}\int_\Omega (h_0(u_1^k)-h_0(\overline{u}_1^k))\,\mathrm{d}x + \frac{\varepsilon_1(\delta)\tau}{\delta_0^2}\int_\Omega |\nabla u_2^k|^2\,\mathrm{d}x \\ &\leqslant \frac{\varepsilon\tau}{2}\int_\Omega \big(h_0'(\overline{u}_1^k)^2+\delta_0^{-2}(u_2^*)^2\big)\,\mathrm{d}x + \frac{\tau}{\delta_0}\int_\Omega (f(u_1^k)-\alpha u_2^k)(u_2^k-u_2^*)\,\mathrm{d}x.\end{aligned}$$

https://doi.org/10.1017/S0956792516000346 Published online by Cambridge University Press

Adding these equations over k and using the notation as in the proof of Theorem 1 for $u_i^{(\tau)}$, we obtain

$$\begin{aligned}
&\int_\Omega (h_0(u_1^{(\tau)}(t)) - h_0(u_1^0))\,\mathrm{d}x - h_0'(u_1^*)\int_\Omega (u_1^{(\tau)}(t) - u_1^0)\,\mathrm{d}x \\
&\quad + \frac{1}{2\delta_0}\int_\Omega \big((u_2^{(\tau)}(t) - u_2^*)^2 - (u_2^0 - u_2^*)^2\big)\,\mathrm{d}x \\
&\quad + \frac{\varepsilon_1(\delta)}{c_S}\int_0^t\int_\Omega \big(h_0(u_1^{(\tau)}) - h_0(\overline{u}_1^{(\tau)})\big)\,\mathrm{d}x\,\mathrm{d}s + \frac{\varepsilon_1(\delta)}{\delta_0^2}\int_0^t\int_\Omega |\nabla u_2^{(\tau)}|^2\,\mathrm{d}x\,\mathrm{d}s \\
&\leqslant \frac{\varepsilon}{2}\int_0^t\int_\Omega \big(h_0'(\overline{u}_1^{(\tau)})^2 + \delta_0^{-2}(u_2^*)^2\big)\,\mathrm{d}x\,\mathrm{d}s + \frac{1}{\delta_0}\int_0^t\int_\Omega (f(u_1^{(\tau)}) - \alpha u_2^{(\tau)})(u_2^{(\tau)} - u_2^*)\,\mathrm{d}x\,\mathrm{d}s.
\end{aligned} \tag{3.15}$$

Step 3 (The limit $(\varepsilon,\tau)\to 0$) Because of the L^∞ bound for $(u_1^{(\tau)})$, it follows that, for a subsequence, $u_1^{(\tau)} \rightharpoonup^* u_1$ weakly* in $L^\infty(0,T;L^1(\Omega))$ and thus, as $(\varepsilon,\tau)\to 0$,

$$\int_\Omega (u_1^{(\tau)}(t) - u_1^0)\,\mathrm{d}x = \int_\Omega (u_1^{(\tau)}(t) - u_1^*)\,\mathrm{d}x \to \int_\Omega (u_1(t) - u_1^*)\,\mathrm{d}x = 0,$$

since $\overline{u}_1(t) = u_1^*$ for $t > 0$, by Step 1. The weak convergence of $(\nabla u_2^{(\tau)})$ to ∇u_2 in $L^2(0,T;L^2(\Omega))$ implies that

$$\liminf_{\tau\to 0}\int_0^t\int_\Omega |\nabla u_2^{(\tau)}|^2\,\mathrm{d}x\,\mathrm{d}s \leqslant \int_0^t\int_\Omega |\nabla u_2|^2\,\mathrm{d}x\,\mathrm{d}s.$$

Furthermore, by the strong convergence $u_1^{(\tau)}\to u_1$ in $L^2(0,T;L^2(\Omega))$, up to a subsequence, $u_1^{(\tau)}\to u_1$ a.e. in $Q_T = \Omega\times(0,T)$ and $h_0(u_1^{(\tau)})\to h_0(u_1)$ a.e. in Q_T. Then, the L^∞ bound of $(u_1^{(\tau)})$ implies that $h_0(u_1^{(\tau)})\to h_0(u_1)$ strongly in $L^p(0,T;L^p(\Omega))$ for any $p<\infty$. Furthermore, we know that $u_2^{(\tau)}\to u_2$ strongly in $L^2(0,T;L^2(\Omega))$, see (3.13). Therefore, the limit $(\varepsilon,\tau)\to 0$ in (3.15) leads to

$$\begin{aligned}
&\int_\Omega (h_0(u_1(t)) - h_0(u_1^0))\,\mathrm{d}x + \frac{1}{2\delta_0}\int_\Omega \big((u_2(t) - u_2^*)^2 - (u_2^0 - u_2^*)^2\big)\,\mathrm{d}x \\
&\quad + \frac{\varepsilon_1(\delta)}{c_S}\int_0^t\int_\Omega \big(h_0(u_1) - h_0(u_1^*)\big)\,\mathrm{d}x\,\mathrm{d}s + \frac{\varepsilon_1(\delta)}{\delta_0^2}\int_0^t\int_\Omega |\nabla u_2|^2\,\mathrm{d}x\,\mathrm{d}s \\
&\leqslant \frac{1}{\delta_0}\int_0^t\int_\Omega (f(u_1) - \alpha u_2)(u_2 - u_2^*)\,\mathrm{d}x\,\mathrm{d}s.
\end{aligned}$$

Now, we estimate the right-hand side. Because of $f(u_1^*) = \alpha u_2^*$ and the Lipschitz continuity of f with Lipschitz constant $c_L > 0$, we infer that (recall (2.6) for the definition

https://doi.org/10.1017/S0956792516000346 Published online by Cambridge University Press

of $h_0(u_1|u_1^*)$)

$$\begin{aligned}
&\int_\Omega \big(h_0(u_1(t)|u_1^*)\,\mathrm{d}x - h_0(u_1(0)|u_1^*)\big)\,\mathrm{d}x + \frac{1}{2\delta_0}\int_\Omega \big((u_2(t)-u_2^*)^2 - (u_2(0)-u_2^*)^2\big)\,\mathrm{d}x \\
&\quad + \frac{\varepsilon_1(\delta)}{c_S}\int_0^t\int_\Omega h_0(u_1(s)|u_1^*)\,\mathrm{d}x\,\mathrm{d}s \\
&\leqslant \frac{1}{\delta_0}\int_0^t\int_\Omega (f(u_1)-f(u_1^*))(u_2-u_2^*)\,\mathrm{d}x\,\mathrm{d}s - \frac{\alpha}{\delta_0}\int_0^t\int_\Omega (u_2-u_2^*)^2\,\mathrm{d}x\,\mathrm{d}s \\
&\leqslant \frac{1}{2\delta_0\alpha}\int_0^t\int_\Omega (f(u_1)-f(u_1^*))^2\,\mathrm{d}x\,\mathrm{d}s - \frac{\alpha}{2\delta_0}\int_0^t\int_\Omega (u_2-u_2^*)^2\,\mathrm{d}x\,\mathrm{d}s \\
&\leqslant \frac{c_L^2}{2\alpha\delta_0}\int_0^t\int_\Omega (u_1-u_1^*)^2\,\mathrm{d}x\,\mathrm{d}s - \frac{\alpha}{2\delta_0}\int_0^t\int_\Omega (u_2-u_2^*)^2\,\mathrm{d}x\,\mathrm{d}s.
\end{aligned}$$

Since $\overline{u}_1 = u_1^*$, a Taylor expansion and the assumption $1/h_0''(u_1) = g(u_1) \leqslant \gamma$ give

$$\begin{aligned}
\int_0^t\int_\Omega h_0(u_1|u_1^*)\,\mathrm{d}x\,\mathrm{d}s &= \int_0^t\int_\Omega (h_0(u_1) - h_0(u_1^*)\,\mathrm{d}x\,\mathrm{d}s \\
&= \int_0^t\int_\Omega \left(h_0'(u_1^*)(u_1-u_1^*) + \frac{1}{2}h_0''(\xi)(u_1-u_1^*)^2\right)\,\mathrm{d}x\,\mathrm{d}s \qquad (3.16)\\
&\geqslant \frac{1}{2\gamma}\int_0^t\int_\Omega (u_1-u_1^*)^2\,\mathrm{d}x\,\mathrm{d}s,
\end{aligned}$$

where ξ is a number between u_1 and u_1^*. We conclude that

$$\begin{aligned}
&\int_\Omega h_0(u_1(t)|u_1^*)\,\mathrm{d}x + \frac{1}{2\delta_0}\int_\Omega (u_2(t)-u_2^*)^2\,\mathrm{d}x + \left(\frac{\varepsilon_1(\delta)}{c_S} - \frac{\gamma c_L^2}{\alpha\delta_0}\right)\int_0^t\int_\Omega h_0(u_1(s)|u_1^*)\,\mathrm{d}x\,\mathrm{d}s \\
&\quad + \frac{\alpha}{2\delta_0}\int_0^t\int_\Omega (u_2-u_2^*)^2\,\mathrm{d}x\,\mathrm{d}s \leqslant \int_\Omega h_0(u_1(0)|u_1^*)\,\mathrm{d}x + \frac{1}{2\delta_0}\int_\Omega (u_2(0)-u_2^*)^2\,\mathrm{d}x,
\end{aligned}$$

and recalling the notation $h(u|U) = h_0(u_1|u_1^*) + (u_2-u_2^*)^2/(2\delta_0)$,

$$\int_\Omega h(u(t)|U)\,\mathrm{d}x + \min\left\{\frac{\varepsilon_1(\delta)}{c_S} - \frac{\gamma c_L^2}{\alpha\delta_0}, \alpha\right\}\int_0^t h(u|U)\,\mathrm{d}s \leqslant \int_\Omega h(u(0)|U)\,\mathrm{d}x.$$

