



HAL
open science

Computer-assisted proofs for some nonlinear diffusion problems

Maxime Breden

► **To cite this version:**

Maxime Breden. Computer-assisted proofs for some nonlinear diffusion problems. 2021. hal-03168195v1

HAL Id: hal-03168195

<https://hal.science/hal-03168195v1>

Preprint submitted on 12 Mar 2021 (v1), last revised 22 Nov 2023 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Computer-assisted proofs for some nonlinear diffusion problems

Maxime Breden*

January 28, 2021

Abstract

In the last three decades, powerful computer-assisted techniques have been developed in order to validate a posteriori numerical solutions of semilinear elliptic problems of the form $\Delta u + f(u, \nabla u) = 0$. By studying a well chosen fixed point problem defined around the numerical solution, these techniques make it possible to prove the existence of a solution in an explicit (and usually small) neighborhood the numerical solution. In this work, we develop a similar approach for a broader class of systems, including nonlinear diffusion terms of the form $\Delta\Phi(u)$. In particular, this enables us to obtain new results about steady states of a cross-diffusion system from population dynamics: the (non-triangular) SKT model. We also revisit the idea of automatic differentiation in the context of computer-assisted proof, and propose an alternative approach based on differential-algebraic equations.

1 Introduction

1.1 Context

Diffusion is a key mechanism in many spatially extended system coming from Physics, Chemistry or Biology. Starting from the prototypical mathematical model used to describe such phenomena that is the heat equation

$$\partial_t u = \Delta u,$$

many more general models have been introduced, for instance in order to take into account nonlinear diffusion effects

$$\partial_t u = \Delta\Phi(u). \tag{1}$$

Some prime examples which have already been studied extensively in the mathematical literature are the fast-diffusion equation and the porous medium equation, which correspond to $\Phi(u) = u^m$ with $m < 1$ and $m > 1$ respectively [2, 37]. More general nonlinearities are also of interest, for instance $\Phi(u) = \log(u)$ which corresponds to the Ricci flow in two dimensions [12, 34].

Nonlinear diffusion also plays a crucial role for systems, where $u = (u_1, \dots, u_d)$, especially when the diffusion rate of one component can be affected by the other components. This phenomenon is often referred to as cross-diffusion, and typical mathematical models used in this case are of the form

$$\partial_t u = \operatorname{div}(A(u)\nabla u). \tag{2}$$

In some cases, namely when the matrix $A(u)$ is the derivative of some map Φ , these systems can also be written under the form (1).

For some of these nonlinear diffusion problems, like the fast-diffusion equation or the porous medium equation, the long-time behavior of the solutions is well understood, see for instance the monograph [36]

*CMAP, École Polytechnique, route de Saclay, 91120 Palaiseau, France.

and the references therein. For more general nonlinear diffusion equations, and especially for systems, the situation is more complex, but in some cases entropy method can be used to guarantee at least the existence of global in time solutions, see e.g. [15].

However, in most applications the models also include nonlinear reaction terms, that is, (1) is replaced by an equation of the form

$$\partial_t u = \Delta \Phi(u) + R(u).$$

In this case the dynamics can be way more complex, therefore the study of the long time behavior is usually significantly more complicated. Even in some specific cases where one can prove that the solution converges to a steady state, see e.g. the early work [16], solving the corresponding stationary problem

$$0 = \Delta \Phi(u) + R(u), \tag{3}$$

is still a demanding task, especially if one is not only interested in existence results but also in quantitative information about the solution(s).

In this work, we develop a methodology to get quantitative existence results about equations of the form (3) on bounded domains, based on computer-assisted proofs. Before giving more details about the scope and the limitation of our results, let us briefly review some past computer-assisted works on elliptic equations upon which we build.

1.2 Computer-assisted proofs for elliptic equations

Computer-assisted proofs for elliptic equations originate from the pioneering works of Nakao [23] and Plum [28], and were then further developed and popularized by them and many others, see e.g. [26, 39, 40, 1, 10, 33, 35, 38, 11].

Most of these works share the same general approach, which consists in validating a posteriori an approximate solution \bar{u} (for an alternative approach, having a topological rather than functional analytic flavor, see [40]). In order to conduct this *a posteriori validation*, one considers the elliptic PDE as an $F(u) = 0$ problem on some well chosen function space, and then turns it into a fixed point problem $T(u) = u$, where the fixed point operator T should be a contraction¹ in a neighborhood of \bar{u} . In practice, the two essential steps are the definition of a suitable fixed point operator T , and the derivation of explicit estimates allowing to apply Banach's fixed point theorem to T in a neighborhood of \bar{u} . So as to consistently get a contraction, T is usually defined as something close to

$$T(u) = u - DF(\bar{u})^{-1}F(u),$$

where DF denotes the Frechet derivative of F . Provided \bar{u} is a sufficiently good approximate solution, meaning that $\|F(\bar{u})\|$ is small enough, the key requirement for applying Banach's fixed point theorem to T is to get an explicit control on $\|DF(\bar{u})^{-1}\|$. At this stage, there are three main options.

1. If the problem is defined on Hilbert spaces, eigenvalues bounds can be used to directly estimate the norm of $DF(\bar{u})^{-1}$ [24, Part II]. While this approach is typically more demanding in practice than the other two presented below, it is also more versatile, and can for instance handle unbounded domains.
2. On bounded domains, an alternative approach consists in studying $DF(\bar{u})^{-1}$ by introducing a well chosen *approximate inverse* A of $DF(\bar{u})$, which is obtained using a finite dimensional projection. In that case, it is more or less equivalent to directly work with T defined by

$$T(u) = u - AF(u).$$

¹Historically, some of the early works used a slightly different strategy based on Schauder's fixed point theorem. However, it seems most of the techniques gradually evolved and aligned towards the usage of a contraction mapping, if only because that is often the easiest way to enforce the inclusion assumption needed for Schauder's fixed point theorem.

Since A is then defined rather explicitly, computing $\|A\|$ becomes straightforward, but the difficulty is transferred to proving that A is a good enough approximate inverse, meaning that $\|I - ADF(\bar{u})\| < 1$, in order for T to still be a contraction.

3. Still on bounded domains, a third option was recently introduced [31]. While it aims at directly studying $DF(\bar{u})^{-1}$ itself like the first option, it does so by also introducing a finite dimensional projection to decompose $DF(\bar{u})$ and then using the Schur complement. Therefore, in practice this approach ends up being close to the second one, as it requires very similar estimates, but those are combined together slightly differently in the end.

Of course, in order to prove that T is a contraction one also has to control the nonlinear terms, but since such control is only required on a small neighborhood of \bar{u} , this is usually not the main difficulty. The whole procedure boils down to a kind of Newton-Kantorovich theorem [27], an example of which will be presented in details in Section 2.2.

As we already mentioned, the general strategy and its variations that we just described were successfully applied to several elliptic problems. However, in all these studies, the leading differential operator is always simply Δu , or $\Delta^2 u$. The only exceptions seem to be [5] and [6], where some very specific non-constant coefficients in front of the Laplacian are considered.

1.3 Generalization to some nonlinear diffusion problems

In this paper, we generalize the framework presented above, and more specifically option 2, in order to treat problems of the form

$$\begin{cases} \Delta\Phi(u) + R(u) = 0 & \text{on } \Omega \\ \frac{\partial u}{\partial n} = 0 & \text{on } \partial\Omega \end{cases} \quad (4)$$

on a bounded domains. The key step is to introduce an approximate inverse A with a different structure than those which are typically used for linear diffusion.

Before going further, let us discuss the assumptions we are going to make in this work, and specify how relevant they are.

- We restrict our attention to rectangular domains, i.e. of the form $\prod[a_i, b_i]$. This allows us to use a discretization based on Fourier series, which makes the derivation of some of the estimates easier. However, we do not consider this assumption to be essential: the same ideas could be adapted to a discretization based on finite elements, in order to handle more general domains.
- Similarly, we consider Neumann boundary conditions because they are natural for the main application we have in mind (namely the SKT system which describe the density of several species in a bounded environment), but the analysis would be similar with Dirichlet boundary conditions.
- To finish with the domain, we rely crucially on the fact that the domain is bounded, and a significantly different approach would be needed for unbounded domains. While we did not explore this possibility thoughtfully, we believe that option 1 described in Section 1.2 could be generalized to handle nonlinear diffusion terms as in (4).
- We chose not to include a first order term ∇u in (4) in order to keep the presentation as simple as possible, but adding such a term presents no essential difficulty.
- Still in order to limit technicalities and to focus on the main novelty of this work, namely the treatment of the nonlinear diffusion, we only consider examples where the reaction terms R are at most quadratic. This can also be easily generalized, and we make some remarks in that regard later on.

- We also emphasize already that our method does not require specific assumptions on the nonlinearity Φ , except smoothness. We do have some non-degeneracy condition, but it will appear as an a posteriori information, rather than as an a priori requirement. That is, if we manage to validate a solution u to (4), it will mean that $D\Phi(u)$ is nonsingular, but we do not need to know this in advance for the method to work. For instance, it could be that $D\Phi(x)$ is singular for some values of x , but that the solution u ends up never taking these values. In some sense, this means the problem (4) could have been rewritten as

$$\Delta v + R(\Phi^{-1}(v)) = 0,$$

but our analysis never requires us to deal with Φ^{-1} . Finally, let us point out that, as soon as we know that $D\Phi(u)$ is nonsingular, we can rewrite the boundary conditions in (4) in a form which directly involves fluxes, namely $\frac{\partial \Phi(u)}{\partial n} = 0$.

Remark 1.1. *The approach developed in this work is also suitable for uniformly elliptic problems of the form $\Delta(au) + R(u) = 0$, where a is a non constant coefficient (or matrix in the case of systems), which already fall outside of the usual framework for computed-assisted proofs for equations of the form $\Delta u + R(u) = 0$.*

The remainder of this paper is organized as follows. In Section 2 we start with a simple case, namely a scalar equation with $\Phi(u) = u^2$, in order to expose the main ideas without too many technicalities. In Section 3 we then focus on the SKT system, and present some new results about its steady states. Finally, we discuss in Section 4 how to handle more complicated, non-polynomial, diffusion terms, and propose an alternative to the usual automatic-differentiation technique.

All the Matlab codes used for the computer-assisted parts of the proofs are available at [3].

2 A simple case first

In this section, we consider a scalar problem on a one-dimensional domain of the form

$$\begin{cases} \Delta \Phi(u) + R(u) = 0 & \text{on } (0, 1), \\ \frac{\partial u}{\partial n} = 0 & \text{on } \{0, 1\}, \end{cases} \quad (5)$$

where $\Phi(u) = u^2$ and $R(u) = \alpha u - \beta u^2 + g$.

2.1 Notations and sequence spaces

The material presented in this subsection is standard, and mostly included for the sake of fixing notations.

We look for solutions as Fourier series, i.e.

$$u(x) = u_0 + 2 \sum_{n=1}^{\infty} u_n \cos(n\pi x) = \sum_{n \in \mathbb{Z}} u_{|n|} \cos(n\pi x),$$

and denote by $\mathbf{u} = (u_n)_{n \geq 0}$ the sequence of Fourier coefficients associated to a function u . In the sequel, we always use this convention that a bold symbol denotes the sequence of Fourier coefficients associated to the function of the same name. For instance $\mathbf{1}$ is the sequence $(1, 0, \dots, 0, \dots)$ representing the constant function equal to 1.

For any $\nu \geq 1$, we consider

$$\begin{aligned} \|\mathbf{u}\|_{\nu} &= |u_0| + 2 \sum_{n=1}^{\infty} |u_n| \nu^n \\ &= \sum_{n \geq 0} |u_n| \xi_n(\nu), \end{aligned}$$

where

$$\xi_n(\nu) = \begin{cases} 1 & n = 0, \\ 2\nu^n & n \geq 1, \end{cases}$$

and the associated Banach space

$$\ell_\nu^1 = \left\{ \mathbf{u} = (u_n)_{n \geq 0}, \|\mathbf{u}\|_\nu < \infty \right\}.$$

Notice that, as soon as $\nu > 1$, if $\mathbf{u} \in \ell_\nu^1$ then the coefficients u_k decay geometrically, and therefore the associated function u is smooth.

The product of two functions u and v in function space gives rise to the discrete convolution product $*$ in sequence space:

$$(\mathbf{u} * \mathbf{v})_n = \sum_{k \in \mathbb{Z}} u_{|k|} v_{|n-k|}.$$

We recall that ℓ_ν^1 is a Banach algebra under the convolution product: for all \mathbf{u}, \mathbf{v} in ℓ_ν^1 ,

$$\|\mathbf{u} * \mathbf{v}\|_\nu \leq \|\mathbf{u}\|_\nu \|\mathbf{v}\|_\nu.$$

In particular, since Φ is assumed to be a polynomial, for any \mathbf{u} in ℓ_ν^1 one can readily define $\Phi(\mathbf{u})$, which also belongs to ℓ_ν^1 , and similarly for R .

