



**HAL**  
open science

# Asymptotic and numerical analysis of models for heterogeneous cell populations

Jules Guilberteau

► **To cite this version:**

Jules Guilberteau. Asymptotic and numerical analysis of models for heterogeneous cell populations. Dynamical Systems [math.DS]. Sorbonne Université, 2023. English. NNT : 2023SORUS316 . tel-04310209

**HAL Id: tel-04310209**

**<https://theses.hal.science/tel-04310209v1>**

Submitted on 27 Nov 2023

**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.

**THÈSE DE DOCTORAT DE SORBONNE UNIVERSITE**

Au sein du  
**Laboratoire Jacques-Louis Lions**

**Ecole Doctorale 386**  
Sciences Mathématiques de Paris Centre

**Spécialité : Mathématiques appliquées**

**Jules GUILBERTEAU**

---

**Analyse asymptotique et numérique de modèles  
pour les populations cellulaires hétérogènes**

---

Après avis des rapporteurs :

**José CARILLO**  
**Jérôme COVILLE**

Devant le jury composé de :

<b>Matthieu ALFARO</b>	<b>Président du jury</b>
<b>Pierre-Alexandre BLIMAN</b>	<b>Directeur de thèse</b>
<b>Jérôme COVILLE</b>	<b>Rapporteur</b>
<b>Sepideh MIRRAHIMI</b>	<b>Examinatrice</b>
<b>Benoît PERTHAME</b>	<b>Examineur</b>
<b>Camille POUCHOL</b>	<b>Co-directeur de thèse</b>
<b>Nastassia POURADIER DUTEIL</b>	<b>Co-directrice de thèse</b>
<b>Jean CLAIRAMBAULT</b>	<b>Invité</b>



## Résumé de la thèse en français

Dans cette thèse, nous nous intéressons à des modèles déterministes visant à mieux comprendre les mécanismes biologiques et chimiques à l'origine de l'hétérogénéité phénotypique de populations cellulaires. Cette étude est motivée par des travaux d'oncologie montrant que l'hétérogénéité au sein des populations de cellules cancéreuses est un facteur jouant un rôle clé dans la croissance et le développement des métastases.

La première partie de notre travail est consacrée à l'étude de systèmes EDO (Équations Différentielles Ordinaires) modélisant les interactions entre différentes molécules au sein du cytoplasme d'une cellule. Ces systèmes, appelés Réseaux de Régulation Génique, sont très utilisés dans le domaine de la biologie des systèmes et visent à modéliser les mécanismes à l'origine de la plasticité cellulaire, autrement dit de l'évolution d'un trait phénotypique au cours du temps. Comprendre le comportement asymptotique, et plus précisément le nombre et le type des attracteurs d'un système permet d'en évaluer la pertinence en le confrontant aux observations biologiques. Nous étudions ici des cas particuliers de ces modèles appelés Boucles de rétroaction cycliques, et déterminons, pour un large éventail de paramètres, le nombre exact de points d'équilibre stables, ainsi que l'existence éventuelle d'orbites.

En partant de ces EDO, nous construisons un modèle EDP (Équation aux dérivées partielles) structuré en phénotype modélisant le comportement d'une population de cellules au cours du temps. L'équation considérée comporte un terme d'advection, modélisant la plasticité cellulaire, et un terme de sélection, modélisant la croissance et la compétition entre les différentes cellules. Nous prouvons que cette équation admet deux régimes de convergence, qui dépendent des paramètres des termes d'advection et de sélection : la convergence vers une masse de Dirac (nous parlons alors de phénomène de concentration), ou la convergence vers une fonction régulière. Biologiquement, ces deux régimes correspondent respectivement à la sélection d'un unique trait phénotypique, et à la préservation d'un ensemble continu de traits au cours du temps. Ce second scénario, qui correspond au maintien d'une population hétérogène, est possible grâce à l'action conjointe des phénomènes d'advection et de sélection ; en effet, des études antérieures prouvent que, en considérant seulement l'un de ces deux termes, la solution converge nécessairement vers une somme de masses de Dirac.

Nous proposons ensuite un modèle plus complet prenant en compte, en surcroît de l'advection et de la sélection, l'instabilité phénotypique, par le biais d'un terme intégral. Afin d'approcher les solutions de ce modèle, nous développons un schéma numérique de la famille des méthodes particulières. Ce type de schéma, qui repose sur la résolution d'EDO et un processus de régularisation, a pour avantage d'être facilement adaptable en cas de modification du modèle. Nous prouvons la convergence du schéma, et évaluons son ordre de convergence, notamment dans le cas où l'advection est non locale. Nous nous intéressons également à son comportement en temps long, en exhibant des conditions sous lesquelles le schéma préserve le comportement asymptotique de la solution, et des exemples où il ne le préserve pas.

L'équation et le schéma sont enfin utilisés dans le cadre d'une collaboration avec des biologistes des systèmes de l'Institut des Sciences de Bangalore : en se focalisant sur un phénomène précis, la plasticité épithélio-mésenchymateuse, pour laquelle différentes EDO ont déjà été développées, nous proposons une EDP structurée en phénotype pour une population de cellules. Nous montrons que ce modèle préserve plusieurs des propriétés clés des modèles EDO, et évaluons l'effet des différents paramètres sur la population, et en particulier la manière dont ils influencent son hétérogénéité.





# Remerciements

Cette thèse n'aurait pas pu voir le jour sans le soutien de mes co-directeurs de thèse, Nastassia Pouradier Duteil et Camille Pouchol. Je les remercie chaleureusement de s'être autant investis au cours de ces trois années. J'ai toujours pu compter sur leur aide, tant du point de vue mathématique que pour la rédaction des résultats. Travailler avec eux a été un réel plaisir, et ils ont su rapidement me faire découvrir le monde de la recherche et me transmettre leur grande curiosité scientifique au cours de nos nombreuses discussions. Mais ce sont leurs qualités humaines qui m'ont été le plus précieuses, et c'est avant tout leur grande sympathie et leur continuelle prévenance que je tiens à souligner. Je remercie également mon directeur de thèse Pierre-Alexandre Bliman de m'avoir donné l'opportunité de réaliser cette thèse sous sa direction, et d'avoir suivi son avancement malgré ses nombreux projets parallèles.

Je remercie José Carillo et Jérôme Coville de l'intérêt qu'ils ont porté à mon travail lors de discussions puis en acceptant d'être rapporteurs de ma thèse. De même, je remercie Sepideh Mirrahimi et Matthieu Alfaro d'avoir accepté de participer à mon jury de thèse.

Il est impossible de mentionner tous les membres du Laboratoire Jacques-Louis Lions, mais je tiens à remercier en premier lieu le personnel présent et passé de l'administration, Salima Lounici, Malika Larcher, Corentin Maday, Erika Loyson, Nathalie Bergame, sollicités à de nombreuses reprises et sans qui le laboratoire ne pourrait pas fonctionner. Je remercie également Khashayar Dadras pour la gestion du système informatique et pour la rapidité de ses interventions à chaque problème. Les réunions hebdomadaires de la feuue équipe MAMBA ont été très formatrices et m'ont donné, au fil des semaines, un aperçu de la diversité des mathématiques appliquées en biologie. Je remercie chacun de ses membres, et tout particulièrement Marie Doumic d'avoir assuré la responsabilité de l'équipe, Luis Almeida pour ses amicaux conseils sur le doctorat en fin de M2 et sa participation à mon comité de suivi, et Benoît Perthame d'avoir toujours accordé un intérêt à mes résultats pourtant si modestes en comparaisons à l'ensemble de ses travaux, et pour son aide précieuse à plusieurs moments au cours du doctorat. Merci à lui d'avoir accepté d'être membre de mon jury. Jean Clairambault a également porté un regard attentif à mon travail depuis mon stage de master. Ses connaissances encyclopédiques en oncologie m'ont permis d'aiguiller mon travail et les échanges avec lui ont toujours été passionnants : c'est un plaisir pour moi qu'il assiste au titre d'invité à ma soutenance. Je remercie enfin l'ensemble des doctorants du laboratoire croisés au cours de ces trois années : j'ai été impressionné par les connaissances mathématiques et la sincère passion qui animait la plupart d'entre eux. Ne pouvant citer tout le monde, je remercie tout spécialement Justine, Lise, David, Igor, Nelly, Ioanna-Maria, Guillaume et Robin qui ont partagé avec moi le bureau 15-16-221, pour leur sympathique présence au quotidien.

Je tiens à remercier Frank Ernesto Alvarez, avec qui j'ai eu la chance de travailler : ce manuscrit lui doit beaucoup pour l'ensemble du Chapitre 5. Je retiendrai son affabilité et nos agréables discussions, notamment sur Cuba.

I would like to warmly thank Mohit Kumar Jolly for allowing me to come to his

laboratory in Bangalore, for doing everything possible to make my stay as pleasant as possible, for kindly inviting me to share convivial moments and for the invaluable tourist and gastronomic advice he gave me. Working in his laboratory was very instructive and rewarding, and I would like to thank his entire team for making me feel so welcome. I especially thank Paras Jain for his daily solicitude and indispensable help throughout my stay, and everything he told me about Jain and Indian culture. I have unforgettable memories of Shravanabelagola, where I would probably never have gone on my own. It was a pleasure to work with him and I learnt a lot, both biologically and numerically. Thanks to Hemanya for the time spent birdwatching, for helping me discover so many birding spots and for his help with identification.

Merci à ma famille, mes parents, mes grand-mères, mon frère Baptiste et Juliette pour les séjours à Sallanches, Loublande ou Cholet, et leur soutien constant malgré mes perpétuels atermoiements. Merci enfin à mes amis Manon, Lucas M., Erwan, Yann, Victor, Aymar et Lucas S. pour leur amour des oiseaux, des montagnes, des églises romanes, des films et des livres et leur salutaire pouvoir de m'extraire, le temps d'un soir, d'un week-end ou d'une semaine de vacances, du monde parfois harassant des mathématiques.





# Contents

<b>1</b>	<b>General Introduction</b>	<b>1</b>
1.1	ODE models for single cells: Gene Regulatory Networks . . . . .	1
1.1.1	General statements and mathematical approach . . . . .	1
1.1.2	Cyclic feedback loops . . . . .	4
1.2	PDE models for cell populations . . . . .	8
1.2.1	Construction of the models . . . . .	8
1.2.2	Long-time behaviour of the advection-selection advection . . . . .	11
1.2.3	A particle method for advection-selection-mutation equations . . . . .	17
1.3	Application to Epithelial-Mesenchymal Plasticity . . . . .	20
<b>2</b>	<b>Monostability and bistability of biological switches</b>	<b>29</b>
2.1	Introduction . . . . .	30
2.2	Preliminary results . . . . .	33
2.2.1	Ordering of stable points . . . . .	33
2.2.2	Basins of attraction . . . . .	34
2.3	A criterion which ensures at most bi-stability . . . . .	36
2.3.1	A priori bounds on the number of fixed points . . . . .	36
2.3.2	Properties of $\gamma^{1/2}$ -convex functions . . . . .	37
2.4	Determining the parameters for mono/bistability . . . . .	44
2.4.1	General framework . . . . .	44
2.4.2	Application to n=1: determining the parameters for mono/bistability	45
2.4.3	Application to n=2 . . . . .	49
2.5	Discussion . . . . .	50
<b>3</b>	<b>Bistability and oscillatory behaviours of cyclic feedback loops</b>	<b>53</b>
3.1	Introduction . . . . .	54
3.2	Characterisation of fixed points . . . . .	56
3.3	Even number of decreasing functions . . . . .	58
3.4	Odd number of decreasing functions . . . . .	61
<b>4</b>	<b>Long-time behaviour of an advection-selection equation</b>	<b>63</b>
4.1	Introduction . . . . .	64
4.1.1	Advection-selection equation . . . . .	64
4.1.2	Main results . . . . .	65
4.1.3	Discussion . . . . .	66
4.2	Framework and reminders . . . . .	68
4.2.1	The space of Radon measures . . . . .	68
4.2.2	General statement regarding the characteristics curves . . . . .	69
4.3	Resolution method . . . . .	70
4.3.1	Semi-explicit expression of the solution . . . . .	70

4.3.2	ODE satisfied by the population size . . . . .	71
4.4	Results in the one-dimensional case . . . . .	75
4.4.1	Asymptotic behaviour of the carrying capacities . . . . .	75
4.4.2	Applications . . . . .	87
4.4.3	More than two equilibria . . . . .	93
4.5	Some results in higher dimensions . . . . .	95
4.5.1	Limit support . . . . .	96
4.5.2	Stationary solutions . . . . .	98
<b>5</b>	<b>A particle method for an advection-selection-mutation equation</b>	<b>103</b>
5.1	Introduction . . . . .	104
5.2	The problem . . . . .	108
5.2.1	Some bounds over the characteristics . . . . .	110
5.2.2	Existence of solution for smooth initial data . . . . .	112
5.2.3	Existence of solution for more general initial data . . . . .	115
5.3	Particle Method . . . . .	120
5.4	Convergence of the numerical solution towards a weak solution . . . . .	126
5.4.1	Convergence on a finite time interval . . . . .	127
5.4.2	Asymptotic preserving properties . . . . .	131
5.5	Simulations . . . . .	142
<b>6</b>	<b>An advection-selection-mutation model for EMP</b>	<b>147</b>
6.1	Introduction . . . . .	148
6.2	Advection equation and hysteresis . . . . .	149
6.2.1	ODE model and corresponding advection equation . . . . .	149
6.2.2	Results . . . . .	150
6.3	Effect of epimutations and growth . . . . .	150
6.3.1	Model . . . . .	150
6.3.2	Focus on epimutations . . . . .	152
6.3.3	Focus on growth . . . . .	153
<b>A</b>	<b>Appendix of Chapter 6</b>	<b>158</b>
A.1	Supplementary figures . . . . .	159
A.2	Parameters for ODE (6.1) . . . . .	165
<b>B</b>	<b>Appendix of Chapter 2</b>	<b>166</b>
B.1	Complement of Proposition 2.5 . . . . .	166
B.2	Complement of Proposition 12 . . . . .	168
B.3	Other sigmoid functions which are $\gamma - 1/2$ -convex . . . . .	170
<b>C</b>	<b>Appendix of Chapter 5</b>	<b>172</b>
C.1	Proof of the results over the characteristics . . . . .	172
C.2	Existence of solution for a system of ODEs . . . . .	176
C.3	A result from approximation theory . . . . .	177
C.4	Proofs of convergence results from ODE theory . . . . .	178
	<b>Bibliography</b>	<b>181</b>

# Glossary

Throughout this document, we denote

- $\mathbb{R}_+$  the set non-negative numbers, *i.e.*  $[0; +\infty)$ .
- $\mathbb{R}_+^*$  the set of positive numbers, *i.e.*  $(0, +\infty)$ .
- For all  $m, n \in \mathbb{N}$ , with  $m < n$ ,  $\llbracket m, n \rrbracket$  the set of integers between  $m$  and  $n$ , *i.e.*  $\{m, m+1, \dots, n\}$ .
- $\nabla \cdot$  the divergence operator.
- For any set  $S \subset \mathbb{R}^d$ ,  $\bar{S}$  its closure,  $S^\circ$  its interior and  $S^c$  its complement.
- For any function  $u : \mathbb{R}^d \rightarrow \mathbb{R}$ ,  $\text{supp}(u)$  its support, *i.e.*  $\overline{\{x \in \mathbb{R}^d : u(x) \neq 0\}}$ .
- $\mathcal{C}_c(\mathbb{R}^d)$  the set of continuous functions from  $\mathbb{R}^d$  onto  $\mathbb{R}$  with a compact support.
- $\mathcal{M}(\mathbb{R}^d)$  the set of Radon measures on  $\mathbb{R}^d$ , *i.e.* measures  $\mu$  which are
  - (i) Inner regular, *i.e.* for all open set  $O \subset \mathbb{R}^d$ ,  
 $\mu(O) = \sup\{\mu(K), K \text{ compact set included in } O\}$ .
  - (ii) Outer regular, *i.e.* for all Borel set  $B \subset \mathbb{R}^d$ ,  
 $\mu(B) = \inf\{\mu(U), U \text{ open set containing } B\}$ .
  - (iii) Locally finite, *i.e.* for all  $x \in \mathbb{R}^d$ , there exists  $V_x$  a neighbourhood of  $x$  such that  
 $\mu(V_x) < +\infty$ .
- $\rightharpoonup$  the weak- $\star$  convergence in the space of Radon measures: if  $(\mu_k)_{k \in \mathbb{N}}$  is a sequence of  $\mathcal{M}(\mathbb{R}^d)$ , and  $\bar{\mu} \in \mathcal{M}(\mathbb{R}^d)$  then  $\mu_k \xrightarrow[k \rightarrow +\infty]{} \bar{\mu}$  if and only if for all  $\phi \in \mathcal{C}_c(\mathbb{R}^d)$

$$\int_{\mathbb{R}^d} \phi(x) d\mu_k(x) \xrightarrow[k \rightarrow +\infty]{} \int_{\mathbb{R}^d} \phi(x) \bar{\mu}(x).$$

- $G : s \mapsto \frac{1}{\sqrt{2\pi}} e^{-\frac{s^2}{2}}$  the Gaussian function.
- $K := 10^3$ .





# Chapter 1

## General Introduction

### 1.1 ODE models for single cells: Gene Regulatory Networks

#### 1.1.1 General statements and mathematical approach

In a same environment, genetically identical cells may assume different functional fates [6]. This process, known as *cellular decision-making*, provides an explanation to the epigenetic heterogeneity of cell populations, and has therefore been the subject of extensive study for over half a century. This notion of *cell-fate decision* cannot be explained with the former view of the cells as a passive system which changes states only in response to environmental stimuli [87, 120]. The necessity to consider internal regulations between molecules within cells has thus arisen, giving birth to *gene regulatory networks* [37, 45, 65, 121].

Due to the lack of a clear definition of the term *cell type*, we will avoid using it and prefer the term ‘*cell state*’ to designate cell characteristics and functions [36, 155, 163]. The state of a cell is usually characterised by the level of certain proteins in its cytoplasm. The purpose of *gene regulatory networks* is to describe how the levels of these proteins evolve over time, potentially leading to cell state transition. Such networks, which are generally modelled by Ordinary Differential Equations (ODEs) are notably built to explain the following two empirical observations:

- In a same environment, two cells can adopt different states.
- A transient change of the environment may produce a stable change in cell state [152].

Let us start with a very naive model sufficient to justify why one should consider regulation between proteins.

**Independent molecule.** As a starting point, we consider a molecule of interest  $B$  present in the cytoplasm of a cell, with a synthesis rate of  $g > 0$  which is assumed to depend solely on the external environment. The level of  $B$  does not increase indefinitely: it is subject to spontaneous decay due to degradation, which is assumed to be proportional to the concentration of this substance itself. Thus, by denoting  $y(t)$  the concentration of  $B$  at time  $t \geq 0$ , the *net rate of synthesis* of  $B$  at time  $t$  is given by  $g - ky(t)$  (with  $k > 0$  a kinetic constant), and  $y$  thus solves

$$\dot{y}(t) = g - ky(t) \quad t \geq 0. \tag{1.1}$$

The solution of this equation is explicitly given by

$$y(t) = e^{-kt} \left( y^0 - \frac{g}{k} \right) + \frac{g}{k},$$

where  $y^0 = y(0)$  is the initial concentration of  $B$ . It is clear that  $y$  converges toward  $\frac{g}{k}$ , regardless of the initial level  $y^0$ , which means that the long-time behaviour of the level  $B$  is entirely determined by the cell environment. Therefore, this simplistic model does not account for phenotypic diversity observed in a given environment, nor for the fact that a transient change in the environment can permanently affect the cell state [152].

**Principle of regulation.** As mentioned above, the idea of gene regulatory networks is to consider, in addition to the environment seen as an external input, internal regulation between several kinds of molecules. Regulation means that the level of a given molecule (say  $A$ ) may impact the synthesis rate of another one (say  $B$ ). With the formalism of (1.1), if  $x$  denotes the concentration of  $A$ , and  $y$  that of  $B$ , this means that  $g$  is a function of  $x$ . Regulation can be positive (if  $A$  promotes the synthesis of  $B$ ) in which case  $g$  is increasing with respect to  $x$ , or negative (if  $A$  inhibits the synthesis of  $B$ ) in which case  $g$  is decreasing with respect to  $x$ .

At this stage, a few reminders about gene expression are in order. A gene is a segment of DNA composed of a nucleotide sequence which codes for a specific protein. Gene expression takes place in two main steps [43, 44]:

- *Transcription:* A messenger RNA (mRNA) is synthesised from the DNA sequence corresponding to the gene.
- *Translation:* The mRNA serves as a model for the synthesis of a specific protein.

Thus, regulatory functions are essentially of two types:

- *Linear regulation*, which notably covers translational regulation. Indeed, the number of synthesised proteins is usually assumed to be proportional to the number of its associated mRNA. Thus,  $g(x)$  writes  $g(x) = \alpha x$ , where  $\alpha$  corresponds to the quantity of protein produced by mRNA. These linear regulations are sometimes omitted to simplify models when they do not alter their qualitative behaviour.
- *Non-linear regulation*, which is usually used to model transcriptional regulation. Molecules (usually proteins), called transcription factors, regulate the rate of transcription by binding to specific DNA sequences [99]. This regulation can be positive (in which case they are called activators) or negative (in which case they are called repressors) [96, 98]. This type of regulation is usually of little effect under a certain threshold, and reaches a plateau for high concentrations [150]. This observation has led to use step functions, giving rise to models known as *piecewise linear differential equations*, and, in an even more schematic way, to *boolean models* [149, 151]. Nevertheless, a more common and relevant choice derives from the Hill–Langmuir equation [63, 75] which leads to model transcriptional regulations with so-called *Hill functions*, *i.e.* functions of the shape

$$H^-(x) = \frac{\theta^r}{\theta^r + x^r} \quad \text{or} \quad H^+(x) = \frac{x^r}{\theta^r + x^r}, \quad (1.2)$$

with  $\theta > 0, r > 1$ , or more generally, *shifted Hill functions*

$$H^\lambda(x) = H^-(x) + \lambda H^+(x) = \frac{\theta^r + \lambda x^r}{\theta^r + x^r}, \quad (1.3)$$

which offer an additional degree of freedom via the variable  $\lambda \geq 0$  [3, 35]. The function  $g$  then writes  $g(x) = \alpha H(x)$ , where  $H$  is one of these three functions, and

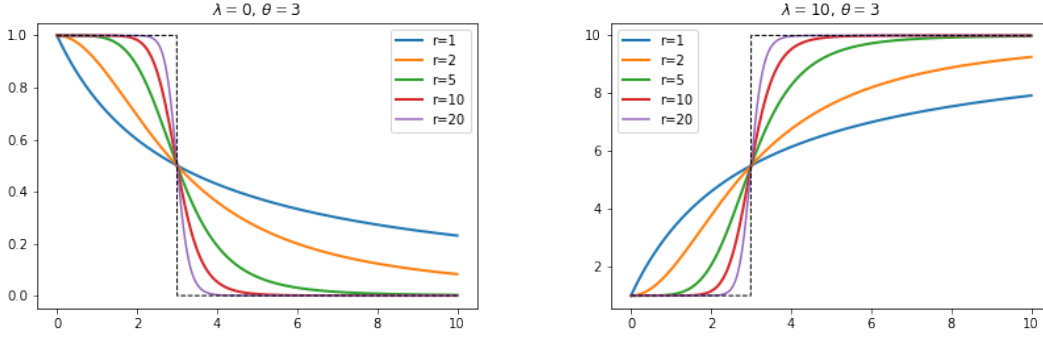


Figure 1.1: Effect of the Hill coefficients on the ‘stiffness’ of Hill functions. The step function to which they tend when  $r$  goes to  $+\infty$  is represented by a black dashed line.

$\alpha > 0$  is a positive parameter such that  $\alpha H(0)$  is the rate of synthesis of  $B$  in the absence of  $A$ .

Note that  $H^-$  is a particular case of  $H^\lambda$  with  $\lambda = 0$ , and that  $H^+ = 1 - H^-$ . Thus, we will only describe the properties of  $H^\lambda$ , for  $\lambda \geq 0$  from which we can deduce those of  $H^-$  and  $H^+$ . First, we note that  $H^\lambda$  is increasing if  $\lambda > 1$ , which thus corresponds to a positive regulation ( $A$  is an activator), decreasing if  $\lambda < 1$  ( $A$  is a repressor), and constant if  $\lambda = 1$  ( $A$  does not regulate  $B$ ). Moreover, except in this last case, if  $r > 1$ , then  $H^\lambda$  is a sigmoid function *i.e.* it is bounded and has a unique inflection point, which is given by  $x_{infl} = \left(\frac{r-1}{r+1}\right)^{1/r} \theta$ . Hence,  $\theta$  linearly influences the value of the inflection point, and is therefore interpreted as an action threshold of transcription factors [152]. The shift parameter  $\lambda$  reflects the maximal action the regulation has, since  $\lim H^\lambda(x) = \lambda$  as  $x \rightarrow +\infty$ . Lastly, the so-called Hill coefficient  $r$  determines the ‘stiffness’ of the function, as illustrated by Figure 1.1: for  $r = 1$ ,  $H^\lambda$  is concave; conversely, for any fixed  $\theta > 0$ ,  $\lambda \geq 0$ ,  $H^\lambda$  converges pointwise to the step function  $\mathbb{1}_{[0,\theta)} + \frac{1+\lambda}{2} \mathbb{1}_{\{\theta\}} + \lambda \mathbb{1}_{(\theta,+\infty)}$  as  $r$  goes to  $+\infty$ , which means that the larger  $r$  is, the closer Hill functions are to step functions.

**Multiregulatory network.** Of course, the synthesis rate of a molecule may depend on the level of several other molecules, and even on its own level [8]. Thus, by considering a network of  $d$  kinds of molecules  $A_1, \dots, A_d$ , and by denoting  $x_1(t), \dots, x_d(t)$  their respective concentrations at time  $t \geq 0$  (and  $x(t)$  the vector  $(x_1(t), \dots, x_d(t))$ ), we obtain the general regulatory network

$$\forall i \in \{1, \dots, d\}, t \geq 0, \quad \dot{x}_i(t) = g_i(x(t)) - k_i x_i(t). \quad (1.4)$$

Such systems can stabilise in different ways, depending on the initial conditions. Each cell state is identified with one possible long-time behaviour of the system. In order to define these notions rigorously, some mathematical formalism is required.

Let  $F : \mathbb{R}^d \rightarrow \mathbb{R}^d$  a vector field, and let us consider the ODE

$$\dot{x}(t) = F(x(t)), \quad t \geq 0 \quad (1.5)$$

which is a generalisation of (1.4). We assume that there exists a non-empty set  $E \subset \mathbb{R}^d$  such that for any initial condition  $x^0 \in \mathbb{R}^d$ , ODE (1.5) has a unique solution, which is defined on  $\mathbb{R}_+$  and remains in  $E$ , that we denote  $X(\cdot, x^0)$  ( $X$  is thus the flow of  $F$ ).

For biological relevance, we assume that  $E \subset \mathbb{R}_+^d$  (since the level of each protein is non-negative), and that there exists a compact attractor set, *i.e.* a compact set  $K \subset E$  such that for any  $x^0 \in E$ , there exists  $T \geq 0$  such that  $X(t, x^0) \in K$  for all  $t \geq T$ .

For any  $x^0 \in E$ , we define the  $\omega$ -limit set of  $x^0$  as

$$\omega(x^0) = \bigcap_{t \geq 0} \overline{X([t, +\infty), x^0)} = \bigcap_{t \geq 0} \overline{\bigcup_{s \geq t} \{X(s, x^0)\}},$$

which is a non-empty, connected and compact subset of  $K$  [130].

Lastly, we say that a set  $S \subset K$  is an *attractor*<sup>1</sup> for (1.5) if there exists  $x^0 \in E$  such that  $S = \omega(x^0)$ . Moreover, we call *basin of attraction* of  $S$  the set

$$\{x^0 \in E : \omega(x^0) = S\},$$

and we say that  $S$  is an *asymptotically stable* attractor if there exists a neighbourhood of  $S$  included in its basin of attraction and that it is a *non-negligible attractor* if its basin of attraction has a positive Lebesgue measure. Note that an asymptotically stable attractor is a non-negligible attractor, but that the converse assertion is false. However, in the case of hyperbolic equilibrium points, these two notions are equivalent.

The mathematical study of gene regulatory networks essentially consists of determining the number, nature and stability (or more generally if they are negligible or not) of the attractors of the ODE system (1.4). Indeed, the number of non-negligible attractors corresponds to the number of possible cell states the model has. In particular, *multistability* (*i.e.* existence of at least two stable attractors) is likely to explain distinct cell-fate decisions. The nature of the attractors is also of great significance: theoretically, they may be of various kinds, and notably be so-called strange attractors when the dimension  $d$  is greater than two [66]. Nevertheless, essentially two kinds of attractors have a biological relevance: equilibrium points and periodic orbits:

- A point  $\bar{x} \in K$  is an attractor for (1.5) if and only if it satisfies  $F(\bar{x}) = 0$ . It is then called an *equilibrium point*. The basin of attraction of an equilibrium point corresponds to the set of  $x^0$  for which  $t \mapsto X(t, x^0)$  converges to  $\bar{x}$ . Biologically speaking, equilibrium points correspond to stable cell states, for which the level of the molecules under consideration shows little fluctuation.
- A *periodic orbit*, also called *cycle*, is the trajectory of a non-constant periodic solution of (1.5), *i.e.* a set of the form

$$X([0, T], x^0),$$

with  $x^0 \in K$ ,  $T > 0$  such that  $X(0, x^0) = X(T, x^0)$ . Periodic orbits correspond to cell states whose molecule levels undergo continuing undamped oscillations [25, 37, 52].

