

Computer-assisted proofs for the many steady states of a chemotaxis model with local sensing

Maxime Breden ^{*} Maxime Payan [†]

November 22, 2023

Abstract

We study the steady states of a system of cross-diffusion equations arising from the modeling of chemotaxis with local sensing, where the motility is a decreasing function of the concentration of the chemical. In order to capture the many different equilibria that sometimes co-exist, we use computer-assisted proofs: Given an approximate solution obtained numerically, we apply a fixed-point argument in a small neighborhood of this approximate solution to prove the existence of an exact solution nearby. This allows us to rigorously study the steady states of this cross-diffusion system much more extensively than what previously possible with purely pen-and-paper techniques. Our computer-assisted argument makes use of Fourier series decomposition, which is common in the literature, but usually restricted to systems with polynomial nonlinearities. This is not the case for the model considered in this paper, and we develop a new way of dealing with some nonpolynomial nonlinearities in the context of computer-assisted proofs with Fourier series.

1 Introduction

Chemotaxis is one of the critical mechanisms through which the behavior of bacteria and other small organisms can be understood, and the patterns they form explained. Following the seminal papers of Patlak [31], Keller and Segel [19, 20], chemotaxis models have been intensively studied in the mathematical literature. However, even the simplest patterns generated by such models, i.e. (nonhomogeneous) stationary solutions, often remains very challenging to rigorously analyze.

The purpose of this paper is to provide a detailed study of the (positive) steady states of the following

^{*}CMAP, CNRS, École polytechnique, Institut Polytechnique de Paris, 91120 Palaiseau, France.
maxime.breden@polytechnique.edu

[†]CMAP, CNRS, École polytechnique, Institut Polytechnique de Paris, 91120 Palaiseau, France.
maxime.payan@polytechnique.edu

chemotaxis model:

$$\begin{cases} \partial_t u = \Delta(\gamma(v)u) + \sigma u(1-u) & \text{in } (0, \infty) \times \Omega, \\ \partial_t v = d\Delta v + u - v & \text{in } (0, \infty) \times \Omega, \\ \partial_n(\gamma(v)u) = 0 = \partial_n v & \text{on } (0, \infty) \times \partial\Omega. \end{cases} \quad (1.1)$$

Here, $u = u(t, x)$ denotes the bacteria or cell density, $v = v(t, x)$ the density of a chemical produced by the bacteria, and Ω is a bounded domain of \mathbb{R}^N . Throughout the paper, we assume that γ is smooth and positive on $(0, \infty)$, $d > 0$ and $\sigma \geq 0$. The most interesting situation, at least if one wants to maybe observe pattern formation, is the one where the bacteria are in some sense attracted by the chemical. In (1.1), this can be modeled by taking γ decreasing. We note that, while we will only consider examples where γ is indeed decreasing, this assumption is in fact not needed for our results (and neither is the positivity of σ).

System (1.1) is in fact a specific case of one of the original models introduced by Keller and Segel in [20] (up to the logistic reaction term), where the equation on u writes

$$\partial_t u = \nabla \cdot (\mu(v)\nabla u - \chi(v)u\nabla v). \quad (1.2)$$

The diffusive term in (1.1) is then obtained by assuming that the diffusivity μ and the chemosensitivity χ are related by $\mu' = -\chi$, which corresponds to assuming that the bacteria are only sensitive to the concentration of v , but not to its gradient (*local sensing*). For more details on these modeling interpretations, we refer to [10] and the references therein. System (1.1) was also reintroduced in [12], as an attempt to explain the stripe patterns observed in [23], and it has become the subject of recent scrutiny from the mathematical community.

For $\sigma > 0$, global well-posedness for the initial value problem associated to (1.1) was first established in [17] in dimension at most 2, and under some further assumptions on γ , encompassing prototypical examples such as

$$\gamma(v) = \frac{a}{(1+bv)^c}, \quad \gamma(v) = \frac{a}{(1+e^{bv})^c} \quad \text{or} \quad \gamma(v) = 1 - \frac{v}{\sqrt{1+v^2}}, \quad (1.3)$$

for $a, b, c > 0$. Global well-posedness was then also obtained in higher dimension in [24, 43], under an additional assumption on σ , which has to be sufficiently large.

The case $\sigma = 0$ is more delicate, as the decay of γ at infinity seems to play a critical role. Indeed, for simple examples of γ , global existence could be established when γ decays algebraically, whereas a critical mass threshold leading to blow-ups was identified for exponentially decaying γ 's. More precisely, assuming $\gamma(v) = \frac{a}{v^c}$ with a small enough, the global existence of bounded solutions was obtained in [47] in any dimension, whereas global weak solutions are obtained in [10] for $\gamma(v) = \frac{a}{1+v^c}$, without any smallness assumption on a but with restrictions on c depending on the dimension. On the other hand, when $\gamma(v) = e^{-bv}$, blow-up can occur if the initial mass is large enough [18], although only

in infinite time [7, 13], which contrasts with the famous behavior for the minimal Keller-Segel model (i.e. when taking μ and χ constant in (1.2)), where blow-up occurs in finite time [16].

Regarding steady states of system (1.1), the first trivial observation is that $(0, 0)$ and $(1, 1)$ are the only homogenous steady states when $\sigma > 0$, whereas (c, c) is a steady state for any $c \in \mathbb{R}$ when $\sigma = 0$. Assuming $\sigma > 0$, a straightforward linear stability analysis of the positive steady state $(1, 1)$ yields the following necessary conditions for its linear instability (which are also sufficient, up to choosing a compatible domain):

$$\begin{cases} \sigma d + \gamma(1) + \gamma'(1) < 0, \\ (\sigma d + \gamma(1) + \gamma'(1))^2 - 4\sigma d\gamma(1) > 0. \end{cases} \quad (1.4)$$

In particular, one immediately sees that, when the product σd is large enough, the trivial steady state $(1, 1)$ is linearly stable. Using a priori bounds and the Leray-Schauder index, one can go further and show that whenever σ is large enough (with a threshold depending on d), or whenever d is large enough (with a threshold depending on σ), there exists no other positive steady state of (1.1), see [25].

Regardless of the (nonnegative) value of σd , one also sees that, if γ is such that $\gamma(1) + \gamma'(1) \geq 0$, then $(1, 1)$ must be linearly stable. On the other hand, if $\gamma(1) + \gamma'(1) < 0$, condition (1.4) does hold for at least some values of σ and d . One can then use local bifurcation theory to predict the apparition of branches of nonhomogeneous steady states, and further usage of the Leray-Schauder index and of other global bifurcation tools can then provide some further insight on the existence of nontrivial steady states. Such bifurcation analyses were carried out in [25, 44], using σ as a bifurcation parameter. These works do provide existence results regarding non trivial steady states, with typical statements of the form: “if σ belongs to a given interval (depending on the other parameters), there exists at least one non trivial steady state of (1.1)”. However, such results are still very far from providing a full description of the steady states of (1.1), especially when contrasted with numerical experiments, which suggest the existence of intricate bifurcation diagrams of steady states, featuring many co-existing steady states for some given value of σ (see Figure 1).

When $\sigma = 0$, the linear instability condition for a given steady state (c, c) reduces to $\gamma(c) + \gamma'(c) < 0$ (see [10] for the details). Here again, there are already results going beyond this linear analysis. First negative ones, assuming that γ decay sublinearly, or more precisely, that $v \mapsto v\gamma(v)$ is increasing. Under this hypothesis, it is shown in [11] that the system (1.1) admits a Lyapunov functional, which prevents the existence of any nontrivial steady state. Second, positive results, when $\gamma(c) + \gamma'(c) < 0$, which rely again on global bifurcation theory [44]. However, in the $\sigma = 0$ case also, the available results fail to encompass the variety of steady states that can be observed numerically.

In this paper, we develop a computer-assisted methodology providing a rigorous description of the steady states of (1.1), which is essentially as complete and as precise as what can be observed numerically.

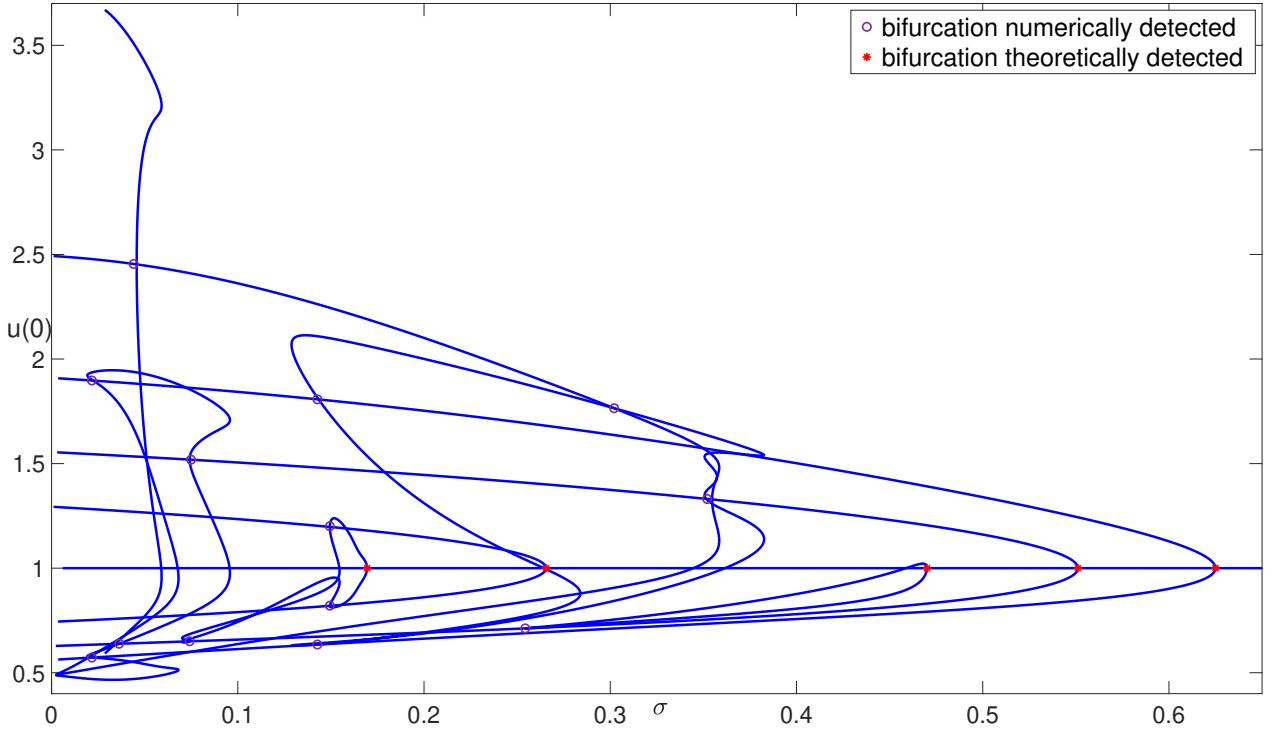


Figure 1: Numerical bifurcation diagram of steady states of the system (1.1), with $\Omega = (0, 3\pi)$, $d = 1$ and $\gamma(x) = \frac{1}{1+x^9}$.

ically. More precisely, for any candidate approximate steady state denoted (\bar{u}, \bar{v}) of (1.1), we provide sufficient conditions ensuring the existence of a true steady state in a small and explicit neighborhood of (\bar{u}, \bar{v}) . Given any (\bar{u}, \bar{v}) , typically obtained numerically, we can then explicitly check whether these conditions hold, and if they do conclude to the existence of a nearby true steady state of (1.1).

Remark 1.1. *In this paper, we only focus on the one dimensional case, which is already very rich, and write $\Omega = (a, b)$, but higher spatial dimensions could also be investigated (in a very similar way for rectangular domains, but with more challenging adaptations for more complicated geometries).*

As an example, we show below the kind of output that can be expected from this procedure, for two different choices of γ .

Theorem 1.1. *Let \bar{u}, \bar{v} be the functions depicted on Figure 2 and whose precise description in terms of Fourier coefficients can be found at [32]. There exists a smooth steady states (u, v) of the system 1.1 with the parameters $\sigma = 0.053$, $d = 1$, $\Omega = (0, 3\pi)$ and $\gamma(x) = \frac{1}{1+x^9}$, such that*

$$\sup_{x \in \Omega} |u(x) - \bar{u}(x)| + \sup_{x \in \Omega} |v(x) - \bar{v}(x)| \leq 2.5197 \times 10^{-8}.$$

Notice that we not only get an existence statement regarding steady states of (1.1), but also very quantitative information regarding the obtained steady state, as we have an explicit approximation together with small error bounds.

Remark 1.2. *The error bound in Theorem 1.1 is stated in the C^0 norm for the sake of simplicity, but is actually first obtained in a stronger norm, which will be introduced in Section 2.*

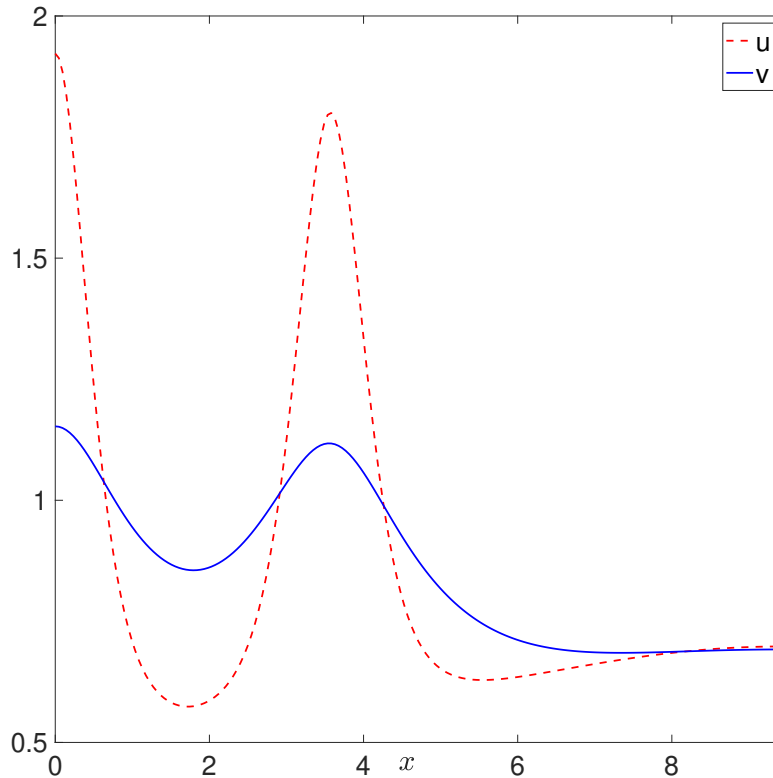


Figure 2: An approximate steady state of (1.1), associated to Theorem 1.1.

Here is another example, for a different choice of motility function γ .

Theorem 1.2. *Let \bar{u}, \bar{v} be the functions depicted on Figure 3 and whose precise description in terms of Fourier coefficients can be found at [32]. There exists a smooth steady states (u, v) of the system 1.1 with the parameters $\sigma = 0.6$, $d = 1$, $\Omega = (0, 4\pi)$ and $\gamma(x) = \frac{1}{1 + e^{9(x-1)}}$, such that*

$$\sup_{x \in \Omega} |u(x) - \bar{u}(x)| + \sup_{x \in \Omega} |v(x) - \bar{v}(x)| \leq 1.6956 \times 10^{-12}.$$

Remark 1.3. *Theorem 1.2 corroborates the results of [44, Fig.3], where an approximate steady state resembling the one of our Figure 3 was obtained numerically. The techniques presented in this paper allow to rigorously establish the existence of an exact solution near this approximation.*

We can of course repeat such an argument for different approximate solutions, and for different choices of parameters. Putting all these results together, we indeed obtain a rigorous description of the full picture, or at least of the full picture that was found numerically: we can essentially validate any solution that we found numerically, but cannot exclude the existence of other solutions that may have escaped our numerical investigation.