Remark 2.1. *In order for $R(\mathbf{u})$ (or $\Phi(\mathbf{u})$) to be well defined in ℓ_ν^1 , it is sufficient to assume that R is analytic on a disk of radius larger than $\|\mathbf{u}\|_\nu$. However, in some cases one might avoid the technicalities associated with having non polynomial terms by introducing extra variables and using automatic differentiation techniques, see e.g. [18]. A new alternative approach is also presented in Section 4.*

Given $\mathbf{u} \in \ell_\nu^1$, we denote by $M(\mathbf{u}) : \ell_\nu^1 \rightarrow \ell_\nu^1$ the associated multiplication operator, i.e. for all v in ℓ_ν^1 , $M(\mathbf{u})v = \mathbf{u} * v$.

Let L be a linear operator on ℓ_ν^1 . L can be identified with an *infinite dimensional matrix* $(L_{k,n})_{k,n \geq 0}$ where, for any \mathbf{u} in ℓ_ν^1 ,

$$(L\mathbf{u})_k = \sum_{n=0}^{\infty} L_{k,n} u_n \quad \forall k \geq 0.$$

We recall that the operator norm of L can be easily expressed in terms of the coefficients $L_{k,n}$, namely

$$\|L\|_\nu := \sup_{\substack{\mathbf{u} \in \ell_\nu^1 \\ \mathbf{u} \neq 0}} \frac{\|L\mathbf{u}\|_\nu}{\|\mathbf{u}\|_\nu} = \sup_{n \geq 0} \frac{1}{\xi_n(\nu)} \sum_{k=0}^{\infty} |L_{k,n}| \xi_k(\nu).$$

We point out that, for any $\mathbf{u} \in \ell_\nu^1$, the operator norm of the multiplication operator $M(\mathbf{u})$ is nothing but the norm of the sequence \mathbf{u} , i.e. $\|M(\mathbf{u})\|_\nu = \|\mathbf{u}\|_\nu$. We also recall that the $\|\cdot\|_\nu$ norm of the coefficients controls the $\|\cdot\|_{C^0}$ norm of the associated function. That is, for any $\nu \geq 1$ and any \mathbf{u} in ℓ_ν^1 ,

$$\|u\|_{C^0} := \sup_{x \in \mathbb{R}} |u(x)| \leq \|\mathbf{u}\|_\nu. \quad (6)$$

We slightly abuse the notation Δ and also use it on a sequence of Fourier coefficients. That is, $\Delta \mathbf{u}$ is the sequence defined by

$$(\Delta \mathbf{u})_n = -(n\pi)^2 u_n \quad n \geq 0.$$

Similarly $\Delta^{-1} \mathbf{u}$ is the sequence defined by

$$(\Delta^{-1} \mathbf{u})_n = \begin{cases} 0 & n = 0, \\ -\frac{1}{(n\pi)^2} u_n & n \geq 1. \end{cases}$$

Remark 2.2. *The Laplacian with Neumann boundary conditions is only invertible if we restrict ourselves to functions having zero mean, or in terms of Fourier coefficients if we only deal with modes $n \geq 1$. This will be enough for our purposes, and justifies the notation Δ^{-1} above.*

Finally, given $N \geq 1$ we introduce the projection onto the N first modes Π_N , where

$$(\Pi_N \mathbf{u})_n = \begin{cases} u_n & 0 \leq n < N, \\ 0 & n \geq N, \end{cases}$$

and the associated subspace $\Pi_N \ell_\nu^1$ of ℓ_ν^1 . In the sequel, we frequently identify a vector $\bar{\mathbf{u}} = (u_0, \dots, u_{N-1})$ of \mathbb{R}^N and its injection $\bar{\mathbf{u}} = (u_0, \dots, u_{N-1}, 0, \dots, 0, \dots)$ in $\Pi_N \ell_\nu^1$.

2.2 $F = 0$ into fixed point problem

Rewriting (5) in Fourier space, our aim is to find \mathbf{u} in ℓ_ν^1 satisfying

$$F_n(\mathbf{u}) := -(n\pi)^2 \Phi_n(\mathbf{u}) + R_n(\mathbf{u}) = 0 \quad \forall n \geq 0,$$

or, in a more compact form

$$F(\mathbf{u}) := \Delta \Phi(\mathbf{u}) + R(\mathbf{u}) = 0. \quad (7)$$

We now assume that we have an approximate zero $\bar{\mathbf{u}} \in \Pi_N \ell_\nu^1$ of F , which in practice can be obtained by numerically solving the finite dimensional problem $\Pi_N F \Pi_N = 0$.

In order to define a suitable approximate inverse A of $DF(\bar{\mathbf{u}})$, we first consider $\bar{\mathbf{w}} \in \Pi_N \ell_\nu^1$ such that $\bar{\mathbf{w}} * \Phi'(\bar{\mathbf{u}}) \approx \mathbf{1}$.

Remark 2.3. *In practice, such $\bar{\mathbf{w}}$ is easily obtained by numerically solving the linear system*

$$\Pi_N M (\Phi'(\bar{\mathbf{u}})) \Pi_N \bar{\mathbf{w}} = \mathbf{1},$$

identifying $\Pi_N M (\Phi'(\bar{\mathbf{u}})) \Pi_N$ with an $N \times N$ matrix, and $\bar{\mathbf{w}}$ and $\mathbf{1}$ with vectors in \mathbb{R}^N .

Next, we consider an $N \times N$ matrix \bar{A} , a numerically computed approximation of $\Pi_N (DF(\bar{\mathbf{u}})^{-1}) \Pi_N$, which we identify with an operator on $\Pi_N \ell_\nu^1$, and define the crucial operator A as follows.

$$A = \bar{A} + (M(\bar{\mathbf{w}})\Delta^{-1} - \Pi_N (M(\bar{\mathbf{w}})\Delta^{-1}) \Pi_N). \quad (8)$$

In order to better understand this operator, one might think of it as an *infinite matrix*, which writes

$$\left(\begin{array}{cccc|cccc} & & & & 0 & 0 & 0 & \dots \\ & & & & \frac{-\bar{w}_{N-1}}{(N\pi)^2} & 0 & 0 & \\ & & & & \vdots & \frac{-\bar{w}_{N-1}}{((N+1)\pi)^2} & 0 & \ddots \\ & & & & \frac{-\bar{w}_2}{(N\pi)^2} & \vdots & \frac{-\bar{w}_{N-1}}{((N+2)\pi)^2} & \ddots \\ & & & & \frac{-\bar{w}_1}{(N\pi)^2} & \frac{-\bar{w}_2}{((N+1)\pi)^2} & \vdots & \ddots \\ \hline 0 & \frac{-\bar{w}_{N-1}}{\pi^2} & \dots & \frac{-\bar{w}_2}{((N-2)\pi)^2} & \frac{-\bar{w}_1}{((N-1)\pi)^2} & \frac{-\bar{w}_0}{(N\pi)^2} & \frac{-\bar{w}_1}{((N+1)\pi)^2} & \frac{-\bar{w}_2}{((N+2)\pi)^2} \\ 0 & 0 & \frac{-\bar{w}_{N-1}}{(2\pi)^2} & \dots & \frac{-\bar{w}_2}{((N-1)\pi)^2} & \frac{-\bar{w}_1}{(N\pi)^2} & \frac{-\bar{w}_0}{((N+1)\pi)^2} & \frac{-\bar{w}_1}{((N+1)\pi)^2} & \ddots \\ 0 & 0 & 0 & \frac{-\bar{w}_{N-1}}{(3\pi)^2} & \dots & \frac{-\bar{w}_2}{(N\pi)^2} & \frac{-\bar{w}_1}{((N+1)\pi)^2} & \frac{-\bar{w}_0}{((N+2)\pi)^2} & \ddots \\ \vdots & & \ddots & \ddots & \ddots & & \ddots & \ddots & \ddots \end{array} \right),$$

or in a more compact form,

$$A = \left(\begin{array}{c|c} \bar{A} & \\ \hline & M(\bar{w})\Delta^{-1} \end{array} \right).$$

Remark 2.4. When there is only linear diffusion, \bar{A} is usually defined by numerically computing the inverse of $\Pi_N DF(\bar{\mathbf{u}})\Pi_N$. In our situation this won't necessarily be good enough. This is related to the fact that the extra-diagonal terms in $DF(\bar{\mathbf{u}})$ cannot be neglected, even outside of the finite block corresponding to the projection on $\Pi_N \ell_\nu^1$. For more details and an explicit computation, see Appendix A.

Instead, in order to get a good enough approximation of $\Pi_N (DF(\bar{\mathbf{u}})^{-1}) \Pi_N$, we numerically compute the inverse of a larger block, and then project it back onto $\Pi_N \ell_\nu^1$, e.g.

$$\bar{A} \approx \Pi_N (\Pi_{2N} DF(\bar{\mathbf{u}})\Pi_{2N})^{-1} \Pi_N,$$

where the \approx sign is only here to emphasize that the inverse does not have to be computed exactly.

Now that the important operator A has been defined, we are ready to consider the fixed-point operator $T : \ell_\nu^1 \rightarrow \ell_\nu^1$ defined by

$$T(u) = u - AF(u), \tag{9}$$

and to give the classical sufficient conditions for T to admit a fixed-point near $\bar{\mathbf{u}}$. This Newton-Kantorovich-like theorem tells us that A has to be a *sufficiently good* approximate inverse of $DF(\bar{\mathbf{u}})$, and that $\bar{\mathbf{u}}$ has to be a *sufficiently good* approximate zero of F , makes it precise what *sufficiently good* means (see (11)), and then gives an explicit error bound for $\bar{\mathbf{u}}$ (see (12)).

Theorem 2.5. *With the notations introduced in this section, assume there exist constants Y , Z_1 and Z_2 satisfying*

$$\|AF(\bar{\mathbf{u}})\|_\nu \leq Y \quad (10a)$$

$$\|I - ADF(\bar{\mathbf{u}})\|_\nu \leq Z_1 \quad (10b)$$

$$\|AD^2F(\mathbf{u})\|_\nu \leq Z_2 \quad \forall \mathbf{u} \in \ell_\nu^1, \quad (10c)$$

and

$$Z_1 < 1 \quad (11a)$$

$$2YZ_2 < (1 - Z_1)^2. \quad (11b)$$

Then, for any r satisfying

$$\frac{1 - Z_1 - \sqrt{(1 - Z_1)^2 - 2YZ_2}}{Z_2} \leq r < \frac{1 - Z_1}{Z_2}, \quad (12)$$

there exists a unique fixed-point \mathbf{u}^* of T in $\mathcal{B}_\nu(\bar{\mathbf{u}}, r)$, the closed ball of center $\bar{\mathbf{u}}$ and radius r in ℓ_ν^1 .

Assume further that $\bar{\mathbf{w}}$ (which plays a role in the definition of A), is such that

$$\|\mathbf{1} - \bar{\mathbf{w}} * \Phi'(\bar{\mathbf{u}})\|_\nu < 1. \quad (13)$$

Then \mathbf{u}^* is the unique zero of F in $\mathcal{B}_\nu(\bar{\mathbf{u}}, r)$.

Proof. The first part of the proof, which consists in showing that T defined in (9) admits a unique fixed point in $\mathcal{B}_\nu(\bar{\mathbf{u}}, r)$ is standard, so we only sketch it. For any \mathbf{u} in $\mathcal{B}_\nu(\bar{\mathbf{u}}, r)$ and r satisfying (12), we have

$$\|T(\mathbf{u}) - \bar{\mathbf{u}}\|_\nu \leq Y + Z_1r + \frac{1}{2}Z_2r^2 \leq r,$$

and

$$\|DT(\mathbf{u})\|_\nu \leq Z_1 + Z_2r < 1,$$

therefore T is a contraction on $\mathcal{B}_\nu(\bar{\mathbf{u}}, r)$, and there is a unique fixed-point \mathbf{u}^* of T in $\mathcal{B}_\nu(\bar{\mathbf{u}}, r)$.

The second part of the proof consists in showing that A is injective, so that this fixed point of T indeed corresponds to a zero of F . From (10b) and (11a), we know that $ADF(\bar{\mathbf{u}})$ is invertible, therefore we have at least that A is surjective. Furthermore, by (13) we have that

$$\|I - M(\bar{\mathbf{w}})M(\Phi'(\bar{\mathbf{u}}))\|_\nu = \|\mathbf{1} - \bar{\mathbf{w}} * \Phi'(\bar{\mathbf{u}})\|_\nu < 1,$$

therefore $M(\bar{\mathbf{w}})$ is invertible ($M(\bar{\mathbf{w}})$ and $M(\Phi'(\bar{\mathbf{u}}))$ commute, because $*$ is associative). Since Δ^{-1} is a Fredholm operator of index 0, so is $M(\bar{\mathbf{w}})\Delta^{-1}$. However, A is just a compact perturbation of $M(\bar{\mathbf{w}})\Delta^{-1}$, therefore it is also Fredholm of index 0, and since we have already shown that A is surjective, it must be injective. \square

Remark 2.6. *With a Fourier discretization and linear diffusion, the second part of the above proof is usually trivial, but ideas similar to the ones used here already appeared in the context of computer-assisted proofs, see e.g. the proof of [24, Theorem 6.2].*

*While this second part of the proof seems to require an extra assumption, namely (13), the term $\|\mathbf{1} - \bar{\mathbf{w}} * \Phi'(\bar{\mathbf{u}})\|_\nu$ is going to naturally appear in our estimate for Z_1 , and therefore (13) will be automatically satisfied as soon as (11a) is, see (14).*

Finally let us mention that (10c) does not really need to hold for all \mathbf{u} in ℓ_ν^1 , but only in a small neighborhood of $\bar{\mathbf{u}}$. In the present case it does not matter, since we took Φ and R as polynomials of degree two and therefore D^2F is constant. In a more general case one may introduce an a priori radius r^ , and restrict the whole analysis to $\mathcal{B}_\nu(\bar{\mathbf{u}}, r^*)$, in order to only have to estimate $D^2F(\mathbf{u})$ locally.*

2.3 Derivation of the bounds

It remains to derive computable estimates Y , Z_1 and Z_2 satisfying (10), and to check that (11) also holds.