The theoretical study of the attractors is out of reach for general networks such as (1.4): we will therefore focus on a particular case: the *cyclic feedback loop*.

### 1.1.2 Cyclic feedback loops

We call *Cyclic feedback loops* [152] the gene regulatory network (1.4), with the additional hypothesis that for all  $i \in \{0, \dots, d\}$ ,  $g_i$  only depends on  $x_i$  and  $x_{i-1}$ , with the convention

$$x_0 = x_d$$

---

<sup>1</sup>There are various definitions of attractors [119]; we opt for this one because it allows to count them, contrarily to other definitions where the union of several attractors is itself an attractor.

which will be used throughout this section. In other words, the *cyclic feedback loops* writes

$$\forall i \in \{0, \dots, d\}, t \geq 0, \quad \dot{x}_i(t) = g_i(x_i(t), x_{i-1}(t)) - k_i x_i(t). \quad (1.6)$$

This model is of special interest because it often serves as the building block for more complex models. It also relies on the notion of *biological feedback*, which is involved in a large number of phenomena and seems to play a central role in gene expression [64,152,158]. Biological feedback refers to the fact that molecules regulate their own synthesis rate. This feedback is twofold: direct, since  $g_i$  depends on  $x_i$ , and indirect, since  $A_i$  regulates  $A_{i+1}$  which regulates  $A_{i+2}$ ... and so on, until finally  $A_{i-1}$  regulates  $A_i$ .

To fully understand the phenomena due to indirect feedback and the ones due to direct feedback, it is useful to study *simple feedback loops*, which corresponds to the cyclic feedback loop without direct regulation, *i.e.*

$$\forall i \in \{0, \dots, d\}, t \geq 0, \quad \dot{x}_i(t) = g_i(x_{i-1}(t)) - k_i x_i(t), \quad (1.7)$$

where for all  $i \in \{1, \dots, d\}$ ,  $g_i$  is decreasing if the regulation of  $x_{i-1}$  on  $x_i$  is negative, and increasing if it is positive.

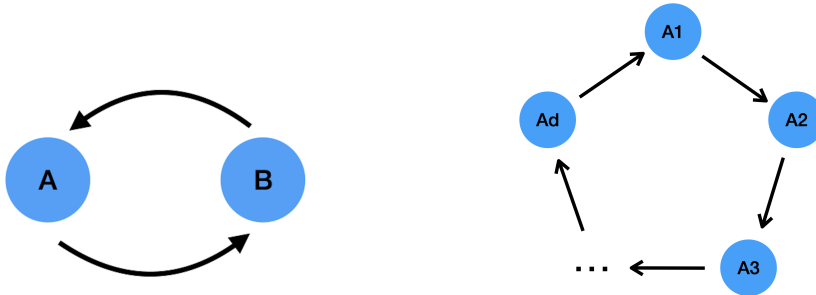


Figure 1.2: Schematic representation of a simple feedback loop: the blue disks represents the molecules of the network, and the arrow interactions between them. On the left, the two-dimensional case, which is the subject of Chapter 2, and, on the right, the general network with  $d$  molecules, studied in Chapter 3.

Despite its simplicity, this model has sparked interest because it suffices to explain experimental observations in *Escherichia coli*: oscillatory behaviours [53] as well as bistability [60].

Our contribution is a theoretical study of this model, for a large range of functions  $g_1, \dots, g_d$ , which encompass several previous studies made in particular cases with a reduced number of non-linear functions [7,102,122,143,146,147,153,161]. Our work is divided in two parts: we first propose a complete analysis of the model in the two-dimensional case, in Chapter 2, before applying the method developed for any superior dimension, in Chapter 3.

In the two dimensional case  $d = 2$ , the Poincaré-Bendixson theorem and the Poincaré-Dulac theorem ensure that any attractor is an equilibrium point [130]. When  $d \geq 3$ , these theorems cannot be applied, and the existence of oscillatory behaviours has been uncovered for large ranges of parameters [57,154]. Nevertheless, two important results have been established, and provide a kind of equivalence. The first one, due to Mallet-Paret and Smith [113], states that if the functions  $g_i$  are monotonous, then the attractors of (1.7) are either equilibrium points or periodic orbits. In the case where  $d$  is even, this result can be completed by that of [147], which shows that the basins of attraction of the periodic orbits are negligible.

Thus, the number of the equilibrium points depends on  $d$  (the number of variables) and  $n$  (the number of negative regulations).

Following the formalism of Cherry and Adler [26], we study the system

$$\forall i \in \{1, \dots, d\}, t \geq 0, \quad \dot{x}_i(t) = \alpha_i f_i(x_{i-1}(t)) - x_i(t), \quad (1.8)$$

where  $\alpha = (\alpha_1, \dots, \alpha_d) \in (\mathbb{R}_+^*)^d$  and  $f_1, \dots, f_d$  are non-negative, strictly monotonous and at least one of them is bounded. By taking, for all  $i \in \{1, \dots, d\}$ ,  $\alpha_i := 1/k_i$ , this system has the same equilibrium point as the simple cyclic feedback loop (1.7).

We introduce the quantity

$$D := \prod_{i=1}^d \sup_{x>0} \left| \frac{x f'_i(x)}{f_i(x)} \right| \in (0, +\infty], \quad (1.9)$$

which will be used to characterise the attractors the system has. Note that this quantity can usually be explicitly computed: if  $l$  is a linear (or more generally affine) and non-constant function, then

$$\sup_{x>0} \left| \frac{x l'(x)}{l(x)} \right| = 1,$$

and if  $H^-, H^+$  and  $H$  are the Hill functions defined by (1.2) and (1.3), then

$$\sup_{x>0} \left| \frac{x H^{-'}(x)}{H^-(x)} \right| = \sup_{x>0} \left| \frac{x H^{+'}(x)}{H^+(x)} \right| = r,$$

and

$$\sup_{x>0} \left| \frac{x H'(x)}{H(x)} \right| = r \left| \frac{1 - \sqrt{\lambda}}{1 + \sqrt{\lambda}} \right|.$$

**Even number of decreasing functions.** As mentioned above, if there is an even number of decreasing functions in the network, then the only possible non-negligible attractors are equilibrium points. Thus, it is enough to compute the solutions of the system

$$\forall i \in \{1, \dots, d\}, \quad \alpha_i f_i(\bar{x}_{i-1}) = \bar{x}_i, \quad (1.10)$$

which is equivalent to

$$\begin{cases} \tilde{f}(\bar{x}_d) := \alpha_d f_d \circ \dots \circ \alpha_1 f_1(\bar{x}_d) = \bar{x}_d \\ \forall i \in \{1, \dots, d-1\}, \quad \alpha_i f_i(\bar{x}_{i-1}) = \bar{x}_i \end{cases}, \quad (1.11)$$

and to determine their stability. The stability study of the ODE thus consists in determining the number of fixed points of the function  $\tilde{f}$ .

The first part of our study is to bound the possible number of fixed points of (1.7), *i.e.* the number of solutions of (1.11). For this purpose, we introduce a notion that we call  $\gamma^{1/2}$ -convexity: We say that a function  $f \in \mathcal{C}^3(\mathbb{R}_+)$  is (strictly)  $\gamma^{1/2}$ -convex if  $|f'| > 0$  and  $|f'|^{-1/2}$  is (strictly) convex. We easily check that linear (or more generally affine) functions and Hill functions are  $\gamma^{1/2}$ -convex, and that Hill functions with Hill coefficients greater than one are strictly  $\gamma^{1/2}$ -convex, as well as most of the usual sigmoid functions. We show  $\gamma^{1/2}$ -convex functions have two interesting properties:

- They are stable by composition: in other words, if  $f$  and  $g$  are two  $\gamma^{1/2}$ -convex functions, then  $f \circ g$  is also  $\gamma^{1/2}$ -convex. Moreover, if (at least) one of them is strictly  $\gamma^{1/2}$ -convex, then  $f \circ g$  is strictly  $\gamma^{1/2}$ -convex.

- A strictly  $\gamma^{1/2}$ -convex function has at most three fixed points.

Thus, combining these two properties ensures that if  $f_1, \dots, f_d$  are  $\gamma^{1/2}$ -convex, then  $\tilde{f}$  has at most three fixed points, and thus that (1.7) has at most three equilibrium points. Further analyses show that, except in rare cases which occur only for a set of parameters of measure zero in the  $\alpha$ -parameter space, the system is either *monostable*, *i.e.* has a unique equilibrium point which is globally asymptotically stable, or *bistable*, *i.e.* has exactly three equilibrium points among which two are asymptotically stable, and one is hyperbolic and asymptotically unstable, and thus negligible.

Moreover, we identify a set of values for  $\alpha$  for which this system is monostable, and the set for which it is bistable. The union of these two set has of complement of Lebesgue measure zero in  $(\mathbb{R}_+^*)^d$ .

The result obtained can be summarised as follows:

**Theorem 1.** *We assume that  $f_1, \dots, f_d \in \mathcal{C}^3(\mathbb{R}_+)$  are non-negative, monotonous and  $\gamma^{1/2}$ -convex functions, and that at least one of them is bounded. Moreover, we assume that an even number of these function is decreasing. With  $D$  defined by (1.9), we have the following alternative:*

- (i) *If  $D < 1$ , then for any  $\alpha \in (\mathbb{R}_+^*)^d$  system (1.7) has a unique equilibrium point which is globally asymptotically stable.*
- (ii) *If  $D > 1$ , then there exists a non-empty set  $A_{bis} \subset (\mathbb{R}_+^*)^d$  such that*
  - *If  $\alpha \in A_{bis}$ , then system (1.7) has exactly two asymptotically stable equilibria, and the union of their basins of attraction is an open subset of  $\mathbb{R}_+^d$  with a complement of Lebesgue measure zero.*
  - *If  $\alpha \in \overline{A_{bis}}^C$ , then system (1.7) has a unique equilibrium point which is globally asymptotically stable.*

We complete this result by proving that, if  $d \geq 5$  and  $D > \cos(\frac{2\pi}{d})^{-d}$ , then there exists a non-empty set  $A_{per} \subset A_{bis}$  such that if  $\alpha \in A_{per}$ , then system (1.7) has (negligible) periodic orbits. Moreover,  $A_{bis}$  and  $A_{per}$  can be explicitly expressed as a function of  $f_1, \dots, f_d$ .

**Odd number of decreasing functions.** In the case where  $n$ , the number of decreasing functions, is odd, we easily check that system (1.7) has a unique equilibrium point: indeed, the function  $\tilde{f}$  as defined in (1.11) is then decreasing, with  $\tilde{f}(0) > 0$ . If  $d \geq 3$ , the ‘Poincaré-Bendixson theorem for monotone cyclic feedback systems’ [113] yields the existence of an asymptotically stable periodic orbit when this equilibrium point is asymptotically unstable. Determining the number of orbits seems out of reach, even if numerical studies suggest that there exists only one which is globally asymptotically stable in this case. Likewise, in the case where this point is asymptotically stable, we have not been able to exclude the existence of periodic orbits, although these have never been detected in this case, up to our knowledge.

For given functions  $f_1, \dots, f_d$ , the stability of the unique equilibrium point depends on the values of  $D$  and  $\alpha$ , as summarised in this second theorem.

**Theorem 2.** *We assume that  $f_1, \dots, f_d \in \mathcal{C}^3(\mathbb{R}_+)$  are non-negative, monotonous and  $\gamma^{1/2}$ -convex functions, and that at least one of them is bounded. Moreover, we assume that an odd number of these function is decreasing, and we recall that  $D$  refers to (1.9). We have the following alternative:*



(i) If  $D < \cos(\frac{\pi}{d})^{-d}$ , then for any  $\alpha \in (\mathbb{R}_+^*)^d$ , this equilibrium point is asymptotically stable, and all the solutions of (1.7) either converge to this point or to a periodic orbit.

(ii) If  $D > \cos(\frac{\pi}{d})^{-d}$ , then there exists  $A_{unst} \subset (\mathbb{R}_+^*)^d$  a non-empty set such that

- If  $\alpha \in A_{unst}$ , then this equilibrium point is asymptotically unstable, and there exists a finite number of periodic solutions, among which at least one is asymptotically stable. Moreover, the set of initial conditions for which the solution converges to a periodic solution is an open subset of  $\mathbb{R}_+^d$ , and its complement, which is the set of initial conditions for which the solution converges to the equilibrium point, has Lebesgue measure zero.
- If  $\alpha \in \overline{A_{unst}}^C$ , this equilibrium point is asymptotically stable, and all the solutions of (1.7) either converge to this point or to a periodic orbit.

The condition  $D \geq \cos(\frac{\pi}{d})^{-d}$  seems necessary for the existence of orbits: as an example, one of the simplest models with three variables and one decreasing regulatory function ( $d = 3$ ,  $n = 1$ ), with two linear functions and one decreasing Hill function ( $x \mapsto \frac{1}{1+x^r}$ ) has a cyclic behaviour only for  $r \geq 8 = \cos(\frac{\pi}{3})^{-3}$  [126, 137].

## 1.2 PDE models for cell populations

In the first section of this introduction, we explained how ODE systems could account for the phenotypic evolution of a cell. The purpose now is to develop Partial Differential Equations (PDEs) describing the dynamics of a population over time. These PDE models are structured by a continuous variable, which is typically a phenotypic trait  $x \in \mathbb{R}^d$  such as the concentration of several molecules of interest, as in the first part of this introduction. We thus study the evolution of a function  $x \mapsto u(t, x)$  over time, which represents the density of cells of trait  $x$ .

Throughout this section, we will assume that the evolution of the phenotypic trait of a single cell is driven by a given ODE

$$\dot{x} = f(x). \quad (1.12)$$

Starting from this model, let us see how we can build a PDE model for a cell population.

### 1.2.1 Construction of the models

**Advection equation.** Let us consider a population of cells initially composed of  $m_1$  cells of phenotypic trait  $x^{1,0} \in \mathbb{R}^d$ ,  $m_2$  cells of phenotypic trait  $x^{2,0} \in \mathbb{R}^d, \dots$ , and  $m_N$  cells of phenotypic trait  $x^{N,0} \in \mathbb{R}^d$ . We describe this initial population with the sum of weighted Dirac masses

$$\mu^0 := \sum_{i=1}^N m_i^0 \delta_{x^{i,0}}. \quad (1.13)$$

Thus, the location of each Dirac mass indicates a possible phenotypic trait, and each associated weight the number of cells of this phenotypic trait.

We momentarily assume that these cells evolve independently of each other, and only according to ODE (1.12). This population can then be described at any time  $t \geq 0$  by the measure

$$\mu^t := \sum_{i=1}^N m_i \delta_{x^i(t)}, \quad (1.14)$$

where, for all  $i \in \{1, \dots, N\}$ ,  $x^i$  is the solution of

$$\begin{cases} \dot{x}^i(t) = f(x^i(t)) & t \geq 0 \\ x^i(0) = x^{i,0} \end{cases}.$$

Another way to characterise  $(\mu^t)_{t \geq 0}$  can be obtained by noting that it satisfies, for all  $t \geq 0$

$$\forall \phi \in \mathcal{C}_c^1(\mathbb{R}^d), \quad \frac{d}{dt} \int_{\mathbb{R}^d} \phi(x) d\mu^t(x) = \int_{\mathbb{R}^d} f(x) \cdot \nabla \phi(x) d\mu^t(x), \quad (1.15)$$

(where  $\mathcal{C}_c^1(\mathbb{R}^d)$  denotes the set of functions in  $\mathcal{C}^1(\mathbb{R}^d)$  with a compact support) together with the initial condition  $\mu^0$  (1.13).

This characterisation allows to define solutions in a much larger space. Indeed, if  $f \in \mathcal{C}^1(\mathbb{R}^d, \mathbb{R}^d)$ , solutions of (1.15) can be found in the dual space of  $\mathcal{C}_c(\mathbb{R}^d)$ , which corresponds to the space of Radon measures [15, 55].

We are particularly interested in the case where  $\mu^0$  has a density in  $L^1(\mathbb{R}^d) \cap \mathcal{C}^1(\mathbb{R}^d)$  with respect to the Lebesgue measure. In this case, problem (1.15) is equivalent to the *advection equation* defined by

$$\begin{cases} \partial_t u(t, x) + \nabla \cdot (f(x)u(t, x)) = 0 & t \geq 0, x \in \mathbb{R}^d \\ u(0, \cdot) = u^0 \end{cases}, \quad (1.16)$$

in the sense that  $u(t, \cdot)$  is a solution of (1.15) if and only if it is a solution of (1.16).

**Advection-selection equation.** The advection equation does not take into account other phenomena than internal regulations described by ODE (1.12). As such, it does not model anything more than this ODE applied to a large number of cells distributed according to a density.

We would now like to take into account cell growth. It means that in the discrete description  $\mu^t = \sum_{i=1}^N m_i(t) \delta_{x^i(t)}$ , the terms ‘ $m_i$ ’ are not longer constant, but also solve an ODE, which generically writes

$$\forall i \in \{1, \dots, N\}, t \geq 0, \quad \dot{m}_i(t) = R(x^i(t), m(t))m_i(t), \quad (1.17)$$

where for all  $x \in \mathbb{R}^d$ ,  $R(x, m)$  denotes the *net growth rate* of the cells of trait  $x$ . This term does not only depend on  $x$  but also on  $m = (m_1, \dots, m_N)$ , which means that the growth rates are influenced by the population size. Such a choice is motivated by the fact that cells are in competition for resources. Hence,  $R(x, m)$  can be interpreted as the fitness of individuals with trait  $x$  inside the environment created by the total population. A model usually used for the population size of homogeneous population is *Verhulst’s logistic equation* [116, 160]: denoting  $U$  the size of the population, the latter is described by the ODE

$$\dot{U}(t) = r \left( 1 - \frac{U(t)}{K} \right) U(t), \quad (1.18)$$

where  $r > 0$  is called the *birth rate* of the population and  $K > 0$  the *carrying capacity*, which corresponds the maximum number of individuals that the environment can support. The net growth rate of this model at any time  $t \geq 0$  is given by  $r \left( 1 - \frac{U(t)}{K} \right)$ , and corresponds to the birth rate  $r$  from which we subtract a death rate  $\frac{r}{K}U(t)$ , which is assumed to be proportional to the size of the population.

Based on this model, the net growth rate for a heterogeneous population typically writes

$$R(x, m) = r(x) - d \sum_{j=1}^N m_j, \quad (1.19)$$

where for all  $x \in \mathbb{R}^d$ ,  $r(x)$  represents the growth rate of cells of trait  $x$ , and  $d \sum_{j=1}^N m_j$  the death rate, which is proportional to the total population size given by  $\sum_{j=1}^N m_j$ .

In order to more precisely quantify the contribution of each phenotypic trait to the death rate, this model can be generalised with a growth term of the shape

$$R(x, m) = r(x) - \sum_{j=1}^N \psi(x^j) m_j, \quad (1.20)$$

where the function  $\psi$  is positive and bounded. Note that (1.19) is a particular case of (1.20), with  $\psi \equiv d$ .

To summarise, with this model, a cell population initially represented by the measure  $\mu^0$  (1.13) is described at any time  $t \geq 0$  by the measure

$$\mu^t = \sum_{i=0}^N m_i(t) \delta_{x^i(t)}, \quad (1.21)$$

where for all  $i \in \{1, \dots, N\}$ ,  $x^i$  and  $m_i$  are the solutions of

$$\begin{cases} \dot{x}^i(t) = f(x^i(t)) \\ \dot{m}_i(t) = \left( r(x^i(t)) - \sum_{j=0}^N \psi(x^j(t)) m_j(t) \right) m_i(t), \quad t \geq 0 \end{cases} \quad (1.22)$$

As previously, one can check that  $(\mu^t)_{t \geq 0}$  solves

$$\forall \phi \in \mathcal{C}_c(\mathbb{R}^d), \quad t \geq 0, \quad \begin{cases} \frac{d}{dt} \int_{\mathbb{R}^d} \phi(x) d\mu^t(x) = \int_{\mathbb{R}^d} (f(x) \cdot \nabla \phi(x) + r(x) - \rho_\psi(t)) d\mu^t(x) \\ \rho_\psi(t) = \int_{\mathbb{R}^d} \psi(x) d\mu^t(x) \end{cases} \quad (1.23)$$

Once again, if the initial condition  $\mu^0 = u^0 \in L^1(\mathbb{R}^d) \cap \mathcal{C}(\mathbb{R}^d)$ , then  $u$  is a solution of (1.23) if and only if it is a solution of the *advection-selection equation* defined by

$$\begin{cases} \partial_t u(t, x) + \nabla \cdot (f(x) u(t, x)) = (r(x) - \rho_\psi(t)) u(t, x) \\ \rho_\psi(t) = \int_{\mathbb{R}^d} \psi(x) u(t, x) dx \\ u(0, \cdot) = u^0 \end{cases} \quad (1.24)$$

**Advection-selection-mutation equation.** The advection-selection model does not consider mutations. In other words, when cells divide, the daughter cells are assumed to adopt the same phenotype as the mother cell. Now, let us assume that this is not the case, and let us denote, when a cell of trait  $x \in \mathbb{R}^d$  divides

- $\nu(x) \in [0, 1]$  the probability that the daughter cells keeps the phenotype  $x$ .
- $y \mapsto M(x, y)$  the density corresponding to the probability that the daughter cells are of phenotype  $y$ , given that it is not of phenotype  $x$ .

It is now necessary to sum cells from other traits which mutate and adopt trait  $x$  in models (1.24), in addition to the cells that do not mutate [23]. Thus, we add a integral term to get the *Advection-selection-mutation equation*

$$\begin{cases} \partial_t u(t, x) + \nabla \cdot (f(x)u(t, x)) = (\nu(x)r(x) - \rho_\psi(t)) u(t, x) + \int_{\mathbb{R}^d} (1 - \nu(y))r(y)M(x, y)u(t, y)dx \\ \rho(t) = \int_{\mathbb{R}^d} \psi(x)u(t, x)dx \\ u(0, \cdot) = u^0 \end{cases} \quad (1.25)$$

This thesis contributes to the understanding of these models in two ways:

- By the theoretical study of the long-time behaviour of the advection-selection equation (1.24), detailed in Chapter 4.
- By developing a numerical scheme that belongs to the family of particle methods, allowing to approach the solutions of advection-selection-mutation equations of the type (1.25). This method is developed in Chapter 5.

### 1.2.2 Long-time behaviour of the advection-selection advection

We are interested in the long-time behaviour (when  $t \rightarrow +\infty$ ) of the solutions of equation (1.24). In order to reduce the number of parameters, let us note that  $(t, x) \mapsto u(t, x)$  is a solution of this equation if and only if  $(t, x) \mapsto v(t, x) := \psi(x)u(t, x)$  satisfies the problem

$$\begin{cases} \partial_t v(t, x) + \nabla \cdot (f(x)v(t, x)) = \left( r(x) + \frac{\nabla \psi(x)}{\psi(x)} \cdot \nabla f(x) - \rho(t) \right) v(t, x) \\ \rho(t) = \int_{\mathbb{R}^d} v(t, x)dx \\ v(0, \cdot) = \psi(\cdot)u^0 \end{cases} \quad (1.26)$$

Thus, the function  $\psi$  can be incorporated into the growth term, and we will therefore limit ourselves to studying the problem

$$\begin{cases} \partial_t u(t, x) + \nabla \cdot (f(x)u(t, x)) = (r(x) - \rho(t)) u(t, x) \\ \rho(t) = \int_{\mathbb{R}^d} u(t, x)dx \\ u(0, \cdot) = u^0 \end{cases}, \quad (1.27)$$

which depends on three parameters: the initial distribution  $u^0$ , the advection function  $f$ , and the growth function  $r$ .

In particular, we are interested in elucidating whether solutions converge to a sum of weighted Dirac masses (in which case we will speak of *concentration*), which means that only a finite number of trait is selected, or if, on the contrary, they converge to a regular measure, which corresponds to a continuous set of traits being preserved over time.

Since we expect to obtain either one of these two regimes, the space of Radon measures is a natural setting. More precisely, we will use the notion of weak-\* convergence in this space (denoted  $\rightharpoonup$ ).

In order to develop some intuition as to the role played by the advection term and the growth term, we will recall known results about the long-time behaviour of simpler models considering only one of these two phenomena: the advection equation, which writes  $\partial_t u(t, x) + \nabla \cdot (f(x)u(t, x)) = 0$ , and the selection equation, which corresponds to PDE (1.27) with  $f \equiv 0$ .

**Long-time behaviour of the advection equation.** We recall that the advection equation writes

$$\begin{cases} \partial_t u(t, x) + \nabla \cdot (f(x)u(t, x)) = 0 \\ u(t, \cdot) = u^0 \end{cases} \quad (1.28)$$

and we assume that  $f$  is a globally Lipschitz continuous vector field, and that  $u^0 \in \mathcal{C}^1(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$ ,  $u^0 \geq 0$ .

As shown by (1.15), this advection equation is the PDE equivalent to the ODE (1.12). Thus, we expect to find the same type of asymptotic behaviour. The link between these equations is done by means of the so-called *method of characteristics* [51]: by denoting the flow (also called characteristic curves) of (1.12) by  $X$ , defined as the solution of

$$\begin{cases} \dot{X}(t, y) = f(X(t, y)) \\ X(0, y) = y, \end{cases} \quad (1.29)$$

and by computing

$$\frac{d}{dt} u(t, X(t, y)) = -\nabla \cdot f(X(t, y)) u(t, X(t, y)),$$

we get an explicit expression for  $u$ , given for all  $y \in \mathbb{R}^d$ ,  $t \geq 0$  by

$$u(t, X(t, y)) = e^{-\int_0^t \nabla \cdot f(X(s, y)) ds} u^0(y).$$

Here, we will focus on the convergence towards equilibrium points, and leave aside the other types of attractors: thus, we make the hypothesis that for almost every  $y \in \text{supp}(u^0)$ ,  $X(\cdot, y)$  either converges, or  $\|X(t, y)\|$  goes to  $+\infty$  when  $t$  goes to  $+\infty$ . This assumption always holds in dimension 1 ( $x \in \mathbb{R}$ ), since any solution of (1.29) is then monotonous (possibly constant). Moreover, we assume that  $f$  has a finite number of roots. Then, by denoting  $\bar{x}^1, \dots, \bar{x}^m$  the roots of  $f$ ,  $B(\bar{x}^1), \dots, B(\bar{x}^m)$  their respective basins of attraction, and for all  $i \in \{1, \dots, M\}$ ,  $\bar{m}_i := \int_{B(\bar{x}^i)} u^0(y) dy$ , one can prove that

$$u(t, \cdot) \xrightarrow[t \rightarrow +\infty]{} \sum_{i=1}^M \bar{m}_i \delta_{\bar{x}^i}.$$

In other words,  $u(t, \cdot)$  concentrates around the equilibrium points of (1.12) having a basin of attraction of positive Lebesgue measure.

This result is expected biologically: indeed, this model only includes the differentiation phenomenon described by the vector field  $f$ : thus, cells contained in the basins of attraction of each of the equilibrium points (whose quantity is given by the associated ' $\bar{m}_i$ ') move toward this point.

**Long-time behaviour of the selection equation.** The selection equation writes

$$\begin{cases} \partial_t u(t, x) = (r(x) - \rho(t)) u(t, x) \\ \rho(t) = \int_{\mathbb{R}^d} u(t, x) dx \\ u(0, \cdot) = u^0 \end{cases}, \quad (1.30)$$

and corresponds to model (1.27) in the absence of advection ( $f \equiv 0$ ), which means that the cell trait of a given cell is neither influenced by the environment nor by internal regulations.

Yet, cells of different traits are nevertheless in competition with the whole population, *via* the non-local term  $\rho$ , and we thus expect that only the most efficient phenotypic traits

survive over time. This is indeed what happens, since the solution concentrates around the points where  $r$  reaches its maximum [131]. This problem has been studied in detail in [106], and the following result has been proved: denoting  $r_M$  the maximum of  $r$  over  $\text{supp}(u^0)$  and assuming that there exist  $\bar{x}^1, \dots, \bar{x}^N$  such that  $\underset{\text{supp}(u^0)}{\text{argmax}}(r) = \{\bar{x}^1, \dots, \bar{x}^N\}$

and for all  $i \in \{1, \dots, N\}$ ,  $\bar{x}^i$  is nondegenerate,

$$u(t, \cdot) \underset{t \rightarrow +\infty}{\rightharpoonup} r_M \sum_{i=1}^N a_i \delta_{\bar{x}^i},$$

where the weights  $a_1, \dots, a_N$  sum to 1 and depend on the value of  $u^0$  and on the concavity of  $r$  at the points  $\bar{x}^1, \dots, \bar{x}^N$ , respectively.