Theorem 1.3. *Each point of the Figure 4 represents a steady state for (1.1), with $\Omega = (0, 3\pi)$, $d = 1$ and $\gamma(x) = \frac{1}{1 + x^9}$, for the σ plotted in x -coordinates.*

Remark 1.4. *While it is beyond the scope of this work, let us mention that one may try to go one step beyond Theorem 1.3, and really validate the continuous bifurcation diagram depicted in Figure 1, instead of only the discreteized version of Figure 4. This would require combining the point-wise (in the*

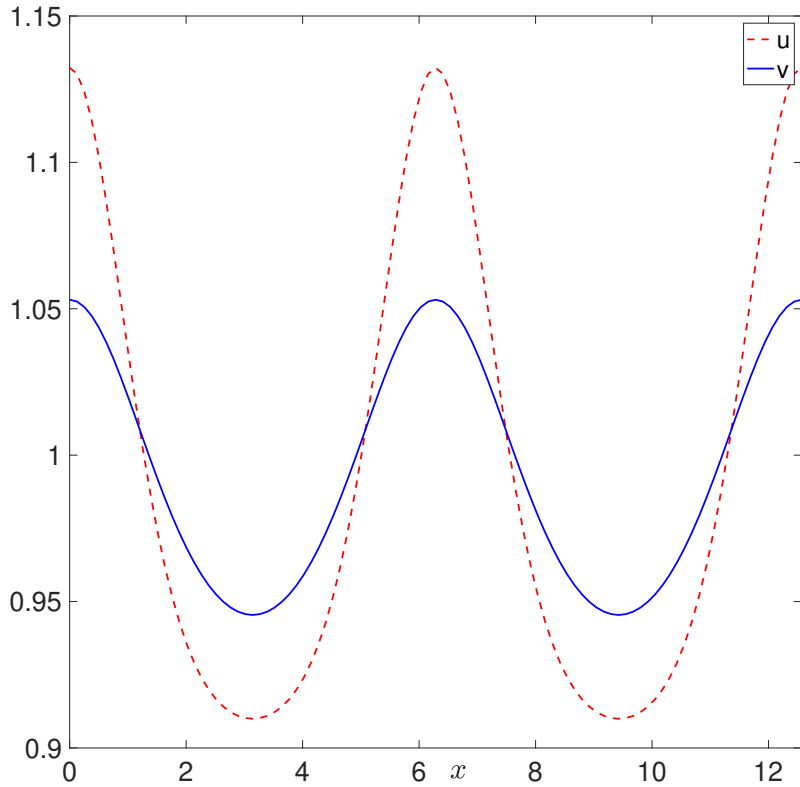


Figure 3: An approximate steady state of (1.1), associated to Theorem 1.2.

parameters) theorems obtained here with two extra ingredients. First, rigorous continuation techniques (see e.g. [5, 6, 42, 45]), in order to validate entire branches of steady states. Second, a desingularization procedure allowing to handle the neighborhood of the pitchfork bifurcation points (see e.g. [1, 22, 36]).

Computer-assisted techniques for studying elliptic PDEs go back to the pioneering works of Nakao [28] and Plum [33], and have been more and more widely used since then [2, 9, 14, 29, 30, 38, 40, 46, 48]. However, their applicability was limited to semilinear problems until very recently, and we will deal with the quasilinear structure of (1.1) by using the techniques introduced in [4].

Several of our estimates are based on Fourier series decomposition of functions, which has proven to be a very efficient framework for conducting computer-assisted proofs. However, this means that (1.1) present an additional difficulty, namely the fact that the typically γ 's we are interested in cannot be polynomial (from a modeling point of view, reasonable choices of γ should go to 0 at $+\infty$). In the context of computer-assisted proofs using Fourier series, nonpolynomial terms are usually dealt with with tools based on automatic differentiation or polynomial embeddings [15, 21], which have proven successful but still suffer from several shortcomings. In particular, they are usually limited to one-dimensional problems, they require enlarging the system with additional equations and unknowns, and for elliptic equations like the steady state problem associated to (1.1), they destroy the second order structure of the system, which in turn makes the computer-assisted proof harder. In this work, we develop an alternate approach to deal with several types of nonlinear terms, including rational functions, exponential functions, and compositions thereof, which does not necessitate the introduction

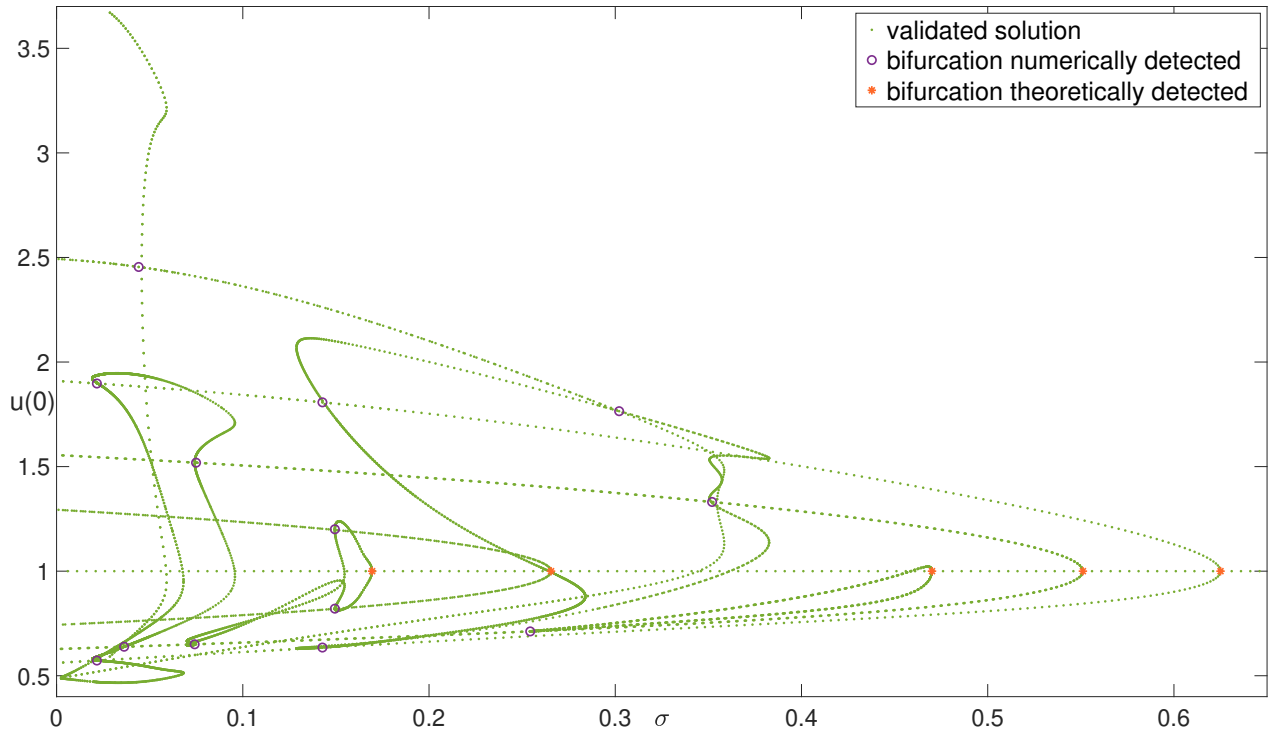


Figure 4: Validated steady states of the system (1.1), with $\Omega = (0, 3\pi)$, $d = 1$ and $\gamma(x) = \frac{1}{1+x^9}$, associated to Theorem 1.3

of new variables and preserves the second order structure of the system.

The remainder of the paper is organized as follows. In Section 2, we introduce some notations and spaces in which we will study the problem, together with a version of the Newton-Kantorovich theorem that will allow us to prove the existence of an exact solution near a numerical approximation. In Section 3, we explain how to deal with the nonpolynomial nonlinearities γ considered in this work. The bounds required to apply the Newton-Kantorovich theorem are then obtained in Section 4. Finally, Section 5 contains some details regarding the proofs of Theorem 1.1, Theorem 1.2 and Theorem 1.3. All the computations presented in this paper have been implemented in MATLAB, using the INTLAB package [37] for interval arithmetic. The computer-assisted parts of the proofs can be reproduced using the code available at [32].

2 Zero-finding problem and the Newton-Kantorovich Theorem

In this section, we introduce the first useful notations to start building a bridge between the numerical aspects (to get numerical solutions) and the theoretical ones (to obtain theorems). To that end, we rewrite our problem as a zero-finding problem, for a single function F defined on a well chosen Banach space. Then, we introduce the Newton-Kantorovich, Theorem 2.1, and its hypotheses. This theorem is the tool which we rely on to show the existence of solutions.

2.1 Theoretical framework and associated notations

Let u be a smooth function defined on $\Omega = (a, b) \subset \mathbb{R}$, $a < b$. We will describe all the objects useful for our demonstration. Since we are in dimension one with homogeneous Neumann conditions at the boundary, we look for solutions as Fourier cosine series, i.e.

$$u(x) = \mathbf{u}_0 + 2 \sum_{n=1}^{+\infty} \mathbf{u}_n \cos\left(\frac{n\pi}{b-a}(x-a)\right).$$

For any smooth function u , we denote by \mathbf{u} the sequence of its Fourier coefficients $(\mathbf{u}_n)_{n \in \mathbb{N}}$. In particular, it will help to use $\mathbf{0} = (0, 0, \dots)$ and $\mathbf{1} = (1, 0, \dots)$. We denote by \mathcal{F} the Fourier transform, i.e. the map such that $\mathcal{F}(u) = \mathbf{u}$.

In order to estimate our objects in the right spaces, we have to choose a norm. We take a weighted ℓ^1 norm with exponential decay on our sequences, i.e., for any $\nu \geq 1$,

$$\|\mathbf{u}\|_\nu = |\mathbf{u}_0| + 2 \sum_{n=1}^{+\infty} |\mathbf{u}_n| \nu^n = \sum_{n=0}^{+\infty} |\mathbf{u}_n| \xi_n(\nu),$$

where

$$\xi_n(\nu) = \begin{cases} 1, & \text{if } n = 0, \\ 2\nu^n, & \text{otherwise.} \end{cases}$$

The Banach space we choose is

$$\ell_\nu^1 = \left\{ \mathbf{u} \in \mathbb{R}^{\mathbb{N}}, \|\mathbf{u}\|_\nu < \infty \right\}.$$

This space represents smooth functions, as soon as $\nu > 1$. There is a clear correspondence between the space ℓ_ν^1 and a functional space of analytic functions (see e.g. [35, p.98]). The Banach space ℓ_ν^1 is furthermore a (unital commutative) Banach algebra with the ‘‘product’’ $*$, the usual discrete convolution,

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

and $\|\cdot\|_\nu$ is an algebraic norm,

$$\|\mathbf{u} * \mathbf{v}\|_\nu \leq \|\mathbf{u}\|_\nu \|\mathbf{v}\|_\nu, \quad \|\mathbf{1}\|_\nu = 1.$$

Remark 2.1. For $\nu < 1$, the space ℓ_ν^1 is no longer a Banach algebra.

Let $L \in \mathcal{L}(\ell_\nu^1)$ a linear operator. It can be seen as a *infinite matrix* $(L_{k,n})_{(k,n) \in \mathbb{N}^2}$ where

$$(L\mathbf{u})_k = \sum_{n=0}^{+\infty} L_{k,n} \mathbf{u}_n.$$

The associated operator norm on $\mathcal{L}(\ell_\nu^1)$ can be expressed as follows,

$$\|L\|_\nu := \sup_{\|\mathbf{u}\|_\nu=1} \|L\mathbf{u}\|_\nu = \sup_{n \in \mathbb{N}} \frac{1}{\xi_n(\nu)} \|L_{(\cdot,n)}\|_\nu.$$

$L_{(\cdot, n)}$ corresponds to the n -th column vector of L , from a matrix point of view.

We denote by $M(\mathbf{u}) \in \mathcal{L}(\ell_\nu^1)$ the multiplication operator associated to the convolution by \mathbf{u} , i.e.

$$\begin{aligned} M(\mathbf{u}) : \ell_\nu^1 &\longrightarrow \ell_\nu^1 \\ \mathbf{v} &\longmapsto \mathbf{u} * \mathbf{v} \end{aligned}$$

Notice that $\|M(\mathbf{u})\|_\nu = \|\mathbf{u}\|_\nu$, and that $\|u\|_{C^0} := \sup_{x \in [a, b]} |u(x)| \leq \|\mathbf{u}\|_\nu$.

We also define the norm of a vector $(\mathbf{u}, \mathbf{v}) \in \ell_\nu^1 \times \ell_\nu^1$:

$$\|(\mathbf{u}, \mathbf{v})\|_\nu = \|\mathbf{u}\|_\nu + \|\mathbf{v}\|_\nu.$$

In the sequel, we also have to deal with linear operators L acting on $\ell_\nu^1 \times \ell_\nu^1$, for which we use a block-notation:

$$L = \begin{pmatrix} L^{11} & L^{12} \\ L^{21} & L^{22} \end{pmatrix}, \text{ with } L^{ij} \in \mathcal{L}(\ell_\nu^1), i, j \in \{1, 2\}.$$

The corresponding operator norm can then be expressed as follows:

$$\begin{aligned} \|L\|_\nu &= \sup_{(\ell_\nu^1)^2 \ni (\mathbf{u}, \mathbf{v}) \neq (\mathbf{0}, \mathbf{0})} \frac{\|L^{11}\mathbf{u} + L^{12}\mathbf{v}\|_\nu + \|L^{21}\mathbf{u} + L^{22}\mathbf{v}\|_\nu}{\|\mathbf{u}\|_\nu + \|\mathbf{v}\|_\nu} \\ &= \sup_{n \in \mathbb{N}} \frac{1}{\xi_k(\nu)} \max \left(\|L_{(\cdot, n)}^{11}\|_\nu + \|L_{(\cdot, n)}^{21}\|_\nu, \|L_{(\cdot, n)}^{12}\|_\nu + \|L_{(\cdot, n)}^{22}\|_\nu \right). \end{aligned}$$

We need the Laplacian Δ and its inverse Δ^{-1} on Fourier sequences. The sequences $\Delta \mathbf{v}$ and $\Delta^{-1} \mathbf{v}$ are defined as follow :

$$\begin{aligned} (\Delta \mathbf{v})_n &= \left(- \left(\frac{n\pi}{b-a} \right)^2 \mathbf{v}_n \right)_n, \quad n \geq 0, \\ (\Delta^{-1} \mathbf{v})_n &= \begin{cases} 0, & n = 0, \\ - \left(\frac{b-a}{n\pi} \right)^2 \mathbf{v}_n, & n \geq 1. \end{cases} \end{aligned}$$

Finally, we will slightly abuse notation and denote by $\gamma(\mathbf{v})$ the sequence of Fourier's coefficients of the function $\gamma \circ v$. In other words, $\gamma(\mathbf{v}) = \mathcal{F}(\gamma \circ \mathcal{F}^{-1}(\mathbf{v}))$. We insist on this way of looking at γ to show that we want to express $\gamma(\mathbf{v})$ in terms of \mathbf{v} . This is the whole purpose of Section 3.

Remark 2.2. *We do not need to assume a priori that $\gamma(\mathbf{v})$ is well defined for an arbitrary element \mathbf{v} in ℓ_ν^1 , but we will obtain as a byproduct of our proof that $\gamma(\mathbf{v})$ is indeed well defined, and belongs to ℓ_ν^1 , for all \mathbf{v} in a neighborhood of the approximate solutions that have been validated.*

2.2 From the PDE to the zero-finding problem

We use our notations and operations to rewrite the steady state problem associated to system (1.1), namely

$$\begin{cases} \Delta(\gamma(v)u) + \sigma u(1-u) = 0, & \text{in } (a, b), \\ d\Delta v + u - v = 0, & \text{in } (a, b), \\ \partial_n u = 0, \quad \partial_n v = 0, & \text{on } \{a, b\}. \end{cases} \quad (2.1)$$

Thanks to the Fourier's transform, we can write directly the system (2.1) in Fourier space:

$$\begin{cases} \Delta(\gamma(\mathbf{v}) * \mathbf{u}) + \sigma \mathbf{u} * (\mathbf{1} - \mathbf{u}) = \mathbf{0}, \\ d\Delta \mathbf{v} + \mathbf{u} - \mathbf{v} = \mathbf{0}. \end{cases} \quad (2.2)$$

Then, let us denote

$$\mathbf{U} = \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix}, \quad F(\mathbf{U}) = \begin{pmatrix} \Delta & 0 \\ 0 & \Delta \end{pmatrix} \Phi(\mathbf{U}) + R(\mathbf{U}),$$

with

$$\Phi \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} = \begin{pmatrix} \gamma(\mathbf{v}) * \mathbf{u} \\ d\mathbf{v} \end{pmatrix}, \quad R \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} = \begin{pmatrix} \sigma(\mathbf{u} - \mathbf{u} * \mathbf{u}) \\ \mathbf{u} - \mathbf{v} \end{pmatrix}.$$

Thus, the system (2.2) can be written as

$$F(\mathbf{U}) = \mathbf{0}. \quad (2.3)$$

So, with this formalism and with $\nu > 1$, a ℓ_ν^1 -solution $\mathbf{U} = (\mathbf{u}, \mathbf{v})$ of (2.3) gives a C^∞ -solution of (2.1).