Then, Gronwall's lemma implies that

$$H(u(t)|U) = \int_\Omega h(u(t)|U)\,\mathrm{d}x \leqslant e^{-\chi(\delta)t}H(u(0)|U), \quad t \geqslant 0,$$

where $\chi(\delta)$ is defined in (2.8). Finally, taking into account (3.16), we estimate

$$h(u|U) \geqslant \frac{1}{2\gamma}(u_1-u_1^*)^2 + \frac{1}{2\delta}(u_2-u_2^*)^2,$$

which shows (2.9) and finishes the proof.

https://doi.org/10.1017/S0956792516000346 Published online by Cambridge University Press

4 Analytical bifurcation analysis – proofs

In this section, we are going to prove Theorem 3. The proofs follow closely ideas presented for similar systems in [9, 39, 45], which are fundamentally based upon an application of results of Crandall and Rabinowitz [10, 11]; see also [31] for a detailed exposition of the these results. Recall that we defined the spaces $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{Y}_0$ in (2.12) and the mapping

$$\mathcal{F} : \mathcal{X} \times \mathcal{X} \times \mathbb{R} \to \mathcal{Y}_0 \times \mathcal{Y} \times \mathbb{R}$$

in (2.14). A first step is to investigate the Fredholm and differentiability properties of $\mathcal{F}$.

Lemma 6 *The mapping $\mathcal{F}$ satisfies the following properties:*

(L1) $\mathcal{F}(u^*, \delta) = 0$ *for all* $\delta \in \mathbb{R}$.

(L2) $\mathcal{F}(u_1, u_2, \delta) = 0$ *implies that* (u_1, u_2) *solves* (2.10).

(L3) $\mathcal{F}$ *is* C^1*-smooth with Fréchet derivative* $\mathrm{D}_u\mathcal{F}$ *given by* (2.15).

(L4) *If* $\tilde{u}(x) \equiv (\tilde{u}_1, \tilde{u}_2)$ *is a homogeneous state and* $\delta g(\tilde{u}_1) \neq -\kappa$, *then* $\mathrm{D}_u\mathcal{F}(\tilde{u}_1, \tilde{u}_2, \delta)$ *is a Fredholm operator with index zero.*

Proof For (L1) recall that $u^* = (u_1^*, u_2^*)$ was the notation for a homogeneous steady state. Regarding (L2), observe that the first two components of $\mathcal{F}$ are just the steady-state equations (2.10). Statement (L3) follows from a direct calculation. The problem is to show (L4). We follow the argument given in [9, 45] and consider

$$\mathrm{D}_u\mathcal{F}(\tilde{u}_1, \tilde{u}_2, \delta)(U_1, U_2)^\top = \mathcal{B}_1(U_1, U_2)^\top + \mathcal{B}_2(U_1, U_2)^\top, \tag{4.1}$$

where $\mathcal{B}_1 : \mathcal{X} \times \mathcal{X} \to \mathcal{Y}_0 \times \mathcal{Y} \times \mathbb{R}$ is defined by

$$\mathcal{B}_1 \begin{pmatrix} U_1 \\ U_2 \end{pmatrix} = \begin{pmatrix} \Delta U_1 - \mathrm{div}[g'(\tilde{u}_1)(\nabla \tilde{u}_2)U_1 + g(\tilde{u}_1)\nabla U_2] \\ \delta \Delta U_1 + \kappa \Delta U_2 - \alpha U_2 + f'(\tilde{u}_1)U_1 \\ 0 \end{pmatrix}, \tag{4.2}$$

and the mapping $\mathcal{B}_2 : \mathcal{X} \times \mathcal{X} \to \mathcal{Y}_0 \times \mathcal{Y} \times \mathbb{R}$ is given by

$$\mathcal{B}_2 \begin{pmatrix} U_1 \\ U_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \int_\Omega U_1(x)\, \mathrm{d}x \end{pmatrix}. \tag{4.3}$$

We observe easily that $\mathcal{B}_2 : \mathcal{X} \times \mathcal{X} \to \mathcal{Y}_0 \times \mathcal{Y} \times \mathbb{R}$ is linear and compact. We need an ellipticity condition and $\mathcal{B}_1$ should satisfy Agmon's condition [39]. We have ellipticity for $\mathcal{B}_1$ (in the sense of Petrovskii [27, 39]) if

$$\det \left[\begin{pmatrix} 1 & -g(\tilde{u}_1) \\ \delta & \kappa \end{pmatrix} \xi \cdot \xi \right] \neq 0, \tag{4.4}$$

https://doi.org/10.1017/S0956792516000346 Published online by Cambridge University Press

for all $\xi = (\xi_1, \xi_2, \ldots, \xi_d) \in \mathbb{R}^d \backslash \{0\}$. Computing the determinant this condition just yields

$$0 \neq (\xi_1^2 + \cdots + \xi_d^2)(\kappa + \delta g(\tilde{u}_1)) \quad \text{if and only if} \quad -\kappa \neq \delta g(\tilde{u}_1)$$

and ellipticity in the sense of Petrovskii follows. Moreover, we need to verify Agmon's condition at a fixed angle $\theta \in [-\pi, \pi)$. Using [39, Remark 2.5] with $\theta = \pi/2$, one verifies computing a shifted determinant similar to the previously computed one in (4.4) that Agmon's condition holds for all values of κ. In particular, the ellipticity condition gives a restriction on the parameters for the bifurcation analysis and not Agmon's condition. By applying [39, Thm. 3.3], we infer that

$$\mathcal{B}_1 : \mathcal{X} \times \mathcal{X} \to \mathcal{Y} \times \mathcal{Y} \times \{0\}$$

is a Fredholm operator of index zero. Hence, $\mathcal{Y}_0 \times \mathcal{Y} \times \{0\} = \mathcal{R}(\mathcal{B}_1) \oplus W$, where $\mathcal{R}(\mathcal{B}_1)$ is the range of $\mathcal{B}_1$ and W is a closed subspace of $\mathcal{Y} \times \mathcal{Y} \times \mathbb{R}$ with $\dim W = \dim \mathcal{N}(\mathcal{B}_1) < \infty$. Consequently, since the first component of $\mathcal{B}_1$ is in $\mathcal{Y}_0$, we have

$$\mathcal{Y}_0 \times \mathcal{Y} \times \mathbb{R} = \mathcal{R}(\mathcal{B}_1) \oplus W_0 \oplus \operatorname{span}\{(0,0,1)^\top\},$$

where $W_0 = \{(H_1, H_2, H_3) \in W | \int_0^L H_1(x)dx = 0\}$ and $W = W_0 + \operatorname{span}\{(1,0,0)\}$. Then, $\dim W = \dim W_0 + 1$. Thus, the co-dimension of $\mathcal{R}(\mathcal{B}_1)$ in $\mathcal{Y}_0 \times \mathcal{Y} \times \mathbb{R}$ is equal to $\dim W = \dim \mathcal{N}(\mathcal{B}_1)$. Hence, $\mathcal{B}_1 : \mathcal{X} \times \mathcal{X} \to \mathcal{Y}_0 \times \mathcal{Y} \times \mathbb{R}$ is a Fredholm operator of index zero for $\delta g(\tilde{u}_1) \neq -\kappa$. Therefore, $\mathrm{D}_u \mathcal{F}$ is a Fredholm operator of index zero as $\mathcal{B}_2$ is a compact perturbation. Hence, the result (R1) in Theorem 3 follows. □

It seems difficult to improve the result to include the degenerate cases when $\kappa = -\delta g(u_1^*)$ as this would require to deal with bifurcation problems with non-elliptic operators. The next goal is to apply [39, Theorem 4.3]. To do so, we need some additional properties of $\mathcal{F}$. In particular, in order that bifurcations occur from the homogeneous steady state $u^* = (u_1^*, u_2^*)$ we need that the implicit function theorem fails. For the following lemma, we need to be in the case where each eigenvalue μ_n of the negative Neumann Laplacian on Ω eigenvalue is simple. For the one-dimensional case, this always holds, whilst for generic d-dimensional domains the eigenvalues are also simple [41].