One key point to keep in mind in this subsection is that, since $\bar{\mathbf{w}} \in \Pi_N \ell_\nu^1$, the infinite matrix $M(\bar{\mathbf{w}})$ is a banded matrix, with a bandwidth of at most N . That is, as soon as $|k - n| \geq N$, $(M(\bar{\mathbf{w}}))_{k,n} = 0$. Since we assumed $\Phi(u) = u^2$ and $R(u) = \alpha u - \beta u^2 + g$, the same is true for $M(\Phi'(\bar{\mathbf{u}}))$ and $M(R'(\bar{\mathbf{u}}))$.

Remark 2.7. *If R or Φ were polynomials of higher order, $M(\Phi'(\bar{\mathbf{u}}))$ and $M(R'(\bar{\mathbf{u}}))$ would simply have a larger bandwidth (equal to $d(N - 1) + 1$ if the polynomial is of degree d), and all the estimates to come would have to be adapted in a straightforward manner.*

If R (or Φ) were to be merely analytic, one could split $R'(\bar{\mathbf{u}})$ between a finite sequence and a remainder whose $\|\cdot\|_\nu$ norm is small and could be estimated, deal as above with the finite part, and keep track of the extra terms that are produced by the remainder. See also Section 4 for a different approach.

Still for the sake of simplicity, in the sequel we also assume that $\mathbf{g} \in \Pi_N \ell_\nu^1$, but more general g could be handled in a similar way, as soon as the remainder can be estimated explicitly.

2.3.1 The bound Y

Since Φ and R are polynomials, and $\bar{\mathbf{u}}$ and \mathbf{g} belong to $\Pi_N \ell_\nu^1$, $F(\bar{\mathbf{u}})$ only has a finite number of non-zero coefficients. Similarly, since $\bar{\mathbf{w}}$ belongs to $\Pi_N \ell_\nu^1$, each column of A only has a finite number of non-zero coefficients. Therefore $AF(\mathbf{u})$ and then $\|AF(\mathbf{u})\|_\nu$ can be computed exactly, up to rounding errors, and the rounding errors can be controlled using interval arithmetic. In our implementation we make use of the Intlab package [29] for Matlab. The upper bound of this computation will be our bound Y , and satisfies (10a).

2.3.2 The bound Z_1

Denoting $B = I - ADF(\bar{\mathbf{u}})$ and $B_{(\cdot,n)}$ the n -th column of B , we split the estimation of the norm of B as follows.

$$\begin{aligned} \|B\|_\nu &= \sup_{n \geq 0} \frac{1}{\xi_n(\nu)} \|B_{(\cdot,n)}\|_\nu \\ &= \max \left(\max_{0 \leq n \leq 2N-2} \frac{1}{\xi_n(\nu)} \|B_{(\cdot,n)}\|_\nu, \sup_{n \geq 2N-1} \frac{1}{\xi_n(\nu)} \|B_{(\cdot,n)}\|_\nu \right) \\ &\leq \max \left(Z_1^{finite}, Z_1^{tail} \right). \end{aligned}$$

Z_1^{finite} can be computed explicitly. Indeed, the n -th column of B is nothing but the n -th column of $DF(\bar{\mathbf{u}})$ multiplied by A , and similarly to the situation for the Y bounds, the norm of any individual column can be computed exactly (or, to be more precise, enclosed rigorously using interval arithmetic).

Regarding the estimate for Z_1^{tail} , we use the fact that $M(\Phi'(\bar{\mathbf{u}}))$ and $M(R'(\bar{\mathbf{u}}))$ have bandwidth N . Hence for any column of $DF(\bar{\mathbf{u}}) = \Delta M(\Phi'(\bar{\mathbf{u}})) + M(R'(\bar{\mathbf{u}}))$ having index $n \geq 2N - 1$, its N first coefficients must be zero, and the multiplication of such a column by A does not depend on \bar{A} . This enables us to easily estimate by hand $\|B_{(\cdot,n)}\|_\nu$ for any $n \geq 2N - 1$. Equivalently, we can start by rewriting

$$\sup_{n \geq 2N-1} \frac{1}{\xi_n(\nu)} \|B_{(\cdot,n)}\|_\nu = \sup_{\substack{\mathbf{u} \in (I - \Pi_{2N-1}) \ell_\nu^1 \\ \mathbf{u} \neq 0}} \frac{\|B\mathbf{u}\|}{\|\mathbf{u}\|}.$$

Then, since $\Phi'(\bar{\mathbf{u}})$ and $R'(\bar{\mathbf{u}})$ belong to $\Pi_N \ell_\nu^1$, for any $\mathbf{u} \in (I - \Pi_{2N-1}) \ell_\nu^1$ we have that $\Phi'(\bar{\mathbf{u}}) * \mathbf{u}$ and

$R'(\bar{\mathbf{u}}) * \mathbf{u}$ belong to $(I - \Pi_N)\ell_\nu^1$, and so does $DF(\bar{\mathbf{u}})\mathbf{u}$. Therefore, for any $\mathbf{u} \in (I - \Pi_{2N-1})\ell_\nu^1$,

$$\begin{aligned} B\mathbf{u} &= \mathbf{u} - ADF(\bar{\mathbf{u}})\mathbf{u} \\ &= \mathbf{u} - \bar{\mathbf{w}} * \Delta^{-1}DF(\bar{\mathbf{u}})\mathbf{u} \\ &= \mathbf{u} - \bar{\mathbf{w}} * \Phi'(\bar{\mathbf{u}}) * \mathbf{u} - \bar{\mathbf{w}} * \Delta^{-1}(R'(\bar{\mathbf{u}}) * \mathbf{u}), \end{aligned}$$

and

$$\|B\mathbf{u}\|_\nu \leq \|\mathbf{1} - \bar{\mathbf{w}} * \Phi'(\bar{\mathbf{u}})\|_\nu \|\mathbf{u}\|_\nu + \|\bar{\mathbf{w}}\|_\nu \|\Delta^{-1}(R'(\bar{\mathbf{u}}) * \mathbf{u})\|_\nu.$$

Finally, since $R'(\bar{\mathbf{u}}) * \mathbf{u}$ belongs to $(I - \Pi_N)\ell_\nu^1$ we have $\|\Delta^{-1}(R'(\bar{\mathbf{u}}) * \mathbf{u})\|_\nu \leq \frac{1}{(N\pi)^2} \|R'(\bar{\mathbf{u}}) * \mathbf{u}\|_\nu$, and we can take

$$Z_1^{tail} = \|\mathbf{1} - \bar{\mathbf{w}} * \Phi'(\bar{\mathbf{u}})\|_\nu + \frac{1}{(N\pi)^2} \|\bar{\mathbf{w}} * R'(\bar{\mathbf{u}})\|_\nu. \quad (14)$$

We then just take $Z_1 = \max(Z_1^{finite}, Z_1^{tail})$, and it satisfies (10b).

2.3.3 The bound Z_2

For any $\mathbf{u}_1, \mathbf{u}_2$ in ℓ_ν^1 , remembering that $\Phi(u) = u^2$ and $R(u) = \alpha u - \beta u^2 + g$, we have

$$D^2F(\mathbf{u})(\mathbf{u}_1, \mathbf{u}_2) = 2\Delta(\mathbf{u}_1 * \mathbf{u}_2) - 2\beta\mathbf{u}_1 * \mathbf{u}_2,$$

and therefore

$$\begin{aligned} \|AD^2F(\mathbf{u})\|_\nu &= \sup_{\substack{\mathbf{u}_1, \mathbf{u}_2 \in \ell_\nu^1 \\ \mathbf{u}_1, \mathbf{u}_2 \neq 0}} \frac{\|AD^2F(\mathbf{u})(\mathbf{u}_1, \mathbf{u}_2)\|_\nu}{\|\mathbf{u}_1\|_\nu \|\mathbf{u}_2\|_\nu} \\ &\leq 2\|A\Delta\|_\nu + 2|\beta|\|A\|_\nu. \end{aligned} \quad (15)$$

We thus take

$$Z_2 = 2\|A\Delta\|_\nu + 2|\beta|\|A\|_\nu.$$

Similarly to what we have done for the Z_1 bound, the norm of A and $A\Delta$ can be computed by splitting between an finite number of column and the rest, and we get

$$\|A\|_\nu = \max\left(\max_{0 \leq n \leq N-1} \frac{1}{\xi_n(\nu)} \|A_{(\cdot, n)}\|_\nu, \frac{1}{(N\pi)^2} \|\bar{\mathbf{w}}\|_\nu\right) \quad (16a)$$

$$\|A\Delta\|_\nu = \max\left(\max_{0 \leq n \leq N-1} \frac{1}{\xi_n(\nu)} \|(A\Delta)_{(\cdot, n)}\|_\nu, \|\bar{\mathbf{w}}\|_\nu\right). \quad (16b)$$

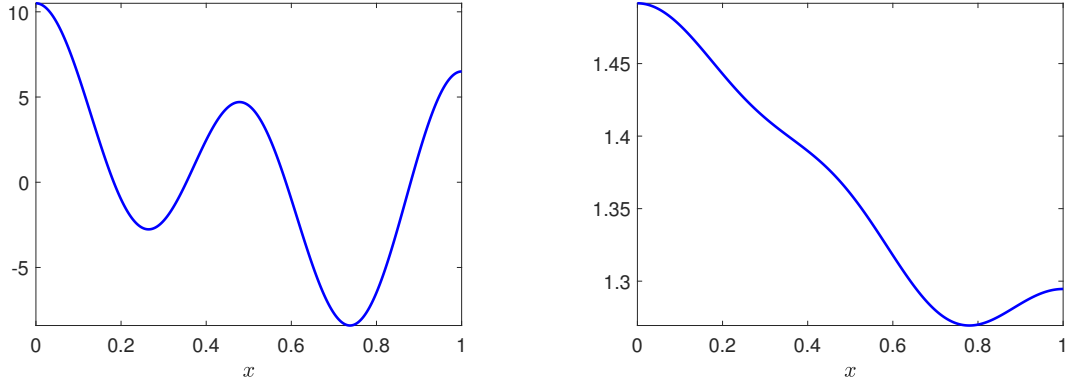
Combining (15) and (16) gives us a computable Z_2 bound satisfying (10c).

2.4 Example and results

In this subsection, we consider an explicit case, namely $\Phi(u) = u^2$ and $R(u) = u - u^2 + g$ (i.e. we take $\alpha = \beta = 1$), and g given by $g(x) = \frac{1}{2} + 3\cos(\pi x) + 2\cos(2\pi x) - \cos(3\pi x) + 6\cos(4\pi x)$ (see Figure 1a), for which we check that Theorem 2.5 is applicable.

We computed an approximate solution $\bar{\mathbf{u}}$ (see Figure 1b), and then validated it using the procedure described in this whole section, which yield the following theorem.

Theorem 2.8. *Let \bar{u} be the function whose Fourier coefficients $\bar{\mathbf{u}}$ can be downloaded at [3], and which is represented in Figure 1b. There exists a strong solution u of (5), with Φ and R as given just above, such that $\|u - \bar{u}\|_{C^0} \leq 2 \times 10^{-10}$.*



(a) The forcing term g used in the examples.

(b) The approximate solution \bar{u} of (5) with $\Phi(u) = u^2$, $R(u) = u - u^2 + g$ and g as in Figure 1a, which has been validated in Theorem 2.8.

Figure 1: Data associated to Theorem 2.8.

Proof. We apply the validation procedure with $N = 20$ and $\nu = 1.1$. That is, we define some $\bar{\mathbf{w}}$ and \bar{A} , then evaluate the bounds Y , Z_1 and Z_2 obtained in Section 2.3, with interval arithmetic to take rounding errors into accounts. We get that assumption (10) is satisfied with

$$Y = 1.3 \times 10^{-10}, \quad Z_1 = 0.0002 \quad \text{and} \quad Z_2 = 2.2.$$

Therefore (11) holds, and so does (12) with $r = 2 \times 10^{-10}$. As mentioned in Remark 2.6, formula (14) show that (13) also holds because (11a) does.

We can then apply Theorem 2.5, and there exists a unique zero \mathbf{u} of F in $\mathcal{B}_\nu(\bar{\mathbf{u}}, r)$. Since $\nu > 1$ the associated function u is smooth, and is therefore a strong solution of (5). The announced error estimate in C^0 norm follows directly from (6).