2.3 Newton-Kantorovitch Theorem

We have to solve the system $F(\mathbf{U}) = \mathbf{0}$. In order to prove the existence of a zero F , we use a fixed point method based on the Newton-Kantorovitch Theorem, for which we follow [4]. There are other versions and applications of the Newton-Kantorovitch Theorem, see for instance [3, 8, 9, 34, 41, 46].

Let $N \in \mathbb{N}$ and $\bar{\mathbf{U}} = (\bar{\mathbf{u}}, \bar{\mathbf{v}}) \in (\ell_\nu^1)^2$ be an approximation of a stationary solution, this is our starting point for the fixed point method. Practically, $\bar{\mathbf{U}}$ is a finite sequence of Fourier coefficients obtained numerically. Let A be an approximate inverse of $DF(\bar{\mathbf{U}})$, the Fréchet derivative of F at $\bar{\mathbf{U}}$. It will be explicitly built and discussed in Section 4.

Theorem 2.1 (Newton-Kantorovich Theorem). (See [4, Theorem 2.5]). *With the notations introduced in this section, $r^* > 0$ and $\nu > 1$, assume there exist constants Y , Z_1 , and Z_2 satisfying :*

$$\|AF(\bar{\mathbf{U}})\|_\nu \leq Y, \quad (2.4a)$$

$$\|I - ADF(\bar{\mathbf{U}})\|_\nu \leq Z_1, \quad (2.4b)$$

$$\|AD^2F(\mathbf{U})\|_\nu \leq Z_2, \quad \forall \mathbf{U} \in \mathcal{B}_\nu(\bar{\mathbf{U}}, r^*), \quad (2.4c)$$

and

$$Z_1 < 1, \tag{2.5a}$$

$$2YZ_2 < (1 - Z_1)^2. \tag{2.5b}$$

Then, for any r satisfying

$$\frac{1 - Z_1 - \sqrt{(1 - Z_1)^2 - 2YZ_2}}{Z_2} \leq r < \min(r^*, \frac{1 - Z_1}{Z_2}), \tag{2.6}$$

there exists a unique fixed-point \mathbf{U}^* of $\mathbf{U} \mapsto I - A\mathbf{F}(\mathbf{U})$ in $\mathcal{B}_\nu(\bar{\mathbf{U}}, r)$ the closed ball of radius r centered at $\bar{\mathbf{U}}$ in $\ell_\nu^1 \times \ell_\nu^1$. If moreover, A is injective, then \mathbf{U}^* is a solution of $F(\mathbf{U}) = 0$. The functions (u^*, v^*) , described by the vector of Fourier coefficients $(\mathbf{u}^*, \mathbf{v}^*)$, are then smooth solutions of (2.1).

Thus, \mathbf{U}^* is a theoretical solution of our problem, and we have a numerical description $\bar{\mathbf{U}}$ with a known margin of error. The proof of this theorem, for the specific choice of A described in Section 4, can be found in [4].

Following the statement of Theorem 2.1, our work in the next sections will be to exhibit bounds Y , Z_1 and Z_2 . The expression of each bound must be computable, and accurate enough to satisfy the conditions (2.5) of Theorem 2.1. In Section 3, we describe the tools and initial results needed to that end. In Section 4, we then derive each of the bounds satisfying conditions (2.4).

3 Approximation tools and technical lemmas

We now have to work out how we fit into the assumptions and statements of Theorem 2.1. In this section, we introduce the notations and the concepts for matching theoretical and numerical objects. Then, to verify the hypotheses of the Newton-Kantorovitch Theorem 2.1, we need to explore those ideas and compute some preliminary results. In particular, we have to control the terms linked to γ and show how γ is seen in ℓ_ν^1 .

3.1 Notations for the computations of approximations

The following notations are motivated by the finiteness of the computer, which can only handle finite sequences. Let $N \in \mathbb{N}$, we denote by $\ell_{\nu, N}^1$ the subspace that contains finite sequences of size N . More precisely, $\mathbf{u} \in \ell_\nu^1$ belongs to $\ell_{\nu, N}^1$ if and only if $\mathbf{u}_n = 0$ for all $n > N$. For any \mathbf{u} in ℓ_ν^1 , we denote by $\mathbf{u}|_N$ its truncation in $\ell_{\nu, N}^1$, that is,

$$(\mathbf{u}|_N)_n = \begin{cases} \mathbf{u}_n & n \leq N, \\ 0 & n > N. \end{cases}$$

Since $\ell_{\nu,N}^1$ is not stable by convolution, we will have to adapt N when we need to remove truncation errors, using that

$$\mathbf{u} \in \ell_{\nu,N_1}^1, \mathbf{v} \in \ell_{\nu,N_2}^1 \implies \mathbf{u} * \mathbf{v} \in \ell_{\nu,N_1+N_2}^1.$$

Remark 3.1. *We will keep in mind that $\ell_{\nu,N}^1$ is isomorphic to \mathbb{R}^{N+1} , indeed we start counting from 0 in the sequences.*

In practice, while it is fine to use $\ell_{\nu,N}^1$ spaces for constructing approximate solutions, when conducting the proof we will have to rigorously control quantities like $\gamma(\mathbf{v})$, which, for typical γ 's of the form (1.3), will no belong to any $\ell_{\nu,N}^1$ space even if \mathbf{v} does. In principle, the truncation operator introduced above provides a natural way to approximate an arbitrary element \mathbf{v} of ℓ_{ν}^1 by one which can be stored and manipulated on the computer, but in practice it requires exact knowledge of the first coefficients $\gamma(\mathbf{v})_n$ for $n \leq N$, and an explicit error bound for the truncation, which are not always easy to obtain. Therefore, we will introduce different ways of practically approximating elements of ℓ_{ν}^1 , and specifically those of the form $\gamma(\mathbf{v})$, for a class of functions γ , and assuming \mathbf{v} is an element of ℓ_{ν}^1 on which we have some control. In the sequel, whenever \mathbf{v} is an element of ℓ_{ν}^1 , $\underline{\mathbf{v}}$ denotes an element of $\ell_{\nu,N}^1$ (for some N which we will not always specify), which is meant to approximate \mathbf{v} . We also denote by $\mathbf{e}_{\mathbf{v}}$ the corresponding error, namely $\mathbf{e}_{\mathbf{v}} = \mathbf{v} - \underline{\mathbf{v}}$, and by $\epsilon_{\mathbf{v}}$ an explicit error bound, i.e. a computable number such that $\|\mathbf{e}_{\mathbf{v}}\|_{\nu} \leq \epsilon_{\mathbf{v}}$. Practically, we adapt our choice of $\underline{\mathbf{v}}$ to the problem at hand, as described in the remainder of this section.

Remark 3.2. *In practice, depending on the procedure we use to obtain an approximation $\underline{\mathbf{v}}$ and an error bound $\epsilon_{\mathbf{v}}$, the N for which $\underline{\mathbf{v}}$ belongs to $\ell_{\nu,N}^1$ might be quite large. This can be detrimental if we are to re-use $\underline{\mathbf{v}}$ in other computations. However, we can easily replace $\underline{\mathbf{v}}$ by another approximation having a truncation level $N' < N$:*

$$\mathbf{v} = \underline{\mathbf{v}}|_{N'} + \underline{\mathbf{v}} - \underline{\mathbf{v}}|_{N'} + \mathbf{e}_{\mathbf{v}},$$

where $\underline{\mathbf{v}}|_{N'}$ becomes our new approximation, and the new error can be bounded by

$$\underbrace{\|\underline{\mathbf{v}} - \underline{\mathbf{v}}|_{N'}\|_{\nu}}_{\text{computable}} + \epsilon_{\mathbf{v}}.$$

We extend the notion of truncation to operators on $\mathcal{L}(\ell_{\nu}^1)$. Given an operator L in $\mathcal{L}(\ell_{\nu}^1)$, we denote by $L|_N$ its restriction to $\ell_{\nu,N}^1$, both in terms of domain and co-domain. That is, for any \mathbf{v} in $\ell_{\nu,N}^1$, $(L|_N)\mathbf{v} = (L\mathbf{v})|_N$.

The restriction to N extends to the operators of $\mathcal{L}(\ell_{\nu}^1 \times \ell_{\nu}^1)$ in a natural way, i.e.

$$L|_N = \left(\begin{array}{c|c} L^{11}|_N & L^{12}|_N \\ \hline L^{21}|_N & L^{22}|_N \end{array} \right).$$

3.2 How to compute in the Banach algebra with errors ?

Let us now focus on the main difficulty of this paper, namely controlling $\gamma(\mathbf{v})$ for functions γ of the form: $\gamma(x) = \frac{f(x)}{g(x)}$, $x \in \Omega \subset \mathbb{R}$. We assume that f and g are analytic, but g may vanish. We propose some methods to determine an approximate element $\underline{\gamma(\mathbf{v})}$ of $\gamma(\mathbf{v})$, and a corresponding error bound $\epsilon_{\gamma(\mathbf{v})}$, for any \mathbf{v} belonging to ℓ_ν^1 with a finite approximation $\underline{\mathbf{v}} \in \ell_{\nu,N}^1$ and an associated error bound $\epsilon_{\mathbf{v}}$. We will use the propositions below in two special ways in the context of Theorem 2.1, where we choose $\underline{\mathbf{v}} = \bar{\mathbf{v}}$, a numerical solution of (1.1):

- Assuming $\epsilon_{\mathbf{v}} = 0$, we obtain a description of an approximate element $\underline{\gamma(\underline{\mathbf{v}})}$ and an error $\epsilon_{\gamma(\underline{\mathbf{v}})}$ due to the operations we apply.
- Assuming $\epsilon_{\mathbf{v}} = r^*$ (from Theorem 2.1), we control the element $\gamma(\mathbf{v})$ described in ℓ_ν^1 by a finite element $\underline{\gamma(\underline{\mathbf{v}})}$ and an error $\epsilon_{\gamma(\mathbf{v})}$ that is provided by γ .

All this will be useful to determine the bounds Y , Z_1 and Z_2 in the hypotheses (2.4) of Theorem 2.1. We will give some examples of our results on entire functions, rational fractions, and compositions thereof.

3.2.1 Division and product

We introduce those two operations in our Banach algebra ℓ_ν^1 . We specify the bounds in each case and how we can control the errors of the result of the operation.

Lemma 3.1. *Let $\mathbf{z} \in \ell_\nu^1$ and $\mathbf{a} \in \ell_\nu^1$ such that $\|\mathbf{a} * \mathbf{z} - \mathbf{1}\|_\nu < 1$. Then, \mathbf{z} is invertible and we have*

$$\mathbf{z}^{-1} = \mathbf{a} * \sum_{j=0}^{+\infty} (-1)^j (\mathbf{a} * \mathbf{z} - \mathbf{1})^{*j},$$

$$\|\mathbf{z}^{-1} - \mathbf{a}\|_\nu \leq \|\mathbf{a}\|_\nu \frac{\|\mathbf{a} * \mathbf{z} - \mathbf{1}\|_\nu}{1 - \|\mathbf{a} * \mathbf{z} - \mathbf{1}\|_\nu}.$$

Proof. The existence of the inverse is guaranteed by the criterion of the Neumann series. Writing down the difference and applying the triangle inequality gives the result above. ■

Remark 3.3. *In practice, \mathbf{a} should of course be taken as a numerically computed approximation of \mathbf{z}^{-1} and will therefore be finite. If \mathbf{z} is itself finite, then $\|\mathbf{a} * \mathbf{z} - \mathbf{1}\|_\nu$ amounts to a finite computation and can therefore be evaluated exactly (up to rounding errors, which are taken care of using interval arithmetic [26, 39]). If not, as long as we have a decomposition $\mathbf{z} = \underline{\mathbf{z}} + \mathbf{e}_z$ with a computable error bound ϵ_z for \mathbf{e}_z , we can use that*

$$\|\mathbf{a} * \mathbf{z} - \mathbf{1}\|_\nu \leq \|\mathbf{a} * \underline{\mathbf{z}} - \mathbf{1}\|_\nu + \|\mathbf{a}\|_\nu \epsilon_z,$$

the right-hand side now being computable.

Corollary 3.1. *Let us take any $\mathbf{z} \in \ell_\nu^1$, let $N \in \mathbb{N}$ and $\underline{\mathbf{z}} \in \ell_{\nu,N}^1$ an approximation of \mathbf{z} with error bound $\epsilon_{\mathbf{z}} \geq 0$. Let $\mathbf{a} \in \ell_{\nu,N}^1$ such that $\|\mathbf{a} * \underline{\mathbf{z}} - \mathbf{1}\|_\nu + \|\mathbf{a}\|_\nu \epsilon_{\mathbf{z}} < 1$.*

Then, \mathbf{z} is invertible in ℓ_ν^1 and we can take

$$\underline{\mathbf{z}}^{-1} = \mathbf{a}$$

$$\epsilon_{\mathbf{z}^{-1}} = \|\mathbf{a}\|_\nu \frac{\|\mathbf{a} * \underline{\mathbf{z}} - \mathbf{1}\|_\nu + \|\mathbf{a}\|_\nu \epsilon_{\mathbf{z}}}{1 - \|\mathbf{a} * \underline{\mathbf{z}} - \mathbf{1}\|_\nu - \|\mathbf{a}\|_\nu \epsilon_{\mathbf{z}}}.$$

Proof. The function $x \mapsto \frac{x}{1-x}$ is increasing and we have $\|\mathbf{a} * \mathbf{z} - \mathbf{1}\|_\nu \leq \|\mathbf{a} * \underline{\mathbf{z}} - \mathbf{1}\|_\nu + \|\mathbf{a}\|_\nu \epsilon_{\mathbf{z}}$. Hence, we have the expected result. ■

Remark 3.4. *Lemma 3.1 shows how to compute naturally in the Banach algebra. Corollary 3.1 shows how we get an explicit expression of an error bound depending on the known errors. Even if the Corollary contains the Lemma, the two results express two ways of thinking.*

Lemma 3.2. *Let $\mathbf{x}, \mathbf{y} \in \ell_\nu^1$, $N \in \mathbb{N}$. Let $\underline{\mathbf{x}}, \underline{\mathbf{y}} \in \ell_{\nu,N}^1$ be approximations and $\epsilon_{\mathbf{x}}, \epsilon_{\mathbf{y}} \geq 0$ be error bounds of \mathbf{x} and \mathbf{y} respectively. We can take*

$$\underline{\mathbf{x} * \mathbf{y}} = \underline{\mathbf{x}} * \underline{\mathbf{y}},$$

$$\epsilon_{\mathbf{x} * \mathbf{y}} = \|\underline{\mathbf{x}}\|_\nu \epsilon_{\mathbf{y}} + \|\underline{\mathbf{y}}\|_\nu \epsilon_{\mathbf{x}} + \epsilon_{\mathbf{x}} \epsilon_{\mathbf{y}}.$$

Combining Corollary 3.1 with the straightforward statement above, we have all the required ingredients to understand Banach algebra calculus with error handling.