Lemma 7 *Suppose the eigenvalues of the negative Neumann Laplacian on $\Omega \subset \mathbb{R}^d$ are simple and $\delta g(u_1^*) \neq -\kappa$. Then, there exist bifurcation points at $\delta = \delta_{\mathrm{b}}^n$ such that the map $\mathcal{F}$ satisfies the following properties:*

(L5) *the null space $\mathcal{N}[\mathrm{D}_u \mathcal{F}(u^*, \delta_{\mathrm{b}}^n)]$ is one-dimensional, i.e.,* $\operatorname{span}[e_{\mathrm{b}}^n] = \mathcal{N}[\mathrm{D}_u \mathcal{F}(u^*, \delta_{\mathrm{b}}^n)]$*;*

(L6) *the non-degenerate crossing condition holds, i.e.,*

$$\mathrm{D}_{\delta u} \mathcal{F}(u^*, \delta_{\mathrm{b}}^n) e_{\mathrm{b}}^n \notin \mathcal{R}[\mathrm{D}_u \mathcal{F}(u^*, \delta_{\mathrm{b}}^n)]. \tag{4.5}$$

https://doi.org/10.1017/S0956792516000346 Published online by Cambridge University Press

Proof We start by proving (L5). By (4.2), the null space of $\mathrm{D}_u\mathcal{F}(u^*,\delta)$ consists of solutions for

$$\begin{aligned} \Delta U_1 - g(u_1^*)\Delta U_2 &= 0,\\ \delta\Delta U_1 + \kappa\Delta U_2 - \alpha U_2 + f'(u_1^*)U_1 &= 0,\\ \int_\Omega U_1(x)\,\mathrm{d}x &= 0, \end{aligned} \tag{4.6}$$

with no-flux conditions on $\partial\Omega$. For any pair $u=(u_1,u_2)\in\mathcal{X}\times\mathcal{X}$, we can expand u_1 and u_2 as a series of mutually orthogonal eigenfunctions of the following system:

$$\left\{\begin{array}{rcll} -\Delta u &=& \mu u & \text{in } \Omega,\\ \frac{\partial u}{\partial \nu} &=& 0 & \text{on } \partial\Omega \end{array}\right. \tag{4.7}$$

multiplied by constants vectors. Let $\mu_n>0$ be a simple eigenvalue of (4.7) and e_{μ_n} is the eigenfunction corresponding to μ_n normalized by $\int_\Omega (e_{\mu_n})^2\,\mathrm{d}x=1$. Then, we define

$$\bar{U}_1 := \int_\Omega u_1(x)e_{\mu_n}(x)\,\mathrm{d}x, \qquad \bar{U}_2 := \int_\Omega u_2(x)e_{\mu_n}(x)\,\mathrm{d}x.$$

We obtain

$$\begin{aligned} \int_\Omega e_{\mu_n}\Delta u_1\,\mathrm{d}x &= -\mu_n\int_\Omega u_1 e_{\mu_n}\,\mathrm{d}x = -\mu_n\bar{U}_1,\\ \int_\Omega e_{\mu_n}\Delta u_2\,\mathrm{d}x &= -\mu_n\int_\Omega u_2 e_{\mu_n}\,\mathrm{d}x = -\mu_n\bar{U}_2. \end{aligned} \tag{4.8}$$

Now, by multiplying the first two equations of (4.6) by e_{μ_n} and integrating over Ω, using the boundary condition and (4.8), we arrive at the following algebraic system for $\bar{U}_1$ and $\bar{U}_2$:

$$\begin{aligned} \bar{U}_1 - g(u_1^*)\bar{U}_2 &= 0,\\ (\kappa\mu_n+\alpha)\bar{U}_2 - (f'(u_1^*)-\delta\mu_n)\bar{U}_1 &= 0. \end{aligned} \tag{4.9}$$

If $\delta > f'(u_1^*)/\mu_n$, then the system (4.9) has only the zero solution. In this case, we would have $\mathcal{N}[\mathrm{D}_u\mathcal{F}(u^*,\delta)]=0$ for all δ. In order to have existence of a non-homogeneous solution, we necessarily require $\delta\leqslant f'(u_1^*)/\mu_n$. In this case, the system (4.9) has a non-zero solution if and only if

$$\delta =: \delta_{\mathrm{b}}^n = -\frac{\kappa}{g(u_1^*)} + \frac{1}{\mu_n}\left[f'(u_1^*) - \frac{\alpha}{g(u_1^*)}\right] = \delta_{\mathrm{d}} + \frac{1}{\mu_n}\left[f'(u_1^*) - \frac{\alpha}{g(u_1^*)}\right]. \tag{4.10}$$

Taking $\delta=\delta_{\mathrm{b}}^n$, we can rewrite the first two equations of (4.6) as the system:

$$\begin{pmatrix}\Delta U_1\\ \Delta U_2\end{pmatrix} = \frac{1}{\kappa+\delta_{\mathrm{b}}^n g(u_1^*)}\begin{pmatrix}-g(u_1^*)f'(u_1^*) & g(u_1^*)\alpha\\ -f'(u_1^*) & \alpha\end{pmatrix}\begin{pmatrix}U_1\\ U_2\end{pmatrix} =: A\begin{pmatrix}U_1\\ U_2\end{pmatrix}. \tag{4.11}$$

Using (4.10) and computing the determinant and the trace of the matrix A, we find that its eigenvalues are $\lambda_1=0$ and $\lambda_2=-\mu_n$, where $\mu_n>0$ is a single eigenvalue of the problem (4.7). Let T be the matrix whose columns are the eigenvectors corresponding to

https://doi.org/10.1017/S0956792516000346 Published online by Cambridge University Press

λ_1 and λ_2, respectively:

$$T = \begin{pmatrix} \alpha & g(u_1^*) \\ f'(u_1^*) & 1 \end{pmatrix}.$$

We have

$$T^{-1}AT = \begin{pmatrix} 0 & 0 \\ 0 & \mu_n \end{pmatrix}.$$

Then, by considering the transformation

$$\begin{pmatrix} p \\ q \end{pmatrix} = T^{-1} \begin{pmatrix} U_1 \\ U_2 \end{pmatrix}, \tag{4.12}$$

it follows that the first two equations of (4.6) can be uncoupled and we find that

$$\begin{aligned} &\Delta p = 0 \qquad \text{in } \Omega, \\ &\Delta q = \mu_n q \qquad \text{in } \Omega, \\ &\alpha \int_\Omega p(x)\,\mathrm{d}x + g(u_1^*) \int_\Omega q(x)\,\mathrm{d}x = 0, \\ &\nabla p \cdot \nu = \nabla q \cdot \nu = 0 \qquad \text{on } \partial\Omega, \end{aligned} \tag{4.13}$$

where the genericity condition $-\kappa \neq \delta_{\mathrm{b}}^n g(u_1^*)$ is used to obtain zero Neumann boundary conditions. Recall that μ_n is a simple eigenvalue of (4.7) with eigenfunction e_{μ_n}. Observe that $\int_\Omega e_{\mu_n}(x)\,\mathrm{d}x = 0$, which implies that $p = 0$ and $q = Ce_{\mu_n}$ for some constant C are the solutions of (4.13). Therefore, it follows that

$$(U_1, U_2)^\top = Ce_{\mu_n}(g(u_1^*), 1)^\top. \tag{4.14}$$

This shows that $\mathcal{N}[\mathrm{D}_u\mathcal{F}(u^*, \delta_{\mathrm{b}}^n)] = \mathrm{span}[e_{\mu_n}(g(u_1^*), 1)^\top] =: \mathrm{span}[e_{\mathrm{b}}^n]$. In particular, the nullspace is one-dimensional and the result (L5) follows.

To prove (L6), we argue by contradiction and suppose that (4.5) is not satisfied. Hence, by computing $\mathrm{D}_{\delta u}\mathcal{F}(u^*, \delta_{\mathrm{b}}^n)$, it follows there exists (p, q) such that

$$\begin{aligned} \Delta p - g(u_1^*)\Delta q &= \mu_n g(u_1^*) e_{\mu_n} \quad \text{in } \Omega, \\ \kappa\Delta q + \delta_{\mathrm{b}}^n \Delta p - \alpha q + f'(u_1^*)p &= 0 \quad \text{in } \Omega, \\ \int_\Omega p(x)\,\mathrm{d}x &= 0, \\ \nabla p \cdot \nu = \nabla q \cdot \nu &= 0 \quad \text{on } \partial\Omega. \end{aligned} \tag{4.15}$$

As in the first part of the proof, it is helpful to consider a suitable projection and we define P and Q as

$$P := \int_\Omega p(x) e_{\mathrm{b}}^n(x)\,\mathrm{d}x, \qquad Q := \int_\Omega q(x) e_{\mathrm{b}}^n(x)\,\mathrm{d}x.$$

https://doi.org/10.1017/S0956792516000346 Published online by Cambridge University Press

Multiplying the first two equations (4.15) by e_b^n and integrating over Ω and using boundary conditions, one obtains an algebraic system for P and Q given by

$$\begin{cases} P - g(u_1^*)Q &= -g(u_1^*), \\ (f'(u_1^*) - \delta_b^n \mu_n)P - (\kappa\mu_n + \alpha)Q &= 0. \end{cases} \tag{4.16}$$

By the definition of δ_b^n, the determinant of the matrix of coefficients on the left-hand side of the system (4.16) is zero. This implies that the inhomogeneous linear system has no solution. Hence, the system (4.15) has no solutions and the result (4.5) in (L6) follows. □

Note that (L5) and (L6) are just the results (R2) and (R3) claimed in Theorem 3. By applying [39, Theorem 4.3], we obtain the existence of a non-trivial branch of solutions. Therefore, the local dynamics of the problem already shows that the entropy method cannot provide exponential decay to a distinguished steady state for all parameter values.