The computational parts of the proof can be reproduced by running the Matlab code `script_PM.m` from [3], together with Intlab [29]. \square

Remark 2.9. *If needed, one could also get error estimates in different function spaces, since the ℓ_ν^1 norm of \mathbf{u} controls the norm of any derivative of u , as soon as $\nu > 1$ (with constants depending on ν).*

In ℓ_ν^1 , we not only have existence of a zero \mathbf{u} of F , but also local uniqueness. This local uniqueness carries over to the solutions of (5), provided we know a priori that any solution of (5) must be smooth enough for its Fourier coefficients to belong to ℓ_ν^1 . Otherwise, we only have local uniqueness among smooth enough functions.

For problems where analyticity is too much too ask for, be it because the solution itself is not analytic, or because we care about local uniqueness but proving the analyticity a priori is hard, it should be noted that one can easily replicated the whole procedure presented here in a sequence space corresponding to functions of lower regularity, see e.g. [17].

3 The SKT system

We now use the techniques introduced in Section 2 to study steady states of the SKT system

$$\begin{cases} \partial_t u^{(1)} = \Delta \left((d_1 + d_{11}u^{(1)} + d_{12}u^{(2)}) u^{(1)} \right) + (r_1 - a_1u^{(1)} - b_1u^{(2)}) u^{(1)} = 0 & \text{on } \Omega, \\ \partial_t u^{(2)} = \Delta \left((d_2 + d_{21}u^{(1)} + d_{22}u^{(2)}) u^{(2)} \right) + (r_2 - b_2u^{(1)} - a_2u^{(2)}) u^{(2)} = 0 & \text{on } \Omega, \\ \frac{\partial u^{(1)}}{\partial n} = 0 = \frac{\partial u^{(2)}}{\partial n} & \text{on } \partial\Omega, \end{cases} \quad (17)$$

which is exactly of the form (1), with

$$u = \begin{pmatrix} u^{(1)} \\ u^{(2)} \end{pmatrix}, \quad \Phi(u) = \begin{pmatrix} \Phi^{(1)}(u) \\ \Phi^{(2)}(u) \end{pmatrix} = \begin{pmatrix} (d_1 + d_{11}u^{(1)} + d_{12}u^{(2)}) u^{(1)} \\ (d_2 + d_{21}u^{(1)} + d_{22}u^{(2)}) u^{(2)} \end{pmatrix},$$

and

$$R(u) = \begin{pmatrix} R^{(1)}(u) \\ R^{(2)}(u) \end{pmatrix} = \begin{pmatrix} (r_1 + a_1u^{(1)} + b_1u^{(2)}) u^{(1)} \\ (r_2 + b_2u^{(1)} + a_2u^{(2)}) u^{(2)} \end{pmatrix}.$$

3.1 Presentation of the model

The SKT system was introduced in the seminal paper [32] in order to model the dynamics of two competing species, whose densities are denoted here by $u^{(1)}$ and $u^{(2)}$. The reactions terms are standard Lotka-Volterra terms with signs indicating intra-specific and inter-specific competitions. The key feature of this model is the presence of nonlinear diffusion terms: the diffusion rate of each species depends on the local density of both species. These nonlinear diffusion terms can give rise to a repulsive effect which leads to *spatial segregation*: the two species co-exist but mostly concentrate in different regions of the domain Ω . Mathematically, this corresponds to the existence of non-homogeneous steady states of (17) exhibiting this type of patterns.

The existence and stability analysis of such steady states has been studied extensively, through a wide variety of techniques such as bifurcation theory, singular perturbation theory or fixed point index theory, see e.g. [21, 22, 19, 30], which provide a qualitative understanding on the conditions required on the many parameters of (17) for non-homogeneous steady states to exist. Some asymptotic parameter regimes have also been scrutinized, especially when one or both of the cross-diffusion coefficients d_{12} and d_{21} become very large, and in such cases additional information about the shape of the obtained solutions is also available, see e.g. [25, 20]. However, numerical studies like [13, 14] show that the steady states of (17) can be very diverse, already in the one dimensional case, and that many of them can co-exists for a give set of parameter values. Proving the existence, and characterizing the shape, of all these different steady states seems very hard by purely analytic means, but computer-assisted techniques provide a valuable tool to attack such questions.

The computer-assisted study of the steady states of the SKT system started first via a system with linear diffusion approximating (17) in [9], and then for (17) itself in [4], but only in the so-called *triangular case*. This case corresponds to taking the *self-diffusion* coefficients d_{11} and d_{22} equal to 0, and one of the cross-diffusion coefficient (say d_{21}) equal to 0. While having only $d_{12}uv$ as a nonlinear diffusion term is already sufficient to produce many interesting steady states, the mathematical analysis is somewhat simpler in that case. In particular, the work [4] takes advantage of the fact that, in the triangular case, Φ^{-1} can be computed explicitly, and therefore one can directly transform (17) into a system with only linear diffusion. However, this feature is specific to the triangular case, and the general case where $d_1, d_2, d_{12}, d_{21}, d_{11}$ and d_{22} all are possibly non-zero could not be handled in [4].

With the techniques introduced in this paper, we get a new approach which is not only more general (i.e. not restricted to the triangular case), but also more efficient compared to the ad hoc technique used in [4] for the triangular case (see Remark 3.2). Building upon the work done in Section 2, we introduce

the necessary supplementary notations and spaces in Section 3.2, the fixed-point setup in Section 3.3, derive the estimates needed for the validation in Section 3.4, and present some examples in Section 3.5, which include the a posteriori validation of some interesting solutions first found numerically in [8].

3.2 Notations

We consider the product space $\mathcal{X}_\nu := \ell_\nu^1 \times \ell_\nu^1$, and slightly abuse all the notations introduced in Section 2.1 by applying them component-wise. That is,

$$\text{for any } \mathbf{u} = \begin{pmatrix} \mathbf{u}^{(1)} \\ \mathbf{u}^{(2)} \end{pmatrix}, \quad \Pi_N \mathbf{u} := \begin{pmatrix} \Pi_N \mathbf{u}^{(1)} \\ \Pi_N \mathbf{u}^{(2)} \end{pmatrix}, \quad \|\mathbf{u}\|_\nu := \begin{pmatrix} \|\mathbf{u}^{(1)}\|_\nu \\ \|\mathbf{u}^{(2)}\|_\nu \end{pmatrix}, \quad M(\mathbf{u}) := \begin{pmatrix} M(\mathbf{u}^{(1)}) \\ M(\mathbf{u}^{(2)}) \end{pmatrix}.$$

The norm on \mathcal{X}_ν is then defined as

$$\|\mathbf{u}\|_{\mathcal{X}_\nu} = \|\|\mathbf{u}\|_\nu\|_1 = \|\mathbf{u}^{(1)}\|_\nu + \|\mathbf{u}^{(2)}\|_\nu,$$

where $|\cdot|_1$ denotes the 1-norm on \mathbb{R}^2 .

Given a linear operator L on \mathcal{X}_ν , written

$$L = \left(\begin{array}{c|c} L^{(1,1)} & L^{(1,2)} \\ \hline L^{(2,1)} & L^{(2,2)} \end{array} \right),$$

where each $L^{(i,j)}$ is an operator on ℓ_ν^1 , we also abuse $\|\cdot\|_\nu$ notation in a similar way, that is

$$\|L\|_\nu = \begin{pmatrix} \|L^{(1,1)}\|_\nu & \|L^{(1,2)}\|_\nu \\ \|L^{(2,1)}\|_\nu & \|L^{(2,2)}\|_\nu \end{pmatrix}.$$

For the operator norm of L , it is then straightforward to show that

$$\begin{aligned} \|L\|_{\mathcal{X}_\nu} &\leq \max \left(\|L^{(1,1)}\|_\nu + \|L^{(2,1)}\|_\nu, \|L^{(1,2)}\|_\nu + \|L^{(2,2)}\|_\nu \right) \\ &= \|\|L\|_\nu\|_1. \end{aligned}$$

3.3 $F = 0$ into fixed point problem

We look for a zero $\mathbf{u} = (\mathbf{u}^{(1)}, \mathbf{u}^{(2)})$ in \mathcal{X}_ν of

$$F = \begin{pmatrix} F^{(1)} \\ F^{(2)} \end{pmatrix}$$

defined as

$$\begin{cases} F_n^{(1)}(\mathbf{u}) = -(N\pi)^2 \Phi_n^{(1)}(\mathbf{u}) + R_n^{(1)}(\mathbf{u}) \\ F_n^{(2)}(\mathbf{u}) = -(N\pi)^2 \Phi_n^{(2)}(\mathbf{u}) + R_n^{(2)}(\mathbf{u}) \end{cases} \quad \forall n \geq 0,$$

or, in a more condensed form,

$$F(\mathbf{u}) = \Delta \Phi(\mathbf{u}) + R(\mathbf{u}),$$

with Φ and R as in Section 3.1.

We now assume that we have computed an approximate zero $\bar{\mathbf{u}}$ in $\Pi_N \mathcal{X}_\nu$ that we want to validate a posteriori. We also assume that we have computed

$$\bar{\mathbf{w}} = \begin{pmatrix} \bar{\mathbf{w}}^{(1,1)} & \bar{\mathbf{w}}^{(1,2)} \\ \bar{\mathbf{w}}^{(2,1)} & \bar{\mathbf{w}}^{(2,2)} \end{pmatrix} \in \mathcal{M}_{2,2}(\Pi_N \ell_\nu^1), \quad \text{such that} \quad \bar{\mathbf{w}} * D\Phi(\bar{\mathbf{u}}) \approx \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{pmatrix},$$

where

$$D\Phi(\bar{\mathbf{u}}) = \begin{pmatrix} d_1 + 2d_{11}\bar{\mathbf{u}}^{(1)} + d_{12}\bar{\mathbf{u}}^{(2)} & d_{12}\bar{\mathbf{u}}^{(1)} \\ d_{21}\bar{\mathbf{u}}^{(2)} & d_2 + d_{21}\bar{\mathbf{u}}^{(1)} + 2d_{22}\bar{\mathbf{u}}^{(2)} \end{pmatrix},$$

and $\bar{\mathbf{w}} * D\Phi(\bar{\mathbf{u}})$ must be understood as

$$\begin{pmatrix} \bar{\mathbf{w}}^{(1,1)} * (d_1 + 2d_{11}\bar{\mathbf{u}}^{(1)} + d_{12}\bar{\mathbf{u}}^{(2)}) + \bar{\mathbf{w}}^{(1,2)} * d_{21}\bar{\mathbf{u}}^{(2)} & \bar{\mathbf{w}}^{(2,1)} * (d_1 + 2d_{11}\bar{\mathbf{u}}^{(1)} + d_{12}\bar{\mathbf{u}}^{(2)}) + \bar{\mathbf{w}}^{(2,2)} * d_{21}\bar{\mathbf{u}}^{(2)} \\ \bar{\mathbf{w}}^{(1,1)} * d_{12}\bar{\mathbf{u}}^{(1)} + \bar{\mathbf{w}}^{(1,2)} * (d_2 + d_{21}\bar{\mathbf{u}}^{(1)} + 2d_{22}\bar{\mathbf{u}}^{(2)}) & \bar{\mathbf{w}}^{(2,1)} * d_{12}\bar{\mathbf{u}}^{(1)} + \bar{\mathbf{w}}^{(2,2)} * (d_2 + d_{21}\bar{\mathbf{u}}^{(1)} + 2d_{22}\bar{\mathbf{u}}^{(2)}) \end{pmatrix}.$$

Next, we consider \bar{A} , a numerically computed approximation of $\Pi_N (DF(\bar{\mathbf{u}})^{-1}) \Pi_N$ (Remark 2.4 is also relevant here). Identifying \bar{A} with a $2N \times 2N$ matrix, we separate it into four $N \times N$ blocks

$$\bar{A} = \left(\begin{array}{c|c} \bar{A}^{(1,1)} & \bar{A}^{(1,2)} \\ \hline \bar{A}^{(2,1)} & \bar{A}^{(2,2)} \end{array} \right).$$

The operator A is then defined as

$$A = \left(\begin{array}{c|c} \boxed{\bar{A}^{(1,1)}} & \boxed{\bar{A}^{(1,2)}} \\ \hline M(\bar{\mathbf{w}}^{(1,1)}) \Delta^{-1} & M(\bar{\mathbf{w}}^{(1,2)}) \Delta^{-1} \\ \hline \boxed{\bar{A}^{(2,1)}} & \boxed{\bar{A}^{(2,2)}} \\ \hline M(\bar{\mathbf{w}}^{(2,1)}) \Delta^{-1} & M(\bar{\mathbf{w}}^{(2,2)}) \Delta^{-1} \end{array} \right),$$

or, in a more compact form,

$$A = \bar{A} + (M(\bar{\mathbf{w}})\Delta^{-1} - \Pi_N (M(\bar{\mathbf{w}})\Delta^{-1}) \Pi_N). \quad (18)$$

Now that A is defined, we can again state sufficient conditions for the validation, which are the same as the one in Theorem 2.5, up to the slightly different space and norm.