Theorem 3.1. *Let $\mathbf{x}, \mathbf{y} \in \ell_\nu^1$. Let $N \in \mathbb{N}$, $\underline{\mathbf{x}}, \underline{\mathbf{y}} \in \ell_{\nu,N}^1$ be approximations and $\epsilon_{\mathbf{x}}, \epsilon_{\mathbf{y}} \geq 0$ be error bounds of \mathbf{x} and \mathbf{y} respectively. Let $\mathbf{a} \in \ell_{\nu,N}^1$ such that $\|\mathbf{a} * \underline{\mathbf{y}} - \mathbf{1}\|_\nu + \|\mathbf{a}\|_\nu \epsilon_{\mathbf{y}} < 1$. We can take*

$$\underline{\mathbf{x} * \mathbf{y}^{-1}} = \underline{\mathbf{x}} * \mathbf{a},$$

$$\epsilon_{\mathbf{x} * \mathbf{y}^{-1}} = \|\mathbf{a}\|_\nu \frac{\|\underline{\mathbf{x}}\|_\nu \left(\|\mathbf{a} * \underline{\mathbf{y}} - \mathbf{1}\|_\nu + \|\mathbf{a}\|_\nu \epsilon_{\mathbf{y}} \right) + \epsilon_{\mathbf{x}}}{1 - \|\mathbf{a} * \underline{\mathbf{y}} - \mathbf{1}\|_\nu - \|\mathbf{a}\|_\nu \epsilon_{\mathbf{y}}}.$$

Proof. Apply Corollary 3.1 to $\mathbf{z} = \mathbf{y}$, $(\underline{\mathbf{y}}^{-1}) = \mathbf{a}$. Then \mathbf{y} is invertible and belongs to ℓ_ν^1 . We have $(\underline{\mathbf{y}}^{-1}) = \mathbf{a}$ and $\epsilon_{\mathbf{y}^{-1}}$ associated.

Then we choose, according to Lemma 3.2

$$\underline{\mathbf{x} * \mathbf{y}^{-1}} = \underline{\mathbf{x}} * \mathbf{a},$$

$$\epsilon_{\mathbf{x} * \mathbf{y}^{-1}} = \|\underline{\mathbf{x}}\|_\nu \epsilon_{\mathbf{y}^{-1}} + \|\mathbf{a}\|_\nu \epsilon_{\mathbf{x}} + \epsilon_{\mathbf{y}^{-1}} \epsilon_{\mathbf{x}}.$$

We simplify and obtain the result. ■

3.2.2 Analytic functions

In the previous subsection, we have seen how to deal with fractions in the Banach algebra ℓ_ν^1 , with rigorous error control, provided we had a description of both the numerator and the denominator in

terms of an approximation and an error bound. In this subsection, we discuss how such approximations and error bounds for the numerator and the denominator can be obtained, provided those are analytic functions.

Let f be an analytic function. We consider $\mathbf{v} \in \ell_\nu^1$ and for $N \in \mathbb{N}$, $\underline{\mathbf{v}} \in \ell_{\nu,N}^1$ an approximation of \mathbf{v} and $\epsilon_{\mathbf{v}} \geq 0$ an associated error. We will first use the analytic description of f to obtain an explicit expression of $\underline{f(\underline{\mathbf{v}})}$ and $\epsilon_{f(\underline{\mathbf{v}})}$. Secondly, we will use the analytic properties of f to get an explicit bound $\epsilon_{f(\mathbf{v})}$ on the norm of $\mathbf{e}_{f(\mathbf{v})} = f(\mathbf{v}) - \underline{f(\underline{\mathbf{v}})}$. This means that we choose, as an approximation of $f(\mathbf{v})$, a nearby element $\underline{f(\underline{\mathbf{v}})}$. It belongs to some $\ell_{\nu,N'}^1$, $N' \in \mathbb{N}$, with a computable description.

Since we work in a (commutative) Banach algebra, we can still make use of Taylor expansions (see e.g. [27, Chapter 12]), and the Fourier transform carries all the calculations from \mathbb{R} to ℓ_ν^1 in a straightforward way.

Theorem 3.2. *Let $K \in \mathbb{N} \setminus \{0\}$ and f a function of class $\mathcal{C}^K(\mathbb{R}; \mathbb{R})$. Then, the function defined by*

$$f : \begin{cases} \ell_\nu^1 & \longrightarrow \ell_\nu^1 \\ \mathbf{v} & \longmapsto \mathcal{F}(f \circ \mathcal{F}^{-1}(\mathbf{v})) := f(\mathbf{v}) \end{cases} \text{ is } K\text{-differentiable.}$$

We have the following Taylor expansion. Let $\mathbf{v}_0 \in \ell_\nu^1$,

$$\forall \mathbf{v} \in \ell_\nu^1, \left\| f(\mathbf{v}) - \sum_{k=0}^{K-1} \frac{1}{k!} f^{(k)}(\mathbf{v}_0) * (\mathbf{v} - \mathbf{v}_0)^{*k} \right\|_\nu \leq \frac{\|\mathbf{v} - \mathbf{v}_0\|_\nu^K}{K!} \sup_{z \in [\mathbf{v}_0, \mathbf{v}]} \|f^{(K)}(z)\|_\nu.$$

We can apply this result to any entire function and go even further.

Proposition 3.1. *Let f be an entire function, written $f(x) = \sum_{k=0}^{+\infty} a_k x^k$. Moreover, we denote $|f'|$*

the entire function defined as $|f'|(x) = \sum_{k=0}^{+\infty} k|a_k|x^{k-1}$. Let $\mathbf{v} \in \ell_\nu^1$ and $N \in \mathbb{N}$. Let $\underline{\mathbf{v}} \in \ell_{\nu,N}^1$ an approximation and $\epsilon_{\mathbf{v}} \geq 0$ an error bound. Taking

$$\underline{f(\underline{\mathbf{v}})} = \sum_{k=0}^{K-1} a_k \underline{\mathbf{v}}^{*k},$$

we get the following error bound

$$\epsilon_{f(\underline{\mathbf{v}})} = \frac{\|\underline{\mathbf{v}}\|_\nu^K}{K!} \sup_{z \in [\underline{\mathbf{0}}, \underline{\mathbf{v}}]} \|f^{(K)}(z)\|_\nu.$$

Furthermore, we can take

$$\begin{aligned} \underline{f(\underline{\mathbf{v}})} &= \underline{f(\underline{\mathbf{v}})} \\ \epsilon_{f(\mathbf{v})} &= \epsilon_{f(\underline{\mathbf{v}})} + |f'|(\|\underline{\mathbf{v}}\|_\nu + \epsilon_{\mathbf{v}})\epsilon_{\mathbf{v}}. \end{aligned}$$

Proof. We just write $f(\mathbf{v}) = \underline{f(\underline{\mathbf{v}})} + \mathbf{e}_{f(\underline{\mathbf{v}})} + f(\mathbf{v}) - \underline{f(\underline{\mathbf{v}})}$ and apply the triangular inequality. We have $\|\mathbf{e}_{f(\underline{\mathbf{v}})}\|_\nu \leq \epsilon_{f(\underline{\mathbf{v}})}$ thanks to Theorem 3.2. The last thing to bound is $\|f(\mathbf{v}) - \underline{f(\underline{\mathbf{v}})}\|_\nu$. To do this we

apply Theorem 3.2 with $K = 1$, $\mathbf{v}_0 = \underline{\mathbf{v}}$. We have

$$\begin{aligned}
\|f(\mathbf{v}) - f(\underline{\mathbf{v}})\|_\nu &\leq \sup_{\mathbf{z} \in [\underline{\mathbf{v}}, \mathbf{v}]} \|f'(\mathbf{z})\|_\nu \|\mathbf{v} - \underline{\mathbf{v}}\|_\nu \\
&\leq \sup_{\mathbf{z} \in \mathcal{B}_\nu(\underline{\mathbf{v}}, \epsilon_\mathbf{v})} \|f'(\mathbf{z})\|_\nu \epsilon_\mathbf{v} \\
&\leq \sup_{\mathbf{z} \in \mathcal{B}_\nu(\underline{\mathbf{v}}, \epsilon_\mathbf{v})} \left\| \sum_{k=0}^{+\infty} k a_k \mathbf{z}^{*k-1} \right\|_\nu \epsilon_\mathbf{v} \\
&\leq \sup_{\mathbf{z} \in \mathcal{B}_\nu(\underline{\mathbf{v}}, \epsilon_\mathbf{v})} \sum_{k=0}^{+\infty} k |a_k| \|\mathbf{z}\|_\nu^{k-1} \epsilon_\mathbf{v},
\end{aligned}$$

but $\|\mathbf{z}\|_\nu \leq \|\underline{\mathbf{v}}\|_\nu + \epsilon_\mathbf{v}$ for any $\mathbf{z} \in \mathcal{B}_\nu(\underline{\mathbf{v}}, \epsilon_\mathbf{v})$ then,

$$\begin{aligned}
\|f(\mathbf{v}) - f(\underline{\mathbf{v}})\|_\nu &\leq \sum_{k=0}^{+\infty} k |a_k| (\|\underline{\mathbf{v}}\|_\nu + \epsilon_\mathbf{v})^{k-1} \epsilon_\mathbf{v} \\
&\leq |f'|(\|\underline{\mathbf{v}}\|_\nu + \epsilon_\mathbf{v}) \epsilon_\mathbf{v}.
\end{aligned}$$

Hence, the result holds. ■

Remark 3.5. *In practice, we need to be able to evaluate $\epsilon_{f(\underline{\mathbf{v}})}$ and $\epsilon_{f(\mathbf{v})}$ (or at least to get computable upper-bounds for them). This must be done on a case by case basis, depending on the function f and the available information we have on its derivatives (or on the asymptotic behavior of the coefficients a_k), and might prove difficult for some analytic functions. However, among the functions γ that are usually considered for (1.1), the typical entire nonlinearities are based on the exponential function, for which these computations can be carried out. See Section 3.2.3 for some explicit examples.*

Remark 3.6. *We stated Proposition 3.1 for entire functions for the sake of convenience, but the result generalizes in a straightforward way to functions f which are merely analytic, provided the disk of center 0 and radius $\|\underline{\mathbf{v}}\|_\nu + \epsilon_\mathbf{v}$ is contained in the domain of analyticity of f . A slightly more natural and less restrictive assumption would be to require that the range of $\underline{\mathbf{v}}$ is contained in the domain of analyticity of f , and at a distance at least $\epsilon_\mathbf{v}$ of its boundary. This goes beyond the scope of this work, but will be the subject of further studies.*

Remark 3.7. *When f is a polynomial $P = \sum_{k=0}^{\deg P} a_k X^k$ we have the following result as a special case of Proposition 3.1:*

$$\begin{aligned}
\underline{P(\underline{\mathbf{v}})} &= \sum_{k=0}^{\deg P} a_k \underline{\mathbf{v}}^{*k}, \\
\epsilon_{P(\underline{\mathbf{v}})} &= 0, \\
\underline{P(\mathbf{v})} &= \underline{P(\underline{\mathbf{v}})}, \\
\epsilon_{P(\mathbf{v})} &= |P'|(\|\underline{\mathbf{v}}\|_\nu + \epsilon_\mathbf{v}) \epsilon_\mathbf{v}.
\end{aligned}$$

3.2.3 Examples

In this whole section, we again take $\mathbf{v} \in \ell_\nu^1$, $N \in \mathbb{N}$, $\underline{\mathbf{v}} \in \ell_{\nu, N}^1$ an approximation and $\epsilon_{\mathbf{v}} \geq 0$ an error bound. We will show the choices we made to get a description in ℓ_ν^1 of the functions γ involved in Theorem 1.1, Theorem 1.2 and Theorem 1.3.

The exponential: Let us consider $f = e^{\alpha \cdot}$, for some $\alpha \in \mathbb{R}$. We can apply Proposition 3.1 and use the fact that $\sup_{\mathbf{z} \in [0, \underline{\mathbf{v}}]} \|f^{(K)}(\mathbf{z})\|_\nu \leq |\alpha|^K e^{|\alpha| \|\underline{\mathbf{v}}\|_\nu}$ to get

$$\begin{aligned} \underline{f}(\underline{\mathbf{v}}) &= \sum_{k=0}^{K-1} \frac{\alpha^k}{k!} \underline{\mathbf{v}}^{*k}, \\ \epsilon_{f(\underline{\mathbf{v}})} &= \frac{|\alpha|^K \|\underline{\mathbf{v}}\|_\nu^K}{K!} e^{|\alpha| \|\underline{\mathbf{v}}\|_\nu}, \end{aligned}$$

and

$$\begin{aligned} \underline{f}(\mathbf{v}) &= \underline{f}(\underline{\mathbf{v}}), \\ \epsilon_{f(\mathbf{v})} &= \epsilon_{f(\underline{\mathbf{v}})} + |\alpha| e^{|\alpha| (\|\underline{\mathbf{v}}\|_\nu + \epsilon_{\mathbf{v}})} \epsilon_{\mathbf{v}}, \end{aligned}$$

where the error bounds are now computable.

Remark 3.8. *Using some specific properties of the exponential function, we could also get a sharper error bound for $\|f(\mathbf{v}) - \underline{f}(\underline{\mathbf{v}})\|_\nu$. Indeed, when in the proof of Proposition 3.1 we need to estimate $\sup_{\mathbf{z} \in \mathcal{B}_\nu(\underline{\mathbf{v}}, \epsilon_{\mathbf{v}})} \|f'(\mathbf{z})\|_\nu$, we do have*

$$\sup_{\mathbf{v} \in \mathcal{B}_\nu(\underline{\mathbf{v}}, \epsilon_{\mathbf{v}})} \|f'(\mathbf{v})\|_\nu \leq \left(\|\underline{f}'(\underline{\mathbf{v}})\|_\nu + \epsilon_{f'(\underline{\mathbf{v}})} \right) e^{|\alpha| \epsilon_{\mathbf{v}}},$$

which could be significantly smaller than $|\alpha| e^{|\alpha| (\|\underline{\mathbf{v}}\|_\nu + \epsilon_{\mathbf{v}})} \epsilon_{\mathbf{v}}$.

A fraction involving an exponential: In the literature [44], the function $\gamma(v) = \frac{1}{1 + e^{9(v-1)}}$ is considered in system (1.1). So let us manage it with our new tools. We first have to consider $g = 1 + e^{9(\cdot - 1)}$, which is not too far from the first example. We can choose,

$$\begin{aligned} \underline{g}(\underline{\mathbf{v}}) &= \mathbf{1} + \sum_{k=0}^{K-1} \frac{9^k}{k!} (\underline{\mathbf{v}} - \mathbf{1})^k, \\ \epsilon_{g(\underline{\mathbf{v}})} &= \frac{9^K \|\underline{\mathbf{v}} - \mathbf{1}\|_\nu^K}{K!} e^{9 \|\underline{\mathbf{v}} - \mathbf{1}\|_\nu}, \end{aligned}$$

and

$$\begin{aligned} \underline{g}(\mathbf{v}) &= \underline{g}(\underline{\mathbf{v}}), \\ \epsilon_{g(\mathbf{v})} &= \epsilon_{g(\underline{\mathbf{v}})} + 9 e^{9 (\|\underline{\mathbf{v}} - \mathbf{1}\|_\nu + \epsilon_{\mathbf{v}})} \epsilon_{\mathbf{v}}. \end{aligned}$$

We point out that Remark 3.8 also applies here to get a better bound $\epsilon_{g(\mathbf{v})}$.