Hence, both the advection and selection equations, under reasonable assumptions, lead to convergence to sums of weighted Dirac masses. Nevertheless, the points on which the population concentrates can be different. To summarise:

- The advection term pushes the solution toward the roots of  $f$  with a non-negligible basin of attraction (such as asymptotically stable equilibrium points),
- The growth term pushes the solution towards the points where  $r$  reaches its maximum.

Guessing the long term behaviour of the advection-selection advection is hardly obvious since it is difficult to predict which of these two phenomena will prevail. We answer the question by showing that solutions of the advection-selection equation can either converge to a weighted Dirac mass at a root of  $f$ , or to a regular function.

The main results we have developed are only valid in the one-dimensional case: therefore, we now assume that  $x \in \mathbb{R}$ . Each results requires some minimal regularity hypotheses which we omit here but are given in detail in Chapter 4.

**The advection-selection equation with a single root.** Let us start with a simple hypothesis satisfied by  $f$ : assume that there exists a closed and bounded interval of  $\mathbb{R}$ , denoted  $I$ , such that  $\text{supp}(u^0) \subset I$ , that there exists a unique  $\bar{x} \in I$  such that  $f(\bar{x}) = 0$ , and that  $f$  is decreasing in a neighbourhood of  $\bar{x}$ . These hypotheses imply that there exist  $a, b \in \mathbb{R} \setminus \text{supp}(u^0)$ , with  $a < b$  such that  $f > 0$  on  $[a, \bar{x})$  and  $f < 0$  on  $(\bar{x}, b]$ , and thus, in particular, that  $t \mapsto X(t, a)$  is increasing,  $t \mapsto X(t, b)$  decreasing, and that these two functions converge to  $\bar{x}$ .

As for the advection equation, we can solve the problem directly with the method of characteristics in this case. Indeed, by computing

$$\frac{d}{dt} u(t, X(t, y)) = (r(X(t, y)) - f'(X(t, y)) - \rho(t)) u(t, X(t, y)),$$

we get, for any  $x \in \mathbb{R}$ ,

$$u(t, X(t, y)) = e^{\int_0^t r(X(s, y)) - f'(X(s, y)) - \rho(s) ds} u^0(y). \quad (1.31)$$

In particular, since, by hypothesis,  $u^0(y) = 0$  for all  $y \leq a$  and all  $y \geq b$ , and since,  $X(\cdot, a)$  and  $X(\cdot, b)$  are respectively increasing and decreasing, then for all  $t \geq 0$

$$\text{supp}(u(t, \cdot)) \subset [X(t, a), X(t, b)].$$

Since  $X(\cdot, a)$  and  $X(\cdot, b)$  both converge to  $\bar{x}$ , the support of  $u(t, \cdot)$  formally becomes  $\{\bar{x}\}$  when  $t$  goes to  $+\infty$ . Using this argument, we show that

$$u(t, \cdot) \underset{t \rightarrow +\infty}{\rightharpoonup} r(\bar{x}) \delta_{\bar{x}}.$$

Hence, in the presence of a single equilibrium point for (1.12) which is asymptotically stable, the solution concentrates onto this point, regardless of the growth function  $r$ .

**The advection-selection equation on a segment between two roots of  $f$ .** Surprisingly, and contrary to the advection equation, this result no longer holds when adding an unstable equilibrium point. Let us now assume that the function  $f$  has two roots, that we denote  $\bar{x}_u$  and  $\bar{x}_s$ , (with  $\bar{x}_u < \bar{x}_s$ , the case  $\bar{x}_u > \bar{x}_s$  being symmetric), such that

$$f(\bar{x}_u) = f(\bar{x}_s) = 0, \quad f'(\bar{x}_u) > 0, \quad f'(\bar{x}_s) < 0, \quad f > 0 \quad \text{on} \quad (\bar{x}_u, \bar{x}_s), \quad (1.32)$$

which ensures that  $\bar{x}_u$  and  $\bar{x}_s$  are respectively asymptotically unstable and asymptotically stable for ODE (1.12), and let us assume that

$$\text{supp}(u^0) \subset [x_u, x_s]. \quad (1.33)$$

Let us note that the method of characteristics proves that for all  $t \geq 0$ ,  $u(t, \cdot) \subset [x_u, x_s]$ . Moreover, if  $\bar{x}_u$  is not in the support of  $u^0$  ( $\text{supp}(u^0) \subset [x_u + \varepsilon, x_s]$  for some  $\varepsilon > 0$ ), then the reasoning of the previous paragraph can be applied, and  $u(t, \cdot)$  converges to  $r(\bar{x}_s)\delta_{\bar{x}_s}$ . In contrast, if  $\bar{x}_u$  is an equilibrium point of (1.12), *i.e.* for all  $t \geq 0$ ,  $X(t, \bar{x}_u) = \bar{x}_u$ , then  $\bar{x}_u \in \text{supp}(u^0)$  implies that for all  $t \geq 0$ ,  $x_u \in \text{supp}(u(t, \cdot))$ . This does not mean, however, that the solution does not converge to a weighted Dirac mass on  $\bar{x}_s$ . In fact, we show that, depending on the value of  $r$  and  $u^0$  on  $\bar{x}_u$ ,  $u(t, \cdot)$  can either converge to a weighted Dirac mass on  $\bar{x}_s$ , or to a regular function in  $L^1(\bar{x}_u, \bar{x}_s)$ . Precisely, we show the following result:

**Theorem 3.** *Let us assume that conditions (1.32) and (1.33) hold, and that  $u^0(\bar{x}_u) > 0$ . The following alternative holds:*

- (i) *If  $r(\bar{x}_u) - f'(\bar{x}_u) < r(\bar{x}_s)$ , then  $\rho$  converges to  $r(\bar{x}_s)$ , and  $u(t, \cdot) \xrightarrow[t \rightarrow +\infty]{} r(\bar{x}_s)\delta_{\bar{x}_s}$ .*
- (ii) *If  $r(\bar{x}_u) - f'(\bar{x}_u) > r(\bar{x}_s)$ , then  $\rho$  converges to  $r(\bar{x}_u) - f'(\bar{x}_u)$ , and  $u(t, \cdot) \xrightarrow[t \rightarrow +\infty]{} \bar{u}_\infty$  in  $L^1(x_u, x_s)$ , where  $\bar{u}_\infty(x) = D e^{\int_{\bar{x}_u}^x \frac{\tilde{r}(s) - \tilde{r}(\bar{x}_u)}{f(s)} ds}$ , with  $\tilde{r} = r - f'$ , and  $D > 0$  is such that  $\int_{\bar{x}_u}^{\bar{x}_s} \bar{u}_\infty(x) dx = r(\bar{x}_u) - f'(\bar{x}_u)$ .*

If  $\bar{x}_u \in \text{supp}(u^0)$  and  $u^0(\bar{x}_u) = 0$ , *i.e.*  $\bar{x}_u$  is at the boundary of  $\text{supp}(u^0)$ , then we show a similar result: the ‘toggle value’ between the two regimes of convergence, as well as the possible regular function which is reached depends on the asymptotic behaviour of  $u^0$  at the neighbourhood of  $\bar{x}_u$ .

The method used to prove these two theorems is based on the following two ideas:

- (i) Let us denote  $Y$  the inverse of the characteristic curves, *i.e.* for all  $t \geq 0$ ,  $Y(t, \cdot) = X(t, \cdot)^{-1} = X(-t, \cdot)$ , which satisfies the ODE

$$\begin{cases} \dot{Y}(t, x) = -f(Y(t, x)) \\ Y(0, x) = x \end{cases}. \quad (1.34)$$

Then, by replacing  $y$  by  $Y(t, x)$  in (1.31), we get

$$u(t, x) = u^0(Y(t, x)) e^{\int_0^t r(Y(s, x)) - f'(Y(s, x)) ds} e^{-\int_0^t \rho(s) ds}. \quad (1.35)$$

We call this formula *semi-explicit expression* of  $u$  since it relates  $u$  with  $\rho$ , and ensures that determining the long-time behaviour of  $\rho$  yields that of  $u$ .

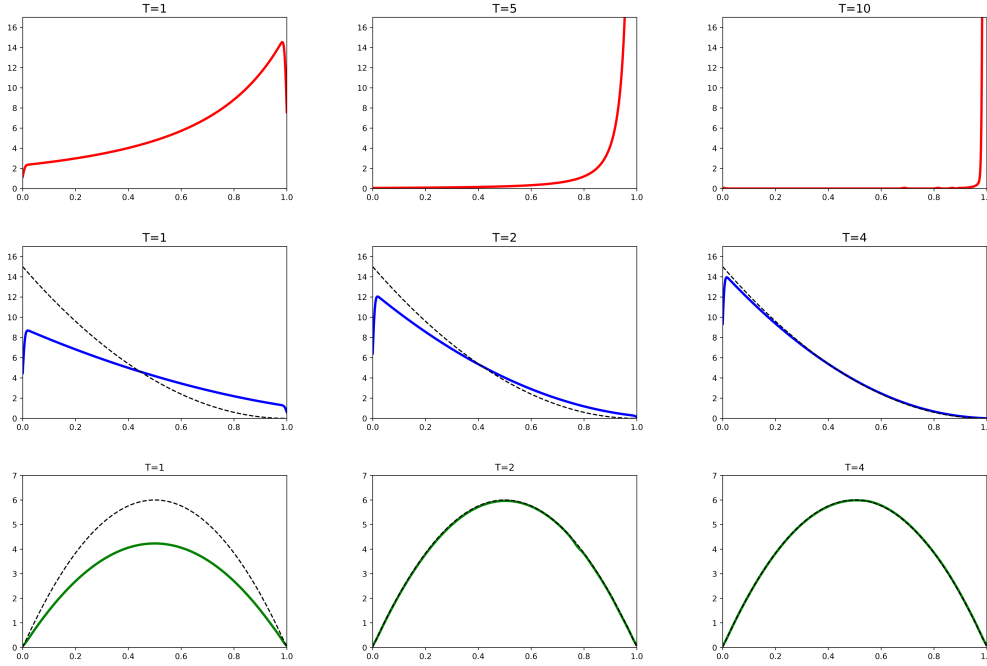


Figure 1.3: Possible regimes of convergence of the solutions of (1.27): In all three cases, we have chosen  $f(x) = x(1-x)$ , and we work on the segment  $[0, 1]$  (hence  $\bar{x}_u = 0$ ,  $\bar{x}_s = 1$ ). The three figures above (in red) show the time evolution of the solution in the case where  $u^0 \equiv 6$  and  $r(x) = 6 - 0.5x$  (and thus  $5.5 = r(1) > r(0) - f'(0) = 5$ ), which implies, according to Proposition 3, that the solution converges to a weighted Dirac mass at 1. The three figures in the middle (in blue) show the time evolution of the solution in the case where  $u^0 \equiv 6$  (and thus  $u^0(0) > 0$ )  $r(x) = 6 - 4x$ , (and thus  $2 = r(1) < r(0) - f'(0) = 5$ ), which implies that the solution converges to a continuous function in  $L^1$ . The three figures below (in green) show the time evolution with the same functions  $r$  as in the previous case, but with  $u^0(x) = x$  (and thus  $u^0(0) = 0$ ), which leads to convergence to a different limit function. In the two last figures, the black dashed curve represents the limit function.

- (ii) By integrating the semi-explicit expression of  $u$  (1.35) on  $(\bar{x}_u, \bar{x}_s)$ , taking the logarithm, and then differentiating on both sides the equality obtained, we show that  $\rho$  satisfies the non-autonomous logistic equation

$$\dot{\rho}(t) = (R(t) - \rho(t)) \rho(t), \quad (1.36)$$

where

$$R(t) = \frac{\dot{S}(t)}{S(t)}, \quad \text{with} \quad S(t) = \int_{\bar{x}_u}^{\bar{x}_s} u^0(Y(t, x)) e^{\int_0^t r(Y(s, x)) - f'(Y(s, x)) ds} dx. \quad (1.37)$$

We call the function  $R$  *time-dependent carrying capacity*, because it plays an analogous role to that of  $K$  in the logistic model (1.18). Indeed, if  $R$  converges to a certain  $\bar{R}$  (with an exponential speed), then  $\rho$  converges to  $\bar{R}$  (with an exponential speed). Because of the definition of  $R$ , computing its limit means studying parameter-dependent integrals. This analysis has been possible in the one-dimensional case, thanks to the changes of variables ‘ $s' = X(s, y)$ ’ and ‘ $s' = Y(s, y)$ ’: Since  $f > 0$  on  $(\bar{x}_u, \bar{x}_s)$ , for all  $y \in (x_u, x_s)$ ,  $s \mapsto X(s, y)$  and  $s \mapsto Y(s, y)$  are indeed  $\mathcal{C}^1$  diffeomorphism from  $(0, t)$  onto  $(y, X(t, y))$ . Unfortunately, this change of variables does not hold in higher dimension, preventing us to generalise the approach.



**The advection-selection equation in the general case.** The method presented in the previous paragraph can be adapted in order to study the long-time behaviour of (1.26) under more general hypotheses.

Let us consider that  $f$  has exactly  $N$  roots, that we denote  $\bar{x}_1, \dots, \bar{x}_N$ . For simplicity, we assume that  $\text{supp}(u^0) \subset [\bar{x}_1, \bar{x}_N]$ , which guarantees that the support of  $u(t, \cdot)$  remains in this segment for all  $t \geq 0$ <sup>2</sup>.

Let us denote, for all  $i \in \{1, \dots, N-1\}$ ,  $\rho_i(t) := \int_{\bar{x}_i}^{\bar{x}_{i+1}} u(t, x) dx$ , which corresponds to the population size in each segment  $[\bar{x}_i, \bar{x}_{i+1}]$ . For all  $i \in \{1, \dots, N-1\}$ ,  $\rho_i$  satisfies the ODE

$$\dot{\rho}_i(t) = (R_i(t) - \rho(t)) \rho_i(t), \quad (1.38)$$

where  $\rho(t) = \int_{\bar{x}_1}^{\bar{x}_N} u(t, x) dx = \sum_{i=1}^N \rho_i(t)$ , and

$$R_i(t) = \frac{\dot{S}_i(t)}{S_i(t)}, \quad \text{with} \quad S_i(t) = \int_{\bar{x}_i}^{\bar{x}_{i+1}} u^0(Y(t, x)) e^{\int_0^t r(Y(s, x)) - f'(Y(s, x)) ds} dx. \quad (1.39)$$

As previously, the limit of the functions ‘ $R_i$ ’ can be computed for a large range of functions, and depends on the values of  $r$ ,  $f'$  and  $u^0$  at the equilibrium points  $\bar{x}_1, \dots, \bar{x}_N$ . In particular, if  $f'(x_i) > 0$ ,  $f'(x_{i+1}) < 0$ , and  $u^0(\bar{x}_i) > 0$ , we find ourselves in the same situation as in Theorem 3, and thus

- If  $r(\bar{x}_i) - f'(\bar{x}_i) < r(\bar{x}_{i+1})$ , then  $R_i$  converges to  $r(\bar{x}_{i+1})$ .
- If  $r(\bar{x}_i) - f'(\bar{x}_i) > r(\bar{x}_{i+1})$ , then  $R_i$  converges to  $r(\bar{x}_i) - f'(\bar{x}_i)$ .

In the cases where  $u^0(\bar{x}_i) = 0$ , a similar alternative holds, which is fully detailed in Chapter 4. The case where  $f'(\bar{x}_i) < 0$  and  $f'(\bar{x}_{i+1}) > 0$ , is symmetric, and the limit is thus obtained by switching ‘ $\bar{x}_i$ ’ and ‘ $\bar{x}_{i+1}$ ’.

Because of the ODE satisfied by the functions ‘ $\rho_i$ ’, if all the functions ‘ $R_i$ ’ converges to a certain  $\bar{R}_i \geq 0$ , and if for all  $i \in \{1, \dots, N-1\}$ ,  $\rho_i(0) > 0$ , we get that

- $\rho$  converges to  $\bar{\rho}_M := \max(\bar{R}_1, \dots, \bar{R}_{N-1})$ .
- If  $\bar{R}_i < \bar{\rho}_M$ , then  $\rho_i$  converges to zero.

Moreover, if all the functions ‘ $R_i$ ’ converge toward their limit with an exponential speed, then  $\rho$  converges to  $\bar{\rho}_M$  with an exponential speed. As a result, we can use the semi-explicit expression of  $u$  (1.35) to determine the long-time behaviour of the solutions of (1.26). We prove that, depending on the values of  $r$ ,  $f$  and  $u^0$  at the roots  $f$ ,  $u(t, \cdot)$  either converges to a weighted Dirac mass at a stable equilibrium point, or to a regular function with a support between two roots of  $f$ .

**Open problems.** The method used to determine the asymptotic behaviour of the solutions of (1.27) is entirely based on the computation of the limit of the functions ‘ $R_i$ ’. For this reason, we did not manage to deal with some limit cases, such as non-hyperbolic equilibria, *i.e.*  $\bar{x} \in \mathbb{R}$  which satisfy  $f(\bar{x}) = f'(\bar{x}) = 0$ , or the equality case  $r(\bar{x}_u) - f'(\bar{x}_u) = r(\bar{x}_s)$  under the hypotheses of Theorem 3.

In the same way, studying this PDE in higher dimension ( $x \in \mathbb{R}^d$ ,  $d \geq 2$ ) has not been possible with this method. As shown by the semi-explicit expression (1.35), the behaviour

<sup>2</sup>In Chapter 4, we show that the long-time behaviour of this equation can be determined even when this assumption is dropped.

of  $u$  is closely linked to that of the solutions of ODE  $\dot{x} = f(x)$ , and we can therefore expect cyclic behaviours, or even chaotic behaviours (if  $d \geq 3$ ) in this case.

Another interesting perspective would be to study this problem with an additional mutation term (such as a Laplacian or as an integral term). Such problem might be tackled with entropy methods [118], contrary to the present work where the absence of mutations means that the entropy does not dissipate.

### 1.2.3 A particle method for advection-selection-mutation equations

The purpose of Chapter 5 is to develop a numerical scheme in order to approach the solutions of equations such as (1.25). Upon relabeling the parameters, this equation can be rewritten

$$\begin{cases} \partial_t u(t, x) + \nabla \cdot (f(x)u(t, x)) = (r(x) - \rho_\psi(t)) u(t, x) + \int_{\mathbb{R}^d} m(x, y)u(t, y)dy \\ \rho(t) = \int_{\mathbb{R}^d} \psi(x)u(t, x)dx \\ u(0, \cdot) = u^0 \end{cases} \quad (1.40)$$

We have chosen to develop and study a scheme from the family of *particle methods*, which are based on solving ODEs in order to approximate the solution of PDEs. Particle methods are particularly adapted for the equation with an advection term, and have been the subject of several studies for problems which mostly come from physics [46, 81].

**Particle approximation.** The main idea of particle method is to seek a sum of weighted Dirac masses in order to approach the solution of a given equation [27]. Such measure is called *particle solution*, which we denote

$$u^N(t) = \sum_{i=1}^N m_i(t) \delta_{x^i(t)}, \quad (1.41)$$

where the weights  $m_i$  and the points  $x^i$  are solutions of a suitably defined ODE system.

To start, let us take as an example the advection-selection equation studied in Chapter 4 (we do not consider mutations for the moment), which writes

$$\begin{cases} \partial_t u(t, x) + \nabla \cdot (f(x)u(t, x)) = (r(x) - \rho_\psi(t)) u(t, x) \\ \rho_\psi(t) = \int_{\mathbb{R}^d} \psi(x)u(t, x)dx \\ u(0, \cdot) = u^0 \end{cases} \quad .$$

We have seen in section 1.2.1 of this introduction that any solution of this equation also satisfies the weak formulation (1.23)

$$\forall \phi \in \mathcal{C}_c(\mathbb{R}^d) \quad \begin{cases} \frac{d}{dt} \int_{\mathbb{R}^d} \phi(x) d\mu^t(x) = \int_{\mathbb{R}^d} (f(x) \cdot \nabla \phi(x) + r(x) - \rho_\psi(t)) d\mu^t(x) \\ \rho_\psi(t) = \int_{\mathbb{R}^d} \psi(x) d\mu^t(x) \end{cases} ,$$

and that any measure of the form

$$\sum_{i=0}^N m_i(t) \delta_{x^i(t)}, \quad (1.42)$$

solves this problem if for all  $i \in \{1, \dots, N\}$ ,  $t \geq 0$ ,  $x^i$  and  $m_i$  satisfy the ODE

$$\begin{cases} \dot{x}^i(t) = f(x^i(t)) \\ \dot{m}_i(t) = \left( r(x^i(t)) - \sum_{j=0}^N \psi(x^j(t)) m_j(t) \right) m_i(t) \end{cases} \quad (1.43)$$

Hence, since  $u$  and  $\mu$  satisfy the same equation, we expect that if  $\mu^0 \simeq u^0$  (in a sense that will be clarified later), then  $\mu^t$  remains close to  $u(t, \cdot)$  for any  $t \geq 0$ .

When considering a mutation term, as in (1.40), things are a little bit more subtle. In order to approximate the term  $\int_{\mathbb{R}^d} m(x, y) u(t, y) dy$ , we need to introduce a third family of ODEs, solved by what we call the volumes  $w_i$  [46], namely

$$\dot{w}_i(t) = \nabla \cdot f(x^i(t)) w_i(t) \quad t \geq 0. \quad (1.44)$$

Indeed, according to Liouville's formula, for any function  $g \in L^1(\mathbb{R}^d)$  and any  $t \geq 0$ , there holds

$$\int_{\mathbb{R}^d} g(x) dx = \int_{\mathbb{R}^d} g(X(t, y)) W(t, y) dy,$$

where for all  $y \in \mathbb{R}^d$ ,  $X$  and  $W$  satisfy

$$\begin{cases} \dot{X}(t, y) = f(X(t, y)) \\ X(0, y) = y \end{cases} \quad \text{and} \quad \begin{cases} \dot{W}(t, y) = \nabla \cdot f(X(t, y)) W(t, y) \\ W(0, y) = 1. \end{cases} \quad (1.45)$$

We can therefore approximate such an integral with the sum

$$\sum_{i=1}^N g(x^i(t)) w_i(t) \approx \int_{\mathbb{R}^d} g(x) dx, \quad (1.46)$$

where for all  $i \in \{1, \dots, N\}$ ,  $t \geq 0$ ,  $x^i$  and  $w_i$  are solutions of

$$\begin{cases} \dot{x}^i(t) = f(x^i(t)) \\ \dot{w}_i(t) = \nabla \cdot f(x^i(t)) w_i(t). \end{cases} \quad (1.47)$$

The weights of the Dirac masses then write  $m_i = \nu_i w_i$ , where  $\nu_i$  satisfies

$$\dot{\nu}_i(t) = \left( r(x^i(t)) - \nabla \cdot f(x^i(t)) - \sum_{j=0}^N \nu_j(t) w_j(t) \right) \nu_i(t) + \sum_{j=1}^N m(x^i(t), x^j(t)) \nu_j(t) w_j(t). \quad (1.48)$$

The initial condition of this ODE must of course be chosen carefully so that they approximate  $u^0$  properly. Assuming that  $u^0$  has a compact support, a way of choosing these values is to take a finite collection of subsets  $\Omega_i^0 \subset \text{supp}(u^0)$  satisfying

$$\Omega_i^0 \cap \Omega_j^0 = \emptyset, \text{ if } i \neq j, \text{ and } \bigcup_{i \in \{1, \dots, N\}} \Omega_i^0 = \text{supp}(u^0),$$

and to take, for any  $i \in \{1, \dots, N\}$

$$x^{i,0} \in \Omega_i^0, \quad w_i^0 = |\Omega_i^0|, \quad \nu_i^0 = u^0(x^{i,0}) \quad \alpha_i = \nu_i w_i.$$

One can check that this choice ensures that for any  $\phi \in \mathcal{C}_c^0(\mathbb{R}^d)$ ,

$$\sum_{i=1}^N \alpha_i^0 \phi(x^{i,0}) \xrightarrow{N \rightarrow +\infty} \int_{\mathbb{R}^d} \phi(x) u^0(x) dx,$$

which means that

$$\sum_{i=1}^N \nu_i^0 w_i^0 \delta_{x^{i,0}} \xrightarrow{N \rightarrow +\infty} u^0 \quad \text{in the space of Radon measures.}$$

**Regularisation of the particle approximation.** We would now like to find an approximation of  $u(t, \cdot)$  which is a regular function rather than a sum of weighted Dirac masses. In order to do so, the particle solution is regularised through convolution with a so-called *cut-off function*  $\varphi_\varepsilon$ . We obtain a regular approximation of  $u(t, \cdot)$  which writes

$$u_\varepsilon^N(t, x) = \sum_{i=1}^N m_i(t) \varphi_\varepsilon(x - x^i(t)). \quad (1.49)$$

The cut-off functions  $\varphi_\varepsilon$  typically writes  $\varphi_\varepsilon = \frac{1}{\varepsilon^d} \varphi\left(\frac{\cdot}{\varepsilon}\right)$ , where  $\varphi$  satisfies some symmetric and regularity properties, and the parameter  $\varepsilon > 0$  is chosen as functions of  $N$ . Determining the optimal  $\varepsilon > 0$  for a given  $N > 0$  is intricate: if  $\varepsilon$  is too large, then the solution is ‘over-regularised’, and the scheme loses its accuracy. Conversely, if  $\varepsilon$  is too small, then some of the particles will be neglected, and the scheme may fail to converge.

**Main results.** In Chapter 5, we study a more general model than (1.40), which takes into account possible non-local terms for all three terms modelling mutations, growth and advection. This model writes

$$\begin{cases} \partial_t u(t, x) + \nabla_x \cdot (a(t, x, \rho_a(t))u(t, x)) = R(t, x, \rho_g(t))u(t, x) + \int_{\mathbb{R}^d} m(t, x, y, \rho_d(t))u(t, y)dy \\ \rho_l(x, t) = \int_{\mathbb{R}^d} \psi_l(t, x, y)u(t, y)dy, \quad l = a, g, d \\ u(0, \cdot) = u^0 \end{cases} \quad (1.50)$$

The particle method we develop for this model is adapted from [46]: the main novelties with respect to that work is

- to allow for non-local terms, their presence posing significant technical difficulties (especially when it comes to the advection),
- to study asymptotic preserving properties of such schemes. These guarantee that, under specific hypotheses, the long-time behaviour of the solution is preserved at the discrete level.

We first show that this problem is well-posed, *i.e.* that is if the parameters satisfy some regularity properties, then this problem has a unique solution in  $\mathcal{C}(\mathbb{R}_+, L^1(\mathbb{R}^d))$ . We use a classical fixed point argument to prove this result with some regular parameters, before considering a more general family of initial conditions. The main difficulty comes from the non-local advection term, which does not allow for an explicit definition of the flow, and therefore requires further estimates. We then prove that the regularity of  $u(t, \cdot)$  is linked to that of  $u^0$ : more precisely, if  $u^0 \in W^{k, \infty}(\mathbb{R}^d)$  with compact support, then  $u \in \mathcal{C}([0, +\infty), W^{k-1, 1}(\mathbb{R}^d))$ . This result can be improved if the advection is local, *i.e.*  $a(t, x, I) = a(t, x)$ : In this case, for any initial condition in  $W^{k, 1}(\mathbb{R}^d)$ , the solution  $u$  is in  $\mathcal{C}([0, +\infty), W^{k, 1}(\mathbb{R}^d))$ .

After defining the ODE satisfied by the  $x^i, w_i$  and  $\nu_i$  in this general case, we prove the convergence of the particle method toward the solution of (1.50). More precisely, we show the following estimate

$$\|u(t, \cdot) - u_\varepsilon^h(t, \cdot)\|_{L^1(\mathbb{R}^d)} \leq C\left(\varepsilon^r + \left(\frac{h}{\varepsilon}\right)^\kappa + h^\kappa\right) \|u^0\|_{W^{\mu, 1}(\mathbb{R}^d)}, \quad \text{for all } 0 \leq t \leq T,$$

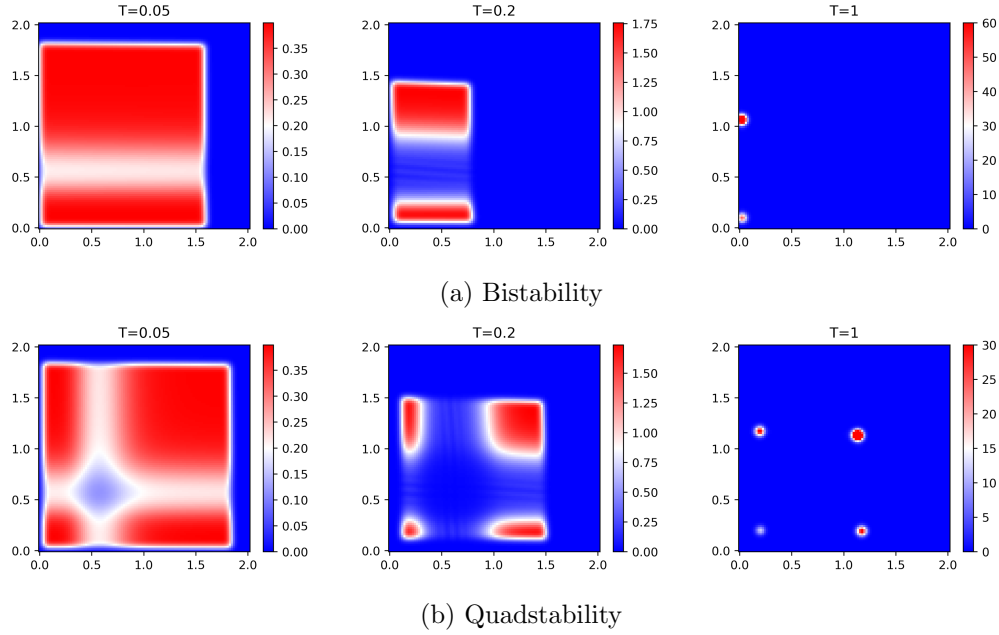


Figure 1.4: Numerical simulations obtained with the particle method for a non-local advection equation studied in [59]. The lines (a) and (b) respectively show the convergence to a single weighted Dirac mass, and a sum of four weighted Dirac masses, and have been obtained by choosing the parameters of Figures 10 and 11 of [59].

which yields a convergence rate of  $h^{\frac{\kappa r}{\kappa+r}}$  after choosing the optimal value of  $\varepsilon(h) \sim h^{\frac{\kappa}{\kappa+r}}$ .