5 Numerical bifurcation analysis – continuation results

In Section 2.1, we proved the existence of a weak solution for $\delta > \delta^* = -\kappa/\delta$ as well as global convergence to a steady state for $\delta > \delta_e$ ($\delta \neq 0$); in addition, δ_e converges to $\delta^* = -\kappa/\gamma$ as $\alpha \to +\infty$ and δ_e converges to $+\infty$ as $\alpha \to 0$. In Section 2.2, we showed the existence of non-trivial solutions for $\delta = \delta_b^n$, where δ_b^n is defined in (4.10) and in particular δ_b^n could be bigger or smaller than $\delta_d = \kappa/g(u_1^*)$ depending on α.

The numerical continuation results presented in this section aim to augment and extend these results. To simplify the comparison to numerical results, we focus on the case

$$\kappa = 1, \qquad g(s) = s(1-s), \qquad f(s) = s(1-s),$$

which yields the condition $\delta > \delta^* = -4$ for the validity of the entropy method for $\alpha \to +\infty$. As already mentioned, the values for δ_b^n depend on α, so we are going to study a case with α sufficiently large (Section 5.2) and the case with α sufficiently small (Section 5.3). Below we are going to define the meaning of sufficiently large and sufficiently small. First, we want to compare the values that we obtain for δ_b^n with the numerical results. The analytical problem did not include the small parameter ρ and the introduction of this term has the effect of shifting the bifurcation points.

5.1 Comparison between the values of δ_b^n

The formula for δ_b^n given in the equation (4.10) does not consider the additional term ρ. Introducing this term, we get a new formula which reads

$$\delta_b^n = \frac{f'(u_1^*)}{\mu} - \frac{(\kappa\mu + \alpha)(\mu + \rho)}{g(u_1^*)\mu^2} = \delta_d + \frac{1}{\mu}\left[f'(u_1^*) - \frac{\kappa\rho + \alpha}{g(u_1^*)} - \frac{\alpha\rho}{g(u_1^*)\mu}\right]. \tag{5.1}$$

We observe that the formulas (4.10) and (5.1), due to the presence of the term ρ, will not give the same values δ_b^n but the two equations correspond if we take $\rho = 0$. We fix the

https://doi.org/10.1017/S0956792516000346 Published online by Cambridge University Press

Table 1. *Comparison between the analytical and numerical bifurcation values. The last two rows compare the numerical and analytical solutions with* $0 < \rho \ll 1$

n	1	2	3	4	5	6	7	8	9
(4.10)	−45.38	−14.45	−8.73	−6.72	−5.80	−5.29	−4.99	−4.79	−4.66
(5.1)	−121.89	−20.81	−10.50	−7.51	−6.24	−5.58	−5.19	−4.94	−4.77
AUTO	−121.89	−20.81	−10.50	−7.51	−6.24	−5.58	−5.19	−4.94	−4.77

following parameter values:

$$(\kappa, \alpha, l, \bar{u}_1, \rho) = (1, 0.2, 20, 0.594, 0.05).$$

We are interested in computing the values of δ_b^n and to observe how the parameter ρ shifts the bifurcation branches.

In Table 1, we reported the bifurcation points δ_b^n for $n \in \{1, 2, \ldots, 9\}$ computed with the two formulas (4.10) and (5.1) in comparison to the numerical continuation results using AUTO. The values detected using AUTO precisely correspond to the values computed with the formula (5.1) as expected whilst the points are shifted in comparison to the values for $\rho = 0$.

5.2 Case 1: α sufficiently large

Recall the formula for δ_b^n given in (4.10):

$$\delta_b^n = \delta_d + \frac{1}{\mu_n}\Big[f'(u_1^*) - \frac{\alpha}{g(u_1^*)}\Big].$$

We observe that if $\alpha > f'(u_1^*)g(u_1^*)$, then $\delta_b^n < \delta_d$ and the branches will approach the limit value δ_d for $n \to \infty$. Since we are using (5.1), the condition on α is

$$\alpha > \mu_n\Big[\frac{f'(u_1^*)g(u_1^*) - \kappa\rho}{\rho + \mu_n}\Big]$$

and, in the case of an interval, we can compute the eigenvalues μ. So, α sufficiently large means

$$\alpha > \Big(\frac{n\pi}{l}\Big)^2\Big[\frac{f'(u_1^*)g(u_1^*) - \kappa\rho}{\rho + (\frac{n\pi}{l})^2}\Big]. \tag{5.2}$$

Figure 4 shows a continuation calculation for fixed parameters

$$(\kappa, \alpha, l, \bar{u}_1, \rho) = (p_2, p_3, p_4, p_5, p_6) = (1, 0.2, 12, 0.594, 0.05)$$

using δ as the primary bifurcation parameter. We observe that the condition on α is satisfied since the right-hand side of (5.2) is negative for all $n \in \mathbb{N}$ and $\alpha = 0.2$. The steady state we start the continuation with is given by

$$(u_1^*, u_2^*) = (\bar{u}_1, f(\bar{u}_1)/\alpha).$$

https://doi.org/10.1017/S0956792516000346 Published online by Cambridge University Press

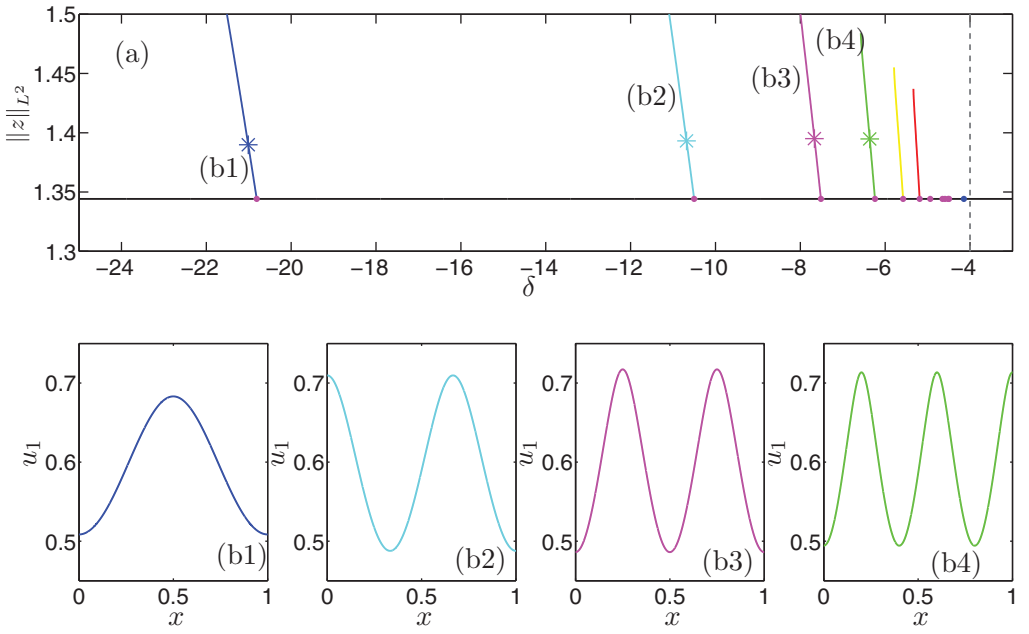

FIGURE 4. Continuation calculation for the system (2.21) with parameter values $(\kappa, \alpha, l, \bar{u}_1, \rho) = (p_2, p_3, p_4, p_5, p_6) = (1, 0.2, 20, 0.594, 0.05)$ and primary bifurcation parameter δ. (a) Bifurcation diagram in $(\delta, \|z\|_{L^2})$-space showing the parameter on the horizontal axis and the solution norm on the vertical axis. Some of the detected bifurcation points are marked as circles (magenta). The last branch point (blue circle) is not a true bifurcation point but results from the degeneracy $\delta = -\kappa/g(u_1^*) =: \delta_\mathrm{d}$. At the other branches points (magenta, filled circles), non-homogeneous solution branches (blue, cyan, magenta, green...) bifurcate via single eigenvalue crossing. The value $\delta^* = -\kappa/\gamma = -4$ is marked by a vertical grey-dashed line. (b) Solutions are plotted for $(x, u_1 = u_1(x))$ at certain points on the non-homogeneous branches; the solutions are marked in (a) using crosses.