Theorem 3.1. *With the notations introduced in this section, assume there exist constants Y , Z_1 and Z_2 satisfying*

$$\begin{aligned} \|AF(\bar{\mathbf{u}})\|_{\mathcal{X}_\nu} &\leq Y \\ \|I - ADF(\bar{\mathbf{u}})\|_{\mathcal{X}_\nu} &\leq Z_1 \\ \|AD^2F(\mathbf{u})\|_{\mathcal{X}_\nu} &\leq Z_2 \quad \forall \mathbf{u} \in \mathcal{X}_\nu, \end{aligned}$$

and

$$\begin{aligned} Z_1 &< 1 \\ 2YZ_2 &< (1 - Z_1)^2. \end{aligned}$$

Then, for any r satisfying

$$\frac{1 - Z_1 - \sqrt{(1 - Z_1)^2 - 2YZ_2}}{Z_2} \leq r < \frac{1 - Z_1}{Z_2},$$

there exists a unique fixed-point \mathbf{u}^* of T in $\mathcal{B}_{\mathcal{X}_\nu}(\bar{\mathbf{u}}, r)$, the closed ball of center $\bar{\mathbf{u}}$ and radius r in \mathcal{X}_ν . Assume further that $\bar{\mathbf{w}}$ (which plays a role in the definition of A), is such that

$$\left\| \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{pmatrix} - \bar{\mathbf{w}} * D\Phi(\bar{\mathbf{u}}) \right\|_{\mathcal{X}_\nu} < 1.$$

Then \mathbf{u}^* is the unique zero of F in $\mathcal{B}_{\mathcal{X}_\nu}(\bar{\mathbf{u}}, r)$.

3.4 Derivation of the bounds

It remains to derive computable estimates Y , Z_1 and Z_2 satisfying the assumptions of Theorem 3.1. Since all the estimations are very similar to the one performed in Section 2.3, we omit the details.

3.4.1 The bound Y

As in Section 2.3.1, there is no pen and paper estimation to make here, $\|AF(\bar{\mathbf{u}})\|_{\mathcal{X}_\nu}$ can be computed explicitly using interval arithmetic.

3.4.2 The bound Z_1

As in Section 2.3.2, we introduce

$$B = \begin{pmatrix} B^{(1,1)} & B^{(1,2)} \\ B^{(2,1)} & B^{(2,2)} \end{pmatrix} = I - ADF(\bar{\mathbf{u}}),$$

and split the norm between a finite number of columns and an estimate for the tail:

$$Z_1 = \max \left(Z_1^{finite}, Z_1^{tail} \right).$$

Z_1^{finite} , which corresponds to

$$\max_{0 \leq n \leq 2N-2} \frac{1}{\xi_n(\nu)} \max \left(\|B_{(\cdot, n)}^{(1,1)}\|_\nu + \|B_{(\cdot, n)}^{(2,1)}\|_\nu, \|B_{(\cdot, n)}^{(1,2)}\|_\nu + \|B_{(\cdot, n)}^{(2,2)}\|_\nu \right),$$

can again be computed explicitly using interval arithmetic. A computation similar to the one performed in Section 2.3.2 shows that the second part, namely

$$\sup_{n \geq 2N-1} \frac{1}{\xi_n(\nu)} \max \left(\|B_{(\cdot, n)}^{(1,1)}\|_\nu + \|B_{(\cdot, n)}^{(2,1)}\|_\nu, \|B_{(\cdot, n)}^{(1,2)}\|_\nu + \|B_{(\cdot, n)}^{(2,2)}\|_\nu \right),$$

can be bounded by

$$Z_1^{tail} = \left\| \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{pmatrix} - \bar{\mathbf{w}} * \Phi'(\bar{\mathbf{u}}) \right\|_\nu + \frac{1}{(N\pi)^2} \|\bar{\mathbf{w}}\|_\nu \|DR(\bar{\mathbf{u}})\|_\nu \Big|_1. \quad (19)$$

Remark 3.2. *In the previous work [4] on computer-assisted proofs for the (triangular) SKT system, the estimate equivalent to Z_1^{tail} scaled only as $\frac{1}{N}$, rather than as $\frac{1}{N^2}$ here (in practice the first term in (19) is small enough that only the second part really matters). Therefore, on top of not being restricted to the triangular case, our approach has the advantage of recovering the $\frac{1}{N^2}$ scaling expected for a second order equation with no first order term. This is comes from the fact that, contrarily to what was done in [4], we do not have to resort to automatic differentiation techniques. This point will be discussed in more details in Section 4.*

3.4.3 The bound Z_2

Again, we proceed as in Section 2.3.3 and estimate, for \mathbf{u} and \mathbf{v} in \mathcal{X}_ν ,

$$\begin{aligned} \|AD^2F(\bar{\mathbf{u}})(\mathbf{u}, \mathbf{v})\|_{\mathcal{X}_\nu} &\leq \left(\|A^{(1,1)}\Delta\|_\nu + \|A^{(2,1)}\Delta\|_\nu \right) \left(2d_{11}\|\mathbf{u}^{(1)}\|_\nu\|\mathbf{v}^{(1)}\|_\nu + d_{12}(\|\mathbf{u}^{(1)}\|_\nu\|\mathbf{v}^{(2)}\|_\nu + \|\mathbf{v}^{(1)}\|_\nu\|\mathbf{u}^{(2)}\|_\nu) \right) \\ &\quad + \left(\|A^{(1,1)}\|_\nu + \|A^{(2,1)}\|_\nu \right) \left(2a_1\|\mathbf{u}^{(1)}\|_\nu\|\mathbf{v}^{(1)}\|_\nu + b_1(\|\mathbf{u}^{(1)}\|_\nu\|\mathbf{v}^{(2)}\|_\nu + \|\mathbf{v}^{(1)}\|_\nu\|\mathbf{u}^{(2)}\|_\nu) \right) \\ &\quad + \left(\|A^{(1,2)}\Delta\|_\nu + \|A^{(2,2)}\Delta\|_\nu \right) \left(2d_{22}\|\mathbf{u}^{(2)}\|_\nu\|\mathbf{v}^{(2)}\|_\nu + d_{21}(\|\mathbf{u}^{(1)}\|_\nu\|\mathbf{v}^{(2)}\|_\nu + \|\mathbf{v}^{(1)}\|_\nu\|\mathbf{u}^{(2)}\|_\nu) \right) \\ &\quad + \left(\|A^{(1,2)}\|_\nu + \|A^{(2,2)}\|_\nu \right) \left(2a_2\|\mathbf{u}^{(2)}\|_\nu\|\mathbf{v}^{(2)}\|_\nu + b_2(\|\mathbf{u}^{(1)}\|_\nu\|\mathbf{v}^{(2)}\|_\nu + \|\mathbf{v}^{(1)}\|_\nu\|\mathbf{u}^{(2)}\|_\nu) \right) \\ &\leq \max \left[2d_{11} \left(\|A^{(1,1)}\Delta\|_\nu + \|A^{(2,1)}\Delta\|_\nu \right) + 2a_1 \left(\|A^{(1,1)}\|_\nu + \|A^{(2,1)}\|_\nu \right), \right. \\ &\quad \left. 2d_{22} \left(\|A^{(1,2)}\Delta\|_\nu + \|A^{(2,2)}\Delta\|_\nu \right) + 2a_2 \left(\|A^{(1,2)}\|_\nu + \|A^{(2,2)}\|_\nu \right), \right. \\ &\quad \left. d_{12} \left(\|A^{(1,1)}\Delta\|_\nu + \|A^{(2,1)}\Delta\|_\nu \right) + b_1 \left(\|A^{(1,1)}\|_\nu + \|A^{(2,1)}\|_\nu \right) \right. \\ &\quad \left. + d_{21} \left(\|A^{(1,2)}\Delta\|_\nu + \|A^{(2,2)}\Delta\|_\nu \right) + b_2 \left(\|A^{(1,2)}\|_\nu + \|A^{(2,2)}\|_\nu \right) \right] \|\mathbf{u}\|_{\mathcal{X}_\nu} \|\mathbf{v}\|_{\mathcal{X}_\nu}, \end{aligned}$$

which yields

$$\begin{aligned} Z_2 = \max &\left[2d_{11} \left(\|A^{(1,1)}\Delta\|_\nu + \|A^{(2,1)}\Delta\|_\nu \right) + 2a_1 \left(\|A^{(1,1)}\|_\nu + \|A^{(2,1)}\|_\nu \right), \right. \\ &2d_{22} \left(\|A^{(1,2)}\Delta\|_\nu + \|A^{(2,2)}\Delta\|_\nu \right) + 2a_2 \left(\|A^{(1,2)}\|_\nu + \|A^{(2,2)}\|_\nu \right), \\ &d_{12} \left(\|A^{(1,1)}\Delta\|_\nu + \|A^{(2,1)}\Delta\|_\nu \right) + b_1 \left(\|A^{(1,1)}\|_\nu + \|A^{(2,1)}\|_\nu \right) \\ &\left. + d_{21} \left(\|A^{(1,2)}\Delta\|_\nu + \|A^{(2,2)}\Delta\|_\nu \right) + b_2 \left(\|A^{(1,2)}\|_\nu + \|A^{(2,2)}\|_\nu \right) \right]. \end{aligned}$$

3.5 Examples and results

We now consider several specific parameter values, and use the presented methodology to obtain qualitative existence theorems about the steady states of the SKT system (17) for those parameter values, with $\Omega = (0, 1)$. The different parameter sets that we study here are given in Table 1, and their respective relevance is explained in the following subsections.

In each case, we only show one or a couple of steady states, but we emphasize that there might exist many more steady states for the exact same parameter values. We refer to [13, 14, 8] for a broader picture and many bifurcation diagrams.

Finally, before getting to each case let us also mention that, even if we only study the existence and the precise description of (some of) the steady states here, their stability is also a very important question. While adapting the computer-assisted techniques presented in this work to prove that a steady state is unstable is relatively straightforward (see e.g. [4] for the triangular case), a computer-assisted approach that could be used to prove (linear) stability for an arbitrary solution of the SKT system is still lacking.

d_1	d_2	d_{12}	d_{21}	d_{11}	d_{22}	r_1	r_2	a_1	a_2	b_1	b_2	
0.005	0.005	3	0	0	0	5	2	3	3	1	1	Section 3.5.1
0.005	0.005	100	100	0	0	15/2	16/7	4	2	6	1	Section 3.5.2
0.05	0.05	3	0	0	0	15	5	1	1	0.5	3	Section 3.5.3
-0.007	-0.007	3	0.002	0.05	0.05	5	2	3	3	1	1	Section 3.5.4

Table 1: The parameter sets of the SKT system (17) considered in this work.

3.5.1 Comparison with the previous setup using automatic differentiation [4]

We first focus on the first parameter set (i.e. the first row) in Table 1. For these parameter values, a complex bifurcation diagram of steady states was first computed in [13], and then mostly validated in [4] using an approach relying on the fact that $d_{11} = d_{22} = d_{21} = 0$. We therefore use this first case as a benchmark, and compare the results of our method with the ones of [4].

Theorem 3.3. *Let \bar{u} be the function represented on Figure 2, and whose precise description in terms of Fourier coefficients can be found at [3]. There exists a smooth steady state u of the SKT system (17), with parameters as in the first row of Table 1 and $\Omega = (0, 1)$, such that $\|u - \bar{u}\|_{C^0} \leq 4 \times 10^{-7}$.*

Proof. We take $N = 50$ and $\nu = 1.01$, compute the estimates Y , Z_1 and Z_2 obtained in Section 3.4, the approximate values for the bounds being

$$Y = 2.7 \times 10^{-7}, \quad Z_1 = 0.25, \quad Z_2 = 14000,$$

and apply Theorem 3.1. The computational parts of the proof can be reproduced by running the Matlab code `script_SKT.m` from [3], together with Intlab [29]. \square

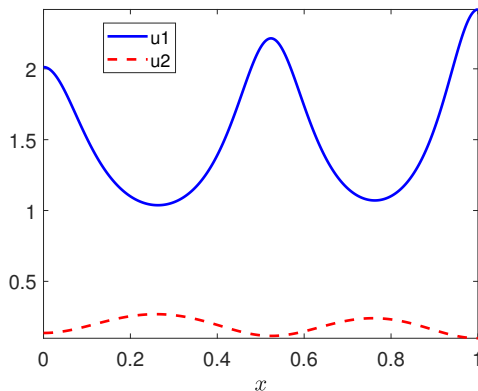


Figure 2: The approximate steady state \bar{u} of the SKT system (17) on $\Omega = (0, 1)$ and with parameters as in the first row of Table 1, which has been validated in Theorem 3.3.

Remark 2.9 also applies here (and for the remainder of Section 3.5). The main point that we want to emphasize with this theorem is not the result itself, which was already obtained in [4], but the fact that our new approach is significantly more efficient. Indeed, with the same parameters ($N = 50$ and $\nu = 1.01$) the estimate corresponding to Z_1^{tail} in [4] would be approximately equal to $1.46 \dots$ and therefore assumption (11a) would not hold. Comparatively, with our approach we get $Z_1^{tail} \approx 0.25$. This means the validation is less costly with our approach, because it works with N smaller, and this comes from the

fact that we avoid the usage of automatic differentiation techniques in (see Remark 3.2). Moreover, with $N = 50$, Z_1 is still significantly away from 1 in our approach, and the main reason why we need close to $N = 50$ modes to validate this solution is only because otherwise Y is not small enough. This is exactly what one would hope for, i.e. that the validation is successful as soon as we have enough modes for the approximate solution to be sufficiently accurate.