Lastly, we investigate in which cases this method is asymptotic preserving, *i.e.*  $u^N(t)$  and  $u(t, \cdot)$  have the same limit when  $t$  goes to  $+\infty$ . We have seen in section 1.2.2 of this introduction that in the absence of mutations, solutions with close initial conditions may have drastically different long-time behaviours. As an example, we have seen that in the one-dimensional case, if the advection function  $f$  has exactly two roots (that we denote  $\bar{x}_u < \bar{x}_s$ ), and if  $f'(\bar{x}_u) > 0$ ,  $f'(\bar{x}_s) < 0$ , and  $r(\bar{x}_s) - f'(\bar{x}_u) > 0$ , then:

- If  $\bar{x}_u \notin \text{supp}(u^0)$ , then  $u(t, \cdot) \xrightarrow[t \rightarrow +\infty]{} r(\bar{x}_s)\delta_{\bar{x}_s}$ .
- If  $u^0(\bar{x}_u) > 0$ , then  $u(t, \cdot)$  converges to a regular function in  $L^1$ .

Consequently, particle methods cannot be expected to preserve the asymptotics in general. We show, however, that in some specific cases, the scheme is asymptotic preserving. Due to the discrete nature of particular solutions, they tend to concentrate toward attractors and thus to converge towards Dirac masses. However, we highlight conditions which ensure the convergence to both  $u(t, \cdot)$  and  $u^N(t)$  (and thus  $u_\varepsilon^N(t, \cdot)$ ) to the same weighted Dirac mass, which means that the scheme is asymptotic-preserving in this case.

### 1.3 Application to Epithelial-Mesenchymal Plasticity

This last section deals with work carried out within a collaboration between mathematicians at Laboratoire Jacques-Louis Lions and systems biologists from the Cancer Systems Biology Laboratory (Indian Institute of Science at Bangalore), led by Mohit Kumar Jolly.

The aim is to propose PDE models for cells undergoing a specific cell differentiation phenomenon, which plays a crucial role during embryonic development, wound healing, tissue regeneration, but also for the spread of metastases: the Epithelial-Mesenchymal Plasticity [94, 124].

**Epithelial-mesenchymal Plasticity.** Epithelial-mesenchymal transition is a cell differentiation process during which epithelial cells (E), characterised by a tight cell-to-cell adhesion and a low mobility, lose these characteristics to become mesenchymal cells (M), characterised by a low cell-to-cell adhesion and a high mobility [103, 142]. During this process called Epithelial-Mesenchymal Transition (EMT), some cells undergo partial transition, which means that they retain some epithelial traits such as cell-to-cell adhesion, and gain mesenchymal traits, such as migratory properties. These cells are referred to as hybrid (E/M), and have special properties such as the ability to move collectively as clusters [93].

The epithelial–mesenchymal transition is a reversible phenomenon, which means that mesenchymal cells can also differentiate into epithelial ones (MET). Nevertheless, during this reverse process, cells directly become epithelial, without passing through the hybrid state. The notion of epithelial-mesenchymal plasticity refers to the combination of these two phenomena [29].

Transitions between epithelial and mesenchymal states play an important role in tumorigenesis: epithelial cells of the primary tumour can lose their cell-to-cell adhesion and apico-basal polarity, hence becoming prone to migrate individually and invade basement membranes and blood vessels. This then allows them to reach some distant organs to seed micrometastases. During seeding, they undergo the reverse of EMT to regain their epithelial characteristics and form secondary tumours [34]. The hybrid phenotype seems to play a key role in the spread of metastases. Indeed, cancer cells can migrate as dispersed individual mesenchymal cells, or collectively, as hybrid cells. Collective migrations obviate the need for all cells to detect extrinsic signal for migration, enabling them to adapt to different microenvironments [58].

**ODE model for EMP.** The team of Mohit Kumar Jolly has developed a comprehensive expertise in designing ODE models for EMT.

A gene regulatory network composed of six molecules, called *Core EMT Network*, has been proposed in order to model EMT at the level of a single cell [110]. Here, we will use a two-dimensional reduction of this model (developed in the supplementary material of the aforementioned article), which considers interactions between a transcription factor ZEB (denoted  $Z$ ) and a micro-RNA miR-200 (denoted  $\mu_{200}$ ). ZEB promotes EMT whereas miR-200 maintains the epithelial phenotype by targeting ZEB molecules. This model writes:

$$\begin{cases} \dot{\mu}_{200} &= g_{\mu_{200}} H_{Z, \mu_{200}}(Z) H_{S, \mu_{200}}(S) - g_{m_Z} H_{Z, m_Z}(Z) H_{S, m_Z}(S) Q(\mu_{200}) - k_{\mu_{200}} \mu_{200} \\ \dot{Z} &= g_Z g_{m_Z} H_{Z, m_Z}(Z) H_{S, m_Z}(S) P(\mu_{200}) - k_Z Z \end{cases}, \quad (1.51)$$

where

- $g_{\mu_{200}}, g_Z, g_{m_Z}, k_{\mu_{200}}, k_Z$  are positive parameters;
- $H_{Z, \mu_{200}}, H_{S, \mu_{200}}, H_{Z, m_Z}, H_{S, m_Z}$  are shifted Hill functions of the form (1.3);
- $Q$  and  $P$  are complex functions defined in Appendix A.

All relevant parameters are given in Appendix A. The variable  $S$  represents a third molecule, SNAIL, which is seen in our case as an external signal characterising the extra-cellular environment. Within this framework, the three cell states (E, E/M, and M) are characterised as follows:

- The epithelial state corresponds to a high level of miR-200, and a low level of ZEB and SNAIL;

- The mesenchymal state corresponds to a low level of miR-200 and high level of ZEB and SNAIL;
- The hybrid state corresponds to a medium level of miR-200, ZEB and SNAIL.

The transcription factor SNAIL ( $S$ ) activates ZEB and inhibits miR-200, and is thus able to induce the epithelial-mesenchymal transition. Depending on the value of  $S$ , system (1.51) can have 1, 2 or 3 stable equilibrium points, as illustrated by the bifurcation diagram of Figure 1.5 A.

This bifurcation diagram highlights two essential properties of cancer cell populations:

- Heterogeneity*: For intermediate values of  $S$ , cells populations can be highly heterogeneous, with distinct subpopulations corresponding to different states [6].
- Hysteresis*: During EMT (which corresponds to the increase of  $S$ ), epithelial cells become hybrid, and then mesenchymal, while mesenchymal cells directly differentiate into epithelial cells during MET (*i.e.* when  $S$  decreases) [109].

The aim of the collaboration is to develop and simulate PDE models for cell populations undergoing EMT. These models must account for heterogeneity and hysteresis, while in addition incorporating growth and epigenetic instability. This in turn contributes to understanding how these different phenomena interact and shape the dynamics and asymptotics of cell populations.

**Hysteresis.** We begin by reproducing the hysteretic behaviour of the ODE model at the scale of a cell population. We temporarily neglect cell growth and epigenetic instability. As seen in the previous section of this introduction, the PDE equivalent of a given ODE ' $\dot{y} = f(y)$ ' is the advection equation

$$\partial_t u(t, y) + \nabla \cdot (f(t, y)u(t, y)) = 0.$$

Thus, by denoting  $F(\mu_{200}, Z, S)$  the right-hand side of (1.51), we take as structure variable  $y = (\mu_{200}, Z)$ , and as advection function  $f(t, y) = F(y, S(t))$ , where  $S$  is a function which linearly increases and then decreases at the same rate, as shown in Figure 1.5 A. Figure 1.5 Ci highlights the asymmetric behaviour of the population. When  $S$  increases, the population, initially uniformly distributed, quickly becomes uniformly epithelial, then becomes hybrid and finally mesenchymal; when  $S$  decreases, the population transits from the mesenchymal state to the epithelial one without passing through the hybrid state.

We then slightly modify our framework in order to take into account heterogeneity within the population, and more specifically the fact that the signal  $S$  can be interpreted in a different way for each cell. In order to do so, we incorporate  $S$  into the structure variable, which becomes  $y = (\mu_{200}, Z, S)$ . The variation of the level of SNAIL is expressed in the advection term, which becomes  $f(t, y) = (F(y), f_S(t))$ , where  $f_S$  is the step function corresponding to the derivative of  $S$  in Figure 1.5 B, implying that the level of  $S$  of each cell increases and decreases at the same rate. As initial population, we take  $u^0(\mu_{200}, Z, S) = \frac{1}{\sigma} G(\frac{S-160K}{\sigma})$ , with  $\sigma = 20K$ , and  $G$  representing the Gaussian function (see Glossary). The results are shown in Figure 1.5: contrarily to the previous simulations, we observe, in addition to heterogeneity in the population, the coexistence of several traits for intermediate values of  $S$ . Hysteresis is still present: when  $S$  increases, we see the co-existence of the three states, whereas no hybrid population appears for low values of  $S$  when it decreases.

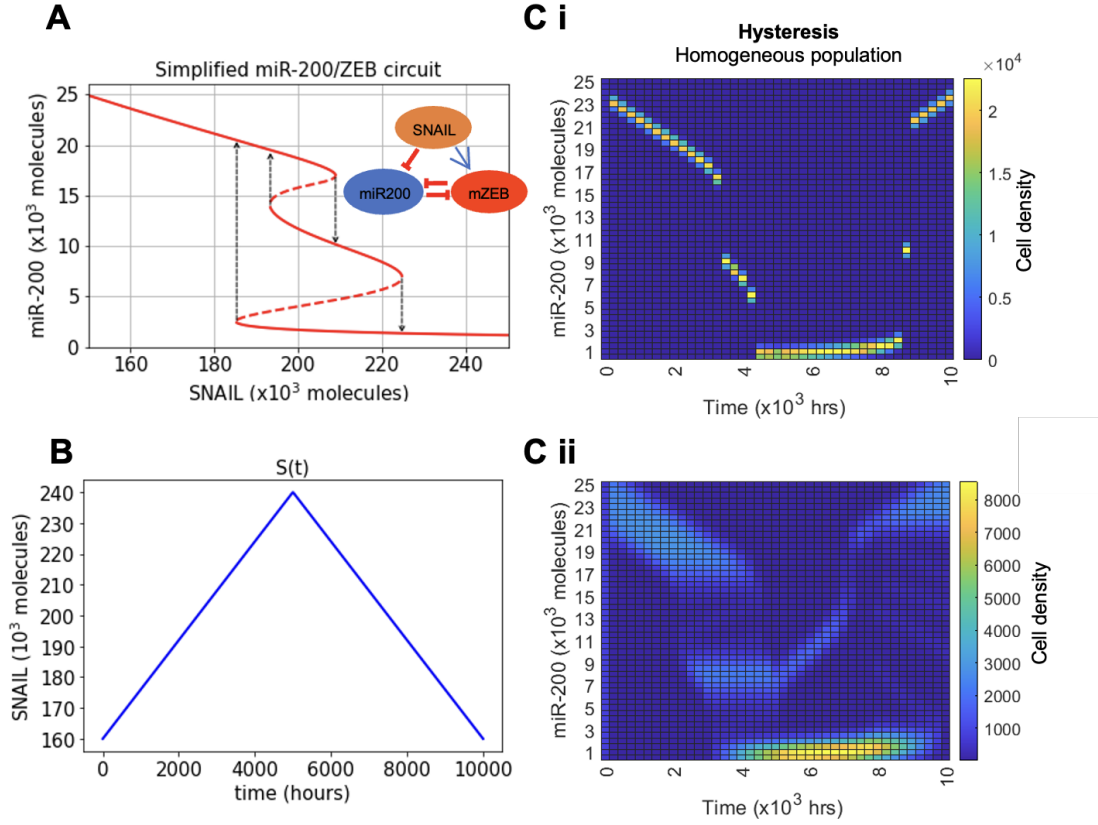


Figure 1.5: The advection equation reproduces hysteresis at the population level. A) EMT gene regulatory network (inset) and the bifurcation of cell states resulting from the network dynamics with increasing input signal (SNAIL) levels, B) Variation in the external input (SNAIL) levels with time to capture hysteresis in C. C) Hysteresis (non-symmetric trajectories) in cell state transition during one cycle of EMT and MET by varying SNAIL levels as shown in Figure 1B while considering – i) Homogeneous population and ii) Heterogeneous population.

**Parameter space analysis for heterogeneous populations.** We then consider a more complete equation in order to model heterogeneity within a population, taking into account growth and epimutations. In order to ease the computational burden of simulations, we use dimension reduction thereby going from 3D to 2D. More precisely, we define a new advection function, which now depends only on  $S$ , and on a new variable that we denote  $x$ , and can which can be seen as an equivalent of  $\mu_{200}$ : indeed, this new advection function, that we denote  $f_r : \mathbb{R}^2 \rightarrow \mathbb{R}$ , has exactly the same bifurcation diagram as  $F$  (Figure 1.5 A).

With the new structural variable writing  $y = (x, S)$ , we are led to the model

$$\begin{cases} \partial_t u(t, y) + \nabla \cdot (f_r(y)u(t, y)) = (r(y) - d(y)\rho(t)) u(t, y) \\ \quad + \int_{\mathbb{R}^2} M(y, z)u(t, z)dz - \int_{\mathbb{R}^2} M(z, y)dz u(t, y) \\ \rho(t) = \int_{\mathbb{R}^2} u(t, y)dy \end{cases}, \quad (1.52)$$

where

- The advection function  $f$  writes  $f(y) = f(x, S) = (f_r(x, S), f_S(S))$ , where  $f_S(S) = \delta(1 - \frac{S}{S_0})$ , with  $S_0 \in [150K, 250K]$  corresponding to the mean of the SNAIL dis-



tribution, and  $\delta := \frac{S_0 \ln(2)}{\alpha}$ , with  $\alpha > 0$  which represents the characteristic time of convergence of SNAIL to the mean  $S_0$ .

- $r(y)$  and  $d(y)$  respectively represent the birth and death rate of a cell of phenotype  $y$ .
- The mutation function  $M$  writes  $M(y, z) = \frac{r(z)}{\tau} P(y - z)$ . Here,  $P(x) = P(y, S) := \frac{1}{\eta_x \eta_S} G(\frac{x}{\eta_x}) G(\frac{S}{\eta_S})$ , where  $G$  is the Gaussian function. Variables  $\eta_x$  and  $\eta_S$  are the standard deviations for  $x$  and  $S$  respectively.

We explore the parameter space for this equation in order to better understand their role in the model. As shown by the bifurcation diagram (Figure 1.5 A), and the definition of  $f_S$ , the value of  $S_0$  determines the phenotypes that can be adopted by the cell population. However, due to the difference in size between the basins of attraction of the three states, the behaviours of the population are highly asymmetric: for low values of  $S_0$  ( $S_0 = 175K$ ), the population is purely epithelial, while for high values of  $S_0$  ( $S_0 = 190K$ ), a part population remains hybrid, even for the smallest values of  $\eta_x$ .

The parameter  $\alpha$ , captures the characteristic time of convergence of SNAIL levels to its equilibrium  $S_0$ . In Figure 1.6, we reproduce epithelial-mesenchymal transition by starting from an epithelial population and taking a high value of  $S_0$  ( $S_0 = 225K$ ) for different values of  $\alpha$ , and simulations show that  $\alpha$  impacts the speed at which the population becomes heterogeneous.

We then investigate the impact of growth on the population, by choosing three different growth functions (Figure 1.7) A ii. We show that, in the case where  $r$  is not constant, the initial population influences the growth of the total population.

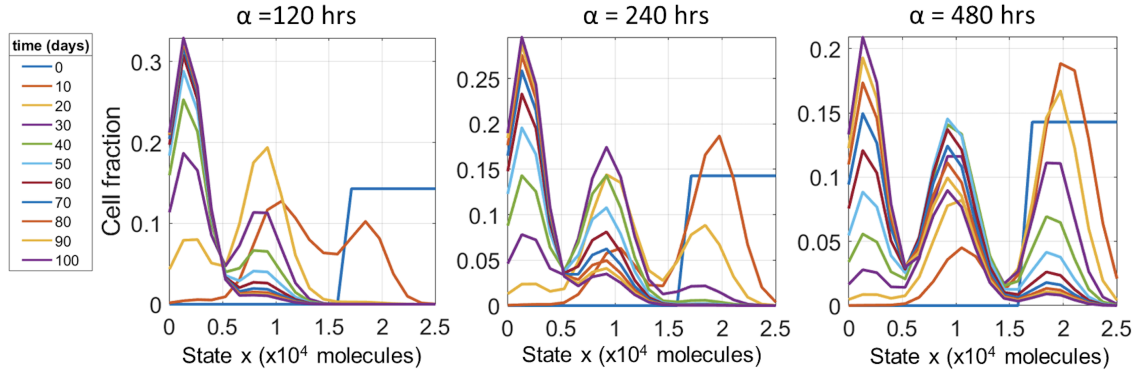


Figure 1.6: Simulations of the advection-selection-mutation (PDE (1.52)) with reduced function  $f_r$ . Temporal changes in cell state distribution of the population for decreasing value of input signal SNAIL's perturbation recovery rate (increasing values of the characteristic time ' $\alpha$ '). Parameters used to generate plots, unless stated otherwise, are  $\eta_x = 1000$ , ini pop Epi, time point 20,  $S_0 = 225K$  molecules, and per-capita growth rate ( $r$ ) of all phenotype (E, hybrid E/M, and M) = 0.0182/hr.

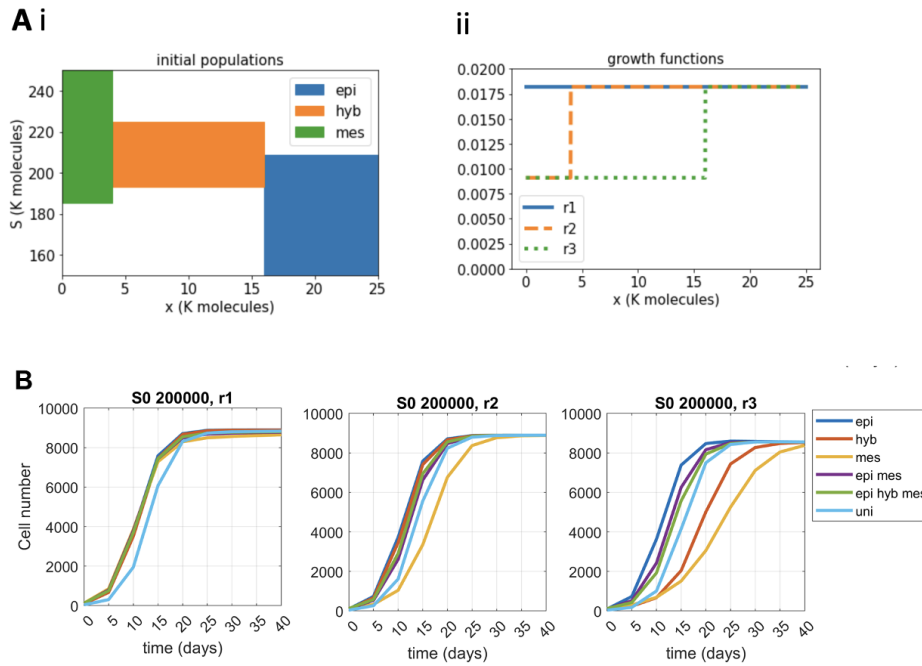


Figure 1.7: Effect of growth rate differences among E, hybrid E/M and M phenotypes on overall growth dynamics. A i) Support of the different initial populations; in all cases, initial population is uniformly distributed on its support, and such that the total population is equal to 100 cells. A ii) Profile of the three growth functions. B) Population growth dynamics for different combinations of growth scenarios and initial population distribution; 'epi hyb mes' corresponds to an initial population supported on the three colored domains of A i), while 'uni' is uniformly distributed on the whole rectangle  $[0, 25K] \times [150K, 250K]$ . The input SNAIL mean ( $S_0$ ) levels used are mentioned for all the individual plots. Other parameters used to generate plots are  $\alpha = 120 \text{ hrs}$ ,  $\eta_x = 1000$ , and per-capita growth rate ( $r$ ) of epithelial phenotype is  $0.0182/\text{hr}$ .

## Publications

### Accepted

- Jules Guilberteau, Camille Pouchol and Nastassia Pouradier Duteil (2021). Monostability and bistability of biological switches. *Journal of Mathematical Biology*, 83(6-7), 65.
- Jules Guilberteau (2023). Bistability and oscillatory behaviours of cyclic feedback loops (forthcoming publication in *Acta Applicandae Mathematicae*).

### Submitted

- Jules Guilberteau, Camille Pouchol and Nastassia Pouradier Duteil (2023). Long-time behaviour of an advection-selection equation. arXiv preprint arXiv:2301.02470.
- Frank Ernesto Alvarez and Jules Guilberteau (2023). A particle method for non-local advection-selection-mutation equations. arXiv preprint arXiv:2304.14210.
- Paras Jain, Jules Guilberteau, Mohit Kumar Jolly, Camille Pouchol and Nastassia Pouradier Duteil. An advection-selection-mutation model for epithelial-mesenchymal plasticity.





## Chapter 2

# Monostability and bistability of biological switches

Cell transitions can be modelled by ordinary differential equations, called gene regulatory networks which describe the behaviour of several molecules in interaction, and for which each stable equilibrium corresponds to a possible state (or ‘biological trait’). In this chapter, we focus on simple ODE systems modelling two molecules which each negatively (or positively) regulate the other. It is well-known that such models may lead to monostability or multistability, depending on the selected parameters. However, extensive numerical simulations have led systems biologists to conjecture that in the vast majority of cases, there cannot be more than two stable points. Our main result is a proof of this conjecture. More specifically, we provide a criterion ensuring at most bistability, which is indeed satisfied by most commonly used functions. This includes Hill functions, but also a wide family of convex and sigmoid functions. We also determine which parameters lead to monostability, and which lead to bistability, by developing a more general framework encompassing all our results.

This work, jointly written with Nastassia Pouradier Duteil and Camille Pouchol, is the subject of a publication in the *Journal of Mathematical biology* [68].

## 2.1 Introduction

A same cell environment may lead to different cell-fate decisions. In most cases, it is considered that the phenotype adopted by a cell is determined by the concentration of several molecules in interaction [67]. It is now well documented that such ‘biological switches’ can be accurately modelled by multistable ordinary differential equations (ODEs), where each stable state represents a possible phenotype [150].

These models have been widely used in order to describe different cellular processes such as the epithelial-mesenchymal transition (EMT) [16, 92, 144], hematopoietic stem cells [84, 97, 138], embryonic stem cells [28] or other cell-fate differentiation phenomena involved in *Xenopus* [54, 125], *Drosophila* [129] or *Escherichia coli* [60, 102, 127].

The development of a relevant ODE model hence benefits from a priori knowledge of the possible number of stable states, and how this number evolves in the parameter space. As an example, the epithelial-mesenchymal transition phenomenon involves three different phenotypes, and it is thus crucial to be able to determine minimal conditions allowing the system to be tristable [110].

A general theoretical answer to finding the number of stable states is certainly out of reach for high-dimensional ODEs with a large number of parameters. Understanding the more simple building blocks of these complex models, however, remains of paramount importance, even more so with the advent of synthetically-built switches where, to some extent, the model may be chosen and kept simple [60].

A widely used starting ODE model writes as follows

$$\begin{cases} \dot{x} = \alpha f(y) - x \\ \dot{y} = \beta g(x) - y, \end{cases} \quad f', g' < 0 \text{ or } f', g' > 0 \quad (2.1)$$

where  $x$  and  $y$  stand for the (normalised) concentrations of the two molecules ( $A$  and  $B$  on Figure 2.1),  $\alpha > 0$  and  $\beta > 0$  their synthesis rates. Here,  $f$  and  $g$  are two monotonic functions which model the interactions between these two molecules, and are both strictly increasing or strictly decreasing, depending on whether the system is cooperative or competitive. A classical choice for  $f$  and  $g$  are Hill functions, *i.e.*, functions of the form

$$x \mapsto \frac{1}{1 + x^r},$$

with  $r \geq 1$ . More generally, molecule interactions are usually considered to behave sigmoidally [150, 152].

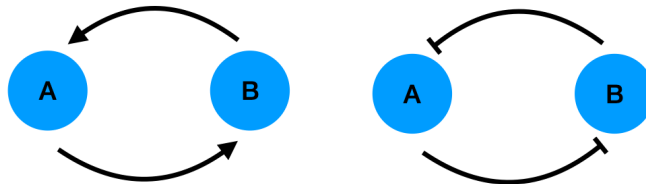


Figure 2.1: Schematic representations of system (2.1): the arrows represent activation ( $f$  and  $g$  increasing) and the bar inhibition ( $f$  and  $g$  decreasing). Under the hypothesis of Theorem 4, we prove that such systems are either monostable or bistable. When bistability holds, the system on the left (cooperative system) has two stable points corresponding to (high A/ high B) and (low A/ low B), while the one on the right (competitive system) leads to (high A/ low B) and (low A/ high B) stable equilibria.

**State of the art.** The seminal paper of Cherry and Adler [26] is the main breakthrough towards understanding when multistability occurs for such models. Under the condition

$$\sup_{y>0} \left( \left| \frac{yf'(y)}{f(y)} \right| \right) \sup_{x>0} \left( \left| \frac{xg'(x)}{g(x)} \right| \right) > 1,$$

it is proved that there exist parameters  $\alpha$  and  $\beta$  such that system (2.1) is multistable. When applied to Hill functions, this shows that multistability will occur for some parameters  $\alpha, \beta$  whenever  $r \in \mathbb{N}$ ,  $r \geq 2$  for  $f$  or  $g$ . Interestingly, the authors noted that the sigmoid shape of  $f$  and  $g$  is not a necessary condition for bistability.

The numerical investigation of systems such as (2.1) suggests that they are in fact always either monostable or bistable. This has led some authors to claim that self-regulation is required in order to get a tristable ODE, *i.e.*, at least one of the cells must have a positive feedback on itself [60, 91, 112].

Up to our knowledge, this conjecture of at most bistability is yet to be proved. Moreover, for given functions  $f$  and  $g$ , determining the exact set of parameters  $(\alpha, \beta)$  for which this system is monostable or bistable remains difficult without the help of numerical simulations. Let us mention the very recent paper [102] which, by means of direct computations, solves the specific case of  $f : x \mapsto \frac{1}{1+x}$  and  $g : x \mapsto \frac{1}{1+x^n}$ .

In the present work, we therefore address the following two key questions.

- Under which conditions is system (2.1) at most bistable?
- For given functions  $f$  and  $g$ , which parameters  $\alpha, \beta$  lead to monostability, and which ones lead to bistability?

A natural way to answer the first question is to note that a point  $(\bar{x}, \bar{y}) \in \mathbb{R}_+$  is an equilibrium of (2.1) if and only if

$$\begin{cases} \alpha f(\beta g(\bar{x})) = \bar{x} \\ \bar{y} = \beta g(\bar{x}) \end{cases}.$$

Hence, studying the equilibria of (2.1) is equivalent to studying the fixed points of  $x \mapsto \alpha f(\beta g(x))$ . The main difficulty lies in the fact that, even if  $f$  and  $g$  are two ‘simple’ functions, there is no reason for  $x \mapsto \alpha f(\beta g(x))$  to be as well. As an example, with Hill functions of integer orders  $n$  and  $m$ , determining the fixed points of  $x \mapsto \alpha f(\beta g(x))$  is equivalent to investigating the positive roots of a polynomial of degree  $nm + 1$ , which proves to be difficult as soon as  $nm \geq 3$ .

**Main results.** Working around this difficulty, our main result is a simple and general result ensuring at most bistability.

**Theorem 4.** *If the functions*

$$\frac{1}{\sqrt{|f'|}} \quad \text{and} \quad \frac{1}{\sqrt{|g'|}}$$

*are strictly convex, then, for any  $\alpha, \beta > 0$ , system (2.1) has at most three equilibria, among which at most two stable equilibria.*

Not only does this result apply to all classically-used functions we are aware of (including Hill and shifted Hill functions), it may easily be checked visually for more involved functions.