We begin the continuation at $\delta = -25$ and we find only one bifurcation point when δ is decreasing, i.e., for $\delta < -25$. This result is expected since $\delta_\mathrm{b}^1 = -121.889$ is the value corresponding to the first eigenvalue. Moreover, we do not detect any bifurcations for $\delta > -4 = \delta^*$. The interesting results in the bifurcation calculation in Figure 4 occur when we increase the primary bifurcation parameter δ. In this case, several branch points are detected, in particular the closer we are to the value δ_d, the more bifurcation points are found. In Figure 4, we have shown the first six branch points detected obtained upon increasing δ. The point detected at $\delta = -20.8116$ corresponds to the second non-trivial bifurcation branch. There are more and more points as we get closer to δ_d. The last point detected (in blue) is not a bifurcation point but corresponds to the degeneracy at

$$\kappa/g(u_1^*) = -1/(0.594(1 - 0.594)) \approx -4.1466.$$

The remaining detected branch points in Figure 4 are true bifurcation points. This numerical result is in accordance with the analytical results on the existence of bifurcations in Theorem 3. In fact, one can carry out the same calculation as in Section 4. At each bifurcation point, a simple eigenvalue crosses the imaginary axis. One can use

https://doi.org/10.1017/S0956792516000346 Published online by Cambridge University Press

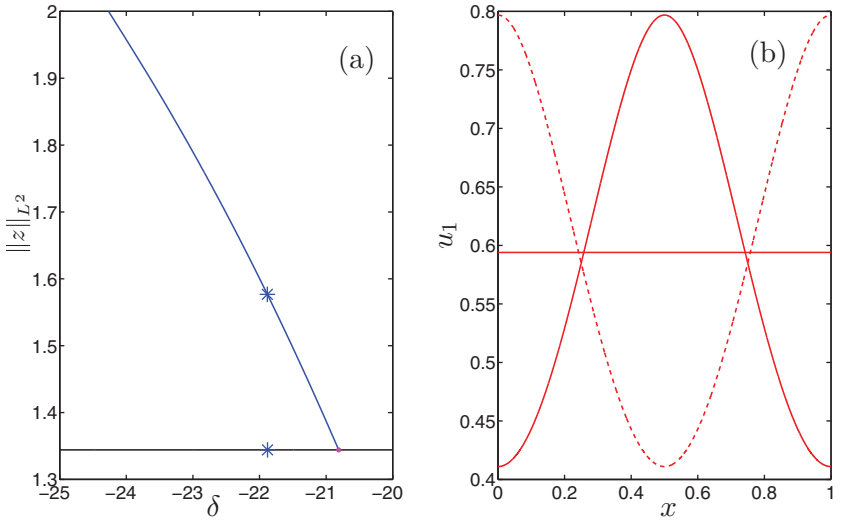

FIGURE 5. Continuation calculation for the system (2.21) as in Figure 4 with a focus on the second bifurcation point (filled circle, magenta). One can show that by using two different local branching directions that two different non-homogeneous solution branches (red) bifurcate via single eigenvalue crossing but the two branches contain solutions with identical L^2-norm for the same parameter value. This is a result of a symmetry in the problem. (b) Three different solutions plotted in $(x, u_1 = u_1(x))$-space at the parameter value $\delta = -21.8819$. The three solutions are marked in (a) using crosses.

the branch switching algorithm implemented in AUTO to compute the non-homogeneous families of solutions as shown for four points in Figure 4(a). In Figure 4(b), we show a representative solution $u_1 = u_1(x)$ on each of the four solution families. The solutions are non-homogeneous steady states and have interface-like behaviour in the spatial variable. Each family has a characteristic number of these interfaces. There are families with even more interfaces than the one shown in Figure 4(b4), which can be found upon increasing δ even further; we are not interested in these highly oscillatory solutions here.

Another observation regarding the continuation run in Figure 4 is reported in more detail in Figure 5 with a focus on the second bifurcation point. It is shown that there are actually two different branches bifurcating at the same point with families of non-homogeneous solutions that are symmetric. In particular, one non-trivial solution branch can be transformed into the other by considering $u \mapsto 1 - u$; as an illustration we refer to three representative numerical solutions on the three branches originating at the second bifurcation point as shown in Figure 5(b).

5.3 Case 2: α sufficiently small

As specified in (M7) when $\alpha < f'(u_1^*)g(u_1^*)$ then $\delta_b^n > \delta_d$ and this means that the branches will approach the limit value δ_d from the right. As pointed out in 5.1, the condition on α

https://doi.org/10.1017/S0956792516000346 Published online by Cambridge University Press

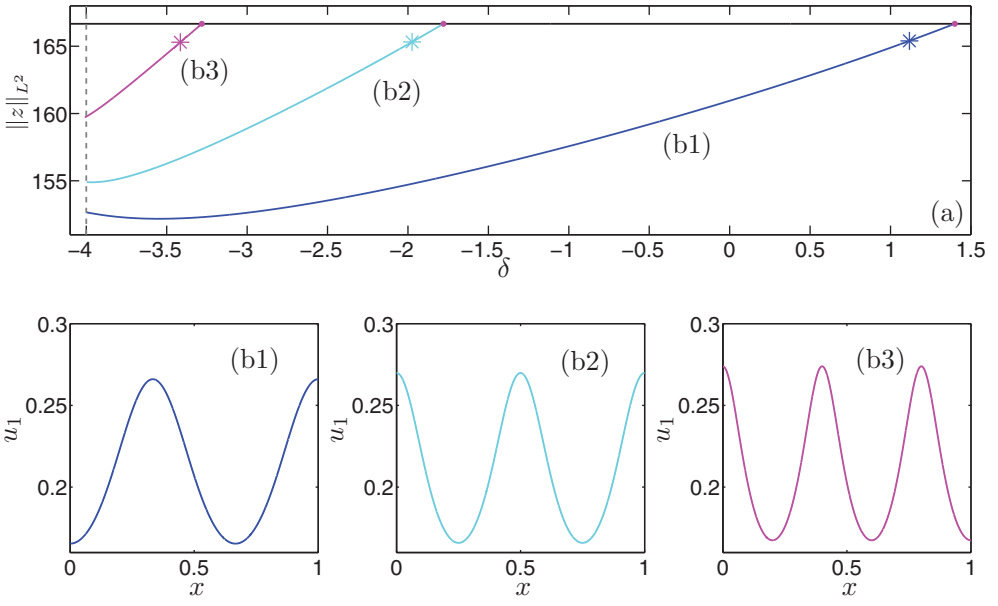

FIGURE 6. Continuation calculation for the system (2.21) with parameter values $(\kappa, \alpha, l, \bar{u}_1, \rho) = (p_2, p_3, p_4, p_5, p_6) = (1, 0.0001, 50, 0.211325, 0.05)$ and primary bifurcation parameter δ. (a) Bifurcation diagram in $(\delta, \|z\|_{L^2})$-space showing the parameter on the horizontal axis and the solution norm on the vertical axis. The detected bifurcation points are marked as circles (magenta). At the three branch points (magenta, filled circles), non-homogeneous solution branches (blue, cyan, magenta) corresponding to $\delta_b^3, \delta_b^4, \delta_b^5$ bifurcate via single eigenvalue crossing. The value $\delta^* = -\kappa/\gamma = -4$ is marked by a vertical grey-dashed line. (b) Solutions are plotted for $(x, u_1 = u_1(x))$ at certain points on the non-homogeneous branches; the solutions are marked in (a) using crosses.

is more complicated since our model contains ρ. The condition on α becomes

$$0 < \alpha < \mu_n \left[\frac{f'(u_1^*)g(u_1^*) - \kappa\rho}{\rho + \mu_n} \right],$$

i.e., we must choose an α which satisfied the inequality for each single μ_n. We fix

$$(\kappa, \alpha, l, \bar{u}_1, \rho) = (1, 0.001, 50, 0.211325, 0.05)$$

for the numerical continuation in this section.

With these values, the condition on α is given by $0 < \alpha < 0.0033827$ which is satisfied. We also find that with our choices

$$\delta_d < \delta_b^n < \delta^* < \delta_b^5 < \delta_b^4 < \delta_b^3 < \delta_b^2 < \delta_b^1 < \delta_e, \quad n \geqslant 6,$$

i.e., there are some bifurcation points which are bigger than δ^* and some which are smaller but all of them are bigger than δ_d. We begin the continuation at $\delta = 3$ and we

https://doi.org/10.1017/S0956792516000346 Published online by Cambridge University Press

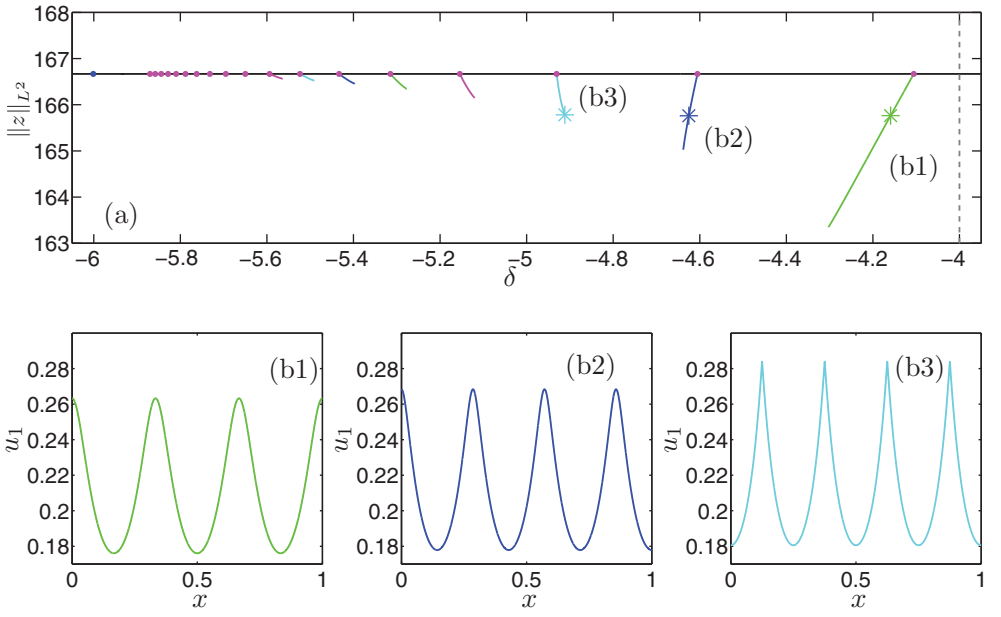

FIGURE 7. Continuation calculation for the system (2.21) with parameter values $(\kappa, \alpha, l, \bar{u}_1, \rho) = (p_2, p_3, p_4, p_5, p_6) = (1, 0.0001, 50, 0.211325, 0.05)$ and primary bifurcation parameter δ. (a) Bifurcation diagram in $(\delta, \|z\|_{L^2})$-space showing the parameter on the horizontal axis and the solution norm on the vertical axis. Some of the detected bifurcation points are marked as circles (magenta). The last branch point (blue circle) is not a true bifurcation point but results from the degeneracy $\delta = -\kappa/g(u_1^*) =: \delta_\mathrm{d}$. At the other branch points (magenta, filled circles), non-homogeneous solution branches (green, blue, cyan) bifurcate via single eigenvalue crossing. The value $\delta^* = -\kappa/\gamma = -4$ is marked by a vertical grey dashed line. (b) Solutions are plotted for $(x, u_1 = u_1(x))$ at certain points on the non-homogeneous branches; the solutions are marked in (a) using crosses.