3.5.2 A first non-triangular case

We now focus on the second parameter set in Table 1. Notice that we are no longer in the so-called triangular case (both d_{12} and d_{21} are non zero), therefore the previous approach from [4] can no longer be used. Moreover, the solutions which are proved to exist for these parameters are of particular interest, because they answer positively a question which was open until recently, about the existence of non homogeneous steady states in the weak competition case ($b_1 b_2 < a_1 a_2$) when both cross diffusion coefficients are large, see [8] and the reference therein for more details. In [8] this question was answered positively by a bifurcation analysis which provides information locally (close to the bifurcation point), and solutions far away from the bifurcation were computed numerically. It is such solutions far away from the bifurcation whose existence we are able to prove here with computer-assistance.

Theorem 3.4. *For each of the functions represented on Figure 3, and whose precise description in terms of Fourier coefficients can be found at [3], there exists a smooth steady state of the SKT system (17), with parameters as in the second row of Table 1 and $\Omega = (0, 1)$, at a distance of at most 8×10^{-12} in C^0 norm.*

Proof. We take $N = 50$ and $\nu = 1.01$, compute the estimates Y , Z_1 and Z_2 obtained in Section 3.4 for each approximate solution, and apply Theorem 3.1. The computational parts of the proof can be reproduced by running the Matlab code `script_SKT.m` from [3], together with Intlab [29], which provides more details such as the precise value of each bound for each solution. \square

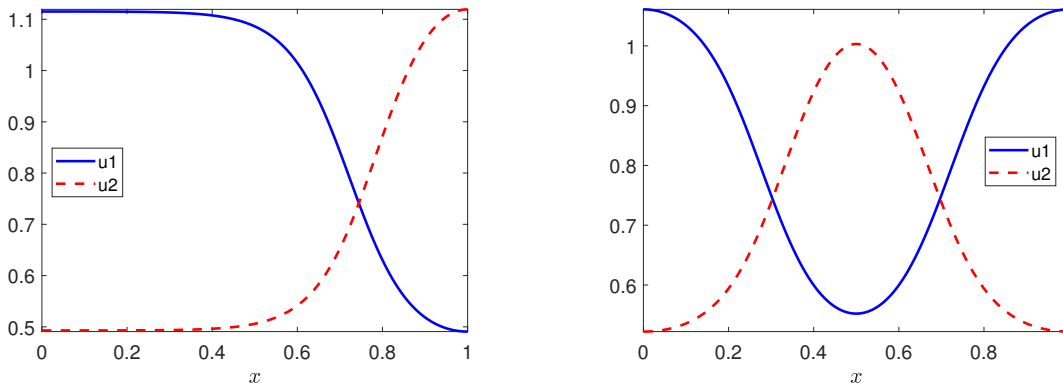


Figure 3: The two approximate steady states of the SKT system (17) on $\Omega = (0, 1)$ and with parameters as in the second row of Table 1, which have been validated in Theorem 3.4.

3.5.3 A new parameter regime

We now focus on the third parameter set in Table 1. In order to understand its interest, it is helpful to briefly consider the homogeneous case, i.e. (17) with only the reaction terms, for which one can easily identify three different parameter regimes with different dynamics, see e.g. [19].

- **Case 1 (weak competition):**

$$\frac{b_1}{a_2} < \frac{r_1}{r_2} < \frac{a_1}{b_2}.$$

Here, any solution with positive initial data converges to a co-existence steady state

$$\left(\frac{r_1 a_2 - r_2 b_1}{a_2 a_1 - b_2 b_1}, \frac{r_2 a_1 - r_1 b_2}{a_2 a_1 - b_2 b_1} \right). \quad (20)$$

- **Case 2 (strong competition):**

$$\frac{a_1}{b_2} < \frac{r_1}{r_2} < \frac{b_1}{a_2}.$$

Here, the above co-existence steady state (20) is still positive, but it is unstable, and generically a solution with positive initial data converges to one of the two extinction states

$$\left(\frac{r_1}{a_1}, 0 \right) \quad \text{or} \quad \left(0, \frac{r_2}{a_2} \right), \quad (21)$$

depending on the initial condition.

- **Case 3:** everything else (except the case $b_1/a_2 = r_1/r_2 = a_1/b_2$ which is degenerate). Here, any solution with positive initial data also converges to one of the two extinction states (21), but all solutions converge to the same one (the one being selected depends on the parameter values, not on the initial data).

The aim of introducing a spatial dimension to model, and more precisely nonlinear diffusion as in (17), is to obtain more interesting and realistic solutions, and in particular non-homogeneous positive steady states. However, it is striking that in the literature on the SKT system, the existence of non-homogeneous steady states is usually studied in cases 1 and 2 only. This might be related to the fact that the analysis in case 3 can be more difficult, because there is no positive steady state that can be used as a kind of starting point: the co-existence steady states (20) still exists (except if $a_1 a_2 = b_1 b_2$), but it is no longer positive in case 3.

However, recent numerical experiments suggested that positive non-homogeneous do also exist in case 3 [8]. The third parameter set in Table 1 corresponds to such a case, and we now rigorously establish the existence of (some of) the solutions that were found numerically in [8].

Theorem 3.5. *For each of the functions represented on Figure 4, and whose precise description in terms of Fourier coefficients can be found at [3], there exists a smooth steady state of the SKT system (17), with parameters as in the third row of Table 1 and $\Omega = (0, 1)$, at a distance of at most 9×10^{-11} in C^0 norm.*

Proof. We take $N = 500$ and $\nu = 1.005$, compute the estimates Y , Z_1 and Z_2 obtained in Section 3.4 for each approximate solution, and apply Theorem 3.1. The computational parts of the proof, can be reproduced by running the Matlab code `script_SKT.m` from [3], together with Intlab [29], which provides more details such as the precise value of each bound for each solution. \square

3.5.4 Full case

We finally focus on the fourth and last parameter set in Table 1, which allows us to showcase that our approach is still successful when all coefficients related to the diffusion term are non zero. As a bonus, we show that positive self-diffusion coefficients d_{11} and d_{22} allow to take negative values of the linear diffusion coefficients d_1 and d_2 and still retain positive and smooth solutions. This was also already noticed numerically in [8].

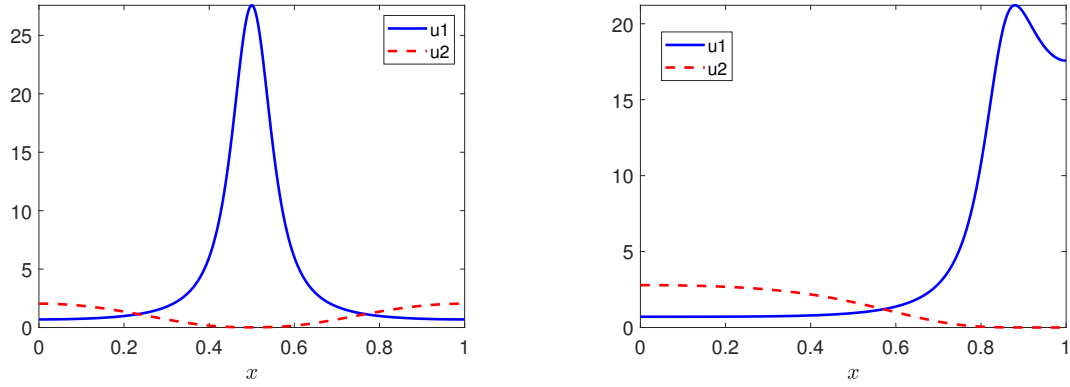


Figure 4: The two approximate steady states of the SKT system (17) on $\Omega = (0, 1)$ and with parameters as in the third row of Table 1, which have been validated in Theorem 3.5.

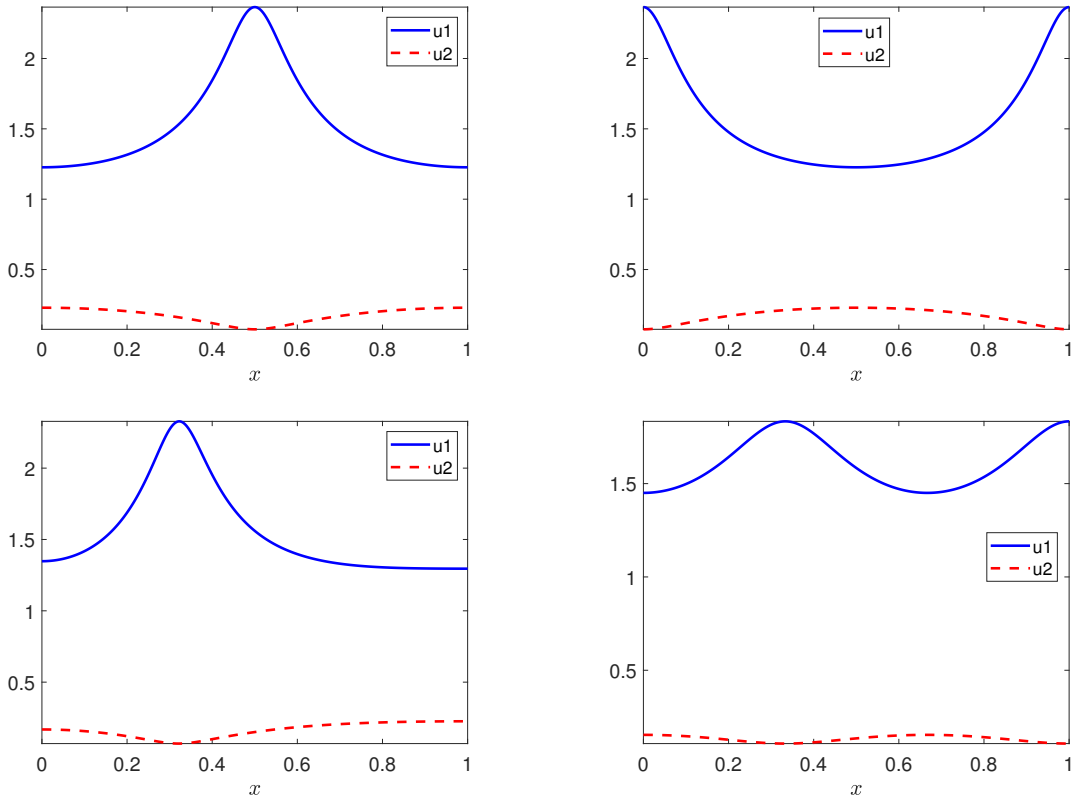


Figure 5: The four approximate steady states of the SKT system (17) on $\Omega = (0, 1)$ and with parameters as in the fourth row of Table 1, which have been validated in Theorem 3.6.

Theorem 3.6. *For each of the functions represented on Figure 5, and whose precise description in terms of Fourier coefficients can be found at [3], there exists a smooth steady state of the SKT system (17), with parameters as in the fourth row of Table 1 and $\Omega = (0, 1)$, at a distance of at most 1×10^{-9} in C^0 norm.*

Proof. We take $N = 100$ and $\nu = 1.01$, compute the estimates Y , Z_1 and Z_2 obtained in Section 3.4 for each approximate solution, and apply Theorem 3.1. The computational parts of the proof, can be reproduced by running the Matlab code `script_SKT.m` from [3], together with Intlab [29], which provides more details such as the precise value of each bound for each solution. \square

4 Beyond polynomial diffusion terms

In this section, we go back to a scalar problem on a one-dimensional domain of the form

$$\begin{cases} \Delta\Phi(u) + R(u) = 0 & \text{on } (0, 1), \\ \frac{\partial u}{\partial n} = 0 & \text{on } \{0, 1\}, \end{cases} \quad (22)$$

with $R(u) = \alpha u - \beta u^2 + g$, but this time with a non-polynomial term Φ of the form $\Phi(u) = \frac{u}{\gamma+u}$. Again, rather than aiming for the most general result possible, we chose to consider an explicit scalar example to simplify the presentation, but the ideas presented in this section can also be applied with different non-polynomial terms, systems, and higher space dimension.

If the solution u under consideration satisfies $\|\mathbf{u}\|_\nu < \gamma$, then $\Phi(\mathbf{u}) = \sum_{k \geq 0} (-1)^k \frac{\mathbf{u}^{k+1}}{\gamma^{k+1}}$ is still well defined in ℓ_ν^1 , and the procedure described in Section 2 could be adapted, at the expense of more pen and paper estimates. However, as we are going to see, there are solutions for which $\|\mathbf{u}\|_\nu > \gamma$.

Even in the case where $\|\mathbf{u}\|_\nu < \gamma$, it might be more convenient to try to get back to a polynomial system, by introducing new variables. Following the ideas of [18], one could consider the additional variable $v = \frac{u}{\gamma+u}$, rewrite (22) as a first order system, and append to it the differential equation satisfied by v . While this approach can be appealing, because the resulting system is again polynomial and therefore fits immediately into the existing framework, it also has a couple of shortcomings. First, having to go back to a system of first order equations is a downside, because all the $\frac{1}{N^2}$ terms in the estimates (like in (14) and (19)) are now replaced by $\frac{1}{N}$ terms, and therefore the dimension N of the projection has to be taken much larger for these terms to be small enough. Second, it does not seem straightforward to generalize these ideas as soon as the space dimension or the number of unknown is greater than one.