Following an approach reminiscent of that of [26], we go further and develop a general method for the identification of which parameters  $\alpha, \beta$  lead to either monostability or



bistability. For this purpose, we develop a framework yielding a condition under which system (2.1) has at most or at least  $n$  equilibria, for all  $n \in \mathbb{N}$ . The obtained criterion is not completely explicit, but the resulting formula makes it numerically straightforward to check if some chosen parameters induce a monostable or a bistable system. We hence bypass any computationally-expensive grid-search through the parameter space. We show that the method in [26] corresponds to the case  $n = 1$  of this general framework. With the same framework, Theorem 4 corresponds to studying the case  $n = 2$ . For higher values of  $n$ , however, we have not been able to apply this theoretical framework as the resulting computations prove to be too intricate.

We also prove that when bistability occurs, the separatrix between the two basins of attraction is a one-dimensional curve. Note that our proof is implicit and does not provide a formula for the curve, unless some specific symmetry assumptions are made.

Taken together, our results show that system (2.1) will generically lead to either one of the pictures of Figure 2.2, *i.e.*, we fall into one of these two cases:

- system (2.1) is monostable, and all the solutions converge to the unique equilibrium,
- system (2.1) is bistable, and, in this case, the basins of attraction of the two stable points are separated by a one-dimensional separatrix which contains the unique other (unstable saddle) equilibrium point.

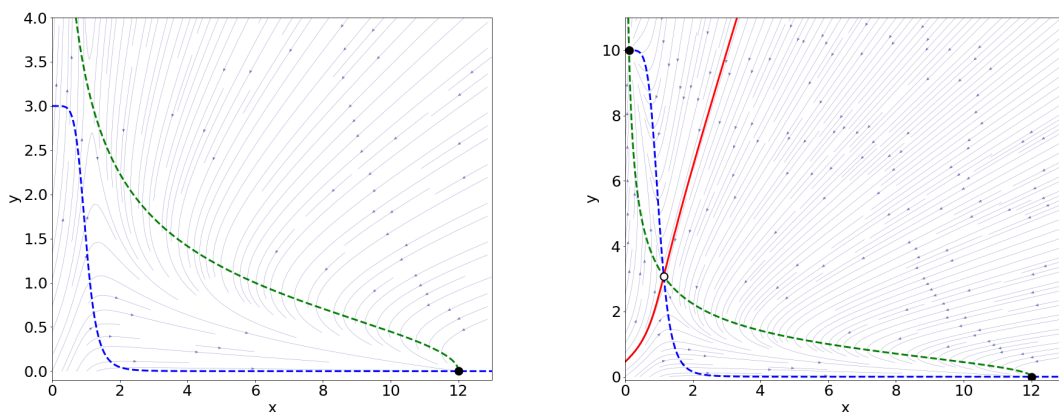


Figure 2.2: Typical phase planes for system (2.1). In both simulations, we have taken  $f : x \mapsto \frac{1}{1+x^2}$  and  $g : x \mapsto \frac{1}{1+x^6}$ , and  $\alpha = 10$ . The value of  $\beta$  is 3 in the left diagram, and 12 in the right one.

**Outline of the chapter.** This chapter is organised as follows. Section 2.2 is devoted to setting the mathematical framework and some general results which prove to be useful throughout. Next, we prove our main Theorem 4, in Section 2.3, which involves defining an appropriate class of functions and studying it in detail. We then turn our attention to finding the parameters for which the system of interest is either monostable or bistable in Section 2.4. We present a generalising framework, and then apply it to the cases of mono and bi-stability. We also compute the parameters for the cases of interest found in the literature, such as the toggle switch of [60].

## 2.2 Preliminary results

Throughout this chapter, we study systems of the form

$$\begin{cases} \dot{x} = \alpha f(y) - x \\ \dot{y} = \beta g(x) - y, \end{cases} \quad (2.2)$$

starting from an initial condition  $(x_0, y_0)$  with  $x_0 \geq 0, y_0 \geq 0$ . Here

- $\alpha, \beta$  are two positive parameters,
- $f, g \in C^1(\mathbb{R}_+, \mathbb{R}_+)$  are two increasing or two decreasing functions, and at least one of these functions is bounded.

If  $f$  and  $g$  are increasing, system (2.2) is called *cooperative*. Examples of cooperative systems include two-species interactions that benefit both species, a kind of interaction often referred to as mutualism [159]. On the other hand, if  $f$  and  $g$  are decreasing, system (2.2) is called *competitive*. A simple example of a competitive system is the well-studied “genetic switch” of two proteins that each repress the synthesis of the other [26, 91].

Without loss of generality, one can assume that  $f, g > 1$  on  $\mathbb{R}_+$  (in the case where  $f$  and  $g$  are increasing) or  $f, g < 1$  on  $\mathbb{R}_+$  (in the case where  $f$  and  $g$  are decreasing). Under these conditions, since  $f$  or  $g$  is bounded, all the solutions of this ODE are bounded, regardless of its initial condition. It is well-known that any solution of such a system converges to an equilibrium point [77].

Thus, the analysis of system (2.2) requires studying its equilibrium points. In this section, we begin by providing important results on how the equilibrium points are ordered, as well as on their basins of attraction. These will be the starting point of our investigation of the system’s multistability.

### 2.2.1 Ordering of stable points

Since the functions  $f$  and  $g$  are one-to-one, it is clear that  $(\bar{x}, \bar{y}) \in \mathbb{R}_+^2$  is an equilibrium point of (2.2) if and only if

$$\begin{cases} \alpha f(\beta g(\bar{x})) = \bar{x} \\ \bar{y} = \beta g(\bar{x}) \end{cases}.$$

Hence, the number of equilibria of (2.2) is equal to the number of fixed points of

$$F : x \mapsto \alpha f(\beta g(x)).$$

Moreover, if (2.2) has a finite number of equilibria, that we denote  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ , with  $x_1 < x_2 < \dots < x_n$ , then

- $y_1 < y_2 < \dots < y_n$  if system (2.2) is cooperative (*i.e.* if  $f$  and  $g$  are increasing);
- $y_1 > y_2 > \dots > y_n$  if system (2.2) is competitive (*i.e.* if  $f$  and  $g$  are decreasing).

Let us give the single notion of stability of an equilibrium point that we shall make use of throughout.

**Definition 1.** We say that an equilibrium point  $(\bar{x}, \bar{y})$  is asymptotically stable if for all  $\varepsilon > 0$  there exists a neighborhood  $U$  of  $(\bar{x}, \bar{y})$  such that if  $(x_0, y_0) \in U$ , the trajectory  $(x(t), y(t))$  starting from  $(x_0, y_0)$  satisfies

$$\forall t \geq 0, |x(t) - \bar{x}| + |y(t) - \bar{y}| \leq \varepsilon \quad \text{and} \quad \lim_{t \rightarrow +\infty} (x(t), y(t)) = (\bar{x}, \bar{y}).$$

We will abusively refer to *stable equilibrium points* when dealing with asymptotic stable points.

From the applicative point of view, in particular, the ordering of equilibrium points means the following. If (2.2) is bistable, the two stable points will be of the type (low  $x$ / low  $y$ ) and (high  $x$ / high  $y$ ) if the system is cooperative, and of the type (low  $x$ / high  $y$ ) and (high  $x$ / low  $y$ ) if it is competitive (see Figure 2.2).

We also observe that for any  $(x, y) \in \mathbb{R}_+$ , the Jacobian matrix of the right-hand side of this ODE in  $(x, y)$  is

$$J_{(x,y)} = \begin{pmatrix} -1 & \alpha f'(y) \\ \beta g'(x) & -1 \end{pmatrix}.$$

Since  $\text{Tr}(J_{(x,y)}) = -2$  and  $\det(J_{(x,y)}) = 1 - \alpha\beta f'(y)g'(x)$ , a fixed point  $(\bar{x}, \bar{y})$  is stable if  $\alpha\beta f'(\bar{y})g'(\bar{x}) < 1$  and unstable if  $\alpha\beta f'(\bar{y})g'(\bar{x}) > 1$ . In other words, under the hypothesis “for any fixed point  $(\bar{x}, \bar{y})$ ,  $\alpha\beta f'(\bar{y})g'(\bar{x}) \neq 1$ ”, the number of stable equilibria of (2.2) is equal to the number of times  $F$  crosses the identity line ‘from above’, and the number of unstable equilibria to the number of times  $F$  crosses the identity line ‘from below’. Therefore, since  $F$  is positive, increasing and bounded, system (2.2) has  $d$  stable equilibria if and only if it has  $d - 1$  unstable equilibria. This result is proved rigorously in the next section.

### 2.2.2 Basins of attraction

Due to the particular shape of the system that we study, we have a precise result regarding the basins of attraction: if the system is monostable, then all solutions converge to the stable point, meaning that it is globally asymptotically stable. If it is bistable, then the two basins of attractions are separated by a separatrix, which is the curve of an increasing function if (2.2) is competitive and of a decreasing function if (2.2) is cooperative (see Figure 2.2). This result relies on the following proposition.

**Proposition 1.** *Let us consider the ODE system*

$$\begin{cases} \dot{x} = F(y) - x \\ \dot{y} = G(x) - y \end{cases}, \quad (2.3)$$

where  $F$  and  $G$  are either both increasing or both decreasing, and at least one of them is bounded. We assume that there exists  $(\bar{x}_u, \bar{y}_u) \in \mathbb{R}_+^2$  an equilibrium point of (2.3) such that

$$G'(\bar{x}_u)F'(\bar{y}_u) > 1$$

Then, the basin of attraction of  $(\bar{x}_u, \bar{y}_u)$  has measure zero in  $\mathbb{R}_+^2$ . More precisely, this basin of attraction is included in a curve of the shape

$$\{(x, \gamma(x)), x \in (a, b)\},$$

where  $a \geq 0$ ,  $b \in \mathbb{R}_+ \cup \{+\infty\}$  and  $\gamma : (a, b) \rightarrow \mathbb{R}_+$  is a continuous function. This function is decreasing if  $F$  and  $G$  are increasing, and increasing if  $F$  and  $G$  are decreasing.

*Proof.* We prove this result only in the competitive case: the arguments can easily be adapted to the cooperative case. First, let us note that (2.3) is strictly competitive, in the sense that when we rewrite (2.3) as  $\dot{z} = \Gamma(z)$ , the vector field  $\Gamma = (\Gamma_1, \Gamma_2)$  underlying the ODE satisfies for all  $z \in \mathbb{R}^2$ ,

$$\frac{\partial \Gamma_1}{\partial z_2}(z) < 0, \quad \frac{\partial \Gamma_2}{\partial z_1}(z) < 0.$$

The strict competitiveness of the system implies that it satisfies the comparison principle, which writes as follows. Let  $z, w$  be two solutions such that

$$\begin{cases} z_1(0) \leq w_1(0) & (\text{resp. } z_1(0) \geq w_1(0)) \\ z_2(0) > w_2(0) & (\text{resp. } z_2(0) < w_2(0)). \end{cases}$$

Then, for any  $t > 0$  such that  $z$  and  $w$  are defined on  $[0, t]$ ,

$$\begin{cases} z_1(t) < w_1(t) & (\text{resp. } z_1(t) > w_1(t)) \\ z_2(t) > w_2(t) & (\text{resp. } z_2(t) < w_2(t)). \end{cases}$$

Let  $(x_u, y_u)$  be a solution of system (2.3) which converges to some equilibrium point  $(\bar{x}_u, \bar{y}_u)$ , and  $(x, y)$  a solution of (2.3) such that

$$\begin{cases} x(0) < x_u(0) \\ y(0) \geq y_u(0). \end{cases}$$

We assume that  $(x, y)$  converges to  $(\bar{x}_u, \bar{y}_u)$ . An application of the comparison principle entails

$$\forall t > 0, \quad w_1(t) := x_u(t) - x(t) > 0, \quad w_2(t) := y(t) - y_u(t) > 0.$$

Hence, we may write

$$\dot{w}_1 = \dot{x}_u - \dot{x} = F(y_u) - x_u - F(y) + x = \frac{F(y_u) - F(y)}{y - y_u} w_2 - w_1.$$

Likewise,

$$\dot{w}_2 = \frac{G(x) - G(x_u)}{x_u - x} w_1 - w_2.$$

Since  $(x, y)$  and  $(x_u, y_u)$  both converge to  $(\bar{x}_u, \bar{y}_u)$ , and  $F'(\bar{y}_u)G'(\bar{x}_u) > 1$ , there exist  $c, d > 0$  which satisfy  $cd > 1$ , and  $T > 0$  such that for all  $t \geq T$ :

$$\frac{F(y_u(t)) - F(y(t))}{y(t) - y_u(t)} > c \quad \text{and} \quad \frac{G(x(t)) - G(x_u(t))}{x_u(t) - x(t)} > d.$$

Therefore, for all  $t \geq T$ ,  $\dot{w}_1(t) \geq c w_2(t) - w_1(t)$  and  $\dot{w}_2(t) \geq d w_1(t) - w_2(t)$ . We now consider

$$W := (1 + d)w_1 + (1 + c)w_2.$$

According to the previous computations, we have, for all  $t \geq T$ ,

$$\begin{aligned} \dot{W}(t) &\geq -(1 + d)w_1(t) + c(1 + d)w_2(t) + d(1 + c)w_1(t) - (1 + c)w_2(t) \\ &= \underbrace{(cd - 1)}_{>0} \underbrace{(w_1(t) + w_2(t))}_{>0} > 0. \end{aligned}$$

Hence,  $W$  does not converge to zero, which contradicts the fact that  $w_1$  and  $w_2$  converge to zero. With the same reasoning, we prove that if  $x(0) > x_u(0)$  and  $y(0) \leq y_u(0)$ , then  $(x, y)$  does not converge to  $(\bar{x}_u, \bar{y}_u)$ .

In particular, for all  $x_0 \in \mathbb{R}_+$ , there exists at most one  $y_0 \in \mathbb{R}_+$  such that  $(x_0, y_0)$  is in the basin of attraction of  $(\bar{x}_u, \bar{y}_u)$ , which proves the existence of  $\gamma$ . Moreover, if  $(x_0, y_0)$  and  $(x'_0, y'_0)$  are two points of this basin of attraction such that  $x_0 < x'_0$ , then  $y_0 < y'_0$ , which shows that  $\gamma$  is increasing. The continuity of  $\gamma$  and the connectedness of the set follow from the continuity of the solutions.  $\square$

## 2.3 A criterion which ensures at most bi-stability

The main purpose of this section is to prove the following theorem:

**Theorem 5.** *Let  $f, g \in C^3(\mathbb{R}_+, \mathbb{R}_+)$  two functions such that  $f' > 0, g' > 0$  or  $f' < 0, g' < 0$  on  $\mathbb{R}_+$ . We recall that system (2.2) refers to*

$$\begin{cases} \dot{x} = \alpha f(y) - x \\ \dot{y} = \beta g(x) - y \end{cases} \quad (2.2)$$

Then

- (i) *If  $\frac{1}{\sqrt{|f'|}}$  and  $\frac{1}{\sqrt{|g'|}}$  are convex (we say that  $f$  and  $g$  are  $\gamma^{1/2}$ -convex), and at least one of these functions is strictly convex, then for any  $\alpha, \beta > 0$  system (2.2) has at most three equilibria, and is either monostable or bistable.*
- (ii) *If  $\frac{1}{\sqrt{|f'|}}$  and  $\frac{1}{\sqrt{|g'|}}$  are concave (we say that  $f$  and  $g$  are  $\gamma^{1/2}$ -concave), and at least one of these functions is strictly concave, then for any  $\alpha, \beta > 0$ , system (2.2) is monostable.*

In order to determine if a given function is  $\gamma^{1/2}$ -convex or  $\gamma^{1/2}$ -concave, we provide different properties about these functions, which are summarised below.

- A function  $f$  is  $\gamma^{1/2}$ -convex (resp.  $\gamma^{1/2}$ -concave) if and only if  $f'f^{(3)} \leq f''^2$  (resp.  $f'f^{(3)} \geq f''^2$ ).
- If  $f$  and  $g$  are  $\gamma^{1/2}$ -convex (resp.  $\gamma^{1/2}$ -concave), then  $f \circ g$  is  $\gamma^{1/2}$ -convex (resp.  $\gamma^{1/2}$ -concave).
- $x \mapsto x^a$  is strictly  $\gamma^{1/2}$ -convex if  $|a| > 1$ , strictly  $\gamma^{1/2}$ -concave if  $0 < |a| < 1$ .
- The only functions which are both  $\gamma^{1/2}$ -convex and  $\gamma^{1/2}$ -concave are the affine and the homographic functions.
- A strictly monotonic function  $f$  is  $\gamma^{1/2}$ -convex if and only if  $f^{-1}$  is  $\gamma^{1/2}$ -concave.

The rest of this section is devoted to proving Theorem 5 and the above properties regarding  $\gamma^{1/2}$ -convexity.

### 2.3.1 A priori bounds on the number of fixed points

In what follows,  $I$  will denote an arbitrary (possibly unbounded) interval of  $\mathbb{R}$ .

The first proposition and its corollary prove an intuitive fact about the fixed points of a function and the sign of its derivative at this point. As explained in the preliminary results, this proposition is the basis for all the results of this section.

**Proposition 2.** *Let  $f \in C^1(I, \mathbb{R})$ . If  $f$  has  $2n$  fixed points or more ( $n \in \mathbb{N}^*$ ), then there exist  $x_1^- < x_2^- < \dots < x_n^-$  and  $x_1^+ < x_2^+ < \dots < x_n^+$  some fixed points of  $f$  such that, for any  $i \in \{1, \dots, n\}$ :*

$$f'(x_i^-) \leq 1 \quad \text{and} \quad f'(x_i^+) \geq 1.$$

*Conversely, if there exist  $n$  fixed points of  $f$ , denoted  $x_1 < x_2 < \dots < x_n$  such that*

$$\forall i \in \{1, \dots, n\}, \quad f'(x_i) < 1 \quad \text{or} \quad \forall i \in \{1, \dots, n\}, \quad f'(x_i) > 1,$$

*then  $f$  has at least  $2n - 1$  fixed points.*

*Proof.* The first implication can clearly be proved by induction. The main difficulty lies in the base case, *i.e.* in the case  $n = 1$ . Let us denote

$$F := f - \text{id}.$$

If  $F$  has a finite number of roots, or more generally, if the set of the roots of  $F$  does not have an accumulation point, the result immediately holds, since  $F$  reaches its roots ‘from above’ and ‘from below’ alternatively. Otherwise, we consider a bounded sequence of roots of  $F$ , denoted  $(c_n)$ , and we assume that for any  $n \in \mathbb{N}$ ,  $F'(c_n) < 0$  (or  $F'(c_n) > 0$ ). Since  $(c_n)$  is bounded, we can extract a convergent subsequence, and we denote  $c$  its limit. According to the continuity of  $F$ ,  $F(c) = 0$  and, according to its differentiability,  $F'(c) = \lim_{n \rightarrow +\infty} \frac{F(c) - F(c_n)}{c - c_n} = 0$ , which proves the result.

The converse implication simply stems from applying the intermediate value theorem to the function  $F$ . □

**Corollary 3.** *Let us assume that  $I = \mathbb{R}$  (or  $\mathbb{R}_+^*$ ) and that  $f$  is positive and bounded. If  $f$  has  $2n$  fixed points or more ( $n \in \mathbb{N}^*$ ), then there exist  $x_1^- < \dots < x_{n+1}^-$  some fixed points of  $f$  such that for any  $i \in \{1, \dots, n+1\}$*

$$f'(x_i^-) \leq 1.$$

*Conversely, if there exist  $n$  fixed points of  $f$  (denoted  $x_1, \dots, x_n$ ) such that*

$$\forall i \in \{1, \dots, n\}, \quad f'(x_i) > 1,$$

*then  $f$  has at least  $2n + 1$  fixed points.*

*Proof.* We note that

$$E := \{x \in \mathbb{R} : f(x) = x\}$$

is a closed and bounded set, and thus that it is compact. Hence, it has a minimum and a maximum element, that we denote  $x^m$  and  $x^M$ . Since for all  $x < x^m$  and all  $x > x^M$ ,  $f(x) > x$ , then  $f'(x^m) \leq 1$ ,  $f'(x^M) \leq 1$ .

Therefore, if  $f$  has at least  $2n$  fixed points, then it has at least  $2n - 2$  fixed points on  $(x^m, x^M)$ . Likewise, if there exist  $x_1, \dots, x_n$  such that  $f'(x_i) > 1$  for all  $i \in \{1, \dots, n\}$ , then  $x_1, \dots, x_n \in (x^m, x^M)$ .

We conclude by applying the previous proposition on  $(x^m, x^M)$ . □

### 2.3.2 Properties of $\gamma^{1/2}$ -convex functions

We now use these two results (in the specific cases  $n = 1$  and  $n = 2$ ) in order to establish a convexity criterion related to the derivative of the function of interest. It ensures that this function cannot have more than three fixed points.

**Proposition 4.** *Let  $f \in C^1(I, \mathbb{R})$ . Let us assume that there exists  $\Psi : f'(I) \rightarrow \mathbb{R}$  a function which satisfies the four following conditions:*

- (i)  $\Psi(1) = 1$
- (ii)  $\forall x \leq 1, \Psi(x) \geq 1$
- (iii)  $\forall x \geq 1, \Psi(x) \leq 1$

(iv)  $\Psi \circ f'$  is strictly convex or strictly concave.

Then,  $f$  has at most three fixed points. In the case where  $\Psi \circ f'$  is strictly concave, if we make the stronger assumption that  $I = \mathbb{R}$  or  $\mathbb{R}_+$  and  $f$  is bounded, then  $f$  has a unique fixed point.

In particular, if  $f' > 0$  and if there exists  $\alpha > 0$  such that

$$\gamma_f^\alpha := \left( \frac{1}{f'} \right)^\alpha$$

is strictly convex or strictly concave, then  $f$  has at most three fixed points.

*Proof.* We show the contrapositive. Let us assume that  $f$  has four fixed points or more, and let  $\Psi$  be a function which satisfies the first three points of the proposition. According to Proposition 2, there exist  $x_1, x_2, y_1, y_2$  four distinct points such that, for  $i \in \{1, 2\}$

$$\begin{aligned} f(x_i) &= x_i & \text{and} & & f(y_i) &= y_i \\ f'(x_i) &\geq 1 & \text{and} & & f'(y_i) &\leq 1. \end{aligned}$$

According to the mean value theorem, there exists  $\theta, \theta' \in (0, 1)$  such that

$$f'(\theta x_1 + (1 - \theta)x_2) = \frac{f(x_1) - f(x_2)}{x_1 - x_2} = 1 \quad \text{and} \quad f'(\theta' y_1 + (1 - \theta')y_2) = \frac{f(y_1) - f(y_2)}{y_1 - y_2} = 1.$$

Thus,

$$\theta \Psi(f'(x_1)) + (1 - \theta) \Psi(f'(x_2)) \leq \theta + (1 - \theta) = 1 = \Psi(f'(\theta x_1 + (1 - \theta)x_2))$$

and

$$\theta' \Psi(f'(y_1)) + (1 - \theta') \Psi(f'(y_2)) \geq \theta' + (1 - \theta') = 1 = \Psi(f'(\theta' y_1 + (1 - \theta')y_2)),$$

which proves that  $\Psi \circ f$  is neither strictly convex nor strictly concave.

When  $f$  is bounded and positive, we argue similarly but with the help of Corollary 3.  $\square$

We recall and extend the definition of  $\gamma_f^\alpha$  for any  $\alpha > 0$  and function  $f \in C^1(I, \mathbb{R})$  such that  $|f'| > 0$ :

$$\begin{aligned} \gamma_f^\alpha &: I \longrightarrow \mathbb{R}_+ \\ x &\longmapsto \left( \frac{1}{|f'(x)|} \right)^\alpha. \end{aligned}$$

The following lemma provides a straightforward way to decide whether a given function  $f$  satisfies one of the ' $\gamma^\alpha$ -properties'.

**Lemma 1.** *Let  $f \in C^3(I, \mathbb{R})$  a strictly monotonic function. Then  $\gamma_f^\alpha$  is convex (resp. concave) if and only if*

$$f' f^{(3)} \leq (\alpha + 1) f''^2 \quad (\text{resp. } f' f^{(3)} \geq (\alpha + 1) f''^2)$$

on  $I$ , and strictly convex (resp. strictly concave) if and only if

$$f' f^{(3)} < (\alpha + 1) f''^2 \quad (\text{resp. } f' f^{(3)} > (\alpha + 1) f''^2)$$

on a dense subset of  $I$ .

*Proof.* First, let us note that for any  $f$  which satisfies the hypotheses, we have  $\gamma_{-f}^\alpha = \gamma_f^\alpha$ . We may hence assume  $f' > 0$  without loss of generality.

We recall that a function  $F \in C^2(I, \mathbb{R})$  is strictly convex (resp. strictly concave) if and only if  $F'' > 0$  (resp.  $F'' < 0$ ) on a dense subset of  $I$ . Computing the second derivative of  $\gamma_f^\alpha$ , we find

$$\gamma_f^{\alpha''} = -\alpha \frac{f' f^{(3)} - (\alpha + 1) f''^2}{f^{\alpha+2}}.$$

The result immediately follows.  $\square$

In general, if  $\gamma_f^\alpha$  and  $\gamma_g^\alpha$  are convex (or concave),  $\gamma_{f \circ g}^\alpha$  has no reason to be. The case  $\alpha = \frac{1}{2}$  stands out since the property is stable under composition, as we shall see. Since  $\gamma_f^{1/2}$  will play a crucial role in what follows, we recall its definition here:

**Definition 2**  $\gamma_f^{1/2}$ -convex function. Let  $f \in C^3(I, \mathbb{R})$  be a strictly monotonic function. We recall that

$$\begin{aligned} \gamma_f^{1/2} : I &\longrightarrow \mathbb{R} \\ x &\longmapsto \frac{1}{\sqrt{|f'(x)|}}. \end{aligned}$$

We say that

- $f$  is  $\gamma_f^{1/2}$ -convex if  $\gamma_f^{1/2}$  is convex ( $\iff f' f^{(3)} \leq \frac{3}{2} f''^2$  on  $I$ ).
- $f$  is strictly  $\gamma_f^{1/2}$ -convex if  $\gamma_f^{1/2}$  is strictly convex ( $\iff f' f^{(3)} < \frac{3}{2} f''^2$  on a dense subset of  $I$ ).
- $f$  is  $\gamma_f^{1/2}$ -concave if  $\gamma_f^{1/2}$  is concave ( $\iff f' f^{(3)} \geq \frac{3}{2} f''^2$  on  $I$ ).
- $f$  is strictly  $\gamma_f^{1/2}$ -concave if  $\gamma_f^{1/2}$  is strictly concave ( $\iff f' f^{(3)} > \frac{3}{2} f''^2$  on a dense subset of  $I$ ).

**Proposition 5.** Let  $g \in C^3(I, \mathbb{R})$ ,  $f \in C^3(g(I), \mathbb{R})$ . Then

$$f, g \text{ } \gamma^{1/2}\text{-convex (resp. } \gamma^{1/2}\text{-concave)} \implies f \circ g \text{ } \gamma^{1/2}\text{-convex (resp. } \gamma^{1/2}\text{-concave)}.$$

Furthermore, if in the above either  $f$  or  $g$  is strictly  $\gamma^{1/2}$ -convex (resp.  $\gamma^{1/2}$ -concave), then  $f \circ g$  is strictly  $\gamma^{1/2}$ -convex (resp.  $\gamma^{1/2}$ -concave).

*Proof.* Let us denote  $h := f \circ g$  and compute its third-order derivative:

$$h^{(3)} = g^{(3)}(f \circ g) + 3g'g''(f' \circ g) + g'^3(f'' \circ g).$$

Thus

$$\begin{aligned} h'h^{(3)} - (\alpha + 1)h''^2 &= g'^4[(f' \circ g)(f^{(3)} \circ g) - (\alpha + 1)(f'' \circ g)^2] \\ &\quad + (f' \circ g)^2[g'g^{(3)} - (\alpha + 1)g''^2] + (3 - 2(\alpha + 1))g'^2g''(f' \circ g)(f'' \circ g). \end{aligned}$$

In the particular case where  $\alpha = \frac{1}{2}$ ,  $3 - 2(\alpha + 1) = 0$ , which shows that the last term vanishes. The previous lemma concludes the proof since the other two terms have the appropriate signs.  $\square$



We now give examples of  $\gamma^{1/2}$ -convex/concave functions. The main point is that a function  $f$  is  $\gamma^{1/2}$ -convex/concave if and only if, for any affine or homographic function  $h$ , the composite functions  $h \circ f$  and  $f \circ g$  are also  $\gamma^{1/2}$ -convex/concave.

**Example.** 1.  $f$  is both  $\gamma^{1/2}$ -convex and  $\gamma^{1/2}$ -concave if and only if  $f$  is an affine function or a homographic function, *i.e.*, if there exist  $a, b, c, d \in \mathbb{R}$  such that

$$f = x \mapsto \frac{ax + b}{cx + d}.$$

2. We consider, for all  $\nu \in \mathbb{R}$ , the function  $x \mapsto x^\nu$  on  $\mathbb{R}_+^*$ .

- If  $|\nu| > 1$ , then  $x \mapsto x^\nu$  is strictly  $\gamma^{1/2}$ -convex on  $\mathbb{R}_+^*$ .
- If  $|\nu| < 1$ , then  $x \mapsto x^\nu$  is strictly  $\gamma^{1/2}$ -concave on  $\mathbb{R}_+^*$ .

3.  $x \mapsto e^x$  is strictly  $\gamma^{1/2}$ -convex on  $\mathbb{R}$ .