detect only two more branches when we increase δ at $\delta = 43.4851$ and $\delta = 9.98041$ which correspond to δ_b^1 and δ_b^2. We focus on the branches for $n \in \{1, 2, 3, 4, 5\}$ such that $\delta_\mathrm{b}^n > \delta^*$. This case is represented in Figure 6.

Numerically, we observe that all the branches stop when they reach the critical value δ^*. Next, we consider $n \geqslant 6$ such that $\delta_\mathrm{d} < \delta_\mathrm{b}^n < \delta^*$ as reported in Figure 7. In this case, there are two critical values: $\delta^* = -4$ (dashed line) and $\delta_\mathrm{d} = -6$ (blue circle). The branches detected for a δ close to δ^* have the same direction as the branches detected for $\delta > \delta^*$; but starting from a certain n, in this case $n = 8$, we notice that the branches change the direction. Probably this behaviour is due to the fact that the branches cannot cross the value $\delta = \delta_\mathrm{d}$. We do no detect any branch for $\delta < \delta_\mathrm{d}$.

In the range between δ_d and δ^*, the branches do not seem to overlap. Numerically, one observes that the branches get shorter and shorter due to the numerical continuation breaking down as the branches approach δ_d. Looking at the shape of the solutions in the different branches, we can observe that they have more and more interfaces as we approach the limiting value δ_d. Moreover, the solutions inside a fixed branch get sharper and sharper peaks along the branch (see for example the cyan branch).

https://doi.org/10.1017/S0956792516000346 Published online by Cambridge University Press

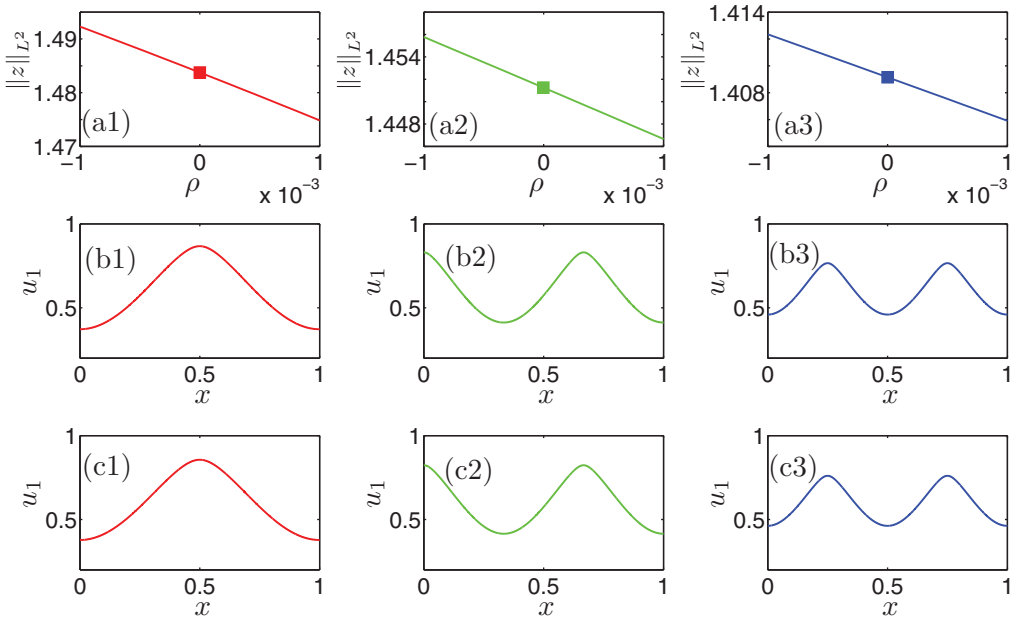

FIGURE 8. Continuation calculation for the system (2.21) starting with the same basic parameter values as in Figure 4 but with $\rho = 0.001$. We stop the continuation at the solution points for a certain δ (as done in Figure 4(a)) and change from δ as a primary continuation parameter to ρ as a primary parameter with the goal to decrease the parameter to $\rho = 0$. The values for δ are $\delta = -16$ for the red branch, $\delta = -9.4$ for the green branch and $\delta = -7$ for the blue one. (a1)–(a3) Bifurcation diagrams in $(\rho, \|z\|_{L^2})$-space. The starting point for the continuation is at the right boundary where $\rho = 0.001$ and then ρ is decreased. (b1)–(b3) Solutions obtained on the bifurcation branches above at the point $\rho = 0$ (points are marked with squares in (a1)–(a3)). (c1)–(c3) Solutions obtained on the bifurcation branches for the initial system with $\rho = 0.001$. We can observe that also for $\rho = 0$ the solutions have a non-trivial herding-type profile.

5.4 Continuation in ρ

The next question is if we can find non-homogeneous steady states also for the original problem with $\rho = 0$. This can be achieved by using a homotopy-continuation idea.

First, we continue the problem in δ and compute the non-homogeneous solution branches. Then, we pick a steady state on the non-homogeneous branch and switch to continuation in ρ whilst keeping δ fixed. The results of this strategy are shown in Figure 8 (for $\alpha = 0.2$) and in Figure 9 (for $\alpha = 0.001$). For the first three solutions shown in Figure 4(b), this strategy works if we start from a very small ρ. Figure 8(c) shows the solution in the branch for a $\rho \neq 0$: We notice that the solutions for the case $\rho = 0$ keep the non-constant profile as for $\rho \neq 0$ yielding relevant herding solutions for applications.

In the case with α sufficiently small, the strategy works better and we indeed find non-homogeneous steady states for $\rho = 0$ as shown in Figure 9(b). Moreover, we can also obtain herding solutions. We use the starting parameter values:

$$(\kappa, \alpha, l, \bar{u}_1, \rho) = (1, 0.001, 50, 0.211325, 0.05).$$

https://doi.org/10.1017/S0956792516000346 Published online by Cambridge University Press

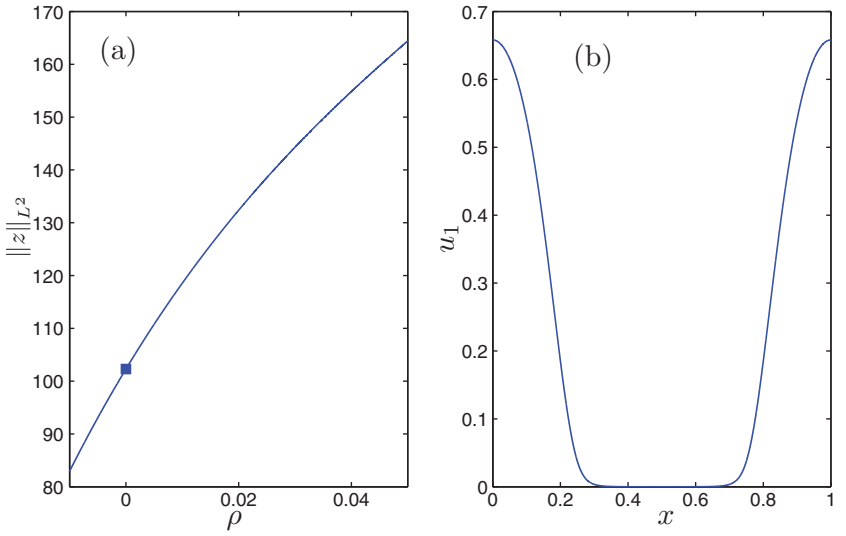

FIGURE 9. Continuation calculation for the system (2.21) starting with the same parameter value and as in Figure 6. We stop the continuation at $\delta = -9$ (as done in Figure 6(a)) and change from δ as a primary continuation parameter to ρ as a primary parameter with the goal to decrease the parameter to $\rho = 0$. (a) Bifurcation diagram in $(\rho, \|z\|_{L^2})$-space. The starting point for the continuation is at the right boundary where $\rho = 0.05$ and then ρ is decreased. (b) Solution on the second branch δ_b^2 of non-homogeneous steady states at $\rho = 0$ (point is marked with squares in (a)).