We can then apply Corollary 3.1, with $\mathbf{a} \in \ell_{\nu, N}^1$ such that $\|\mathbf{a} * \underline{g(\underline{\mathbf{v}})} - \mathbf{1}\|_{\nu} + \|\mathbf{a}\|_{\nu} \epsilon_{g(\underline{\mathbf{v}})} < 1$, both to $\underline{g(\underline{\mathbf{v}})}$ and to $g(\underline{\mathbf{v}})$, which yields,

$$\begin{aligned} \underline{g(\underline{\mathbf{v}})}^{-1} &= \mathbf{a}, \\ \epsilon_{g(\underline{\mathbf{v}})}^{-1} &= \|\mathbf{a}\|_{\nu} \frac{\|\mathbf{a} * \underline{g(\underline{\mathbf{v}})} - \mathbf{1}\|_{\nu} + \|\mathbf{a}\|_{\nu} \epsilon_{g(\underline{\mathbf{v}})}}{1 - \|\mathbf{a} * \underline{g(\underline{\mathbf{v}})} - \mathbf{1}\|_{\nu} - \|\mathbf{a}\|_{\nu} \epsilon_{g(\underline{\mathbf{v}})}}, \end{aligned}$$

and

$$\begin{aligned} \underline{g(\underline{\mathbf{v}})}^{-1} &= \underline{g(\underline{\mathbf{v}})}^{-1}, \\ \epsilon_{g(\underline{\mathbf{v}})}^{-1} &= \|\mathbf{a}\|_{\nu} \frac{\|\mathbf{a} * \underline{g(\underline{\mathbf{v}})} - \mathbf{1}\|_{\nu} + \|\mathbf{a}\|_{\nu} (\epsilon_{g(\underline{\mathbf{v}})} + 9e^{9(\|\underline{\mathbf{v}} - \mathbf{1}\|_{\nu} + \epsilon_{\mathbf{v}})} \epsilon_{\mathbf{v}})}}{1 - \|\mathbf{a} * \underline{g(\underline{\mathbf{v}})} - \mathbf{1}\|_{\nu} - \|\mathbf{a}\|_{\nu} (\epsilon_{g(\underline{\mathbf{v}})} + 9e^{9(\|\underline{\mathbf{v}} - \mathbf{1}\|_{\nu} + \epsilon_{\mathbf{v}})} \epsilon_{\mathbf{v}})}}. \end{aligned}$$

A rational fraction: For a rational fraction $\gamma = \frac{P}{Q}$, we apply Remark 3.7 and Theorem 3.1, with $\mathbf{a} \in \ell_{\nu, N}^1$ such that $\|\mathbf{a} * Q(\underline{\mathbf{v}}) - \mathbf{1}\|_{\nu} + \|\mathbf{a}\|_{\nu} |Q'|(\|\underline{\mathbf{v}}\|_{\nu} + \epsilon_{\mathbf{v}}) \epsilon_{\mathbf{v}} < 1$, which yields

$$\begin{aligned} \underline{\gamma(\underline{\mathbf{v}})} &= P(\underline{\mathbf{v}}) * \mathbf{a}, \\ \epsilon_{\gamma(\underline{\mathbf{v}})} &= \|P(\underline{\mathbf{v}})\|_{\nu} \|\mathbf{a}\|_{\nu} \frac{\|\mathbf{a} * Q(\underline{\mathbf{v}}) - \mathbf{1}\|_{\nu}}{1 - \|\mathbf{a} * Q(\underline{\mathbf{v}}) - \mathbf{1}\|_{\nu}}, \end{aligned}$$

and

$$\begin{aligned} \underline{\gamma(\underline{\mathbf{v}})} &= \underline{\gamma(\underline{\mathbf{v}})}, \\ \epsilon_{\gamma(\underline{\mathbf{v}})} &= \|\mathbf{a}\|_{\nu} \frac{\|P(\underline{\mathbf{v}})\|_{\nu} (\|\mathbf{a} * Q(\underline{\mathbf{v}}) - \mathbf{1}\|_{\nu} + \|\mathbf{a}\|_{\nu} |Q'|(\|\underline{\mathbf{v}}\|_{\nu} + \epsilon_{\mathbf{v}}) \epsilon_{\mathbf{v}}) + |P'|(\|\underline{\mathbf{v}}\|_{\nu} + \epsilon_{\mathbf{v}}) \epsilon_{\mathbf{v}}}{1 - \|\mathbf{a} * Q(\underline{\mathbf{v}}) - \mathbf{1}\|_{\nu} - \|\mathbf{a}\|_{\nu} |Q'|(\|\underline{\mathbf{v}}\|_{\nu} + \epsilon_{\mathbf{v}}) \epsilon_{\mathbf{v}}}. \end{aligned}$$

Remark 3.9. *A rational fraction is an analytic function. Therefore, we might deal with it as discussed in Remark 3.6. However, even in situations where this would prove feasible, we expect the approximation and estimates obtained via Theorem 3.1 to be both sharper and cheaper to compute.*

4 Derivation of the bounds

In this section, we first define the operator A that we are going to use in Theorem 2.1, and then derive bounds Y , Z_1 and Z_2 satisfying assumptions (2.4a), (2.4b) and (2.4c). These estimates rely on the controlled approximations of γ introduced in Section 3. Let $\bar{\mathbf{U}} = (\bar{\mathbf{u}}, \bar{\mathbf{v}}) \in (\ell_{\nu, N}^1)^2$, an approximate solution of (2.3). Thanks to Section 3 with $\underline{\mathbf{v}} = \bar{\mathbf{v}}$, we can calculate $\underline{\gamma(\bar{\mathbf{v}})}$, $\epsilon_{\gamma(\bar{\mathbf{v}})}$, $\underline{\gamma'(\bar{\mathbf{v}})}$, $\epsilon_{\gamma'(\bar{\mathbf{v}})}$, $\underline{\gamma''(\bar{\mathbf{v}})}$ and $\epsilon_{\gamma''(\bar{\mathbf{v}})}$, at least for the functions γ used in Theorem 1.1, Theorem 1.2 and Theorem 1.3.

For any $\mathbf{U} = (\mathbf{u}, \mathbf{v})$ in ℓ_{ν}^1 , we recall that

$$F(\mathbf{U}) = \begin{pmatrix} \Delta(\gamma(\mathbf{v}) * \mathbf{u}) + \sigma \mathbf{u} * (\mathbf{1} - \mathbf{u}) \\ d\Delta \mathbf{v} + \mathbf{u} - \mathbf{v} \end{pmatrix},$$

so that

$$DF(\mathbf{U}) = \begin{pmatrix} \Delta\gamma(\mathbf{v}) + \sigma(\mathbf{1} - 2\mathbf{u}) & \Delta(\gamma'(\mathbf{v}) * \mathbf{u}) \\ \mathbf{1} & d\Delta - \mathbf{1} \end{pmatrix}.$$

For A , we consider a well chosen approximation of the inverse of $DF(\bar{\mathbf{U}})$, combining a numerically approximated inverse of a truncation of $DF(\bar{\mathbf{U}})$ and an infinite rest that correspond to the “main” part of the inverse. Following [4], we therefore take

$$A = \left(\begin{array}{c|c} \frac{A_{11}^{11}}{\phantom{A_{11}^{11}}} & \frac{A_{11}^{12}}{\phantom{A_{11}^{12}}} \\ \hline \frac{A_{11}^{21}}{\phantom{A_{11}^{21}}} & \frac{A_{11}^{22}}{\phantom{A_{11}^{22}}} \end{array} \begin{array}{c} M(\mathbf{w}^{11})\Delta^{-1} \\ M(\mathbf{w}^{12})\Delta^{-1} \\ \hline M(\mathbf{w}^{21})\Delta^{-1} \\ M(\mathbf{w}^{22})\Delta^{-1} \end{array} \right) = \left(\frac{A^{11}}{A^{21}} \middle| \frac{A^{12}}{A^{22}} \right),$$

The block-by-block notation is inherited from the two species system. We use sub-blocks to separate the finite part and the tail of the operator. Each A_{11}^{ij} is a linear operator on $\ell_{\nu,2N}^1$, this choice is explained in Remark 4.1, just below.

The blocks A_{11}^{ij} are truncation of an inverted truncation of $DF(\bar{\mathbf{U}})$,

$$\left(\frac{A_{11}^{11}}{A_{11}^{21}} \middle| \frac{A_{11}^{12}}{A_{11}^{22}} \right) \approx [DF(\bar{\mathbf{U}})|_{4N}]^{-1}|_{2N},$$

and \mathbf{w}^{ij} , $i, j \in 1, 2$, are elements of $\ell_{\nu,N}^1$ such that

$$\begin{cases} \mathbf{w}^{11} \approx \gamma(\bar{\mathbf{v}})^{-1}, \\ \mathbf{w}^{12} \approx -\frac{1}{d}\mathbf{w}^{11} * (\gamma'(\bar{\mathbf{v}}) * \bar{\mathbf{u}}), \\ \mathbf{w}^{21} = \mathbf{0}, \\ \mathbf{w}^{22} = \frac{1}{d}\mathbf{1}. \end{cases}$$

The choices made here are motivated by the main part of the operator $\Delta^{-1}DF(\bar{\mathbf{U}})$ that we want to invert. Morally, we invert by hand the matrix $\begin{pmatrix} \gamma(\bar{v}) & \gamma'(\bar{v})\bar{u} \\ 0 & d \end{pmatrix}$ in ℓ_{ν}^1 in order to get the w^{ij} .

For later use, we also introduce

$$\mathbf{W} = \left(\frac{\mathbf{w}^{11}}{\mathbf{w}^{21}} \middle| \frac{\mathbf{w}^{12}}{\mathbf{w}^{22}} \right).$$

Remark 4.1. *In all the following computations, we have $\bar{\mathbf{u}}, \bar{\mathbf{v}}$ belonging to $\ell_{\nu,N}^1$, with $N \in \mathbb{N}$. We want to keep the total information of the convolution of $\bar{\mathbf{u}}$ with itself or with $\underline{\gamma(\bar{\mathbf{v}})}$. Since we fix $\underline{\gamma(\bar{\mathbf{v}})} \in \ell_{\nu,N}^1$, we choose $\underline{F(\bar{\mathbf{U}})} \in \ell_{\nu,2N}^1 \times \ell_{\nu,2N}^1$. Then, to match, we take $\mathbf{w}^{ij} \in \ell_{\nu,2N}^1$ and $A_{11}^{ij} \in \mathcal{L}(\ell_{\nu,2N}^1)$.*

4.1 The bound Y

The bound Y is derived from the calculation of $\|AF(\bar{\mathbf{U}})\|_\nu$. Let us do the math.

$$\begin{aligned}
AF(\bar{\mathbf{U}}) &= A \begin{pmatrix} \Delta(\gamma(\bar{\mathbf{v}}) * \bar{\mathbf{u}}) + \sigma \bar{\mathbf{u}} * (\mathbf{1} - \bar{\mathbf{u}}) \\ d\Delta \bar{\mathbf{v}} + \bar{\mathbf{u}} - \bar{\mathbf{v}} \end{pmatrix} \\
&= A \begin{pmatrix} \Delta((\underline{\gamma}(\bar{\mathbf{v}}) + \mathbf{e}_{\gamma(\bar{\mathbf{v}})}) * \bar{\mathbf{u}}) + \sigma \bar{\mathbf{u}} * (\mathbf{1} - \bar{\mathbf{u}}) \\ d\Delta \bar{\mathbf{v}} + \bar{\mathbf{u}} - \bar{\mathbf{v}} \end{pmatrix} \\
&= \underbrace{A \begin{pmatrix} \Delta(\gamma(\bar{\mathbf{v}}) * \bar{\mathbf{u}}) + \sigma \bar{\mathbf{u}} * (\mathbf{1} - \bar{\mathbf{u}}) \\ d\Delta \bar{\mathbf{v}} + \bar{\mathbf{u}} - \bar{\mathbf{v}} \end{pmatrix}}_{:= \underline{AF(\bar{\mathbf{U}})}, \text{ completely known and computable exactly}} + A \begin{pmatrix} \Delta(\mathbf{e}_{\gamma(\bar{\mathbf{v}})} * \bar{\mathbf{u}}) \\ \mathbf{0} \end{pmatrix}.
\end{aligned}$$

We then focus on the rest,

$$\begin{aligned}
A \begin{pmatrix} \Delta(\mathbf{e}_{\gamma(\bar{\mathbf{v}})} * \bar{\mathbf{u}}) \\ \mathbf{0} \end{pmatrix} &= \begin{pmatrix} A^{11} & A^{12} \\ A^{21} & A^{22} \end{pmatrix} \begin{pmatrix} \Delta(\mathbf{e}_{\gamma(\bar{\mathbf{v}})} * \bar{\mathbf{u}}) \\ \mathbf{0} \end{pmatrix} \\
&= \begin{pmatrix} A^{11} \Delta(\mathbf{e}_{\gamma(\bar{\mathbf{v}})} * \bar{\mathbf{u}}) \\ A^{21} \Delta(\mathbf{e}_{\gamma(\bar{\mathbf{v}})} * \bar{\mathbf{u}}) \end{pmatrix}.
\end{aligned}$$

So,

$$\begin{aligned}
\left\| A \begin{pmatrix} \Delta(\mathbf{e}_{\gamma(\bar{\mathbf{v}})} * \bar{\mathbf{u}}) \\ \mathbf{0} \end{pmatrix} \right\|_\nu &= \|A^{11} \Delta(\mathbf{e}_{\gamma(\bar{\mathbf{v}})} * \bar{\mathbf{u}})\|_\nu + \|A^{21} \Delta(\mathbf{e}_{\gamma(\bar{\mathbf{v}})} * \bar{\mathbf{u}})\|_\nu \\
&\leq (\|A^{11} \Delta\|_\nu + \|A^{21} \Delta\|_\nu) \|\mathbf{e}_{\gamma(\bar{\mathbf{v}})} * \bar{\mathbf{u}}\|_\nu \\
&\leq (\|A^{11} \Delta\|_\nu + \|A^{21} \Delta\|_\nu) \|\mathbf{e}_{\gamma(\bar{\mathbf{v}})}\|_\nu \|\bar{\mathbf{u}}\|_\nu \\
&\leq (\|A^{11} \Delta\|_\nu + \|A^{21} \Delta\|_\nu) \|\bar{\mathbf{u}}\|_\nu \epsilon_{\gamma(\bar{\mathbf{v}})}.
\end{aligned}$$

Thus we take,

$$\boxed{Y = \|\underline{AF(\bar{\mathbf{U}})}\|_\nu + (\|A^{11} \Delta\|_\nu + \|A^{21} \Delta\|_\nu) \|\bar{\mathbf{u}}\|_\nu \epsilon_{\gamma(\bar{\mathbf{v}})}}.$$

It satisfies (2.4a) in Theorem 2.1, $\|AF(\bar{\mathbf{U}})\|_\nu \leq Y$.