**Remark.** According to Proposition 5, all the functions of the form

$$x \mapsto \left( \frac{ax^\nu + b}{cx^\nu + d} \right)^\mu,$$

with  $\nu, \mu \geq 1$  and  $\nu > 1$  or  $\mu > 1$  are strictly  $\gamma^{1/2}$ -convex. Many common sigmoid functions are  $\gamma^{1/2}$ -convex, as shown in Appendix B.3.

*Proof.* 1. First, let us assume that  $f : x \mapsto \frac{ax+b}{cx+d}$  (which is defined on  $\mathbb{R}$  if  $c = 0$  and on  $(-\infty, -d/c) \cup (-d/c, +\infty)$  otherwise).

Then, for any  $x$  in its domain,

$$f'(x) = \frac{ad - cb}{(cx + d)^2}, \quad f''(x) = -2c \frac{ad - cb}{(cx + d)^3}, \quad f^{(3)}(x) = 6c^2 \frac{ad - cb}{(cx + d)^4}.$$

Thus,

$$f'(x)f^{(3)}(x) - \frac{3}{2}f''(x)^2 = 0.$$

Hence,  $f$  is indeed  $\gamma^{1/2}$ -convex and  $\gamma^{1/2}$ -concave.

Conversely, let us assume that  $f$  is both  $\gamma^{1/2}$ -convex and  $\gamma^{1/2}$ -concave, *i.e.*, that  $\frac{1}{\sqrt{|f'|}}$  is an affine function. Then, there exist  $\alpha, \beta \in \mathbb{R}$  such that for all  $x \geq 0$   $\frac{1}{\sqrt{|f'(x)|}} = \alpha x + \beta$ . Thus,

$$f'(x) = \pm \frac{1}{(\alpha x + \beta)^2}.$$

Integrating, we conclude that  $f$  is of the announced form  $x \mapsto \frac{ax+b}{cx+d}$ .

2. We recall that the function  $x \mapsto x^\beta$  is strictly convex on  $\mathbb{R}_+^*$  if and only if  $\beta > 1$  or  $\beta < 0$  and strictly concave if and only if  $\beta \in (0, 1)$ .

Hence, since the derivative of  $x \mapsto x^\nu$  is  $x \mapsto \nu x^{\nu-1}$ ,  $x \mapsto x^\nu$  is strictly  $\gamma^{1/2}$ -convex if and only if  $|\nu| > 1$ . Likewise,  $x \mapsto x^\nu$  is strictly  $\gamma^{1/2}$ -concave if and only if  $|\nu| < 1$ .

3. Since, for all  $x \in \mathbb{R}$ ,

$$\frac{1}{\sqrt{(e^x)'}} = e^{-\frac{1}{2}x},$$

the result immediately follows, owing to the strict convexity of  $x \mapsto e^{-\frac{1}{2}x}$ . □

**Non-local characterisation.** The  $\gamma^{1/2}$ -convexity/concavity of a function can surprisingly be written under a non-local form. Although this characterisation is generally less suitable to check if a given function is  $\gamma^{1/2}$ -convex/concave, it can be useful to understand why the  $\gamma^{1/2}$ -convexity/concavity is especially adapted to our problem. Indeed, the fact that the ‘ $\gamma^{1/2}$ -properties’ are stable by composition and that they ensure that a function has at most three fixed points can easily be proved by considering this non-local form.

**Proposition 6.** *Let  $f \in C^3(I, \mathbb{R})$ . Then*

- $f$  is  $\gamma^{1/2}$ -convex (resp.  $\gamma^{1/2}$ -concave) if and only if for any  $x, y \in I$ ,  $x \neq y$ ,

$$f'(x)f'(y) \leq \left( \frac{f(x) - f(y)}{x - y} \right)^2 \quad (\text{resp. } f'(x)f'(y) \geq \left( \frac{f(x) - f(y)}{x - y} \right)^2).$$

- $f$  is strictly  $\gamma^{1/2}$ -convex (resp. strictly  $\gamma^{1/2}$ -concave) if and only if for any  $x, y \in I$ ,  $x \neq y$ ,

$$f'(x)f'(y) < \left( \frac{f(x) - f(y)}{x - y} \right)^2 \quad (\text{resp. } f'(x)f'(y) > \left( \frac{f(x) - f(y)}{x - y} \right)^2).$$

*Proof.* We only prove the  $\gamma^{1/2}$ -convex case. Let us assume that  $f$  is  $\gamma^{1/2}$ -convex, and let  $y \in I$ . We define  $I_y^- = \{x \in I : x < y\}$  and  $I_y^+ = \{x \in I : x > y\}$ , and we introduce the homographic function

$$h_y(x) := \frac{f'(y)}{x - f(y)}$$

which is well defined on  $(-\infty, f(y))$  and on  $(f(y), +\infty)$ . As seen in the examples, and according to Proposition 5,  $h_y \circ f$  is  $\gamma^{1/2}$ -convex on  $I_y^-$  and on  $I_y^+$ . In other words,  $\gamma_{h_y \circ f}^{1/2} = \frac{1}{\sqrt{|(h_y \circ f)'}}$  is convex on  $I_y^-$  and on  $I_y^+$ , and for all  $x \in I \setminus \{y\}$ ,

$$\gamma_{h_y \circ f}^{1/2}(x) = \frac{|f(x) - f(y)|}{\sqrt{f'(x)f'(y)}}.$$

We observe that  $\gamma_{h_y \circ f}^{1/2}$  can be extended by continuity at  $y$  by defining  $\gamma_{h_y \circ f}^{1/2}(y) = 0$ . We now assume that  $f' > 0$  (the case  $f' < 0$  is similar). Then, for any  $x \in I \setminus \{y\}$ ,

$$\gamma_{h_y \circ f}^{1/2 \prime}(x) = \text{sgn}(x - y) \left( \sqrt{\frac{f'(x)}{f'(y)}} - (f(x) - f(y)) \frac{f''(x)}{2(f'(x)f'(y))^{3/2}} \right).$$

Thus,

$$\gamma_{h_y \circ f}^{1/2 \prime}(x) \xrightarrow{x \rightarrow y^-} -1 \quad \text{and} \quad \gamma_{h_y \circ f}^{1/2 \prime}(x) \xrightarrow{x \rightarrow y^+} 1.$$

Since  $\gamma_{h_y \circ f}^{1/2}$  is convex on  $I_y^-$  and on  $I_y^+$ ,  $\gamma_{h_y \circ f}^{1/2}$  lies above its tangent at  $y$  on these two intervals, *i.e.*,

- For all  $x \in I_y^+$ ,  $\gamma_{h_y \circ f}^{1/2} \geq -(x - y)$ .
- For all  $x \in I_y^-$ ,  $\gamma_{h_y \circ f}^{1/2} \geq (x - y)$ .

Therefore, for all  $x \in I \setminus \{y\}$ ,

$$\frac{|f(x) - f(y)|}{\sqrt{f'(x)f'(y)}} \geq |x - y|,$$

which proves

$$f'(x)f'(y) \leq \left( \frac{f(x) - f(y)}{x - y} \right)^2.$$

If  $f$  is strictly  $\gamma^{1/2}$ -convex, then these inequalities become strict, which proves the second point.

Now, let us prove the converse assertion. We assume that for any  $x, y \in I$ ,  $x \neq y$ ,

$$f'(x)f'(y) \leq \left( \frac{f(x) - f(y)}{x - y} \right)^2.$$

Then for any  $x > 0, \varepsilon \in \mathbb{R}$  small enough,

$$f'(x)f'(x + \varepsilon) - \left( \frac{f(x + \varepsilon) - f(x)}{\varepsilon} \right)^2 \leq 0.$$

A Taylor-expansion of the left-hand side leads to

$$\varepsilon^2 \left( \frac{1}{6} f'(x) f^{(3)}(x) - \frac{1}{4} f''(x)^2 \right) + o(\varepsilon^2) \leq 0,$$

which implies the inequality

$$f'(x) f^{(3)}(x) \leq \frac{3}{2} f''(x)^2.$$

In the case where, for any  $x \neq y$

$$f'(x)f'(y) < \left( \frac{f(x) - f(y)}{x - y} \right)^2,$$

this inequality becomes strict in a dense subset of  $I$ . Indeed, let us assume that there exists a segment  $J \subset I$  such that

$$f' f^{(3)} = \frac{3}{2} f''^2 \quad \text{on } J.$$

Then, according to the examples,  $f|_J$  is a homographic or an affine function, *i.e.*, there exist  $a, b, c, d \in \mathbb{R}$  such that:

$$\forall x \in J, \quad f(x) = \frac{ax + b}{cx + d}.$$

Then, for all  $x, y \in J$ ,

$$f'(x)f'(y) = \left( \frac{(ad - cb)}{(cx + d)(cy + d)} \right)^2 = \left( \frac{f(x) - f(y)}{x - y} \right)^2,$$

which contradicts the strict inequality.  $\square$

We now use this non-local characterisation in order to prove the last result concerning the  $\gamma^{1/2}$ -convexity/concavity.

**Proposition 7.** *A strictly monotonic function  $f$  is (strictly)  $\gamma^{1/2}$ -convex if and only if  $f^{-1}$  is (strictly)  $\gamma^{1/2}$ -concave.*

*Proof.* By applying the formula for the derivative of an inverse function, we have, for any  $x \neq y$ ,

$$f^{-1}'(x)f^{-1}'(y) = \frac{1}{f'(f^{-1}(x))f'(f^{-1}(y))}.$$

If  $f$  is  $\gamma^{1/2}$ -convex, then, according to the non-local characterisation,

$$f'(f^{-1}(x))f'(f^{-1}(y)) \leq \left( \frac{f(f^{-1}(x)) - f(f^{-1}(y))}{f^{-1}(x) - f^{-1}(y)} \right)^2 = \left( \frac{x - y}{f^{-1}(x) - f^{-1}(y)} \right)^2.$$

Hence,

$$f^{-1}'(x)f^{-1}'(y) \geq \left( \frac{f^{-1}(x) - f^{-1}(y)}{x - y} \right)^2.$$

We prove the converse with the same method.  $\square$

Let us finally come to the proof of our main theorem 5, namely:

**Theorem 5.** *Let  $f, g \in C^3(\mathbb{R}_+, \mathbb{R}_+)$  be two functions such that  $f' > 0$ ,  $g' > 0$  or  $f' < 0$ ,  $g' < 0$  on  $\mathbb{R}_+$ . We recall that system (2.2) refers to*

$$\begin{cases} \dot{x} = \alpha f(y) - x \\ \dot{y} = \beta g(x) - y \end{cases}. \quad (2.2)$$

*Then*

- (i) *If  $f$  and  $g$  are  $\gamma^{1/2}$ -convex, and (at least) one of these functions is strictly  $\gamma^{1/2}$ -convex, then for any  $\alpha, \beta > 0$  system (2.2) has at most three equilibria, and is either monostable or bistable.*
- (ii) *If  $f$  and  $g$  are  $\gamma^{1/2}$ -concave, and (at least) one of these functions is strictly  $\gamma^{1/2}$ -concave, then for any  $\alpha, \beta > 0$ , system (2.2) is monostable.*

*Proof.* (i) As evidenced by the examples, composing with an affine function does not change the  $\gamma^{1/2}$ -convexity/concavity of the functions. Hence, according to Proposition 5, under the first hypothesis,  $x \mapsto \alpha f(\beta g(x))$  is strictly  $\gamma^{1/2}$ -convex. According to Proposition 4, this function has at most three fixed points, which shows that system (2.2) has at most three equilibria.

Moreover, it is well-known that the basins of attraction of stable equilibrium points are open sets. Indeed, if  $\bar{x}_s$  is a stable equilibrium of a differential equation, then, by definition, there exists an open set  $U$  which contains  $\bar{x}_s$  and which is included in the basin of attraction. Let  $x_0$  be a point of this basin. By denoting  $\Phi_T$  the flow associated to this ODE at time  $T$ , we get  $\Phi_T(x_0) \in U$  for  $T$  large enough, and  $\Phi_T^{-1}(U)$  is thus a neighborhood of  $x_0$  which is included in the basin of attraction of  $\bar{x}_s$ .

Hence the system cannot be tristable, and is thus at most bistable.

- (ii) Under this second hypothesis,  $x \mapsto \alpha f(\beta g(x))$  is  $\gamma^{1/2}$ -concave. According to Proposition 4, this proves that  $x \mapsto \alpha f(\beta g(x))$  has a unique equilibrium, and thus that system (2.2) is monostable.  $\square$

## 2.4 Determining the parameters for mono/bistability

In the previous section, we have shown that, under general hypotheses on  $f$  and  $g$ , the system

$$\begin{cases} \dot{x} = \alpha f(y) - x \\ \dot{y} = \beta g(x) - y \end{cases} \quad (2.2)$$

is either monostable or bistable. The purpose of this section is, once the functions  $f, g$  are fixed, to determine which parameters  $\alpha, \beta$  induce a monostable system, and which ones induce a bistable system.

To answer this question, we introduce a general framework that will encompass this particular result. Using this framework, we will also recover the result of the third section, which states that, under some hypotheses, system (2.2) cannot have more than two stable equilibria.

### 2.4.1 General framework

We introduce a family of sets: for  $n \in \mathbb{N}^*$ , let

$$E_n := \left\{ (x, y) \in (\mathbb{R}_+^n)^2 : \forall i \neq j, \frac{x_i}{f(y_i)} = \frac{x_j}{f(y_j)}, \frac{y_i}{g(x_i)} = \frac{y_j}{g(x_j)}, x_i \neq x_j \right\}.$$

Notice that, since  $f$  and  $g$  are one-to-one,  $E_n$  can also be written

$$E_n = \left\{ (x, y) \in (\mathbb{R}_+^n)^2 : \forall i \neq j, \frac{x_i}{f(y_i)} = \frac{x_j}{f(y_j)}, \frac{y_i}{g(x_i)} = \frac{y_j}{g(x_j)}, y_i \neq y_j \right\}.$$

By convention, we take  $E_1 = \mathbb{R}_+^2$ . We also introduce two subsets of the above sets  $E_n$ :

$$E_n^> := \left\{ (x, y) \in E_n : \min_{i \in \{1, \dots, n\}} \left( \frac{x_i g'(x_i)}{g(x_i)} \frac{y_i f'(y_i)}{f(y_i)} \right) > 1 \right\}$$

$$E_n^{\geq} := \left\{ (x, y) \in E_n : \min_{i \in \{1, \dots, n\}} \left( \frac{x_i g'(x_i)}{g(x_i)} \frac{y_i f'(y_i)}{f(y_i)} \right) \geq 1 \right\}$$

We note that, in the particular case  $n = 1$ ,  $E_1^{\geq}$  is the closure of  $E_1^>$ .

We also introduce

$$\begin{aligned} G_n : E_n &\longrightarrow \mathbb{R}_+^2 \\ (x, y) &\longmapsto \left( \frac{x_i}{f(y_i)}, \frac{y_i}{g(x_i)} \right), \end{aligned}$$

for some  $i \in \{1, \dots, n\}$ . Remark that by the definition of  $E_n$ , the function  $G_n$  is well-defined since the choice of  $i$  is arbitrary.

The following property explains how the sets  $E_n^>$  and  $E_n^{\geq}$  and the function  $G_n$  can be used to study the number of equilibria of (2.2).

**Proposition 8.** *Let  $n \in \mathbb{N}^*$ .*

- *If  $(\alpha, \beta) \in G_n(E_n^>)$ , then system (2.2) has at least  $2n + 1$  equilibria.*
- *If system (2.2) has  $2n$  equilibria or more, then  $(\alpha, \beta) \in G_n(E_n^{\geq})$ .*

Here is a direct and useful consequence of this proposition:

**Corollary 9.** *If  $E_n^> = \emptyset$ , then for any choice of  $\alpha, \beta$ , system (2.2) has at most  $2n - 1$  equilibria.*

*Proof of Proposition 8.* Let us assume that  $E_n^> \neq \emptyset$ , and let  $(\alpha, \beta) \in G_n(E_n^>)$ . Then, there exists  $(x, y) \in E_n$  such that, for all  $i \in \{1, \dots, n\}$ ,

$$\frac{x_i g'(x_i)}{g(x_i)} \frac{y_i f'(y_i)}{f(y_i)} > 1, \quad \text{and} \quad \alpha = \frac{x_i}{f(y_i)}, \quad \beta = \frac{y_i}{g(x_i)}.$$

In other words, all the points  $(x_i, y_i)$  are equilibria of the system, and satisfy  $\alpha \beta f'(y_i) g'(x_i) > 1$ . According to Corollary 3, this proves that  $x \mapsto \alpha f(\beta g(x))$  has at least  $2n+1$  fixed points, and thus that system (2.2) has at least  $2n + 1$  equilibria.

Now, let us assume that system (2.2) has at least  $2n$  equilibria. According to Proposition 2, there exist at least  $n$  of these equilibria (denoted  $(\bar{x}_1, \bar{y}_1), \dots, (\bar{x}_n, \bar{y}_n)$ ) which satisfy  $\alpha \beta f'(\bar{y}_i) g'(\bar{x}_i) \geq 1$ . Since each  $(\bar{x}_i, \bar{y}_i)$  is an equilibrium of (2.2), then for all  $i \in \{1, \dots, n\}$ ,

$$\alpha = \frac{\bar{x}_i}{f(\bar{y}_i)}, \quad \beta = \frac{\bar{y}_i}{g(\bar{x}_i)},$$

which proves that  $(\alpha, \beta) \in G_n(E_n^{\geq})$ . □

## 2.4.2 Application to $n=1$ : determining the parameters for mono/bistability

### General statement

In [26], Cherry and Adler state conditions on  $f, g$  which ensure that, for some values of  $\alpha, \beta$ , system (2.2) is multistable. As seen in Section 2.2, under the hypotheses of Theorem 5, this system is at most bistable, which means that multistability actually simplifies to mere bistability.

In the following theorem, we improve this result in order to identify, depending on  $f$  and  $g$ , the spaces of parameters  $(\alpha, \beta)$  for which the system is monostable or bistable. From now on,  $\bar{A}$  denotes the closure of a set  $A \subset \mathbb{R}^2$ , and  $A^C$  its complement.

**Theorem 6.** *Let us define*

$$E_1^> := \left\{ (x, y) \in \mathbb{R}_+^2 : \left( \frac{x g'(x)}{g(x)} \frac{y f'(y)}{f(y)} \right) > 1 \right\}$$

and

$$G_1 : (x, y) \mapsto \left( \frac{x}{f(y)}, \frac{y}{g(x)} \right) \text{ for } (x, y) \in \mathbb{R}_+^2.$$

*Let us assume that  $f$  and  $g$  are both  $\gamma^{1/2}$ -convex, and that at least one of them is strictly  $\gamma^{1/2}$ -convex. We have the following alternative:*

- *If  $(\alpha, \beta) \in G_1(E_1^>)$ , then (2.2) has exactly three equilibria, among which exactly two stable equilibria.*
- *If  $(\alpha, \beta) \in \overline{G_1(E_1^>)}^C$ , then (2.2) has a unique equilibrium, which is stable.*

*Proof.* This is a direct consequence of Proposition 8, applied with  $n = 1$ . □

**Remark.** We note that, if  $E_1^> = \emptyset$ , *i.e.*, if

$$\sup_{x>0} \left( \left| \frac{xg'(x)}{g(x)} \right| \right) \sup_{y>0} \left( \left| \frac{yf'(y)}{f(y)} \right| \right) < 1,$$

then (2.2) is monostable for any  $\alpha, \beta$ , which is nothing but the theorem of Cherry and Adler in [26].

In general, this criterion does not lead to an explicit expression of  $G_1(E_1^>)$ , but it does for  $E_1^>$ . Hence, the set  $G_1(E_1^>)$  is numerically easily derived. These computations may be significantly facilitated by assuming that the system is symmetric, *i.e.*, of the form

$$\begin{cases} \dot{x} = \alpha f(y) - x \\ \dot{y} = \alpha f(x) - y. \end{cases} \quad (2.4)$$

In this very special case, we are led to studying the fixed points of  $x \mapsto \alpha f(\alpha f(x))$ . Since any fixed point of  $\alpha f$  is also a fixed point of  $\alpha f \circ \alpha f$ , this system has at least one ‘diagonal’ equilibrium, *i.e.*, an equilibrium of the form  $(\bar{x}, \bar{x})$ . The search for fixed points is thus made much simpler, and we obtain the following proposition.

**Proposition 10.** *Let us define*

$$E_s^> := \left\{ x \in \mathbb{R}_+ : \left| \frac{xf'(x)}{f(x)} \right| > 1 \right\}$$

and

$$\begin{aligned} G_s : E_s^> &\longrightarrow \mathbb{R}_+ \\ x &\longmapsto \frac{x}{f(x)}. \end{aligned}$$

$G_s$  is increasing if  $f$  is decreasing, and decreasing if  $f$  is increasing. Moreover,

- If  $\alpha \in G_s(E_s^>)$ , then system (2.4) has three equilibria, and is bistable.
- If  $\alpha \in \overline{G_s(E_s^>)}^C$ , then system (2.4) is monostable.

Using Theorem 6, this result can be easily proved:

*Proof.* If  $f$  is strictly decreasing, then  $G_s$  is the product of two positive and increasing functions, and is thus increasing. If  $f$  is strictly increasing, since the derivative of  $G_s$  is equal to

$$x \mapsto \frac{f(x) - xf'(x)}{f(x)^2},$$

the results holds according to the definition of  $E_s^>$ .

Let us now prove the second part of the proof. First, let us note that  $(\alpha, \alpha) \in G_1(E_1^>)$  if and only if  $\alpha \in G_s(E_s^>)$ . Indeed, if  $(\alpha, \alpha) \in G_1(E_1^>)$ , then there exists  $(x, y) \in E_1^>$  such that

$$(\alpha, \alpha) = G_1(x, y).$$

Since  $G_1(x, y) = G_1(y, x)$  for all  $(x, y)$  and  $G_1$  is injective, then  $x = y$ , which proves that

$$(\alpha, \alpha) = G_1(x, x) = (G_s(x, x), G_s(x, x)).$$

The converse inclusion is obvious from this last equality. We conclude by applying Theorem 6. □

**Remark.** In the case where the system is symmetric and cooperative (*i.e.*,  $f$  is increasing in (2.4)), all the equilibria are exactly the points of the form  $(\bar{x}, \bar{x})$ , where  $\bar{x}$  is a root of  $f_\alpha : x \mapsto \alpha f(x)$ . Indeed, let us assume that there exists  $\bar{x}$  a root of  $f_\alpha \circ f_\alpha$  which is not a root of  $f_\alpha$ . Then,  $\bar{y} := f_\alpha(\bar{x}) \neq \bar{x}$  is also a root of  $f_\alpha \circ f_\alpha$ . Let us assume that  $\bar{y} < \bar{x}$  (the case  $\bar{x} < \bar{y}$  is similar). Since  $f_\alpha$  is increasing,  $\bar{x} = f_\alpha(\bar{y}) < f_\alpha(\bar{x}) = \bar{y}$ . This leads to a contradiction, which proves the result.

On the contrary, if the system is competitive, there exists at most one ‘diagonal equilibrium’, since  $\alpha f$  has at most one fixed point if  $f$  is decreasing.

### Examples with Hill functions

We now illustrate the previous two theorems, in the case where  $f$  and  $g$  are two Hill functions (shifted or not), *i.e.*, that  $f$  and  $g$  are defined on  $\mathbb{R}_+$  by

$$f : z \mapsto \frac{1 + \lambda_1 \left(\frac{z}{z_{01}}\right)^a}{1 + \left(\frac{z}{z_{01}}\right)^a} \quad g : z \mapsto \frac{1 + \lambda_2 \left(\frac{z}{z_{02}}\right)^b}{1 + \left(\frac{z}{z_{02}}\right)^b},$$

with  $\lambda_1, \lambda_2 \in [0, 1)$ ,  $a, b \geq 1$ ,  $z_{01}, z_{02} > 0$ .

First, we note that, for all  $\alpha, \beta > 0$ ,

$$\begin{cases} \alpha f(y) - x = 0 \\ \beta g(x) - y = 0 \end{cases} \quad \text{if and only if} \quad \begin{cases} \alpha \tilde{f}\left(\frac{y}{z_{01}}\right) - x = 0 \\ \beta \tilde{g}\left(\frac{x}{z_{02}}\right) - y = 0 \end{cases},$$

where

$$\tilde{f} : z \mapsto \frac{1 + \lambda_1 z^a}{1 + z^a} \quad \text{and} \quad \tilde{g} : z \mapsto \frac{1 + \lambda_2 z^b}{1 + z^b}$$

and that

$$\alpha \beta f'(y) g'(x) = \left(\frac{\alpha}{z_{01}}\right) \left(\frac{\beta}{z_{02}}\right) \tilde{f}'\left(\frac{y}{z_{01}}\right) \tilde{g}'\left(\frac{x}{z_{02}}\right),$$

which shows that we can choose, without loss of generality,  $z_{01} = z_{02} = 1$ .

For conciseness, we present the result only in the case where  $\lambda_1 = \lambda_2 = 0$ : the complete result for general values of  $\lambda_1, \lambda_2$  as well as the proof can be found in Appendix B.1.

**Proposition 11.** *Let us assume that  $\lambda_1 = \lambda_2 = 0$  and let  $\rho := ab$ . If  $\rho < 1$ , then  $E_1^> = \emptyset$ . If  $\rho \geq 1$ , we define*

$$\begin{aligned} x^- &:= \left(\frac{1}{\rho - 1}\right)^{1/b}, & y^- &:= \left(\frac{1}{\rho - 1}\right)^{1/a}, \\ r^-(x) &:= \left(\frac{1 + x^b}{(\rho - 1)x^b - 1}\right)^{1/a}, & s^-(y) &:= \left(\frac{1 + y^a}{(\rho - 1)y^a - 1}\right)^{1/b}. \end{aligned}$$

Then

$$\begin{aligned} E_1^> &= \{(x, y) \in (x^-, +\infty) \times (y^-, +\infty) : y > r^-(x)\} \\ &= \{(x, y) \in (x^-, +\infty) \times (y^-, +\infty) : x > s^-(y)\}. \end{aligned} \tag{2.5}$$

Together with Theorem 6, this allows us to identify where, in the parameter space defined by  $(\alpha, \beta)$ , we have bistability or monostability. These results are illustrated by Figure 2.3 for various choices of  $a, b, \lambda_1$  and  $\lambda_2$ .



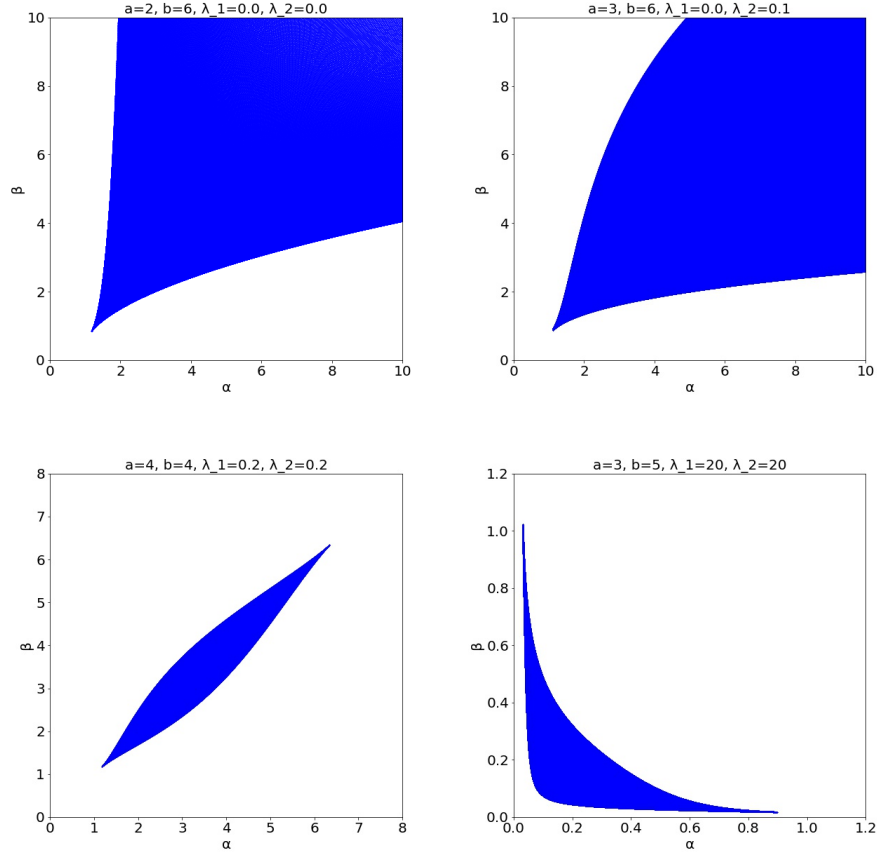


Figure 2.3: The parameter spaces for system (2.2) with Hill functions. The blue fields represent the parameters  $(\alpha, \beta)$  for which the system is bistable, and the white one the parameters for which it is monostable. The problem remains unsolved only for the values at the boundary of the blue field. The parameters of the Hill functions are indicated above each graph.