In order to overcome these limitations, we propose a new approach, which is based on the following observation: Introducing new variables is a good idea, but in the context of computer-assisted proofs there is nothing wrong with differential-algebraic equations (especially when the algebraic equations can be rewritten as polynomial ones), and in particular there is no need to differentiate the algebraic equations. That is, we do also introduce the additional variable $v = \frac{u}{\gamma+u}$, but we then directly work on the system

$$\begin{cases} \Delta v + R(u) = 0 & \text{on } (0, 1), \\ u - (\gamma + u)v = 0 & \text{on } (0, 1), \\ \frac{\partial u}{\partial n} = 0 = \frac{\partial v}{\partial n} & \text{on } \{0, 1\}. \end{cases} \quad (23)$$

Indeed, we show in this section that this system of differential-algebraic equations is amenable to a posteriori validation techniques very similar to the ones presented up to now in this paper.

4.1 $F = 0$ into fixed-point problem

Relabeling the unknowns (u, v) in (23) as $(u^{(1)}, u^{(2)})$, we can directly use all the notations, spaces and norms introduced in Section 3. That is, we look for a zero $\mathbf{u} = (\mathbf{u}^{(1)}, \mathbf{u}^{(2)})$ in \mathcal{X}_ν of

$$F = \begin{pmatrix} F^{(1)} \\ F^{(2)} \end{pmatrix}$$

defined as

$$\begin{cases} F_n^{(1)}(\mathbf{u}) = \mathbf{u}_n^{(1)} - \gamma \mathbf{u}_n^{(2)} - (\mathbf{u}^{(1)} * \mathbf{u}^{(2)})_n \\ F_n^{(2)}(\mathbf{u}) = -(n\pi)^2 \mathbf{u}_n^{(2)} + R_n(\mathbf{u}^{(1)}) \end{cases} \quad \forall n \geq 0,$$

or, in a more condensed form,

$$F(\mathbf{u}) = \begin{pmatrix} \mathbf{u}^{(1)} - \gamma \mathbf{u}^{(2)} - \mathbf{u}^{(1)} * \mathbf{u}^{(2)} \\ \Delta \mathbf{u}^{(2)} + R(\mathbf{u}^{(1)}) \end{pmatrix},$$

with $R(u) = \alpha u - \beta u^2 + g$.

We now assume that we have computed an approximate zero $\bar{\mathbf{u}}$ in $\Pi_N \mathcal{X}_\nu$ that we want to validate a posteriori. We also assume that we have computed $\bar{\mathbf{w}} \in \Pi_N \ell_\nu^1$ such that $\bar{\mathbf{w}} * (\mathbf{1} - \bar{\mathbf{u}}^{(2)}) \approx \mathbf{1}$, and $\bar{\boldsymbol{\sigma}} \in \Pi_N \ell_\nu^1$ such that $\bar{\boldsymbol{\sigma}} \approx \bar{\mathbf{w}} * (\gamma + \bar{\mathbf{u}}^{(1)})$. Next, we consider

$$\bar{A} = \begin{pmatrix} \bar{A}^{(1,1)} & \bar{A}^{(1,2)} \\ \bar{A}^{(2,1)} & \bar{A}^{(2,2)} \end{pmatrix},$$

a numerically computed approximation of $\Pi_N (DF(\bar{\mathbf{u}})^{-1}) \Pi_N$ (Remark 2.4 is also relevant here). The approximate inverse A for this problem is then defined as

$$A = \begin{pmatrix} \boxed{\bar{A}^{(1,1)}} & \boxed{\bar{A}^{(1,2)}} & & \\ & & M(\bar{\mathbf{w}}) & M(\bar{\boldsymbol{\sigma}})\Delta^{-1} \\ \boxed{\bar{A}^{(2,1)}} & \boxed{\bar{A}^{(2,2)}} & & \\ & & 0 & \Delta^{-1} \end{pmatrix}, \quad (24)$$

or more compactly as

$$A = \bar{A} + \left(\tilde{A} - \Pi_N \tilde{A} \Pi_N \right),$$

where \tilde{A} is defined by

$$\tilde{A}\mathbf{u} := \begin{pmatrix} \bar{\mathbf{w}} * \mathbf{u}^{(1)} + \bar{\boldsymbol{\sigma}} * \Delta^{-1} \mathbf{u}^{(2)} \\ \Delta^{-1} \mathbf{u}^{(1)} \end{pmatrix}.$$

Now that A is defined, the validation conditions are again the same as in Theorem 2.5 and 3.1, the only difference being the extra condition ensuring the injectivity of A , which is differs because the structure of A does.

Theorem 4.1. *With the notations introduced in this section, assume there exist constants Y , Z_1 and Z_2 satisfying*

$$\begin{aligned} \|AF(\bar{\mathbf{u}})\|_{\mathcal{X}_\nu} &\leq Y \\ \|I - ADF(\bar{\mathbf{u}})\|_{\mathcal{X}_\nu} &\leq Z_1 \\ \|AD^2F(\mathbf{u})\|_{\mathcal{X}_\nu} &\leq Z_2 \quad \forall \mathbf{u} \in \mathcal{X}_\nu, \end{aligned}$$

and

$$\begin{aligned} Z_1 &< 1 \\ 2YZ_2 &< (1 - Z_1)^2. \end{aligned}$$

Then, for any r satisfying

$$\frac{1 - Z_1 - \sqrt{(1 - Z_1)^2 - 2YZ_2}}{Z_2} \leq r < \frac{1 - Z_1}{Z_2},$$

there exists a unique fixed-point \mathbf{u}^* of T in $\mathcal{B}_{\mathcal{X}_\nu}(\bar{\mathbf{u}}, r)$, the closed ball of center $\bar{\mathbf{u}}$ and radius r in \mathcal{X}_ν . Assume further that $\bar{\mathbf{w}}$ (which plays a role in the definition of A), is such that

$$\|\mathbf{1} - \bar{\mathbf{w}} * (\mathbf{1} - \bar{\mathbf{u}}^{(2)})\|_\nu < 1.$$

Then \mathbf{u}^* is the unique zero of F in $\mathcal{B}_{\mathcal{X}_\nu}(\bar{\mathbf{u}}, r)$.

4.2 Derivation of the bounds

It once again remains to derive computable estimates Y , Z_1 and Z_2 satisfying the assumptions of Theorem 3.1. The required computations are still very similar to the one performed in Sections 2.3 and 3.4.

4.2.1 The bound Y

$\|AF(\bar{\mathbf{u}})\|_{\mathcal{X}_\nu}$ can be computed explicitly using interval arithmetic.

4.2.2 The bound Z_1

Introducing once more

$$B = \begin{pmatrix} B^{(1,1)} & B^{(1,2)} \\ B^{(2,1)} & B^{(2,2)} \end{pmatrix} = I - ADF(\bar{\mathbf{u}}),$$

we again consider $Z_1 = \max(Z_1^{finite}, Z_1^{tail})$, where Z_1^{finite} corresponds to

$$\max_{0 \leq n \leq 2N-2} \frac{1}{\xi_n(\nu)} \max \left(\|B_{(\cdot, n)}^{(1,1)}\|_\nu + \|B_{(\cdot, n)}^{(2,1)}\|_\nu, \|B_{(\cdot, n)}^{(1,2)}\|_\nu + \|B_{(\cdot, n)}^{(2,2)}\|_\nu \right),$$

and

$$\begin{aligned} Z_1^{tail} &= \left\| \begin{pmatrix} \mathbf{1} - \bar{\mathbf{w}} * (\mathbf{1} - \bar{\mathbf{u}}^{(2)}) & \bar{\boldsymbol{\sigma}} - \bar{\mathbf{w}} * (\boldsymbol{\gamma} + \bar{\mathbf{u}}^{(1)}) \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \right\|_\nu + \frac{1}{(N\pi)^2} \left\| \begin{pmatrix} \|\bar{\boldsymbol{\sigma}}\|_\nu \|R'(\bar{\mathbf{u}}^{(1)})\|_\nu & \mathbf{0} \\ \|R'(\bar{\mathbf{u}}^{(1)})\|_\nu & \mathbf{0} \end{pmatrix} \right\|_1 \\ &= \max \left[\|\mathbf{1} - \bar{\mathbf{w}} * (\mathbf{1} - \bar{\mathbf{v}})\|_\nu + \frac{1}{(N\pi)^2} (1 + \|\bar{\boldsymbol{\sigma}}\|_\nu) \|R'(\bar{\mathbf{u}})\|_\nu, \|\bar{\boldsymbol{\sigma}} - \bar{\mathbf{w}} * (\boldsymbol{\gamma} + \bar{\mathbf{u}})\|_\nu \right], \end{aligned} \quad (25)$$

bounds

$$\sup_{n \geq 2N-1} \frac{1}{\xi_n(\nu)} \max \left(\|B_{(\cdot, n)}^{(1,1)}\|_\nu + \|B_{(\cdot, n)}^{(2,1)}\|_\nu, \|B_{(\cdot, n)}^{(1,2)}\|_\nu + \|B_{(\cdot, n)}^{(2,2)}\|_\nu \right).$$

Remark 4.2. The obtained Z_1^{tail} estimate shows that the proposed approach of introducing a system of differential-algebraic equations to deal with (at least some types of) non-polynomial diffusion terms is not only valid but also efficient, because we keep the $\frac{1}{N^2}$ scaling. Even the constant is close to optimal, see the discussion in Appendix B.

4.2.3 The bound Z_2

Finally, remembering that $R(u) = \alpha u - \beta u^2 + g$, an estimation similar to the ones performed in the two previous cases leads to

$$Z_2 = \max \left[\|A^{(1,1)}\|_\nu + \|A^{(2,1)}\|_\nu, 2\beta \left(\|A^{(1,2)}\|_\nu + \|A^{(2,2)}\|_\nu \right) \right].$$

4.3 Example and results

In this subsection, we apply the technique we just presented to a specific example, and again consider $R(u) = u - u^2 + g$ (i.e. we take $\alpha = \beta = 1$), with g as in Section 2.4, but this time with $\Phi(u) = \frac{u}{\gamma+u}$, and two different values of γ .

In each case, we computed an approximate solution \bar{u} (see Figure 6), and then validated it using the procedure described in this whole section, which yield the following theorem.

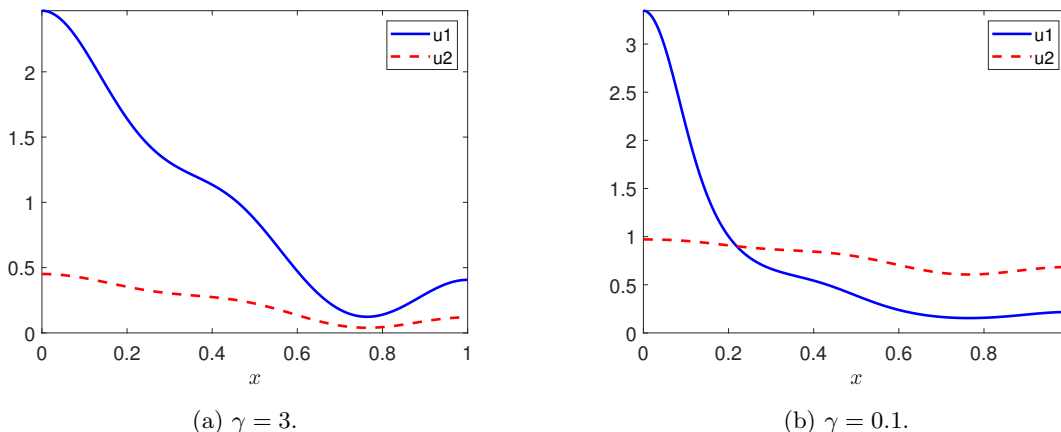


Figure 6: The two approximate solutions of (22) with $\Phi(u) = \frac{u}{\gamma+u}$, $R(u) = u - u^2 + g$ and g as in Figure 1a, which have been validated in Theorem 4.3. u_1 corresponds to the solution u itself, while u_2 corresponds to the additional variable $v = \Phi(u)$ introduced in (23).

Theorem 4.3. *Let $\gamma = 3$, and \bar{u} be the function whose Fourier coefficients $\bar{\mathbf{u}}$ can be downloaded at [3], and which is represented in Figure 6a. There exists a strong solution u of (22), with Φ and R as given just above, such that $\|u - \bar{u}\|_{C^0} \leq 1 \times 10^{-13}$.*

Let $\gamma = 0.1$, and \bar{u} be the function whose Fourier coefficients $\bar{\mathbf{u}}$ can be downloaded at [3], and which is represented in Figure 6b. There exists a strong solution u of (22), with Φ and R as given just above, such that $\|u - \bar{u}\|_{C^0} \leq 8 \times 10^{-6}$.