Remark 4.2. *The computations made with the approximate vectors are said to be computable exactly. This is because they are finite (belong to a certain $\ell^1_{\nu, N}$ with N not too large) and thanks to interval arithmetic [37] these vectors can be calculated exactly.*

4.2 The bound Z_1

The bound Z_1 is derived to control $\|I - ADF(\bar{\mathbf{U}})\|_\nu$. We use the same idea as before. We divide the terms related to γ between parts we know exactly, and small remainders whose norm we can bound. Furthermore, we have to deal with an infinite dimensional operator, so we can only use the computer

for a finite part, and then have to estimate the “tail” by hand. Firstly, as above, we obtain

$$\|I - ADF(\bar{\mathbf{U}})\|_\nu \leq \|I - \underline{ADF}(\bar{\mathbf{U}})\|_\nu + (\|A^{11}\Delta\|_\nu + \|A^{21}\Delta\|_\nu)(\epsilon_{\gamma(\bar{\mathbf{v}})} + \|\bar{\mathbf{u}}\|_\nu \epsilon_{\gamma'(\bar{\mathbf{v}})}),$$

$$\text{with } \underline{DF}(\bar{\mathbf{U}}) = \begin{pmatrix} \frac{\Delta\gamma(\bar{\mathbf{v}}) + \sigma(\mathbf{1} - 2\bar{\mathbf{u}})}{\mathbf{1}} & \Delta(\gamma'(\bar{\mathbf{v}}) * \bar{\mathbf{u}}) \\ \mathbf{1} & d\Delta - \mathbf{1} \end{pmatrix}.$$

We cannot compute exactly $B = I - \underline{ADF}(\bar{\mathbf{U}})$ because A and $\underline{DF}(\bar{\mathbf{U}})$ are built with an infinite “tail”. However, we can compute exactly a finite number of $B_{i,j}$. It leads us to separate between the columns whose norm we will compute numerically and exactly, and those we will bound by hand. With $K \geq 2N - 1$ to be fixed later:

$$\|B\|_\nu = \sup_{k \geq 0} \frac{1}{\xi_k(\nu)} \max \left(\|B_{(\cdot,k)}^{11}\|_\nu + \|B_{(\cdot,k)}^{21}\|_\nu, \|B_{(\cdot,k)}^{12}\|_\nu + \|B_{(\cdot,k)}^{22}\|_\nu \right)$$

$$= \max \left(\sup_{k \leq K} \frac{1}{\xi_k(\nu)} \max \left(\|B_{(\cdot,k)}^{11}\|_\nu + \|B_{(\cdot,k)}^{21}\|_\nu, \|B_{(\cdot,k)}^{12}\|_\nu + \|B_{(\cdot,k)}^{22}\|_\nu \right), \right.$$

$$\left. \sup_{k \geq K+1} \frac{1}{\xi_k(\nu)} \max \left(\|B_{(\cdot,k)}^{11}\|_\nu + \|B_{(\cdot,k)}^{21}\|_\nu, \|B_{(\cdot,k)}^{12}\|_\nu + \|B_{(\cdot,k)}^{22}\|_\nu \right) \right).$$

Let us denote, the “finite” part to compute,

$$Z_{1,finite} := \sup_{k \leq K} \frac{1}{\xi_k(\nu)} \max \left(\|B_{(\cdot,k)}^{11}\|_\nu + \|B_{(\cdot,k)}^{21}\|_\nu, \|B_{(\cdot,k)}^{12}\|_\nu + \|B_{(\cdot,k)}^{22}\|_\nu \right),$$

and the “tail” to bound,

$$\mathcal{T}_{K,\nu}(B) := \sup_{k \geq K+1} \frac{1}{\xi_k(\nu)} \max \left(\|B_{(\cdot,k)}^{11}\|_\nu + \|B_{(\cdot,k)}^{21}\|_\nu, \|B_{(\cdot,k)}^{12}\|_\nu + \|B_{(\cdot,k)}^{22}\|_\nu \right)$$

$$= \sup_{\substack{(\mathbf{z}, \mathbf{z}') \neq (\mathbf{0}, \mathbf{0}) \\ (\mathbf{z}|_K, \mathbf{z}'|_K) = (\mathbf{0}, \mathbf{0})}} \frac{\|B^{11}\mathbf{z} + B^{12}\mathbf{z}'\|_\nu + \|B^{21}\mathbf{z} + B^{22}\mathbf{z}'\|_\nu}{\|\mathbf{z}\|_\nu + \|\mathbf{z}'\|_\nu}. \quad (4.1)$$

4.2.1 The finite part to compute

Here we give an explanation of the calculation of $Z_{1,finite}$ made by the computer. To compute $Z_{1,finite}$, we have to understand correctly the calculation that determines $B_{(\cdot,k)}^{ij}$.

We have for any $k \leq K$ and $(i, j) \in \{1, 2\}$, $B_{(\cdot,k)}^{ij} = \delta_{ij}I_{(\cdot,k)} - (A^{i1}\underline{DF}(\bar{\mathbf{U}})_{(\cdot,k)}^{1j} + A^{i2}\underline{DF}(\bar{\mathbf{U}})_{(\cdot,k)}^{2j})$.

Let $l \in \{1, 2\}$, let us take a look at the column of $\underline{DF}(\bar{\mathbf{U}})^{lj}$ (see Figure 5 for a global view),

$$\underline{DF}(\bar{\mathbf{U}})_{(\cdot,k)}^{lj} = \begin{cases} (\underbrace{* \dots *}_{2N+k} 0 \dots)^T, & \text{if } k \in \{0, \dots, 2N - 1\}, \\ (\underbrace{0 \dots 0}_{k-2N+1} \underbrace{* \dots *}_{4N-1} 0 \dots)^T, & \text{if } k \in \{2N, \dots, K\}. \end{cases}$$

Thus, to be exact in the matrix-vector product $A^{il}\underline{DF}(\bar{\mathbf{U}})_{(\cdot,k)}^{lj}$, it is sufficient to take the first $k + 2N$ rows of A^{il} . These rows have the same structure as the columns $\underline{DF}(\bar{\mathbf{U}})_{(\cdot,k)}^{lj}$, except the transpose operation. Let $p \in \mathbb{N}$, we have the following equality:

$$\left(A^{il} \underline{DF(\bar{\mathbf{U}})}^{lj} \right)_p = \begin{cases} \sum_{q=0}^{k+2N-1} A_{(p,q)}^{il} \underline{DF(\bar{\mathbf{U}})}_{(q,k)}^{lj}, & \text{if } p \leq k + 4N - 2, \\ 0, & \text{if } p \geq k + 4N - 1, \end{cases}$$

because for all $p \geq k + 4N - 1$, $A_{(p,q)}^{il} = 0$, $\forall q \in \{0, \dots, k + 2N - 1\} \subset \{0, \dots, p - 2N\}$.

The situation is summarized on Figure 5. Remember we start counting at 0 for the columns and the rows.

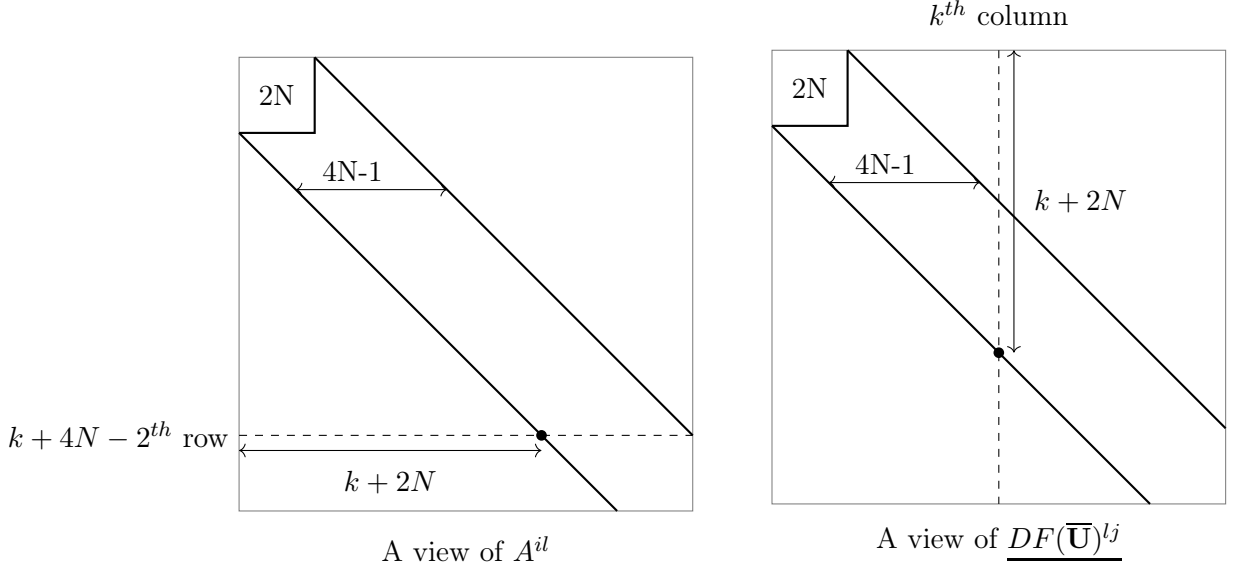


Figure 5: Figure of the matrices

Then, because we are interested in the first K columns of B^{ij} , this is sufficient to take the first K columns and the first $K + 2N$ rows of $\underline{DF(\bar{\mathbf{U}})}^{lj}$, $l \in \{1, 2\}$ and the first $K + 2N$ columns and the first $K + 4N - 1$ rows of A^{il} , $l \in \{1, 2\}$ to compute exactly $Z_{1,finite}$.

4.2.2 The infinite part to bound

Let us bound the “tail” $\mathcal{T}_{K,\nu}(B)$ (4.1). Let $K = 2N - 1$, let $\mathbf{z} \in \ell_\nu^1$ with $\mathbf{z}|_K = 0$ and $\mathbf{z} \neq 0$. This choice of K means we look for the columns just after the finite part, this is sufficient to obtain the following calculations on $Z_{1,tail}$ and so Z_1 .

Because $K \geq 2N - 1$, we simply get

$$B^{ij} \mathbf{z} = \delta_{ij} \mathbf{z} - \left(M(\mathbf{w}^{i1}) \Delta^{-1} \underline{DF(\bar{\mathbf{U}})}^{1j} + M(\mathbf{w}^{i2}) \Delta^{-1} \underline{DF(\bar{\mathbf{U}})}^{2j} \right) \mathbf{z}.$$

Indeed, the top-right block of A^{ij} , namely A_{11}^{ij} only acts on the $2N$ (from 0 to $2N - 1$) first coefficients of \mathbf{z} which are null.

Recall that $\underline{DF}(\bar{\mathbf{U}})^{ij} = \Delta \underline{D\Phi}(\bar{\mathbf{U}})^{ij} + \underline{DR}(\bar{\mathbf{U}})^{ij}$, with

$$\bar{\mathbf{U}} = \begin{pmatrix} \bar{\mathbf{u}} \\ \bar{\mathbf{v}} \end{pmatrix}, \quad \Phi \begin{pmatrix} \bar{\mathbf{u}} \\ \bar{\mathbf{v}} \end{pmatrix} = \begin{pmatrix} \gamma(\bar{\mathbf{v}}) * \bar{\mathbf{u}} \\ d\bar{\mathbf{v}} \end{pmatrix} \quad \text{and} \quad R \begin{pmatrix} \bar{\mathbf{u}} \\ \bar{\mathbf{v}} \end{pmatrix} = \begin{pmatrix} \sigma \bar{\mathbf{u}} * (\mathbf{1} - \bar{\mathbf{u}}) \\ \bar{\mathbf{u}} - \bar{\mathbf{v}} \end{pmatrix}.$$

Let us compute $B^{ij}\mathbf{z}$,

$$\begin{aligned} B^{ij}\mathbf{z} &= \delta_{ij}\mathbf{z} - \left(M(\mathbf{w}^{i1}) \underline{D\Phi}(\bar{\mathbf{U}})^{1j} + M(\mathbf{w}^{i2}) \underline{D\Phi}(\bar{\mathbf{U}})^{2j} \right) \mathbf{z} \\ &\quad - \left(M(\mathbf{w}^{i1}) \Delta^{-1} \underline{DR}(\bar{\mathbf{U}})^{1j} + M(\mathbf{w}^{i2}) \Delta^{-1} \underline{DR}(\bar{\mathbf{U}})^{2j} \right) \mathbf{z}. \end{aligned}$$

Since we have, by associativity, $M(\mathbf{w}^{ij}) \underline{D\Phi}(\bar{\mathbf{U}})^{jk} \mathbf{z} = M(\mathbf{w}^{ij}) (\underline{D\Phi}(\bar{\mathbf{U}})^{jk} * \mathbf{z}) = \mathbf{w}^{ij} * (\underline{D\Phi}(\bar{\mathbf{U}})^{jk} * \mathbf{z}) = (\mathbf{w}^{ij} * \underline{D\Phi}(\bar{\mathbf{U}})^{jk}) * \mathbf{z}$.

Then,

$$B^{ij}\mathbf{z} = \left[\begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{pmatrix} - \mathbf{W} * \underline{D\Phi}(\bar{\mathbf{U}}) \right]^{ij} * \mathbf{z} - \mathbf{w}^{ij} * \left[\Delta^{-1} (\underline{DR}(\bar{\mathbf{U}})^{ij} * \mathbf{z}) \right].$$

Finally, with $(\mathbf{z}, \mathbf{z}')$ such that $(\mathbf{z}|_K, \mathbf{z}'|_K) = (\mathbf{0}, \mathbf{0})$,

$$B \begin{pmatrix} \mathbf{z} \\ \mathbf{z}' \end{pmatrix} = \left(\begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{pmatrix} - \mathbf{W} * \underline{D\Phi}(\bar{\mathbf{U}}) \right) * \begin{pmatrix} \mathbf{z} \\ \mathbf{z}' \end{pmatrix} - \mathbf{W} * \left[\begin{pmatrix} \Delta^{-1} & \mathbf{0} \\ \mathbf{0} & \Delta^{-1} \end{pmatrix} \left(\underline{DR}(\bar{\mathbf{U}}) * \begin{pmatrix} \mathbf{z} \\ \mathbf{z}' \end{pmatrix} \right) \right].$$

So,

$$\begin{aligned} \left\| B \begin{pmatrix} \mathbf{z} \\ \mathbf{z}' \end{pmatrix} \right\|_{\nu} &\leq \left\| \left(\begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{pmatrix} - \mathbf{W} * \underline{D\Phi}(\bar{\mathbf{U}}) \right) \begin{pmatrix} \mathbf{z} \\ \mathbf{z}' \end{pmatrix} \right\|_{\nu} + \left\| \mathbf{W} * \left[\begin{pmatrix} \Delta^{-1} & \mathbf{0} \\ \mathbf{0} & \Delta^{-1} \end{pmatrix} \left(\underline{DR}(\bar{\mathbf{U}}) * \begin{pmatrix} \mathbf{z} \\ \mathbf{z}' \end{pmatrix} \right) \right] \right\|_{\nu} \\ &\leq \left\| \left(\begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{pmatrix} - \mathbf{W} * \underline{D\Phi}(\bar{\mathbf{U}}) \right) \begin{pmatrix} \mathbf{z} \\ \mathbf{z}' \end{pmatrix} \right\|_{\nu} + \|\mathbf{W}\|_{\nu} \left\| \begin{pmatrix} \Delta^{-1} & \mathbf{0} \\ \mathbf{0} & \Delta^{-1} \end{pmatrix} \left(\underline{DR}(\bar{\mathbf{U}}) * \begin{pmatrix} \mathbf{z} \\ \mathbf{z}' \end{pmatrix} \right) \right\|_{\nu}. \end{aligned}$$

Also,

$$\begin{aligned} \left\| \begin{pmatrix} \Delta^{-1} & \mathbf{0} \\ \mathbf{0} & \Delta^{-1} \end{pmatrix} \left(\underline{DR}(\bar{\mathbf{U}}) * \begin{pmatrix} \mathbf{z} \\ \mathbf{z}' \end{pmatrix} \right) \right\|_{\nu} &= \left\| \begin{pmatrix} \Delta^{-1} & \mathbf{0} \\ \mathbf{0} & \Delta^{-1} \end{pmatrix} \begin{pmatrix} \sigma(\mathbf{1} - 2\bar{\mathbf{u}}) * \mathbf{z} \\ \mathbf{z} - \mathbf{z}' \end{pmatrix} \right\|_{\nu} \\ &= \left\| \begin{pmatrix} \sigma \Delta^{-1} [(\mathbf{1} - 2\bar{\mathbf{u}}) * \mathbf{z}] \\ \Delta^{-1} [\mathbf{z} - \mathbf{z}'] \end{pmatrix} \right\|_{\nu} \\ &= \sigma \|\Delta^{-1} [(\mathbf{1} - 2\bar{\mathbf{u}}) * \mathbf{z}]\|_{\nu} + \|\Delta^{-1} [\mathbf{z} - \mathbf{z}']\|_{\nu} \\ &\leq \sigma \frac{(b-a)^2}{\pi^2(K-N+1)^2} \|(\mathbf{1} - 2\bar{\mathbf{u}}) * \mathbf{z}\|_{\nu} + \frac{(b-a)^2}{\pi^2(K+1)^2} \|\mathbf{z} - \mathbf{z}'\|_{\nu} \\ &\leq \frac{(b-a)^2}{\pi^2(K-N+1)^2} (\sigma \|(\mathbf{1} - 2\bar{\mathbf{u}}) * \mathbf{z}\|_{\nu} + \|\mathbf{z}\|_{\nu} + \|\mathbf{z}'\|_{\nu}). \end{aligned}$$

Remark 4.3. Since $\bar{\mathbf{u}} \in \ell_{\nu, 2N}^1$ and $\mathbf{z} \in \ell_{\nu}^1$, $\mathbf{z}|_K = \mathbf{0}$, we have $\bar{\mathbf{u}} * \mathbf{z}|_{K-N} = \mathbf{0}$. Furthermore, for any $\mathbf{x} \in \ell_{\nu}^1$ such that $\mathbf{x}|_{K_{\mathbf{x}}} = \mathbf{0}$ we have $\|\Delta^{-1}\mathbf{x}\|_{\nu} \leq \frac{(b-a)^2}{\pi^2(K_{\mathbf{x}}+1)^2} \|\mathbf{x}\|_{\nu}$. This explains the appearance of the $\frac{(b-a)^2}{\pi^2(K-N+1)^2}$ term in the above estimate.