We now study the symmetric case. Still assuming that  $f$  is a (shifted) Hill function, *i.e.*,

$$f : z \mapsto \frac{1 + \lambda \left(\frac{z}{z_0}\right)^a}{1 + \left(\frac{z}{z_0}\right)^a},$$

with  $z_0 > 0$ ,  $a > 1$ ,  $\lambda \in \mathbb{R}_+ \setminus \{1\}$ , we now consider that the system is symmetric. Then, with the notations of Proposition 10, we can identify  $E_s$  and  $G_s(E_s)$ . Again, for clarity of presentation and because this is the most standard case in applications, we assume that  $\lambda = 0$ : the complete result for the other values of  $\lambda$  and the proof are gathered in Appendix B.2.

**Proposition 12.** *Let us assume that  $\lambda = 0$  and  $a > 1$ . We define:  $\alpha_0^- := z_0 \left(\frac{1}{a-1}\right)^{1/a} \frac{a}{a-1}$ . Then,*

$$G_s(E_s^>) = (\alpha_0^-, +\infty).$$

Figure 2.4 illustrates the symmetric case by showing the monostable and bistable regions of the parameter space  $(\lambda, \alpha)$ , for various choices of  $a$ . Notice that for  $\lambda = 0$ , the stable region for  $\alpha$  is indeed of the form  $(\alpha_0^-, +\infty)$ , as shown in the previous proposition.

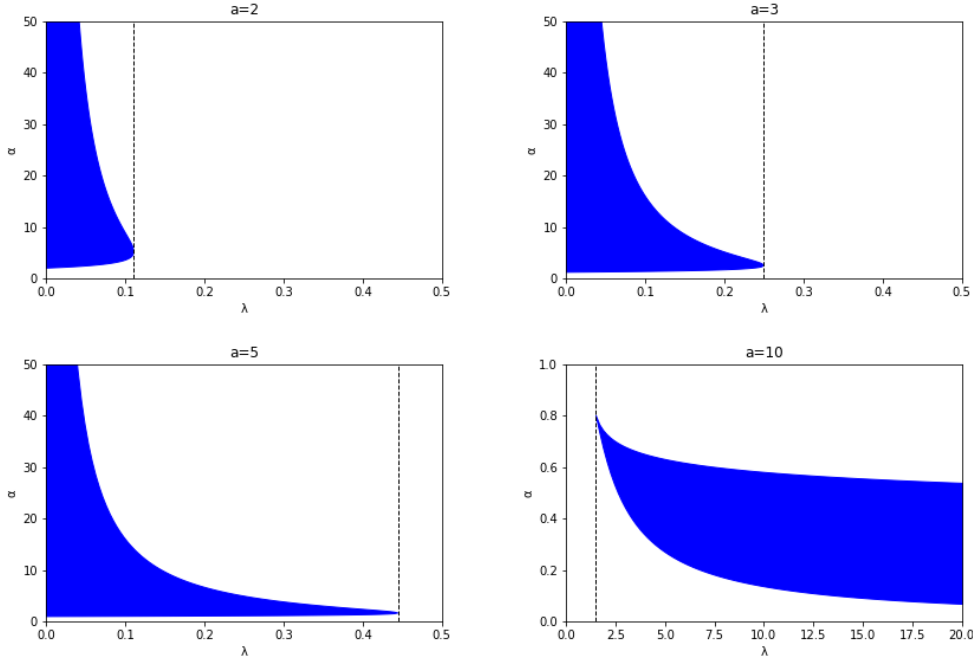


Figure 2.4: The parameter spaces  $(\lambda, \alpha)$  is the case of a symmetric system (2.4) with Hill functions. The blue set stands for the values of  $(\lambda, \alpha)$  which ensure bistability, and the white fields represent the parameters for which this system is monostable. As for the general case, the question of the number of equilibria remains unsolved for the values at the boundary of the blue field. All the Hill functions involved have a parameter  $z_0$  equal to one; the value of the parameter  $a$  is specified above each graph. The vertical dashed line indicates the highest  $\lambda$  for which a symmetric system can be bistable (the smallest one on the fourth graphic, which exemplifies the cooperative case). This value can be explicitly computed, as shown in the Appendix.

### 2.4.3 Application to $n=2$

We now apply Proposition 8 in the specific case where  $n = 2$ , in order to retrieve the result of the second section. When  $n = 2$ , we get

$$E_2 = \left\{ (x_1, x_2, y_1, y_2) \in \mathbb{R}_+^4 : \frac{x_1}{f(y_1)} = \frac{x_2}{f(y_2)}, \frac{y_1}{g(x_1)} = \frac{y_2}{g(x_2)}, x_1 \neq x_2 \right\},$$

and, since for any  $(x_1, x_2, y_1, y_2) \in E_2$ ,

$$\frac{x_i}{f(y_i)} = \frac{x_1 - x_2}{f(y_1) - f(y_2)} \quad \text{and} \quad \frac{y_i}{g(x_i)} = \frac{y_1 - y_2}{g(x_1) - g(x_2)}$$

we can rewrite

$$\begin{aligned} E_2^{\geq} &= \left\{ (x_1, x_2, y_1, y_2) \in E_2 : \frac{|x_1 - x_2|}{|g(x_1) - g(x_2)|} \frac{|y_1 - y_2|}{|f(y_1) - f(y_2)|} \min(g'(x_1)f'(y_1), g'(x_2)f'(y_2)) \geq 1 \right\} \\ &\subset \left\{ (x_1, x_2, y_1, y_2) \in \mathbb{R}_+^4 : \min(g'(x_1)f'(y_1), g'(x_2)f'(y_2)) \geq \frac{|g(x_1) - g(x_2)|}{|x_1 - x_2|} \frac{|f(y_1) - f(y_2)|}{|y_1 - y_2|} \right\}. \end{aligned}$$

Thus, if  $f$  and  $g$  are  $\gamma^{1/2}$ -convex, and one of these functions is strictly  $\gamma^{1/2}$ -convex, as seen with Proposition 6, we get, for any  $(x_1, x_2, y_1, y_2) \in \mathbb{R}_+^4$ ,

$$f'(y_1)f'(y_2) \leq \left( \frac{f(y_2) - f(y_1)}{y_2 - y_1} \right)^2 \quad \text{and} \quad g'(x_1)g'(x_2) \leq \left( \frac{g(x_2) - g(x_1)}{x_2 - x_1} \right)^2,$$

(and one of this inequality is in fact strict) and thus

$$g'(x_1)f'(y_1)f'(y_2)g'(x_2) < \left( \frac{|g(x_1) - g(x_2)|}{|x_1 - x_2|} \frac{|f(y_1) - f(y_2)|}{|y_1 - y_2|} \right)^2,$$

which proves that

$$g'(x_1)f'(y_1) < \frac{|g(x_1) - g(x_2)|}{|x_1 - x_2|} \frac{|f(y_1) - f(y_2)|}{|y_1 - y_2|} \quad \text{or} \quad g'(x_2)f'(y_2) < \frac{|g(x_1) - g(x_2)|}{|x_1 - x_2|} \frac{|f(y_1) - f(y_2)|}{|y_1 - y_2|},$$

and therefore that  $E_2^{\geq} = \emptyset$  under this hypothesis.

## 2.5 Discussion

Ordinary differential equations (ODE) are widely used in order to model ‘biological switches’ representing cell-fate decision phenomena. This work was dedicated to studying ODEs of the form

$$\begin{cases} \dot{x} = \alpha f(y) - x \\ \dot{y} = \beta g(x) - y, \end{cases} \quad f', g' < 0 \text{ or } f', g' > 0 \quad (2.6)$$

which model the behaviour of two molecules which are mutually inhibiting or activating. Such models are essential as basic elements of more complex models for cellular processes, but they have been little studied from a theoretical point of view.

Each stable equilibrium is usually associated with a given cell phenotype. As a result, any estimate for the number of equilibria these models can have is valuable to the field. This seemingly simple question (at least for the case of (2.6)) has remained up to our knowledge a vastly open question, even if numerical simulations all concur to conjecture that they cannot be more than bistable.

The seminal paper dealing with these ODE models from a theoretical point of view is due to Cherry and Adler [26]. The main result is a criterion which ensures that this model is multistable (*i.e.*, has at least two stable equilibria), for *some* parameters  $\alpha$  and  $\beta$ , leaving open the question of which parameters. More importantly, the conjecture of at most bistability was yet to be confirmed. These two problems have only recently been tackled by Li and Zhang [102], in the very particular case where  $f$  and  $g$  are two specific Hill functions.

In this chapter, we have provided a general answer to these questions. Firstly, we have established a criterion on  $f$  and  $g$  ensuring that the system has at most three equilibria, among which at most two are stable. More precisely, we have proved that if the functions  $|f'|^{-1/2}$  and  $|g'|^{-1/2}$  are strictly convex, then, for any  $\alpha, \beta > 0$ , system (2.6) has at most three equilibria, among which at most two are stable. The hypothesis regarding  $f$  and  $g$  is satisfied by all Hill functions, *i.e.*, functions of the shape  $x \mapsto \frac{1}{1+x^r}$ , but also by a wide number of usual functions and sigmoid functions. This result confirms the crucial role of self-regulation in cell differentiation phenomena, in particular when it comes to the co-existence of three different phenotypes or more. Secondly, we have developed a general method allowing to specify the parameters leading to either monostability or bistability.

We have thus proved that, with any standardly used function, system (2.6) cannot have three stable equilibria or more. The property on  $f$  and  $g$  can be used in order to ensure the monostability or bistability of more complex ODEs with three variables or more, even if, in this case, we cannot ensure the convergence of all the trajectories (some of them may be periodic).

We have also developed a framework which subsumes the result of Cherry and Adler and ours. Theoretically, it allows to find some bounds for the number of equilibria of (2.6). Nevertheless, we have not been able to clarify the conditions allowing tristability for even slightly more complex models. With the current framework, system with self-regulation, *i.e.*, systems of the shape:

$$\begin{cases} \dot{x} = \alpha\phi(x)f(y) - x \\ \dot{y} = \beta\psi(y)g(x) - y, \end{cases} \quad (2.7)$$

seem out of reach, even if numerical simulations have proved that such systems can lead to tristability [91]. Finding a criterion ensuring the tristability of (2.7), and determining the maximal number of stable equilibria it can have is certainly a next exciting and relevant step in the understanding of biological switches.



## Chapter 3

# Bistability and oscillatory behaviours of cyclic feedback loops

This chapter is a generalisation, in higher dimensions, of the results of the previous chapter. We study the stability of an Ordinary Differential Equation (ODE) usually referred to as Cyclic Feedback Loop, which typically models a biological network of  $d$  molecules where each molecule regulates its successor in a cycle ( $A_1 \rightarrow A_2 \rightarrow \dots \rightarrow A_{d-1} \rightarrow A_d \rightarrow A_1$ ). Regulations, which can be either positive or negative, are modelled by increasing or decreasing functions. We make a complete analysis of this model for a wide range of functions (including affine and Hill functions) by determining the parameters for which bistability and oscillatory behaviours arise. These results encompass previous theoretical studies of gene regulatory networks, which are particular cases of this model.

This chapter has been submitted for publication.

### 3.1 Introduction

We aim to characterise the stability of the ODE system

$$\begin{cases} \dot{x}_1 = \alpha_1 f_1(x_d) - x_1 \\ \dot{x}_2 = \alpha_2 f_2(x_1) - x_2 \\ \vdots \\ \dot{x}_d = \alpha_d f_d(x_{d-1}) - x_d \end{cases} \quad (3.1)$$

where  $f_1, \dots, f_d \in \mathcal{C}^1(\mathbb{R}_+, \mathbb{R}_+)$  (with  $\mathbb{R}_+ = [0, +\infty)$ ) are non-negative functions, at least one of them is bounded, and  $\alpha_1, \dots, \alpha_d$  are positive parameters. Throughout this chapter, we use the convention  $x_0 = x_d$ , which allows us to write (3.1) under the compacted form

$$\forall i \in \{1, \dots, d\}, \quad \dot{x}_i = \alpha_i f_i(x_{i-1}) - x_i.$$

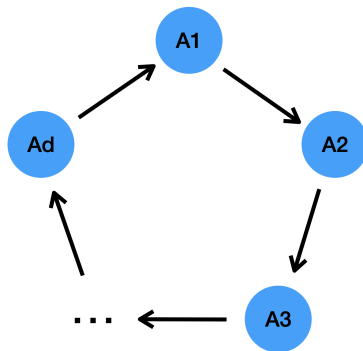


Figure 3.1: Schematic representation of the *cyclic feedback loop* (3.1): The blue circles represent the molecules of the network, and the arrows between them regulation, which can be positive ( $f_i$  increasing) or negative ( $f_i$  decreasing).

This model is a generalisation of a gene regulatory network initially proposed by Goodwin [64, 65], usually referred to as *Cyclic feedback loop*, which represents interactions between genes, mRNAs, enzymes and proteins called repressors which have the ability to inhibit the expression of some genes. In system (3.1),  $x_1, \dots, x_d$  represent the concentration of each of the molecules involved in the network (denoted  $A_1, \dots, A_n$ ), and  $f_1, \dots, f_d$  the regulation between them. The system is assumed to be cyclic ( $A_i$  regulates  $A_{i+1}$  and only  $A_{i+1}$ , as illustrated by Figure 3.1), and each regulation can be positive ( $f_i$  increasing) or negative ( $f_i$  decreasing). The relevance of these cyclic models has been established in [5] and [60] where some theoretical predictions (oscillatory phenomena and bistability) have been observed experimentally. This highlights the importance of understanding the dynamical behaviour of such systems *i.e.* determining the number of stable equilibrium points and their basins of attraction, as well as the possible existence of periodic solutions or chaotic behaviours.

System (3.1) has been, for some specific choices of  $f_i$ , the subject of several theoretical studies [7, 122, 143, 146, 147, 161]. In these papers, restrictions on the functions  $f_i$  were notably imposed by the necessity to compute the value of the equilibrium points of the system, which is intricate when more than two functions are not affine, and are not identical. In the present chapter, we follow a method initiated by Cherry and Adler [26] allowing to avoid explicitly computing the equilibrium points. In the two dimensional case, which writes

$$\begin{cases} \dot{x}_1 = \alpha_1 f_1(x_2) - x_1 \\ \dot{x}_2 = \alpha_2 f_2(x_1) - x_2 \end{cases}, \quad (3.2)$$

the following results have already been established:

1. If  $f_1$  and  $f_2$  are both increasing or both decreasing, and if

$$\sup_{x>0} \left| \frac{x f_1'(x)}{f_1(x)} \right| \sup_{x>0} \left| \frac{x f_2'(x)}{f_2(x)} \right| > 1, \quad (3.3)$$

then there exist values of  $(\alpha_1, \alpha_2) \in \mathbb{R}_+^{*2}$  such that system (3.2) is multistable *i.e.* there exist at least two equilibrium points which are asymptotically stable [26].

2. If  $f_1$  and  $f_2$  are both increasing or both decreasing, and if

$$\frac{1}{\sqrt{|f_1'|}} \quad \text{and} \quad \frac{1}{\sqrt{|f_2'|}} \quad \text{are strictly convex}^1, \quad (3.4)$$

then system (3.2) is either monostable or bistable, *i.e.* there exist exactly one or exactly two equilibrium points which are asymptotically stable. Moreover, it is possible to determine, up to a set of measure zero, the set of parameters for  $(\alpha_1, \alpha_2)$  for which the system is monostable and the set of parameters for which it is bistable [68].

3. All the solutions to system (3.2) converge (even without assuming any monotonicity).

This last result is a direct application of the Poincaré-Bendixson theorem and the Dulac-Bendixson theorem [130].

A natural question at this stage is which of these properties generalise to higher dimensions ( $d \geq 3$ ). A major result was achieved by Mallet-Paret and Smith in [113], showing that the Poincaré-Bendixson theorem can be adapted to monotone feedback systems, including (3.1). Furthermore, a series of results of Hirsch [77–79] generalised and summarised in [147], has shown that in the case where an even number of functions  $f_i$  is decreasing, the solutions of (3.1) converge to an equilibrium point for almost every initial condition. It is well-known that this property does not hold when the number of decreasing functions is odd (in which case the system is often called ‘repressilator’), and that there can there exist stable orbits [19, 122].

In this chapter, we prove the following result, which is a generalisation of the result of Chapter 2:

**Theorem 7.** *Let us assume that  $f_1, \dots, f_d \in \mathcal{C}^3(\mathbb{R}_+, \mathbb{R}_+^*)$  are monotonous, non-negative, and that (at least) one of them is bounded. Moreover, let us assume that the functions  $\frac{1}{\sqrt{|f_1'|}}, \dots, \frac{1}{\sqrt{|f_d'|}}$  are defined and convex, and that (at least) one of them is strictly convex. Lastly, let us denote by  $n$  the number of these functions which are decreasing, and let*

$$D := \prod_{k=1}^d \sup_{x>0} \left| \frac{x f_k'(x)}{f_k(x)} \right| \in (0, +\infty]$$

1. If  $n$  is even, then

(i) *If  $D < 1$ , then for any  $\alpha \in (\mathbb{R}_+^*)^d$  system (3.1) has a unique equilibrium point which is globally asymptotically stable.*

---

<sup>1</sup>It is in fact sufficient to assume that these two functions are convex, and that at least one of them is strictly convex.



- (ii) If  $D > 1$ , then there exists a non-empty set  $A_{bis} \subset (\mathbb{R}_+^*)^d$  such that
- If  $\alpha \in A_{bis}$ , then system (3.1) has exactly two asymptotically stable equilibria, and the union of their basins of attraction is a dense open subset of  $\mathbb{R}_+^d$ , with a complement of Lebesgue measure zero.
  - If  $\alpha \in \overline{A_{bis}}^C$ , then system (3.1) has a unique equilibrium point which is globally asymptotically stable.
- (iii) If  $d \geq 5$  and  $D > \frac{1}{\cos(\frac{2\pi}{d})^d}$ , then there exists a non-empty set  $A_{per} \subset A_{bis}$  such that if  $\alpha \in A_{per}$ , then system (3.1) has periodic solutions.

2. If  $n$  is odd, then system (3.1) has a unique equilibrium point. Moreover, if  $d \geq 3$ , then

- (i) If  $D < \frac{1}{\cos(\frac{\pi}{d})^d}$ , then for any  $\alpha \in (\mathbb{R}_+^*)^d$ , this equilibrium point is asymptotically stable, and all the solutions of (3.1) either converge to this point or to a periodic orbit.
- (ii) If  $D > \frac{1}{\cos(\frac{\pi}{d})^d}$ , then there exists  $A_{unst} \subset (\mathbb{R}_+^*)^d$  a non-empty set such that
- If  $\alpha \in A_{unst}$ , then this equilibrium point is asymptotically unstable, and there exists a finite number of periodic solutions, among which at least one is asymptotically stable. Moreover, the set of initial conditions for which the solution converges to a periodic solution is a dense open subset of  $\mathbb{R}_+^d$ , and its complement, which is the set of initial conditions for which the solution converges to the equilibrium point, has Lebesgue measure zero.
  - If  $\alpha \in \overline{A_{unst}}^C$ , this equilibrium point is asymptotically stable, and all the solutions of (3.1) either converge to this point or to a periodic orbit.

In each of these cases, the sets  $A_{bis}$ ,  $A_{per}$ ,  $A_{unst}$  can be explicitly expressed, as we will show in Sections 3.3 and 3.4. Moreover, note that any Hill function (even shifted), *i.e.* function of the form  $x \mapsto \frac{1+\lambda x^r}{1+x^r}$ , with  $\lambda \in \mathbb{R}_+ \setminus \{1\}$  and  $r \geq 1$ , as well as linear functions, satisfy the convexity hypothesis of this theorem, which means that this result encompasses the other theoretical studies mentioned above [7, 122, 143, 146, 147, 161].

It is worth noting that, when  $n$  is even, the set of initial conditions for which the solution converges to a periodic orbit has Lebesgue measure zero, and can therefore hardly be reached numerically. Nevertheless, we highlight this result since, up to our knowledge, the question of the existence of such periodic solution remained open, as mentioned in [147]. Moreover, the proof of this result, which uses the stable manifold theorem, seems to us non-trivial and worthwhile. We also note that condition 1 (iii) is sufficient, but perhaps not necessary for the existence of periodic solution: in particular, this question for  $d \in \{3, 4\}$  remains open. Lastly, we do not know if periodic solutions do exist under the hypotheses of 2 (i) and in the second point of 2 (ii), but our study does not rule out this possibility.

This article, entirely dedicated to the proof of Theorem 7, is organised as follows: after characterising the equilibrium points of system (3.1) in Section 3.2, we prove the theorem when  $n$  is even (Section 3.3), before dealing with the case where  $n$  is odd, which is simpler, in the last section.

## 3.2 Characterisation of fixed points

Throughout this chapter, we assume that  $f_1, \dots, f_d \in \mathcal{C}^3(\mathbb{R}_+, \mathbb{R}_+^*)$  are non-negative, monotonous and that at least one of these functions is bounded. Note that, according

to the regularity of these functions, the Cauchy-Lipschitz theorem ensures the local existence and the uniqueness of the solution of this equation for any initial condition  $x_0 \in \mathbb{R}_+^d$ . Furthermore, the positivity of the functions guarantees that the solutions remain in  $\mathbb{R}_+^d$ . Lastly, using the fact that one of them is bounded, we easily prove the global existence of solutions and the existence of a compact attractor set, *i.e.* the existence of a compact set  $K \subset \mathbb{R}_+^d$  such that for any initial condition  $x_0 \in \mathbb{R}_+^d$ , there exists  $T \geq 0$  such that  $x(t) \in K$  for all  $t \geq T$ .

We now make the additional assumption that  $f_1, \dots, f_d$  are  $\gamma^{1/2}$ -convex, and that at least one of them is strictly  $\gamma^{1/2}$ -convex, *i.e.* they satisfy the following definition:

**Definition 3** ( $\gamma^{1/2}$ -convexity). Let  $f \in \mathcal{C}^3(\mathbb{R}_+, \mathbb{R}_+)$  a non-negative and monotonous function. We say that  $f$  is (strictly)  $\gamma^{1/2}$ -convex if  $|f'| > 0$  and  $\frac{1}{\sqrt{|f'|}}$  is (strictly) convex.

Note that this definition can be related to the definition of the Schwartzian derivative of  $f$  defined by  $S(f) = \frac{f'''}{f'} - \frac{3}{2} \left( \frac{f''}{f'} \right)^2$  by noting that  $\left( \frac{1}{\sqrt{|f'|}} \right)'' = -\frac{1}{2} \frac{1}{\sqrt{|f'|}} S(f)$ .

We relate here the equilibrium points of system (3.1) to the fixed points of an auxiliary function  $\tilde{f}$ . The  $\gamma^{1/2}$ -convexity of the functions  $f_i$  ensures that the number of equilibrium points cannot exceed three, and provides a criterion which characterises this exact number of equilibria. A similar approach was used, (with the Schwartzian derivative) for a particular case of this system in [122].

We start by recalling some key properties of the  $\gamma^{1/2}$ -convexity, which have been established in Chapter 2.

**Proposition 13.** Let  $f, g$  be two  $\gamma^{1/2}$ -convex functions,  $c > 0$ .

- (i)  $f \circ g$  and  $cf$  are  $\gamma^{1/2}$ -convex. Moreover, if  $f$  or  $g$  is strictly  $\gamma^{1/2}$ -convex, then  $f \circ g$  is strictly  $\gamma^{1/2}$ -convex.

Let us now assume that  $f$  is strictly  $\gamma^{1/2}$ -convex. Then:

- (ii)  $cf$  is strictly  $\gamma^{1/2}$ -convex.
- (iii)  $f$  has at most three fixed points.
- (iv) If all the fixed points  $x$  of  $f$  satisfy  $f'(x) < 1$ , then  $f$  has a unique fixed point.
- (v) If there exists a fixed point of  $f$  (denoted  $x$ ) such that  $f'(x) > 1$ , then  $f$  has exactly three fixed points and the other two fixed points (denoted  $y, z$ ) satisfy  $f'(y) < 1$  and  $f'(z) < 1$ .

**Example.** For any  $r \geq 1$ ,  $a, b, c, d \geq 0$  such that  $ad - bc \neq 0$ , the function  $x \mapsto \frac{ax^r + b}{cx^r + d}$  is  $\gamma^{1/2}$ -convex. Moreover, if  $r > 1$ , then it is strictly  $\gamma^{1/2}$ -convex. In particular, affine functions and Hill functions are  $\gamma^{1/2}$ -convex.<sup>2</sup>

By definition, the point  $\bar{x} = (\bar{x}_1, \dots, \bar{x}_d) \in \mathbb{R}_+^d$  is an equilibrium point of (3.1) if and only if

$$\begin{cases} \bar{x}_1 = \alpha_1 f_1(\bar{x}_d) \\ \bar{x}_2 = \alpha_2 f_2(\bar{x}_1) \\ \vdots \\ \bar{x}_d = \alpha_d f_d(\bar{x}_{d-1}) \end{cases} \iff \begin{cases} \bar{x}_1 = \alpha_1 f_1(\bar{x}_d) \\ \bar{x}_2 = \alpha_2 f_2(\bar{x}_1) \\ \vdots \\ \bar{x}_{d-1} = \alpha_{d-1} f_{d-1}(\bar{x}_{d-2}) \\ \bar{x}_d = \alpha_d f_d \circ \alpha_{d-1} f_{d-1} \circ \dots \circ \alpha_1 f_1(\bar{x}_d) \end{cases} .$$

<sup>2</sup>In the Appendix of Chapter 2, we show that many other usual sigmoid functions are strictly  $\gamma^{1/2}$ -convex.

Thus, the number of equilibrium points of (3.1) is equal to the number of fixed points of  $\tilde{f} := \alpha_d f_d \circ \dots \circ \alpha_1 f_1$ . Since  $\tilde{f}(0) > 0$  and  $\tilde{f}$  is bounded, it proves in particular that (3.1) has at least one equilibrium point. Note that, according to the first two properties of Proposition 13,  $\tilde{f}$  is strictly  $\gamma^{1/2}$ -convex, and a direct computation shows that

$$\tilde{f}'(\bar{x}_d) = \prod_{i=1}^d \alpha_i f'_i(\bar{x}_{i-1}).$$

We can thus apply the fourth and the fifth properties of Proposition 13 to  $\tilde{f}$  to derive the following lemma:

**Lemma 2.** *Let us denote, for all  $\alpha, x \in \mathbb{R}_+^d$ ,  $p_x^\alpha = \prod_{i=1}^d \alpha_i f'_i(x_{i-1})$ .*

- *If all the equilibrium points of system (3.1) satisfy  $p_x^\alpha < 1$ , then this system has a unique equilibrium point.*
- *If there exists an equilibrium point of system (3.1) (denoted  $\bar{x}$ ) such that  $p_{\bar{x}}^\alpha > 1$ , then this system has exactly three equilibrium points, and the other two points (denoted  $\bar{y}, \bar{z}$ ) satisfy  $p_{\bar{y}}^\alpha < 1$ ,  $p_{\bar{z}}^\alpha < 1$ .*

Hence, the value of  $p_{\bar{x}}^\alpha$  characterises the number of fixed points the system has. In the following section, we show that it also determines the dimension of the basin of attraction of  $\bar{x}$ .

### 3.3 Even number of decreasing functions

In the case where  $n$  is even, one easily checks that system (3.1) is an *irreducible type K monotone system* in the sense defined in [147]. As seen in the previous section, this system has a finite number of equilibrium points (at most three), and a compact attractor set: thus, we can apply Theorem 2.5 and Theorem 2.6 of [147] which prove that the union of the basins of attraction of the equilibrium points is dense, and that the complement of this set has Lebesgue measure zero.

In this section, we complete this result in two ways:

- We determine, for given functions  $f_1, \dots, f_d$ , a set of parameters  $A_{\text{bis}}$  such that system (3.1) is bistable (*i.e.* has exactly two asymptotically stable equilibrium points) if  $\alpha \in A_{\text{bis}}$ , and monostable (*i.e.* has exactly one asymptotically stable equilibrium point) if  $\alpha \in \overline{A_{\text{bis}}}^C$ .
- We determine a set  $A_{\text{per}} \subset A_{\text{bis}}$  such that system (3.1) has some periodic solutions if  $\alpha \in A_{\text{per}}$ .

Note that this last point does not mean that periodic solutions do not exist when  $\alpha \notin A_{\text{per}}$ , and that, in all cases, the set of initial conditions for which the solution converges to a periodic solution has Lebesgue measure zero (as a corollary of [147]).