We start from $\delta = 10$ and the first branch we detect is $\delta_b^2 = 9.98041$. Once we are in this branch, we continue in ρ for a fixed δ (in this case $\delta = -9$). For information herding models, solutions which are of particular importance are those with sharp interfaces between the endstates, i.e., the solution is near zero and near one in certain regions with sharp interfaces in between. These solutions represent a herding effect in the sense of sharply split opinions. More precisely, they indicate for which values of the information variable x we observe a herding behaviour, i.e., a concentration of individuals ($u \approx 1$) at certain values of x. Figure 9(b) shows herding in the interval $[0, 0.2] \cup [0.8, 1]$, whilst only a few individuals adopt the information value in $[0.3, 0.7]$.

5.5 Solutions and other parameters

In this section, we focus on the case with α sufficiently small. We are interested in studying, how the solutions change depending on the other parameters κ and l. We fix as starting parameters

$$(\kappa, \alpha, l, \bar{u}_1, \rho) = (1, 0.001, 50, 0.211325, 0.05)$$

and consider the branch δ_b^2. We study the solutions depending on the different parameters. In Figure 10, we show changes along the branch (which bifurcates at $\delta = 9.98041$). We observe that the shape is the same along the branch but the interfaces sharpen as δ is decreased.

In Figure 11, we show how the solution changes with the length of the domain. We consider $l = 20$, $l = 50$, and $l = 100$. The branch δ_b^2 is detected at

https://doi.org/10.1017/S0956792516000346 Published online by Cambridge University Press

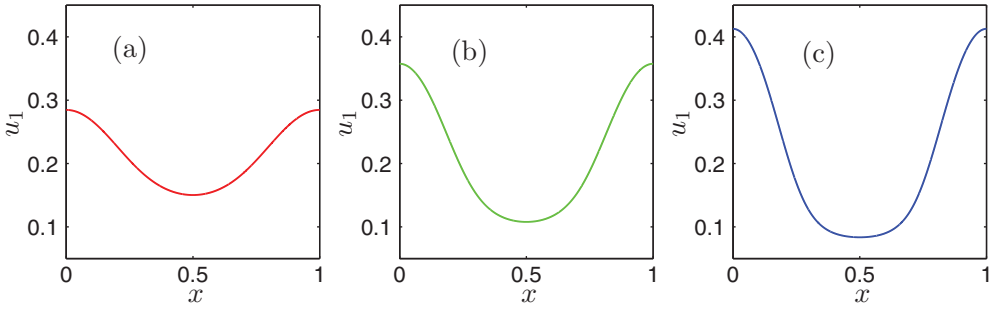

FIGURE 10. Solutions along the branch δ_b^2 for the system (2.21) with parameter values $(\kappa, \alpha, l, \bar{u}_1, \rho) = (p_2, p_3, p_4, p_5, p_6) = (1, 0.001, 50, 0.211325, 0.05)$. (a) Solution of non-homogeneous steady states at $\delta = 8.72901$. (b) Solution of non-homogeneous steady states at $\delta = 5.76477$. (c) Solution of non-homogeneous steady states at $\delta = 1.548$.

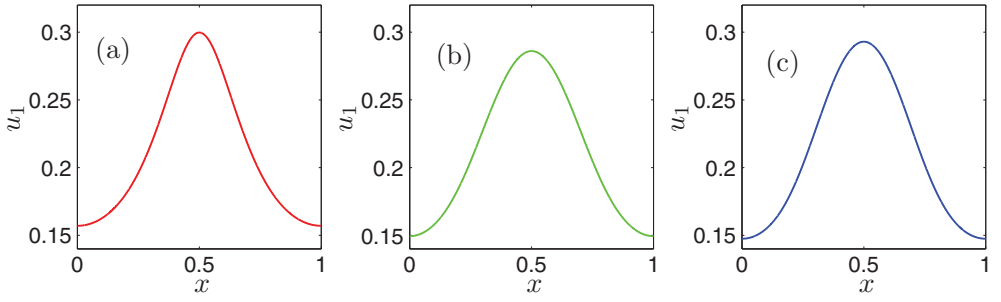

FIGURE 11. Solutions in the branch δ_b^2 for the system (2.21) with parameter values $(\kappa, \alpha, \bar{u}_1, \rho) = (p_2, p_3, p_5, p_6) = (1, 0.001, 0.211325, 0.05)$. (a) Solution of non-homogeneous steady states at $\delta = -3.5154$, $l = 20$. (b) Solution of non-homogeneous steady states at $\delta = 8.93964$, $l = 50$. (c) Solution of non-homogeneous steady states at $\delta = 37.9117$, $l = 100$.

$\delta = -3.28144, 9.98041, 43.4851$, respectively. Since we consider the same branch, the shape does not change and length of the domain shifts the bifurcation points and just scales the solution.

When we change the parameter κ the bifurcation points are also simply shifted. We consider $\kappa = 1$, $\kappa = 5$, and $\kappa = 10$. The branch δ_b^2 is detected at $\delta = 9.98041, -92.2877, -214.999$, respectively. Moreover, for the first case, the branches approach the value δ_d from the right, whilst in the other two cases from the left. As for the previous case, we consider three different solutions with (almost) the same norm (163.863 for the case (a), 163.872 for (b) and 163.911 for (c)).

In summary, we conclude that κ and l do not seem to be the parameters of primary importance in our context as we can re-obtain similar solutions and similar bifurcation structures for different values of κ and l upon varying δ, α as primary parameters.

6 Outlook

So far, relatively little attention has been devoted to the study of the parameter space interfaces of different mathematical methods. In this contribution, we have analysed as

https://doi.org/10.1017/S0956792516000346 Published online by Cambridge University Press

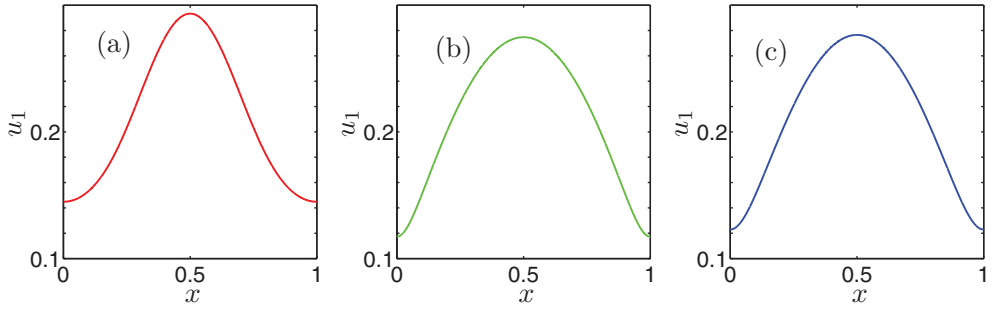

FIGURE 12. Solutions in the branch δ_b^2 for the system (2.21) with parameter values $(\alpha, l, \bar{u}_1, \rho) = (p_3, p_4, p_5, p_6) = (0.001, 50, 0.211325, 0.05)$. (a) Solution of non-homogeneous steady states at $\delta = 8.72901$, $\kappa = 1$. (b) Solution of non-homogeneous steady states at $\delta = -92.2877$, $\kappa = 5$. (c) Solution of non-homogeneous steady states at $\delta = -220.578$, $\kappa = 10$.

an example a cross-diffusion herding model to understand where, and how, the global non-linear analysis approach via entropy variables is connected to bifurcation analysis techniques from dynamical systems. We have shown that both approaches encounter similar problems regarding the degeneracy of the diffusion matrix and we were able to cover different parameter regimes by combining the results of the two methods.

This paper is only a first starting point. Here, we shall just mention a few ideas for future work.

The next step is to analyse the regime $\alpha \to 0$ and to check whether the limitation in (2.8) on α can be improved, or not. In this regard, one also has to consider in which sense the forward problem should be interpreted for moderate and small values of α and for $\delta < \delta_d$. Recent work [35] suggests that one should not only use the notion of Petrovskii ellipticity for the stationary problem [39] but also consider it in the parabolic context; see the classical survey [5].

The next step is to expand the approach to other examples. In particular, many reaction–diffusion systems as well as other classes of PDEs have natural entropies, which can be used to study global existence and convergence properties. In the non-linear case, one frequently can also employ approaches from dynamical systems to understand the dynamics of the PDE. Using a similar approach as we presented here could be illuminating for other examples. For example, it is natural to conjecture that there are examples in applications, which exhibit the following characteristics:

(Z1) There exists one fixed parameter region in which the entropy method yields global decay. Upon variation of a single parameter, the validity boundary of the entropy method coincides precisely with an isolated local supercritical bifurcation point.

(Z2) There exists one fixed parameter region in which the entropy method yields global decay. Upon variation of a single parameter, the validity boundary of the entropy method does not coincide with a local bifurcation point. Instead, the obstruction is a global bifurcation branch in parameter space with a fold point precisely at the validity boundary.

https://doi.org/10.1017/S0956792516000346 Published online by Cambridge University Press

In this work, we apparently found a more complicated case as shown in Figure 1. However, it seems plausible that the cases (Z1) and (Z2) should occur even in classical problems without cross-diffusion, i.e., reaction–diffusion equations with a diagonal positive-definite diffusion matrix. Determining whether this is true for several classical examples from applications is an interesting open problem.