Proof. We take $N = 50$ and $\nu = 1.1$, compute the estimates Y , Z_1 and Z_2 obtained in Section 4.2 for each approximate solution, and apply Theorem 4.1. The computational parts of the proof, can be reproduced by running the Matlab code `script_NP.m` from [3], together with Intlab [29], which provides more details such as the precise value of each bound for each solution. \square

Remark 2.9 also applies here. To tie back these results to the discussion at the start of this section, notice from Figure 6 that, for $\gamma = 3$ we have $\|u\|_{C^0} < \gamma$, so one could hope to prove that $\Phi(u) = \sum_{k \geq 0} (-1)^k \frac{u^{k+1}}{\gamma^{k+1}}$ is converging power series, but that is definitely not true for $\gamma = 0.1$. In that case, u stays positive so we are still safely away from the pole of Φ , but one would have to use a power series expansion around some non zero value depending on the solution (e.g. the middle of the interval of values taken by u).

Appendix

A About the computation of \bar{A}

We give here more details about Remark 2.4, and in particular explain why, no matter how large N is, taking $\bar{A} = (\Pi_N DF(\bar{\mathbf{u}})\Pi_N)^{-1}$ is not necessarily suitable. Indeed, let us consider the case described in Section 2, and have a closer look at the coefficient in position (N, N) of $ADF(\bar{\mathbf{u}})$. The coefficients in A and $DF(\bar{\mathbf{u}})$ that are relevant for this computation are highlighted below:

$$\left(\begin{array}{c|c} \bar{A} & \\ \hline \frac{-\bar{w}_1}{(N\pi)^2} & \frac{-\bar{w}_2}{((N+1)\pi)^2} \cdots \end{array} \right) \left(\begin{array}{c|c} \Pi_N DF(\bar{\mathbf{u}})\Pi_N & \\ \hline -N^2\Phi'_1(\bar{\mathbf{u}}) & \\ -(N+1)^2\Phi'_2(\bar{\mathbf{u}}) & \\ \vdots & \end{array} \right).$$

There should also be terms corresponding to $R'(\bar{\mathbf{u}})$ in $DF(\bar{\mathbf{u}})$, but they do not have a factor N^2 in front, hence for N large enough we can neglect them. Therefore we have

$$(ADF(\bar{\mathbf{u}}))_{N,N} \approx (\bar{A} \Pi_N DF(\bar{\mathbf{u}})\Pi_N)_{N,N} + \sum_{n \geq 1} \bar{w}_n \Phi'_n(\bar{\mathbf{u}}).$$

Now, if we were to take $A \approx (\Pi_N DF(\bar{\mathbf{u}})\Pi_N)^{-1}$, then we would have in particular

$$(\bar{A} \Pi_N DF(\bar{\mathbf{u}})\Pi_N)_{N,N} \approx 1$$

and therefore

$$(ADF(\bar{\mathbf{u}}))_{N,N} \approx 1 + \sum_{n \geq 1} \bar{w}_n \Phi'_n(\bar{\mathbf{u}}).$$

However, $\sum_{n \geq 1} \bar{w}_n \Phi'_n(\bar{\mathbf{u}})$ has no reason to be small in general, so $(ADF(\bar{\mathbf{u}}))_{N,N}$ could be significantly different from 1, which would prevent us from satisfying (11a).

B About the Z_1^{tail} bound in the non-polynomial case

We make some comments about the quality of the bound (25), which in some sense measures how well adapted the transformation of (22) into (23) and the choice of A in (24) are for computer-assisted proofs, and explain why we claim in Remark 4.2 the output is close to optimal.

In order to get a comparison, notice that for this example Φ can be inverted by hand, and therefore (22) can easily be rewritten as an equation with linear diffusion

$$\begin{cases} \Delta v + \tilde{R}(v) = 0 & \text{on } (0, 1), \\ \frac{\partial v}{\partial n} = 0 & \text{on } \{0, 1\}, \end{cases}$$

where $v = \Phi(u)$ and

$$\tilde{R}(v) = R(\Phi^{-1}(v)) = R\left(\frac{\gamma v}{1-v}\right).$$

In this case which is standard since there is no nonlinear diffusion, the Z_1^{tail} estimate would be equal to

$$\frac{1}{(N\pi)^2} \|\tilde{R}'(\bar{\mathbf{v}})\|_\nu = \frac{1}{(N\pi)^2} \left\| \frac{\gamma}{(1-\bar{\mathbf{v}})^2} R'(\bar{\mathbf{u}}) \right\|_\nu.$$

However, since $\bar{\mathbf{w}} \approx \frac{1}{1-\bar{\mathbf{v}}}$, $\gamma + \bar{\mathbf{u}} \approx \gamma \bar{\mathbf{w}}$ and $\bar{\sigma} \approx \bar{\mathbf{w}} * (\gamma + \bar{\mathbf{u}}) \approx \gamma \bar{\mathbf{w}}^2$, hence the above Z_1^{tail} estimate would be roughly equal to

$$\frac{1}{(N\pi)^2} \|\bar{\sigma} R'(\bar{\mathbf{u}})\|_\nu.$$

The estimate obtained in (25) is worst, but not dramatically so. Indeed, in practice the first part of (25) can always be made small enough not to matter, and essentially (25) reduces to

$$\frac{1}{(N\pi)^2} (1 + \|\bar{\sigma}\|_\nu) \|R'(\bar{\mathbf{u}})\|_\nu.$$

The extra 1 could even be removed by choosing an appropriately weighted norm on the product space $\ell_\nu^1 \times \ell_\nu^1$ (see e.g. the discussion about the weights in [7]). Therefore the only thing that we lost, at least concerning the crucial Z_1^{tail} estimate, by going through a system of differential-algebraic equation is that $\|\bar{\sigma} R'(\bar{\mathbf{u}})\|_\nu$ is replaced by $\|\bar{\sigma}\|_\nu \|R'(\bar{\mathbf{u}})\|_\nu$, which in most cases is perfectly acceptable. This shows that, in more complex situations where Φ cannot be explicitly inverted, the approach proposed in this paper provides a good alternative.

References

- [1] G. Arioli and H. Koch. Computer-assisted methods for the study of stationary solutions in dissipative systems, applied to the Kuramoto-Sivashinski equation. *Arch. Ration. Mech. Anal.*, 197(3):1033–1051, 2010.
- [2] D. G. Aronson. The porous medium equation. In *Nonlinear diffusion problems*, pages 1–46. Springer, 1986.
- [3] M. Breden. Matlab code for "Computer-assisted proofs for some nonlinear diffusion problems". <https://sites.google.com/site/maximebreden/research>, 2021.
- [4] M. Breden and R. Castelli. Existence and instability of steady states for a triangular cross-diffusion system: a computer-assisted proof. *Journal of Differential Equations*, 264(10):6418–6458, 2018.
- [5] M. Breden, L. Desvillettes, and J.-P. Lessard. Rigorous numerics for nonlinear operators with tridiagonal dominant linear part. *Discrete & Continuous Dynamical Systems-A*, 35(10):4765, 2015.
- [6] M. Breden and M. Engel. Computer-assisted proof of shear-induced chaos in stochastically perturbed Hopf systems. *arXiv preprint arXiv:2101.01491*, 2021.
- [7] M. Breden and C. Kuehn. Rigorous validation of stochastic transition paths. *Journal de Mathématiques Pures et Appliquées*, 131:88–129, 2019.
- [8] M. Breden, C. Kuehn, and C. Soresina. On the influence of cross-diffusion in pattern formation. *arXiv preprint arXiv:1910.03436*, 2019.
- [9] M. Breden, J.-P. Lessard, and M. Vanicat. Global bifurcation diagrams of steady states of systems of PDEs via rigorous numerics: a 3-component reaction-diffusion system. *Acta applicandae mathematicae*, 128(1):113–152, 2013.
- [10] S. Day, J.-P. Lessard, and K. Mischaikow. Validated continuation for equilibria of PDEs. *SIAM J. Numer. Anal.*, 45(4):1398–1424, 2007.

- [11] J. Gómez-Serrano. Computer-assisted proofs in PDE: a survey. *SeMA Journal*, 76(3):459–484, 2019.
- [12] R. S. Hamilton et al. Three-manifolds with positive Ricci curvature. *J. Differential geom*, 17(2):255–306, 1982.
- [13] M. Iida, M. Mimura, and H. Ninomiya. Diffusion, cross-diffusion and competitive interaction. *Journal of mathematical biology*, 53(4):617–641, 2006.
- [14] H. Izuhara, M. Mimura, et al. Reaction-diffusion system approximation to the cross-diffusion competition system. *Hiroshima Mathematical Journal*, 38(2):315–347, 2008.
- [15] A. Jüngel. The boundedness-by-entropy method for cross-diffusion systems. *Nonlinearity*, 28(6):1963, 2015.
- [16] M. Langlais and D. Phillips. Stabilization of solutions of nonlinear and degenerate evolution equations. *Nonlinear Analysis: Theory, Methods & Applications*, 9(4):321–333, 1985.
- [17] J.-P. Lessard and J. D. Mireles James. Computer assisted fourier analysis in sequence spaces of varying regularity. *SIAM Journal on Mathematical Analysis*, 49(1):530–561, 2017.
- [18] J.-P. Lessard, J. D. Mireles James, and J. Ransford. Automatic differentiation for Fourier series and the radii polynomial approach. *Physica D: Nonlinear Phenomena*, 334:174–186, 2016.
- [19] Y. Lou and W.-M. Ni. Diffusion, self-diffusion and cross-diffusion. *Journal of Differential Equations*, 131(1):79–131, 1996.
- [20] Y. Lou, W.-M. Ni, and S. Yotsutani. On a limiting system in the lotka–volterra competition with cross-diffusion. *Discrete & Continuous Dynamical Systems-A*, 10(1&2):435, 2004.
- [21] M. Mimura and K. Kawasaki. Spatial segregation in competitive interaction-diffusion equations. *Journal of Mathematical Biology*, 9(1):49–64, 1980.
- [22] M. Mimura, Y. Nishiura, A. Tesei, and T. Tsujikawa. Coexistence problem for two competing species models with density-dependent diffusion. *Hiroshima Mathematical Journal*, 14(2):425–449, 1984.
- [23] M. T. Nakao. A numerical approach to the proof of existence of solutions for elliptic problems. *Japan Journal of Applied Mathematics*, 5(2):313, 1988.
- [24] M. T. Nakao, M. Plum, and Y. Watanabe. *Numerical verification methods and computer-assisted proofs for partial differential equations*. Springer, 2019.
- [25] W.-M. Ni. Diffusion, cross-diffusion, and their spike-layer steady states. *Notices of the AMS*, 45(1):9–18, 1998.
- [26] S. Oishi. Numerical verification of existence and inclusion of solutions for nonlinear operator equations. *Journal of Computational and Applied Mathematics*, 60(1-2):171–185, 1995.
- [27] J. M. Ortega. The newton-kantorovich theorem. *The American Mathematical Monthly*, 75(6):658–660, 1968.
- [28] M. Plum. Explicit H²-estimates and pointwise bounds for solutions of second-order elliptic boundary value problems. *Journal of Mathematical Analysis and Applications*, 165(1):36–61, 1992.
- [29] S. M. Rump. INTLAB - INTerval LABoratory. *Developments in Reliable Computing*, Kluwer Academic Publishers, Dordrecht, pp, pages 77–104, 1999.
- [30] K. Ryu and I. Ahn. Coexistence theorem of steady states for nonlinear self-cross diffusion systems with competitive dynamics. *Journal of mathematical analysis and applications*, 283(1):46–65, 2003.

- [31] K. Sekine, M. T. Nakao, and S. Oishi. A new formulation using the schur complement for the numerical existence proof of solutions to elliptic problems: without direct estimation for an inverse of the linearized operator. *Numerische Mathematik*, 146(4):907–926, 2020.
- [32] N. Shigesada, K. Kawasaki, and E. Teramoto. Spatial segregation of interacting species. *Journal of theoretical biology*, 79(1):83–99, 1979.
- [33] A. Takayasu, X. Liu, and S. Oishi. Verified computations to semilinear elliptic boundary value problems on arbitrary polygonal domains. *Nonlinear Theory and Its Applications, IEICE*, 4(1):34–61, 2013.
- [34] P. M. Topping and H. Yin. Sharp decay estimates for the logarithmic fast diffusion equation and the Ricci flow on surfaces. *Annals of PDE*, 3(1):6, 2017.
- [35] J. B. van den Berg and J. F. Williams. Validation of the bifurcation diagram in the 2D Ohta–Kawasaki problem. *Nonlinearity*, 30(4):1584, 2017.
- [36] J. L. Vázquez. *Smoothing and decay estimates for nonlinear diffusion equations: equations of porous medium type*, volume 33. Oxford University Press, 2006.
- [37] J. L. Vázquez. *The porous medium equation: mathematical theory*. Oxford University Press, 2007.
- [38] T. Wanner. Computer-assisted equilibrium validation for the diblock copolymer model. *Discrete & Continuous Dynamical Systems-A*, 37(2):1075, 2017.
- [39] N. Yamamoto. A numerical verification method for solutions of boundary value problems with local uniqueness by Banach’s fixed-point theorem. *SIAM J. Numer. Anal.*, 35:2004–2013, 1998.
- [40] P. Zgliczynski and K. Mischaikow. Rigorous numerics for partial differential equations: The Kuramoto–Sivashinsky equation. *Foundations of Computational Mathematics*, 1(3):255–288, 2001.