Thus, thanks to the precedent calculations and equation (4.1), we have

$$\begin{aligned} \mathcal{T}_{K,\nu}(B) &\leq \mathcal{T}_{K,\nu} \left(\left(\begin{array}{c|c} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{array} \right) - \mathbf{W} * \underline{D\Phi(\bar{\mathbf{U}})} \right) \\ &\quad + \|\mathbf{W}\|_\nu \frac{(b-a)^2}{\pi^2(K-N+1)^2} \sup_{\substack{(\mathbf{z}, \mathbf{z}') \neq (\mathbf{0}, \mathbf{0}) \\ (\mathbf{z}|_K, \mathbf{z}'|_K) = (\mathbf{0}, \mathbf{0})}} \left(\frac{\sigma \|\mathbf{1} - 2\bar{\mathbf{u}}\|_\nu * \|\mathbf{z}\|_\nu + \|\mathbf{z}\|_\nu + \|\mathbf{z}'\|_\nu}{\|\mathbf{z}\|_\nu + \|\mathbf{z}'\|_\nu} \right). \end{aligned}$$

Firstly,

$$\mathcal{T}_{K,\nu} \left(\left(\begin{array}{c|c} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{array} \right) - \mathbf{W} * \underline{D\Phi(\bar{\mathbf{U}})} \right) \leq \left\| \left(\begin{array}{c|c} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{array} \right) - \mathbf{W} * \underline{D\Phi(\bar{\mathbf{U}})} \right\|_\nu.$$

Secondly,

$$\begin{aligned} \sup_{\substack{(\mathbf{z}, \mathbf{z}') \neq (\mathbf{0}, \mathbf{0}) \\ (\mathbf{z}|_K, \mathbf{z}'|_K) = (\mathbf{0}, \mathbf{0})}} \left(\frac{\sigma \|\mathbf{1} - 2\bar{\mathbf{u}}\|_\nu * \|\mathbf{z}\|_\nu + \|\mathbf{z}\|_\nu + \|\mathbf{z}'\|_\nu}{\|\mathbf{z}\|_\nu + \|\mathbf{z}'\|_\nu} \right) &\leq \sup_{\substack{(\mathbf{z}, \mathbf{z}') \neq (\mathbf{0}, \mathbf{0}) \\ (\mathbf{z}|_K, \mathbf{z}'|_K) = (\mathbf{0}, \mathbf{0})}} \left(\frac{\sigma \|\mathbf{1} - 2\bar{\mathbf{u}}\|_\nu * \|\mathbf{z}\|_\nu}{\|\mathbf{z}\|_\nu + \|\mathbf{z}'\|_\nu} \right) + 1 \\ &\leq \sup_{\substack{\mathbf{z} \neq \mathbf{0} \\ \mathbf{z}|_K = \mathbf{0}}} \left(\frac{\sigma \|\mathbf{1} - 2\bar{\mathbf{u}}\|_\nu * \|\mathbf{z}\|_\nu}{\|\mathbf{z}\|_\nu} \right) + 1 \\ &\leq \sigma \|\mathbf{1} - 2\bar{\mathbf{u}}\|_\nu + 1. \end{aligned}$$

Then, we have

$$\begin{aligned} \mathcal{T}_{K,\nu}(B) &\leq \left\| \left(\begin{array}{c|c} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{array} \right) - \mathbf{W} * \underline{D\Phi(\bar{\mathbf{U}})} \right\|_\nu + \frac{(b-a)^2}{\pi^2(K-N+1)^2} \|\mathbf{W}\|_\nu (\sigma \|\mathbf{1} - 2\bar{\mathbf{u}}\|_\nu + 1) \\ &:= Z_{1,tail}. \end{aligned}$$

All of this is *computable exactly*, and we take

$$\boxed{Z_1 = Z_{1,finite} + Z_{1,tail} + (\|A^{11}\Delta\|_\nu + \|A^{21}\Delta\|_\nu)(\epsilon_\gamma(\bar{\mathbf{v}}) + \|\bar{\mathbf{u}}\|_\nu \epsilon_{\gamma'}(\bar{\mathbf{v}}))}.$$

It satisfies (2.4b) in Theorem 2.1, $\|I - ADF(\bar{\mathbf{U}})\|_\nu \leq Z_1$.

Remark 4.4. Notice that the injectivity of A , which is part of the assumptions needed in Theorem 2.1 in order to obtain the existence of a zero of F , is in fact a consequence of the other assumptions in this setting. Indeed, as soon as assumption (2.5a) is satisfied, then $Z_1 < 1$, which implies that A is injective (see e.g. the proof of [4, Theorem 2.5]).

4.3 The bound Z_2

We assume here that we have done the work of Section 3 on γ and its derivatives. To compute Z_2 we need to compute the second Fréchet derivative of F . Let $\mathbf{U} = \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} \in (\ell_\nu^1)^2$.

Let $\begin{pmatrix} \mathbf{u}_1 \\ \mathbf{v}_1 \end{pmatrix}$ and $\begin{pmatrix} \mathbf{u}_2 \\ \mathbf{v}_2 \end{pmatrix}$ belong to $(\ell_\nu^1)^2$,

$$D^2F \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} \left(\begin{pmatrix} \mathbf{u}_1 \\ \mathbf{v}_1 \end{pmatrix}, \begin{pmatrix} \mathbf{u}_2 \\ \mathbf{v}_2 \end{pmatrix} \right) = \Delta D^2\Phi \left(\begin{pmatrix} \mathbf{u}_1 \\ \mathbf{v}_1 \end{pmatrix}, \begin{pmatrix} \mathbf{u}_2 \\ \mathbf{v}_2 \end{pmatrix} \right) + D^2R \left(\begin{pmatrix} \mathbf{u}_1 \\ \mathbf{v}_1 \end{pmatrix}, \begin{pmatrix} \mathbf{u}_2 \\ \mathbf{v}_2 \end{pmatrix} \right),$$

where

$$D^2\Phi \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} \left(\begin{pmatrix} \mathbf{u}_1 \\ \mathbf{v}_1 \end{pmatrix}, \begin{pmatrix} \mathbf{u}_2 \\ \mathbf{v}_2 \end{pmatrix} \right) = \begin{pmatrix} \gamma'(\mathbf{v}) * \mathbf{u}_1 * \mathbf{v}_2 + \gamma'(\mathbf{v}) * \mathbf{v}_1 * \mathbf{u}_2 + \gamma''(\mathbf{v}) * \mathbf{u} * \mathbf{v}_1 * \mathbf{v}_2 \\ 0 \end{pmatrix},$$

and

$$D^2R \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} \left(\begin{pmatrix} \mathbf{u}_1 \\ \mathbf{v}_1 \end{pmatrix}, \begin{pmatrix} \mathbf{u}_2 \\ \mathbf{v}_2 \end{pmatrix} \right) = \begin{pmatrix} -2\sigma \mathbf{u} * \mathbf{u}_1 * \mathbf{u}_2 \\ 0 \end{pmatrix}.$$

Let us denote

$$S \left(\begin{pmatrix} \mathbf{u}_1 \\ \mathbf{v}_1 \end{pmatrix}, \begin{pmatrix} \mathbf{u}_2 \\ \mathbf{v}_2 \end{pmatrix} \right) = \Delta (\gamma'(\mathbf{v}) * \mathbf{u}_1 * \mathbf{v}_2 + \gamma'(\mathbf{v}) * \mathbf{v}_1 * \mathbf{u}_2 + \gamma''(\mathbf{v}) * \mathbf{u} * \mathbf{v}_1 * \mathbf{v}_2) - 2\sigma \mathbf{u} * \mathbf{u}_1 * \mathbf{u}_2.$$

We have

$$\begin{aligned} AD^2F \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} \left(\begin{pmatrix} \mathbf{u}_1 \\ \mathbf{v}_1 \end{pmatrix}, \begin{pmatrix} \mathbf{u}_2 \\ \mathbf{v}_2 \end{pmatrix} \right) &= \begin{pmatrix} A^{11} & A^{12} \\ A^{21} & A^{22} \end{pmatrix} D^2F \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} \left(\begin{pmatrix} \mathbf{u}_1 \\ \mathbf{v}_1 \end{pmatrix}, \begin{pmatrix} \mathbf{u}_2 \\ \mathbf{v}_2 \end{pmatrix} \right) \\ &= \begin{pmatrix} A^{11}S \\ A^{21}S \end{pmatrix}. \end{aligned}$$

So

$$\left\| AD^2F \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} \left(\begin{pmatrix} \mathbf{u}_1 \\ \mathbf{v}_1 \end{pmatrix}, \begin{pmatrix} \mathbf{u}_2 \\ \mathbf{v}_2 \end{pmatrix} \right) \right\|_{\nu} = \left\| A^{11}S \left(\begin{pmatrix} \mathbf{u}_1 \\ \mathbf{v}_1 \end{pmatrix}, \begin{pmatrix} \mathbf{u}_2 \\ \mathbf{v}_2 \end{pmatrix} \right) \right\|_{\nu} + \left\| A^{21}S \left(\begin{pmatrix} \mathbf{u}_1 \\ \mathbf{v}_1 \end{pmatrix}, \begin{pmatrix} \mathbf{u}_2 \\ \mathbf{v}_2 \end{pmatrix} \right) \right\|_{\nu}.$$

We estimate

$$\|A^{ij}S\|_{\nu} = \sup_{\left(\begin{pmatrix} \mathbf{u}_1 \\ \mathbf{v}_1 \end{pmatrix}, \begin{pmatrix} \mathbf{u}_2 \\ \mathbf{v}_2 \end{pmatrix} \right) \neq (0,0)} \frac{\left\| A^{ij}S \left(\begin{pmatrix} \mathbf{u}_1 \\ \mathbf{v}_1 \end{pmatrix}, \begin{pmatrix} \mathbf{u}_2 \\ \mathbf{v}_2 \end{pmatrix} \right) \right\|_{\nu}}{\left\| \begin{pmatrix} \mathbf{u}_1 \\ \mathbf{v}_1 \end{pmatrix} \right\|_{\nu} \cdot \left\| \begin{pmatrix} \mathbf{u}_2 \\ \mathbf{v}_2 \end{pmatrix} \right\|_{\nu}}.$$

$$\begin{aligned} \left\| A^{ij}S \left(\begin{pmatrix} \mathbf{u}_1 \\ \mathbf{v}_1 \end{pmatrix}, \begin{pmatrix} \mathbf{u}_2 \\ \mathbf{v}_2 \end{pmatrix} \right) \right\|_{\nu} &= \left\| A^{ij} (\Delta(\gamma'(\mathbf{v}) * (\mathbf{u}_1 * \mathbf{v}_2 + \mathbf{v}_1 * \mathbf{u}_2) + \gamma''(\mathbf{v}) * \mathbf{u} * \mathbf{v}_1 * \mathbf{v}_2) - 2\sigma \mathbf{u} * \mathbf{u}_1 * \mathbf{u}_2) \right\|_{\nu} \\ &\leq \left\| A^{ij} \Delta(\gamma'(\mathbf{v}) * (\mathbf{u}_1 * \mathbf{v}_2 + \mathbf{v}_1 * \mathbf{u}_2)) \right\|_{\nu} + \left\| A^{ij} \Delta(\gamma''(\mathbf{v}) * \mathbf{u} * \mathbf{v}_1 * \mathbf{v}_2) \right\|_{\nu} \\ &\quad + 2\sigma \left\| A^{ij} \mathbf{u} * \mathbf{u}_1 * \mathbf{u}_2 \right\|_{\nu} \\ &\leq \left\| A^{ij} \Delta \right\|_{\nu} \|\gamma'(\mathbf{v})\|_{\nu} (\|\mathbf{u}_1\|_{\nu} \|\mathbf{v}_2\|_{\nu} + \|\mathbf{v}_1\|_{\nu} \|\mathbf{u}_2\|_{\nu}) \\ &\quad + \left\| A^{ij} \Delta \right\|_{\nu} \|\gamma''(\mathbf{v}) * \mathbf{u}\|_{\nu} \|\mathbf{v}_1\|_{\nu} \|\mathbf{v}_2\|_{\nu} \\ &\quad + 2\sigma \left\| A^{ij} \mathbf{u} \right\|_{\nu} \|\mathbf{u}_1\|_{\nu} \|\mathbf{u}_2\|_{\nu}. \end{aligned}$$

So

$$\begin{aligned} \left\| AD^2F \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} \left(\begin{pmatrix} \mathbf{u}_1 \\ \mathbf{v}_1 \end{pmatrix}, \begin{pmatrix} \mathbf{u}_2 \\ \mathbf{v}_2 \end{pmatrix} \right) \right\|_{\nu} &\leq (\|A^{11}\Delta\|_{\nu} + \|A^{21}\Delta\|_{\nu}) \|\gamma'(\mathbf{v})\|_{\nu} (\|\mathbf{u}_1\|_{\nu} \|\mathbf{v}_2\|_{\nu} + \|\mathbf{v}_1\|_{\nu} \|\mathbf{u}_2\|_{\nu}) \\ &\quad + (\|A^{11}\Delta\|_{\nu} + \|A^{21}\Delta\|_{\nu}) \|\gamma''(\mathbf{v}) * \mathbf{u}\|_{\nu} \|\mathbf{v}_1\|_{\nu} \|\mathbf{v}_2\|_{\nu} \\ &\quad + 2\sigma(\|A^{11}\mathbf{u}\|_{\nu} + \|A^{21}\mathbf{u}\|_{\nu}) \|\mathbf{u}_1\|_{\nu} \|\mathbf{u}_2\|_{\nu}. \end{aligned}$$

Finally,

$$\begin{aligned} \left\| AD^2F \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} \right\|_{\nu} &\leq \max \left((\|A^{11}\Delta\|_{\nu} + \|A^{21}\Delta\|_{\nu}) \|\gamma'(\mathbf{v})\|_{\nu}, \right. \\ &\quad \left. (\|A^{11}\Delta\|_{\nu} + \|A^{21}\Delta\|_{\nu}) \|\gamma''(\mathbf{v}) * \mathbf{u}\|_{\nu}, \right. \\ &\quad \left. 2\sigma(\|A^{11}\mathbf{u}\|_{\nu} + \|A^{21}\mathbf{u}\|_{\nu}) \right). \end{aligned} \quad (4.2)$$

In order to use Theorem 2.1, we have to get such an estimate for all $\begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix}$ the neighbourhood of $\begin{pmatrix} \bar{\mathbf{u}} \\ \bar{\mathbf{v}} \end{pmatrix}$.

Let $r^* > 0$ and $\begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} \in \mathcal{B}_{\nu} \left(\begin{pmatrix} \bar{\mathbf{u}} \\ \bar{\mathbf{v}} \end{pmatrix}, r^* \right)$. We have directly $\|\mathbf{u} - \bar{\mathbf{u}}\|_{\nu} \leq r^*$ and $\|\mathbf{v} - \bar{\mathbf{v}}\|_{\nu} \leq r^*$. We can therefore apply the results of Section 3 with $\underline{\mathbf{v}} = \bar{\mathbf{v}}$ and $\epsilon_{\underline{\mathbf{v}}} = r^*$, and $\underline{\mathbf{u}} = \bar{\mathbf{u}}$ and $\epsilon_{\underline{\mathbf{u}}} = r^*$. In particular, we get a the full description of $\gamma'(\mathbf{v})$ with $\underline{\gamma'(\bar{\mathbf{v}})}$ and $\epsilon_{\gamma'(\mathbf{v})}$ that depends on r^* , and similarly for $\gamma''(\mathbf{v})$.