Regarding the entropy method [8, 13], it would be interesting to investigate in more detail parametric scenarios for its validity regime. For example, the question arises whether it is possible to find criteria for the validity range that are computable for entire classes of PDEs. The entropy approach relies on upper bounds. Although the bounds we present here turn out to be sharp in the sense of global decay dynamics in a suitable singular limit, this may not always be easy to achieve as demonstrated by the $\alpha \to 0$ case discussed above. It would be relevant to estimate *a priori*, which regime in parameter space one fails to cover if certain non-optimal upper bounds are used. As above, carrying this out for several examples could already be very illuminating.

Regarding the analytical and numerical bifurcation analysis, there are multiple strategies to deal with the problem of mass conservation, or more generally with higher dimensional solution manifolds. For example, one may try to compute the entire solution family of steady states parametrized by the mass numerically [17, 23], which yields a numerical continuation problem for higher dimensional manifolds and not only curves. Furthermore, we have focussed on the numerical problem in the one-dimensional setup and computing the two- and three-space dimension cases could be interesting [**?**, 42]. Regarding analytical generalizations, a possible direction is to view δ^* as a singular limit and phrase the problem as a perturbation problem [2, 19, 38].

Acknowledgements

We would like to thank two anonymous referees for very helpful comments and suggestions that led to several improvements.

References

[1] Arnold, A., Abdallah, N. B. & Negulescu, C. (1996) Liapunov functionals and large-time-asymptotics of mean-field nonlinear Fokker-Planck equations. *Transp. Theory Stat. Phys.* **25**(7), 733–751.

[2] Achleitner, F. & Kuehn, C. (2015) On bounded positive stationary solutions for a nonlocal Fisher-KPP equation. *Nonl. Anal. A: Theor. Meth. Appl.* **112**, 15–29.

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

[4] Arnold, A., Markowich, P., Toscani, G. & Unterreiter, A. (2001) On convex Sobolev inequalities and the rate of convergence to equilibrium for Fokker-Planck type equations. *Commun. Partial Differ. Equ.* **26**(1–2), 43–100.

[5] Agranovich, M. S. & Vishik, M. I. (1964) Elliptic problems with a parameter and parabolic problems of general type. *Russ. Math. Surv.* **19**(3), 53–157.

[6] Bakry, D., Gentil, I. & Ledoux, M. (2014) *Analysis and Geometry of Markov Diffusion Operators*, Springer.

[7] Burger, M., Markowich, P. & Pietschmann, J.-F. (2011) Continuous limit of a crowd motion and herding model: Analysis and numerical simulations. *Kinet. Relat. Mod.* **4**, 1025–1047.

[8] Carrillo, J. A., Jüngel, A., Markowich, P. A., Toscani, G. & Unterreiter, A. (2001) Entropy dissipation methods for degenerate parabolic problems and generalized Sobolev inequalities. *Monatshefte für Math.* **133**(1), 1–82.

[9] Chertock, A., Kurganov, A., Wang, X. & Wu, Y. (2012) On a chemotaxis model with saturated chemotactic flux. *Kinet. Relat. Mod.* **5**, 51–95.

[10] Crandall, M. G. & Rabinowitz, P. H. (1971) Bifurcation from simple eigenvalues. *J. Funct. Anal.* **8**(2), 321–340.

[11] Crandall, M. G. & Rabinowitz, P. H. (1973) Bifurcation, perturbation of simple eigenvalues and linearized stability. *Arch. Ration. Mech. Anal.* **52**, 161–180.

[12] Doedel, E. J., Champneys, A., Dercole, F., Fairgrieve, T., Kuznetsov, Y., Oldeman, B., Paffenroth, R., Sandstede, B., Wang, X. & Zhang, C. (2007) Auto 2007p: Continuation and bifurcation software for ordinary differential equations (with homcont). URL: http://cmvl.cs.concordia.ca/auto, accessed July 1st, 2016.

[13] Desvillettes, L. & Fellner, K. (2006) Exponential decay toward equilibrium via entropy methods for reaction-diffusion equations. *J. Math. Anal. Appl.* **319**(1), 157–176.

[14] Desvillettes, L. & Fellner, K. (2007) Entropy methods for reaction-diffusion systems. *Discrete Cont. Dyn. Sys. (suppl.)* 304–312.

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

[16] Delitala, M. & Lorenzo, T. (2014) A mathematical model for value estimation with public information and herding. *Kinet. Relat. Mod.* **7**, 29–44.

[17] Dankowicz, H. & Schilder, F. (2013) *Recipes for Continuation*. SIAM.

[18] Evans, L. C. (2002) *Partial Differential Equations*, AMS.

[19] Fife, P. C. (1973) Semilinear elliptic boundary value problems with small parameters. *Arch. Ration. Mech. Anal.* **52**(3), 205–232.

[20] Gabriel, P. (2012) Long-time asymptotics for nonlinear growth-fragmentation equations. *Commun. Math. Sci.* **10**, 787–820.

[21] Govaerts, W. F. (1987) *Numerical Methods for Bifurcations of Dynamical Equilibria*, SIAM, Philadelphia, PA.

[22] Galiano, G. & Selgas, V. (2014) On a cross-diffusion segregation problem arising from a model of interacting particles. *Nonlin. Anal.: Real World Appl.* **18**, 34–49.

[23] Henderson, M. E. (2002) Multiple parameter continuation: Computing implicitly defined k-manifolds. *Int. J. Bif. Chaos* **12**(3), 451–476.

[24] Hittmeir, S. & Jüngel, A. (2011) Cross diffusion preventing blow up in the two-dimensional Keller-Segel model. *SIAM J. Math. Anal.* **43**, 997–1022.

[25] Horstmann, D. (2011) Generalizing the keller-segel model: Lyapunov functionals, steady state analysis, and blow-up results for multi-species chemotaxis models in the presence of attraction and repulsion between competitive interacting species. *J. Nonlin. Sci.* **21**, 231–270.

[26] Hillen, T. & Painter, K. (2002) Volume filling and quorum sensing in models for chemosensitive movement. *Canad. Appl. Math. Quart.* **10**, 501–543.

[27] Januauskas, A. (1998) Classification of second-order partial differential equation systems elliptic in the petrovskii sense. *Lithuanian Math. J.* **38**, 59–63.

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

[29] Jiang, J. & Zhang, Y. (2009) On convergence to equilibria for a chemotaxis model with volume-filling effect. *Asympt. Anal.* **65**, 79–102.

[30] Keller, H. (1977) Numerical solution of bifurcation and nonlinear eigenvalue problems. In: P. Rabinowitz (editor), *Applications of Bifurcation Theory*, Academic Press, pp. 359–384.

[31] Kielhoefer, H. (2004) *Bifurcation Theory: An Introduction with Applications to PDEs*, Springer.

[32] Krauskopf, B., Osinga, H. M. & Galán-Vique, J. (editors) (2007) *Numerical Continuation Methods for Dynamical Systems: Path Following and Boundary Value Problems*, Springer.

[33] Keller, E. & Segel, S. (1970) Initiation of slime mold aggregation viewed as an instability. *J. Theor. Biol.* **26**, 399–415.
[34] Kuehn, C. (2015). Efficient gluing of numerical continuation and a multiple solution method for elliptic PDEs. Applied Mathematics and Computation, **266**, 656–674.
[35] Lions, P.-L. (2015) Some new classes of nonlinear Kolmogorov equations. Talk at the 16th Pauli Colloquium, Wolfgang-Pauli Institute.
[36] Liero, M. & Mielke, A. (2013) Gradient structures and geodesic convexity for reaction-diffusion systems. *Phil. Trans. Roy. Soc. A* **371**, 20120346.
[37] Lambda, H. & Seaman, T. (2008) Market statistics of a psychology-based heterogeneous agent model. *Intern. J. Theor. Appl. Finance* **11**, 717–737.
[38] Ni, W.-M. (1998) Diffusion, cross-diffusion and their spike-layer steady states. *Not. Amer. Math. Soc.* **45**(1), 9–18.
[39] Shi, J. & Wang, X. (2009) On the global bifurcation for quasilinear elliptic systems on bounded domains. *J. Differ. Equ.* **246**, 2788–2812.
[40] Temam, R. (1997) *Infinite-Dimensional Dynamical Systems in Mechanics and Physics*, Springer.
[41] Uhlenbeck, K. (1972) Eigenfunctions of Laplace operators. *Bull. Amer. Math. Soc.* **78**, 1073–1076.
[42] Uecker, H., Wetzel, D. & Rademacher, J. D. M. (2014) pde2path - A Matlab package for continuation and bifurcation in 2D elliptic systems. *Num. Math.: Th. Meth. Appl.* **7**, 58–106.
[43] Wolansky, G. (2002) Multi-components chemotactic system in the absence of conflicts. *Europ. J. Appl. Math.* **13**, 641–661.
[44] Wrzosek, D. (2004) Global attractor for a chemotaxis model with prevention of overcrowding. *Nonlin. Anal.* **59**, 1293–1310.
[45] Wang, X. & Xu, Q. (2013) Spiky and transition layer steady states of chemotaxis systems via global bifurcation and Helly's compactness theorem. *J. Math. Biol.* **66**(6), 1241–1266.
[46] Zinsl, J., & Matthes, D. (2015). Transport distances and geodesic convexity for systems of degenerate diffusion equations. Calculus of Variations and Partial Differential Equations, **54**(4), 3397–3438.