Thus we have,

$$\begin{aligned} \mathbf{u} &= \bar{\mathbf{u}} + \mathbf{e}_{\mathbf{u}}, \quad \text{with } \|\mathbf{e}_{\mathbf{u}}\|_{\nu} \leq r^*, \\ \gamma'(\mathbf{v}) &= \underline{\gamma'(\bar{\mathbf{v}})} + \mathbf{e}_{\gamma'(\mathbf{v})}, \quad \text{with } \|\mathbf{e}_{\gamma'(\mathbf{v})}\|_{\nu} \leq \epsilon_{\gamma'(\mathbf{v})}, \\ \gamma''(\mathbf{v}) &= \underline{\gamma''(\bar{\mathbf{v}})} + \mathbf{e}_{\gamma''(\mathbf{v})}, \quad \text{with } \|\mathbf{e}_{\gamma''(\mathbf{v})}\|_{\nu} \leq \epsilon_{\gamma''(\mathbf{v})}. \end{aligned}$$

We inject these identities into the inequality (4.2) and use triangle inequality to obtain Z_2 .

To clarify, let us define:

$$\begin{aligned} Z_2^a &= (\|A^{11}\Delta\|_{\nu} + \|A^{21}\Delta\|_{\nu})(\|\underline{\gamma'(\bar{\mathbf{v}})}\|_{\nu} + \epsilon_{\gamma'(\bar{\mathbf{v}})}), \\ Z_2^b &= (\|A^{11}\Delta\|_{\nu} + \|A^{21}\Delta\|_{\nu}) \left[\|\underline{\gamma''(\bar{\mathbf{v}})} * \bar{\mathbf{u}}\|_{\nu} + \|\underline{\gamma''(\bar{\mathbf{v}})}\|_{\nu} r^* + (\|\bar{\mathbf{u}}\|_{\nu} + r^*) \epsilon_{\gamma''(\bar{\mathbf{v}})} \right], \\ Z_2^c &= 2\sigma[\|A^{11}\bar{\mathbf{u}}\|_{\nu} + \|A^{21}\bar{\mathbf{u}}\|_{\nu} + (\|A^{11}\|_{\nu} + \|A^{21}\|_{\nu})r^*]. \end{aligned}$$

Finally we have

$$\boxed{Z_2 = \max(Z_2^a, Z_2^b, Z_2^c).}$$

It satisfies (2.4c) in Theorem 2.1, $\|AD^2F(\mathbf{U})\|_{\nu} \leq Z_2, \forall \mathbf{U} \in \mathcal{B}_{\nu}(\bar{\mathbf{U}}, r^*)$.

5 Proofs of the existence theorems

In this last section, we specify how to apply the techniques presented in this paper to prove Theorem 1.1, Theorem 1.2 and Theorem 1.3. All the computer-assisted parts of the proofs can be reproduced using

the code available at [32].

5.1 Proof of Theorem 1.1

Let $\sigma = 0.053$, $d = 1$, $\Omega = (0, 3\pi)$ and $\gamma(x) = \frac{1}{1+x^9}$. The approximate steady state (\bar{u}, \bar{v}) is depicted in Figure 2, the complete description of \bar{u} and \bar{v} in terms of Fourier coefficients can be found in the file `Datas/U_initial_single_solution.mat` from [32]. Given $r^* = 1 \times 10^{-6}$, $N = 100$ and $\nu = 1.0001$, we compute the bounds Y , Z_1 and Z_2 for this approximate solution according to Section 4 and obtain the following results¹:

$$Y = 2.4051 \times 10^{-8},$$

$$Z_1 = 3.1194 \times 10^{-2},$$

$$Z_2 = 3.6100 \times 10^4.$$

We then check that condition (2.5) is satisfied. Theorem 2.1 therefore applies, and yields the existence of a locally unique zero (\mathbf{u}, \mathbf{v}) of F satisfying

$$\|(\mathbf{u}, \mathbf{v}) - (\bar{\mathbf{u}}, \bar{\mathbf{v}})\|_\nu \leq 2.5197 \times 10^{-8}.$$

The estimate announced in Theorem 1.1 then simply follows from the fact that $\sup_{x \in \Omega} |z(x)| \leq \|\mathbf{z}\|_\nu$. The computational parts of the proof (i.e. the evaluation of Y , Z_1 and Z_2 and the verification of assumption (2.5)) can be reproduced by running the MATLAB code `proof_single_solution.m` from [32], together with INTLAB [37].

5.2 Proof of Theorem 1.2

Let $\sigma = 0.6$, $d = 1$, $\Omega = (0, 4\pi)$ and $\gamma(x) = \frac{1}{1+e^{9(x-1)}}$. The approximated steady state (\bar{u}, \bar{v}) is described in Figure 3, the complete description of \bar{u} and \bar{v} in terms of Fourier coefficients can be found in the file `Datas/U_initial_WX.mat` after a Newton's method done in the proof. Given $r^* = 1 \times 10^{-6}$, $N = 100$ and $\nu = 1.0001$, for this approximate solution according to Section 4 and obtain the following results¹:

$$Y = 1.5327 \times 10^{-12},$$

$$Z_1 = 2.4338 \times 10^{-2},$$

$$Z_2 = 6.4843 \times 10^2.$$

¹The results shown are the upper bounds of each quantity calculated with MATLAB and INTLAB. We write the results in scientific notation to 5 significant digits.

As before, the condition (2.5) is satisfied. Again thanks to Theorem 2.1, there exists (\mathbf{u}, \mathbf{v}) a locally unique zero of F satisfying

$$\|(\mathbf{u}, \mathbf{v}) - (\bar{\mathbf{u}}, \bar{\mathbf{v}})\|_\nu \leq 1.6956 \times 10^{-12}.$$

The computational parts of the proof can be reproduced by running the MATLAB code `proof_WX.m` from [32], together with INTLAB [37].

5.3 Proof of Theorem 1.3

Let $d = 1$, $\Omega = (0, 3\pi)$ and $\gamma(x) = \frac{1}{1+x^9}$. To prove the existence of solution for each point of the discrete diagram Figure 4, we follow the exact same method as the proof of Section 5.1. Let `Datas/UU_discrete_diagram.mat` be all the data of the each solution shown in Figure 4. Let $r^* = 1 \times 10^{-8}$, $N = 150$ and $\nu = 1.0001$, for i from 1 to `length(UU)` we compute the bounds Y , Z_1 and Z_2 associated to the system with the current approximate solution $(\bar{\mathbf{u}}, \bar{\mathbf{v}}) = \text{UU}(2:\text{end}, i)$ with the $\sigma = \text{UU}(1, i)$ associated. We check if the condition (2.5) is satisfied. If so, we say that we have a validated solution, we apply Theorem 2.1: called $(\sigma, \bar{u}, \bar{v})$ the validated approximate solution, we obtain a unique stationary solution (σ, u, v) of (1.1) that lives in a r^* -neighborhood. If not, increase N and perform the proof on the part not yet validated. The computational parts can also be reproduced by running the MATLAB code `proof_discrete_diagram.m` from [32], together with INTLAB [37]. Figure 4 can be obtained at the end of the script.

Declaration of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

MB and MP were supported by the ANR project CAPPS: ANR-23-CE40-0004-01.

References

- [1] Gianni Arioli, Filippo Gazzola, and Hans Koch. Uniqueness and bifurcation branches for planar steady Navier–Stokes equations under Navier boundary conditions. *Journal of Mathematical Fluid Mechanics*, 23(3):1–20, 2021.

- [2] Gianni Arioli and Hans Koch. Computer-assisted methods for the study of stationary solutions in dissipative systems, applied to the kuramoto–sivashinski equation. *Archive for rational mechanics and analysis*, 197(3):1033–1051, 2010.
- [3] Gianni Arioli, Hans Koch, and Susanna Terracini. Two novel methods and multi-mode periodic solutions for the fermi-pasta-ulam model. *Communications in mathematical physics*, 255(1):1–19, 2005.
- [4] Maxime Breden. Computer-assisted proofs for some nonlinear diffusion problems. *Communications in Nonlinear Science and Numerical Simulation*, 109:106292, 2022.
- [5] Maxime Breden. A posteriori validation of generalized polynomial chaos expansions. *SIAM Journal on Applied Dynamical Systems*, 22(2):765–801, 2023.
- [6] Maxime Breden, Jean-Philippe Lessard, and Matthieu 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.
- [7] Martin Burger, Philippe Laurençot, and Ariane Trescases. Delayed blow-up for chemotaxis models with local sensing. *Journal of the London Mathematical Society*, 103(4):1596–1617, 2021.
- [8] Gabriel Caloz and Jacques Rappaz. Numerical analysis for nonlinear and bifurcation problems. *Handbook of numerical analysis*, 5:487–637, 1997.
- [9] S. Day, J.-P. Lessard, and K. Mischaikow. Validated continuation for equilibria of PDEs. *SIAM J. Numer. Anal.*, 45(4):1398–1424, 2007.
- [10] Laurent Desvillettes, Yong-Jung Kim, Ariane Trescases, and Changwook Yoon. A logarithmic chemotaxis model featuring global existence and aggregation. *Nonlinear Analysis: Real World Applications*, 50:562–582, 2019.
- [11] Laurent Desvillettes, Philippe Laurençot, Ariane Trescases, and Michael Winkler. Weak solutions to triangular cross diffusion systems modeling chemotaxis with local sensing. *Nonlinear Analysis*, 226:113153, 2023.
- [12] Xiongfei Fu, Lei-Han Tang, Chenli Liu, Jian-Dong Huang, Terence Hwa, and Peter Lenz. Stripe formation in bacterial systems with density-suppressed motility. *Physical review letters*, 108(19):198102, 2012.
- [13] Kentarou Fujie and Jie Jiang. Comparison methods for a keller–segel-type model of pattern formations with density-suppressed motilities. *Calculus of Variations and Partial Differential Equations*, 60(3):92, 2021.

- [14] Javier Gómez-Serrano. Computer-assisted proofs in PDE: a survey. *SeMA Journal*, 76(3):459–484, 2019.
- [15] Olivier Hénot. On polynomial forms of nonlinear functional differential equations. *Journal of Computational Dynamics*, 8(3):307–323, 2021.
- [16] Miguel A Herrero and Juan JL Velázquez. A blow-up mechanism for a chemotaxis model. *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze*, 24(4):633–683, 1997.
- [17] Hai-Yang Jin, Yong-Jung Kim, and Zhi-An Wang. Boundedness, stabilization, and pattern formation driven by density-suppressed motility. *SIAM Journal on Applied Mathematics*, 78(3):1632–1657, 2018.
- [18] Hai-Yang Jin and Zhi-An Wang. Critical mass on the keller-segel system with signal-dependent motility. *Proceedings of the American Mathematical Society*, 148(11):4855–4873, 2020.
- [19] Evelyn F Keller and Lee A Segel. Initiation of slime mold aggregation viewed as an instability. *Journal of theoretical biology*, 26(3):399–415, 1970.
- [20] Evelyn F Keller and Lee A Segel. Model for chemotaxis. *Journal of theoretical biology*, 30(2):225–234, 1971.
- [21] 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.
- [22] Jean-Philippe Lessard, Evelyn Sander, and Thomas Wanner. Rigorous continuation of bifurcation points in the diblock copolymer equation. *Journal of Computational Dynamics*, 4(1&2):71, 2017.
- [23] Chenli Liu, Xiongfei Fu, Lizhong Liu, Xiaojing Ren, Carlos KL Chau, Sihong Li, Lu Xiang, Hualing Zeng, Guanhua Chen, Lei-Han Tang, et al. Sequential establishment of stripe patterns in an expanding cell population. *Science*, 334(6053):238–241, 2011.
- [24] Zhengrong Liu and Jiao Xu. Large time behavior of solutions for density-suppressed motility system in higher dimensions. *Journal of Mathematical Analysis and Applications*, 475(2):1596–1613, 2019.
- [25] Manjun Ma, Rui Peng, and Zhian Wang. Stationary and non-stationary patterns of the density-suppressed motility model. *Physica D: Nonlinear Phenomena*, 402:132259, 2020.
- [26] Ramon E Moore. *Methods and applications of interval analysis*. SIAM, 1979.
- [27] Joseph Muscat. *Functional analysis: an introduction to metric spaces, Hilbert spaces, and Banach algebras*. Springer, 2014.

- [28] Mitsuhiro T Nakao. A numerical approach to the proof of existence of solutions for elliptic problems. *Japan Journal of Applied Mathematics*, 5(2):313, 1988.
- [29] Mitsuhiro T. Nakao, Michael Plum, and Yoshitaka Watanabe. *Numerical Verification Methods and Computer-Assisted Proofs for Partial Differential Equations*, volume 53 of *Springer Series in Computational Mathematics*. Springer Singapore, 2019.
- [30] Shin’ichi 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.
- [31] Clifford S Patlak. Random walk with persistence and external bias. *The bulletin of mathematical biophysics*, 15:311–338, 1953.
- [32] Maxime Payan. Non linear cross diffusion - github. <https://github.com/MaximePayan/NonLinearCrossDiffu> 2023. Accessed : 2023-11-02.
- [33] Michael 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.
- [34] Michael Plum. Computer-assisted enclosure methods for elliptic differential equations. *Linear Algebra and its applications*, 324(1-3):147–187, 2001.
- [35] Hervé Queffélec and Claude Zuily. *Analyse pour l’agrégation-4e éd.* Dunod, 2013.
- [36] Peter Rizzi, Evelyn Sander, and Thomas Wanner. Cyclic symmetry induced pitchfork bifurcations in the diblock copolymer model. *Discrete and Continuous Dynamical Systems-B*, 2023.
- [37] Siegfried M Rump. Intlab—interval laboratory. In *Developments in reliable computing*, pages 77–104. Springer, 1999.
- [38] Akitoshi Takayasu, Xuefeng Liu, and Shin’ichi Oishi. Verified computations to semilinear elliptic boundary value problems on arbitrary polygonal domains. *Nonlinear Theory and Its Applications, IEICE*, 4(1):34–61, 2013.
- [39] W. Tucker. *Validated numerics: a short introduction to rigorous computations*. Princeton University Press, 2011.
- [40] 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.
- [41] Jan Bouwe van den Berg, Maxime Breden, Jean-Philippe Lessard, and Lennaert van Veen. Spontaneous periodic orbits in the navier–stokes flow. *Journal of nonlinear science*, 31:1–64, 2021.

- [42] Jan Bouwe van den Berg, Jean-Philippe Lessard, and Konstantin Mischaikow. Global smooth solution curves using rigorous branch following. *Mathematics of computation*, 79(271):1565–1584, 2010.
- [43] Jianping Wang and Mingxin Wang. Boundedness in the higher-dimensional keller-segel model with signal-dependent motility and logistic growth. *Journal of Mathematical Physics*, 60(1), 2019.
- [44] Zhi-An Wang and Xin Xu. Steady states and pattern formation of the density-suppressed motility model. *IMA Journal of Applied Mathematics*, 86(3):577–603, 2021.
- [45] Thomas Wanner. Computer-assisted bifurcation diagram validation and applications in materials science. In *Proc. Sympos. Appl. Math. Rigorous Numerics in Dynamics.*, volume 74, pages 123–174. Amer. Math. Soc., 2018.
- [46] Nobito Yamamoto. A numerical verification method for solutions of boundary value problems with local uniqueness by banach’s fixed-point theorem. *SIAM Journal on Numerical Analysis*, 35(5):2004–2013, 1998.
- [47] Changwook Yoon and Yong-Jung Kim. Global existence and aggregation in a keller–segel model with fokker–planck diffusion. *Acta Applicandae Mathematicae*, 149:101–123, 2017.
- [48] Piotr Zgliczynski and Konstantin Mischaikow. Rigorous numerics for partial differential equations: The Kuramoto—Sivashinsky equation. *Foundations of Computational Mathematics*, 1(3):255–288, 2001.