In order to prove these two points, we determine the dimension of the basin of attraction of an equilibrium point  $\bar{x}$ , as a function of  $p_{\bar{x}}^\alpha$ . For any equilibrium point  $\bar{x}$ , we denote its basin of attraction  $B_{\bar{x}}$ . Our reasoning is based on the stable manifold theorem (the proof of which can be found for instance in [130]), that we recall:

**Theorem** (Stable manifold). *Let  $F \in \mathcal{C}^1(\mathbb{R}^d, \mathbb{R}^d)$  a vector field, and let  $\bar{x} \in \mathbb{R}^d$  such that  $F(\bar{x}) = 0_{\mathbb{R}^d}$ . If  $\bar{x}$  is a hyperbolic equilibrium point, *i.e.* if all the eigenvalues of  $\text{Jac } F(\bar{x})$  have a non-zero real part, then the basin of attraction of  $\bar{x}$  is a manifold of dimension  $m$ , where  $m$  is the number of eigenvalues of  $\text{Jac } F(\bar{x})$  with a negative real part.*

In order to apply this theorem, we need to compute the eigenvalues of the Jacobian matrix associated to system (3.1) which writes, at a given point  $x \in \mathbb{R}^d$

$$M_x^\alpha = \begin{pmatrix} -1 & 0 & \cdots & 0 & \alpha_1 f_1'(x_d) \\ \alpha_2 f_2'(x_1) & -1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \alpha_d f_d'(x_{d-1}) & -1 \end{pmatrix}.$$

Thus, the characteristic polynomial of  $M_x^\alpha$  is easily computed to be

$$(-1)^d \left( (\lambda + 1)^d - \prod_{i=1}^d \alpha_i f_i'(x_{i-1}) \right) = (-1)^d \left( (\lambda + 1)^d - p_x^\alpha \right)$$

Since  $n$  is even,  $p_x^\alpha > 0$  and hence the spectrum of  $M_x^\alpha$  is given by

$$\text{Sp}(M_x^\alpha) = \left\{ (p_x^\alpha)^{1/d} e^{2k\pi i/d} - 1, k \in \{0, \dots, d-1\} \right\}.$$

We deduce that

- (i) If  $p_x^\alpha < 1$ , then all the eigenvalues of  $M_x^\alpha$  have a negative real part.
- (ii) If  $d \in \{3, 4\}$  and  $p_x^\alpha > 1$ , or if  $d \geq 5$  and  $p_x^\alpha \in \left(1, \frac{1}{\cos(\frac{2\pi}{d})^d}\right)$ , then  $M_x^\alpha$  has exactly  $d-1$  eigenvalues with a negative real part, and one with a positive real part.
- (iii) If  $d \geq 5$ , and if  $p_x^\alpha > \frac{1}{\cos(\frac{2\pi}{d})^d}$ , then  $M_x^\alpha$  has at most  $d-3$  eigenvalues with a negative real part. Moreover, if  $d \leq 8$ , or  $d \in \llbracket 4j+1, 4j+4 \rrbracket$  ( $j \in \mathbb{N} \setminus \{0, 1\}$ ), and for any  $k \in \{2, \dots, j\}$ ,  $p_x^\alpha \neq \frac{1}{\cos(\frac{2\pi k}{d})^d}$ , then all eigenvalues of  $M_x^\alpha$  have a non-zero real part (where for all  $a, b \in \mathbb{R}$ ,  $a < b$ ,  $\llbracket a, b \rrbracket = [a, b] \cap \mathbb{N}$ .)

Thus, the stable manifold theorem yields

**Lemma 3.** *Let us assume that  $n$  is even, and let  $\bar{x} \in \mathbb{R}^d$  be an equilibrium point of (3.1).*

- (i) *If  $p_{\bar{x}}^\alpha < 1$ , then  $\dim(B_{\bar{x}}) = d$ , i.e.  $B_{\bar{x}}$  is an open set.*
- (ii) *If  $d \in \{3, 4\}$  and  $p_{\bar{x}}^\alpha > 1$ , or if  $d \geq 5$  and  $p_{\bar{x}}^\alpha \in \left(1, \frac{1}{\cos(\frac{2\pi}{d})^d}\right)$ , then  $\dim(B_{\bar{x}}) = d-1$ .*
- (iii) *If  $d \geq 5$ ,  $p_{\bar{x}}^\alpha > \frac{1}{\cos(\frac{2\pi}{d})^d}$ , and  $p_{\bar{x}}^\alpha \notin S_d$ , with*

$$S_d := \begin{cases} \emptyset & \text{if } d \in \{5, 6, 7, 8\} \\ \left\{ \frac{1}{\cos(\frac{2\pi k}{d})^d}, k \in \{2, \dots, j\} \right\} & \text{if } d \in \llbracket 4j+1, 4j+4 \rrbracket, \quad j \geq 2 \end{cases},$$

*then  $\dim(B_{\bar{x}}) \leq d-3$ .*

Moreover, we easily check that, in all cases,  $M_{\bar{x}}^\alpha$  has an odd number of eigenvalues with a positive real part. Thus, since  $n$  is even, we get from the main theorem of [113] that any solution converges to an equilibrium point or to a periodic orbit.

We will now use Lemmas 3 and 4 to prove Theorem 7 in a more precise form which specifies the sets  $A_{\text{bis}}$  and  $A_{\text{per}}$ , in the case where  $n$  is even. Before stating it, we give a last lemma linking the dimension of the basin of attraction of the unstable equilibrium point to the existence of divergent solutions (which thus converge to periodic solutions), in the bistable case.

**Lemma 4.** *Let ‘ $\dot{x} = F(x)$ ’ be an ODE which has exactly three equilibrium points, (denoted  $\bar{x}, \bar{y}, \bar{z}$ ) assumed hyperbolic, and let us assume that for all initial condition  $x_0 \in \mathbb{R}^d$ , the solution of this ODE is defined on  $\mathbb{R}_+$  and converges. If  $\bar{y}, \bar{z}$  are asymptotically stable, then  $\dim(B_{\bar{x}}) = d - 1$ .*

*Proof.* First, let us note that, according to the stable manifold theorem,  $B_{\bar{y}}$  and  $B_{\bar{z}}$  are two open sets, and that  $B_{\bar{x}}$  is a manifold. Since, by hypothesis, all the solutions converge,  $\mathbb{R}_+^d = B_{\bar{x}} \cup B_{\bar{y}} \cup B_{\bar{z}}$ , which means that  $B_{\bar{x}}$  separates  $\mathbb{R}_+^d$  in two open sets. As shown in [85] (Corollary 1 of Theorem IV 4), and by the connectedness of  $\mathbb{R}_+^d$ , this is possible only if  $\dim(B_{\bar{x}}) = d - 1$ .  $\square$

Before stating our theorem, let us introduce the functions  $\Gamma$  and  $G$ , defined for any  $x \in \mathbb{R}_+^d$  by  $\Gamma(x) = \left( \frac{x_1}{f_1(x_d)}, \frac{x_2}{f_2(x_1)}, \dots, \frac{x_d}{f_d(x_{d-1})} \right)$  and  $G(x) = \prod_{i=1}^d \frac{x_i f'(x_i)}{f(x_i)}$ , and the sets  $E_{\text{bis}} = \{x \in \mathbb{R}_+^d : G(x) > 1\}$  and if  $d \geq 5$ ,  $E_{\text{per}} = \left\{ x \in \mathbb{R}_+^d : G(x) > \frac{1}{\cos(\frac{2\pi}{d})^d}, G(x) \notin S_d \right\}$ . Note that  $E_{\text{per}} \subset E_{\text{bis}}$ .

**Theorem 8.** *Let us assume that  $n$  is even.*

- (i) *If  $\alpha \in \overline{\Gamma(E_{\text{bis}})}^C$ , then (3.1) has a unique equilibrium point which is globally asymptotically stable.*
- (ii) *If  $\alpha \in \Gamma(E_{\text{bis}})$ , then (3.1) has exactly three equilibrium points, among which two are asymptotically stable and one is asymptotically unstable. Moreover, the union of the basins of attraction of the two stable equilibria is a dense open subset of  $\mathbb{R}_+^d$ .*
- (iii) *If  $d \geq 5$ , and if  $\alpha \in \Gamma(E_{\text{per}})$ , then there exist periodic solutions of (3.1).*

*Proof.* First, let us note that, according to the definitions of  $\Gamma$  and  $G$ ,  $\bar{x} \in \mathbb{R}_+^d$  is a fixed point of system (3.1) if and only if  $\alpha = \Gamma(\bar{x})$ , and that for any equilibrium point  $\bar{x}$  of (3.1),

$$p_{\bar{x}}^\alpha = p_{\bar{x}}^{\Gamma(\bar{x})} = G(\bar{x}).$$

- (i) Let us assume that  $\alpha \in \overline{\Gamma(E_{\text{bis}})}^C$ , and let  $\bar{x}$  be a fixed point of (3.1). Since  $\Gamma(\overline{E_{\text{bis}}}) \subset \overline{\Gamma(E_{\text{bis}})}$ , and  $\alpha = \Gamma(\bar{x})$ ,  $\bar{x} \in \overline{\Gamma(E_{\text{bis}})}^C$ , which means, by definition of  $E_{\text{bis}}$ , that  $p_{\bar{x}}^\alpha = G(\bar{x}) < 1$ . Since this equality holds for any equilibrium point, we conclude by Lemma 2, that system (3.1) has a unique equilibrium point, which is asymptotically stable, by Lemma 3. Since (3.1) is an irreducible type K monotone system, this unique equilibrium is in fact globally asymptotically stable [147].
- (ii) Let us assume that  $\alpha \in \Gamma(E_{\text{bis}})$ . Then, there exists  $\bar{x} \in E_{\text{bis}}$  such that  $\alpha = \Gamma(\bar{x})$ . The point  $\bar{x}$  is thus an equilibrium point of (3.1) which satisfies  $p_{\bar{x}}^\alpha = G(\bar{x}) > 1$ . Therefore, Lemmas 2, 3 and the result of [147] mentioned at the beginning of the section yield the result.
- (iii) Let us assume that  $\alpha \in \Gamma(E_{\text{per}})$ . Since  $E_{\text{per}} \subset E_{\text{bis}}$ , (3.1) has exactly three equilibrium points, denoted  $\bar{x}, \bar{y}$  and  $\bar{z}$ , with  $\bar{x} \in E_{\text{per}}$ , and  $\bar{y}$  and  $\bar{z}$  which are asymptotically

stable. By definition of  $E_{\text{per}}$ ,  $\dim(B_{\bar{x}}) \leq d - 3$ . By the contrapositive of Lemma 4, system (3.1) has some divergent solutions, which thus converge to a periodic orbit, according to [113].

□

**Remark.** As mentioned above, this theorem is a more precise version of Theorem 7: we find the statement of the latter by defining  $A_{\text{bis}} = \Gamma(E_{\text{bis}})$  and  $A_{\text{per}} = \Gamma(E_{\text{per}})$ , and by noting that  $E_{\text{bis}}$  (resp.  $E_{\text{per}}$ ) is empty if and only if  $D \leq 1$  (resp.  $D \leq \frac{1}{\cos(\frac{2\pi}{d})}$ ).

### 3.4 Odd number of decreasing functions

We now deal with the case where  $n$  is odd. This case is simpler, since the system has a unique equilibrium point under this hypothesis. Nevertheless, we make weaker conclusions regarding the global behaviour of the system, since it is not an irreducible type K monotone system (see [147]). Thus, we simply study the linearised system at the neighbourhood of the equilibrium point, and we conclude with [113], which guarantees that the solutions either converge to this equilibrium point, or to a periodic orbit.

Let us denote  $E_{\text{unst}} := \left\{ x \in \mathbb{R}^d : G(x) < -\frac{1}{\cos(\frac{\pi}{d})^d} \right\}$ .

We get the following result:

**Theorem 9.** *Let us assume that  $n$  is odd, and that  $d \geq 3$ . Then, system (3.1) has a unique equilibrium point. Moreover,*

- (i) *If  $\alpha \in \overline{\Gamma(E_{\text{unst}})}^C$ , then this equilibrium point is asymptotically stable. Moreover, all the solutions of (3.1) either converge to this point or to a periodic orbit.*
- (ii) *If  $\alpha \in \Gamma(E_{\text{unst}})$ , then this equilibrium point is asymptotically unstable. Moreover, the set of initial conditions for which the solution converges to a periodic solution is a dense open subset of  $(\mathbb{R}_+^*)^d$ , and its complement, which is the set of initial conditions for which the solution converges to the equilibrium point, has Lebesgue measure zero.*

*Proof.* We use the same notations as in the previous section. Since  $n$  is odd,  $\tilde{f}$  is decreasing with  $\tilde{f}(0) > 0$ , it has a unique fixed point, which implies that system (3.1) has a unique equilibrium point, that we denote  $\bar{x}$ . This point is asymptotically stable if all the eigenvalues of  $M_{\bar{x}}^\alpha$  have a negative real part, and asymptotically unstable if at least one of these eigenvalues has a positive real part. Since  $n$  is odd,

$$\text{Sp}(M_{\bar{x}}^\alpha) = \left\{ (|p_{\bar{x}}^\alpha|)^{1/d} e^{(2k+1)\pi i/d} - 1, k \in \{0, \dots, d-1\} \right\},$$

which implies that  $\bar{x}$  is stable if  $|p_{\bar{x}}^\alpha| < \frac{1}{\cos(\frac{\pi}{d})^d}$ , and unstable if  $|p_{\bar{x}}^\alpha| > \frac{1}{\cos(\frac{\pi}{d})^d}$ . We conclude

by noting that  $|p_{\bar{x}}^\alpha| = -p_{\bar{x}}^\alpha = -p_{\bar{x}}^{\Gamma(\bar{x})} = -G(\bar{x})$ , and by applying the main theorem of [113] for the first point, and Theorem 4.3 of this same article for the second one.

□

We recover the result of Theorem 7 by defining  $A_{\text{unst}} = \Gamma(E_{\text{unst}})$ .



## Chapter 4

# Long-time behaviour of an advection-selection equation

We study the long-time behaviour of the advection-selection equation

$$\partial_t n(t, x) + \nabla \cdot (f(x)n(t, x)) = (r(x) - \rho(t))n(t, x), \quad \rho(t) = \int_{\mathbb{R}^d} n(t, x) dx \quad t \geq 0, x \in \mathbb{R}^d,$$

with an initial condition  $n(0, \cdot) = n^0$ . In the field of adaptive dynamics, this equation typically describes the evolution of a phenotype-structured population over time. In this case,  $x \mapsto n(t, x)$  represents the density of the population characterised by a phenotypic trait  $x$ , the advection term ' $\nabla \cdot (f(x)n(t, x))$ ' a cell differentiation phenomenon driving the individuals toward specific regions, and the selection term ' $(r(x) - \rho(t))n(t, x)$ ' the growth of the population, which is of logistic type through the total population size  $\rho(t) = \int_{\mathbb{R}^d} n(t, x) dx$ .

In the one-dimensional case  $x \in \mathbb{R}$ , we prove that the solution to this equation can either converge to a weighted Dirac mass or to a function in  $L^1$ . Depending on the parameters  $n^0$ ,  $f$  and  $r$ , we determine which of these two regimes of convergence occurs, and we specify the weight and the point where the Dirac mass is supported, or the expression of the  $L^1$ -function which is reached.

This Chapter has been written with the collaboration of Nastassia Pouradier Duteil and Camille Pouchol, and has been submitted for publication.



## 4.1 Introduction

### 4.1.1 Advection-selection equation

We consider the asymptotic behaviour of the advection-selection equation

$$\begin{cases} \partial_t n(t, x) + \nabla \cdot (f(x)n(t, x)) = (r(x) - \rho(t)) n(t, x), & t \geq 0, x \in \mathbb{R}^d \\ \rho(t) = \int_{\mathbb{R}^d} n(t, x) dx, & t \geq 0 \\ n(0, x) = n^0(x), & x \in \mathbb{R}^d. \end{cases} \quad (4.1)$$

This type of model typically comes up in the field of *adaptive dynamics*. The aim is to understand how, among heterogeneous populations of individuals structured by a so-called continuous *trait* or *phenotype*  $x$ , the distribution of the density  $x \mapsto n(t, x)$  evolves over time, and which phenotypes prevail in large times  $t \rightarrow +\infty$ .

In the model above (4.1), the partial differential equation (PDE) takes into account

- advection via the term  $\nabla \cdot (f(x)n(t, x))$ , whereby individuals follow the flow associated with  $f$ ,
- growth via the term  $(r(x) - \rho(t))n(t, x)$ , which is of logistic type through the total population size  $\rho(t) = \int_{\mathbb{R}^d} n(t, x) dx$ .

The literature concerning so-called phenotype-structured partial differential equations for adaptive dynamics is abundant [1, 10, 14, 20, 21, 30, 42, 50, 105, 107, 131, 132]. These models usually take into account selection, which favors individuals with the most adapted traits in terms of growth, and mutations, which induce a slight phenotypic change upon reproduction. Mutation is often assumed to be rare and small compared to selection, [49, 62, 117]. Models with no mutation at all have also been the subject of several studies [2, 48, 69, 86, 106, 134].

One way to analyse how the population adapts is to study the long-time behaviour for solutions of such PDE models. In particular, determining if the population becomes monomorphic (*i.e.* the solution concentrates around a certain trait, called *Evolutionary Stable Strategy* (ESS) [76]), or if phenotypic diversity is preserved is a fundamental question when studying such models. Broadly speaking, it has been shown that selection leads to concentration (around a finite number of phenotypic traits), while mutations, on the contrary, tend to regularise solutions, and, possibly, their limits [13, 69].

However, less emphasis has been put on studying the effect of advection, except for the recent few examples [30, 31, 104] where most results are of numerical nature, or assume a very specific form of the functions  $r$  and  $f$ .

Yet, considering advection is relevant in various contexts. From the phenomenological point of view, it may represent how the environment drives the individuals towards specific regions, as opposed to more random mutations. It is also the rigorous way to model phenotype changes that are intrinsic to the individual, mediated by an ordinary differential equation (ODE) of the form

$$\dot{x}(t) = f(x(t)), \quad (4.2)$$

where  $x(t) \in \mathbb{R}^d$  denotes the phenotypic trait of the individual at time  $t \geq 0$ . As is well known, the PDE for the density of individuals corresponding to the sole model (4.2) is indeed the advection equation  $\partial_t n(t, x) + \nabla \cdot (f(x)n(t, x)) = 0$ . Our original motivation is that of cell differentiation, for which very refined ODE models have been developed in systems biology (see for instance [67, 150, 157, 165]).

The goal of the present article is to investigate the combined effect of selection and advection, assuming that mutations are absent or sufficiently small to be neglected. We hence study the long-time behaviour of the PDE (4.1), where  $n^0$  is the initial population distribution, and  $\rho(t)$  is the size of the population at time  $t \geq 0$ . The equation incorporates advection with the flow  $f$  of the corresponding ODE, and selection (or growth) through the non-linear and non-local term  $(r(x) - \rho(t))n(t, x)$ . Here,  $r(x) - \rho(t)$  can be interpreted as the fitness of individuals with trait  $x$  inside the environment created by the total population, where the individuals are in a blind competition with all the other ones, regardless of their phenotype. We note that such models can rigorously be derived from stochastic individual based-models, in the limit of large populations [23, 24].

In the absence of differentiation ( $f \equiv 0$ ), the long-time behaviour of this model has been studied in detail by Benoît Perthame [131], Tommaso Lorenzi and Camille Pouchol [106], and it has been proved that, in general, solutions typically concentrate onto a single trait. This result is rather intuitive, since this model does not take mutations into account. Solutions of the advection equation alone are also known to converge to weighted Dirac masses located at the roots of  $f$  which are asymptotically stable for the ODE (4.2) [51]. On the contrary, when considering both selection and advection as in equation (4.1), the long-time behaviour is not known, to the best of our knowledge. Intuitively, two antagonistic effects will compete:

- advection will push the solution towards the asymptotically stable equilibria of ODE (4.1).
- growth will push the solution towards regions where  $r$  is maximised.

When coupling these two phenomena, our aim is to uncover whether the solution of (4.1) converges to a weighted Dirac mass, or if it converges to a smooth function. We show that both phenomena can occur, depending on the parameters  $n^0, f$  and  $r$ . Perhaps surprisingly, the model (4.1) features convergence to smooth functions even in the absence of terms modelling mutations.

Determining which parameters lead to convergence to a continuous function seems rather intricate in full generality. In particular, this problem cannot be addressed with traditional entropy methods as developed in [118], since in the absence of mutations, there is no decrease of entropy.

#### 4.1.2 Main results

In this chapter, we thus develop a different strategy allowing to reduce this problem to the study of parameter-dependent integrals, which is mainly applied to the one-dimensional case ( $x \in \mathbb{R}$ ). In this case, we elucidate the asymptotic behaviour for a large class of parameter values, and we show that there exist many different subcases depending on the number of zeros of the function  $f$ . A general statement encompassing all our results is hence rather convoluted. In order to illustrate our main results, we here focus on a few example cases which highlight the main two parameter regimes encountered for the asymptotic behaviour of (4.1).

**Proposition 14.** *Let us assume that the parameter functions  $f, n^0$  and  $r$  are smooth enough, that  $f$  has a unique root (that we denote  $x_s$ ), and that  $f'(x_s) < 0$  (which means that  $x_s$  is asymptotically stable for ODE (4.2)). Then,  $\rho$  converges to  $r(x_s)$ , and  $n$  converges to a weighted Dirac mass at  $x_s$ , when  $t$  goes to  $+\infty$ .*

Hence, in the presence of a single asymptotically stable equilibrium point for ODE (4.2), the solution of PDE (4.1) converges to a Dirac mass at this point. In other words, the

selection term is dominated by the advection term, which determines the point in which the solution concentrates. As soon as  $f$  has at least two roots, the situation is much more complex and solutions may converge to  $L^1$  functions, as illustrated in Figure 4.1 and exposed in the following proposition:

**Proposition 15.** *Let us assume that the functions  $f$ ,  $n^0$  and  $r$  are smooth enough, that  $f$  has exactly two roots (that we denote  $x_u$  and  $x_s$ , with  $x_u < x_s$ ), such that  $f'(x_u) > 0$  and  $f'(x_s) < 0$ , which means that the points  $x_u$  and  $x_s$  are respectively asymptotically unstable and asymptotically stable for the ODE (4.2). Moreover, let us assume that  $n^0$  has its support in  $[x_u, x_s]$ , and that  $n^0(x_u) > 0$ . Then, the following alternative holds:*

- *If  $r(x_s) > r(x_u) - f'(x_u)$ ,  $n$  converges to a weighted Dirac mass at  $x_s$ , and  $\rho$  converges to  $r(x_s)$ .*
- *If  $r(x_s) < r(x_u) - f'(x_u)$ ,  $n$  converges to a function in  $L^1(x_u, x_s)$ , and  $\rho$  converges to  $r(x_u) - f'(x_u)$ .*

This proposition can be interpreted as follows: since  $f$  is positive on  $(x_u, x_s)$ , the advection term drives the solution towards  $x_s$ . On the other hand, since  $x_u$  is an equilibrium, albeit unstable, it acts as a counterweight by controlling the speed of the transition towards  $x_s$  in the neighbourhood of  $x_u$ . Hence, in the case where  $r(x_u) - f'(x_u)$  is large enough ( $r(x_u) - f'(x_u) > r(x_s)$ ), the growth rate around  $x_u$  is large enough to compensate for the advection term, leading to the convergence of  $n$  to a continuous function. In the other case, the advection term is dominant, and  $n$  converges to a weighted Dirac mass at  $x_s$ . If  $n^0(x_u) = 0$ , the toggle value between the two regimes (*i.e.* the convergence to a smooth function or to a Dirac mass) changes, depending on how  $n^0$  vanishes at  $x_u$ , and other limit functions can be reached: the complete result is detailed in Proposition 22. The method of analysis proposed in this article allows in fact to solve this problem for any function  $f$  with a finite number of roots, as detailed in Proposition 23. The case where  $f$  is equal to zero on a whole interval can also be studied with our method, as highlighted by Proposition 24.

### 4.1.3 Discussion

**Open problems.** Some limit cases of the problem remain unclear: we do not deal with the case of non-hyperbolic equilibria, *i.e.*  $\bar{x} \in \mathbb{R}$  which satisfy  $f(\bar{x}) = f'(\bar{x}) = 0$ , and we are not able to determine what happens in the case where several carrying capacities, as defined in Section 4.3, converge to the same maximum limit. This last case might lead to other asymptotic behaviours, such as convergence to a sum of weighted Dirac masses, or a sum of weighted Dirac masses and  $L^1$ -functions. Lastly, we did not manage to elucidate the equality cases (of the form  $r(x_s) = r(x_u) - f'(x_u)$ ).

Furthermore, even if the framework introduced in Section 4.3 could theoretically be applied in any dimension, computing the limits of the carrying capacities seems out of reach in the multidimensional case. As shown by the semi-explicit expression introduced in Subsection 4.3.1, the behaviour of  $n$  is closely linked to that of the solutions of ODE  $\dot{x} = f(x)$ , which suggests that other asymptotic behaviours, such as convergence to a limit cycle, or chaotic behaviours (if the dimension is greater than or equal to 3) might occur.

These behaviours may be excluded by making specific assumptions regarding the function  $f$ , for example by requiring in the 2D case that ODE  $\dot{x} = f(x)$  be competitive or cooperative. Additionally if the roots of  $f$  are hyperbolic and none of them is a repeller, then  $n$  cannot converge to a  $L^1$ -function (Proposition 26). Nevertheless, the question of the asymptotic limit of  $n$  in this case remains open, and might be, in the presence of

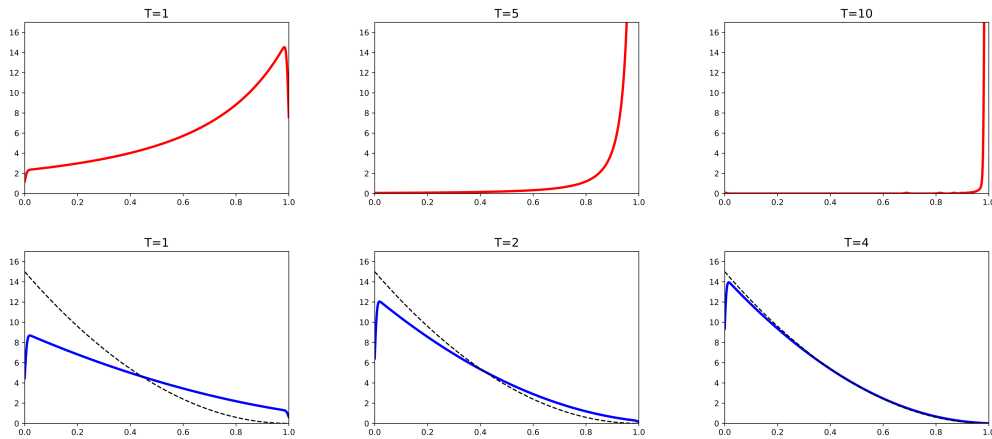


Figure 4.1: The two possible regimes of convergence stated in Proposition 15. In both cases, we have chosen  $f(x) = x(1 - x)$ ,  $n^0 \equiv 6$ , and we work on the segment  $(0, 1)$  (hence  $x_u = 0$ ,  $x_s = 1$ ). The three figures above (in red) show the time evolution of the solution in the case where  $r(x) = 6 - 0.5x$  (and thus  $5.5 = r(1) > r(0) - f'(0) = 5$ ), which implies, according to Proposition 15, that the solution converges to a weighted Dirac mass at 1. The three figures below (in blue) show the time evolution of the solution in the case where  $r(x) = 6 - 4x$ , (and thus  $2 = r(1) < r(0) - f'(0) = 5$ ), which implies that the solution converges to a continuous function in  $L^1$ . The black dashed curve represents this limit function, which can explicitly be computed (see Proposition 22).

a saddle point, a singular measure which is not a sum of weighted Dirac masses. This situation is commonplace for some applications, since toggle switches used to model cell differentiation phenomena are usually competitive or cooperative ODE models.

**Perspectives.** A natural generalisation for the model would be to model mutations, either by means of a Laplacian term or an integral term. Because of their smoothing effect, convergence to Dirac masses will typically be lost. The method developed in this chapter does not seem to handle such cases well. However, it is an interesting perspective to tackle the asymptotic behaviour with entropy methods when mutations are added [118].

From the numerical point of view, we have proved that the solution of this equation could be approximated with a particle method, with which we obtained the plots of Figure 4.1. The details of the scheme, and the proof of its convergence are detailed in Chapter 5.

**Outline of the chapter.** This chapter is organised as follows: Section 4.2 introduces the measure-theoretic framework in which convergence is considered, and includes several important reminders regarding ODE theory which will be used throughout the article. Section 4.3 details the method used to determine the asymptotic behaviour of (4.1), and Section 4.4 corresponds to a direct application of this method to several examples in the one-dimensional case. Lastly, section 4.5 presents two results in higher dimension which allow to determine, in some specific cases, if some initial solution can lead to a convergence to a smooth function or not.

## 4.2 Framework and reminders

We consider the asymptotic behaviour of the integro-differential PDE

$$\begin{cases} \partial_t n(t, x) + \nabla \cdot (f(x)n(t, x)) = (r(x) - \rho(t))n(t, x), & t \geq 0, x \in \mathbb{R}^d \\ \rho(t) = \int_{\mathbb{R}^d} n(t, x) dx, & t \geq 0 \\ n(0, x) = n^0(x), & x \in \mathbb{R}^d. \end{cases} \quad (4.1)$$

All along this chapter, we make the following regularity hypotheses

- $f$  is Lipschitz-continuous, and is in  $C^2(\mathbb{R}^d)$ .
- $r$  is positive, is in  $L^1(\mathbb{R}^d) \cap C^1(\mathbb{R}^d)$ , and goes to zero when  $\|x\|$  goes to  $+\infty$ . Let us note that these hypotheses imply that  $r$  is bounded.
- $n^0$  is in  $C_c^1(\mathbb{R}^d)$  (the space of  $C^1$  functions with a compact support), is non-negative and is not the zero function.

Whenever possible, we will indicate whether these hypotheses can be weakened for a given specific result. If not specified, it will be assumed that these three hypotheses hold.

From the modelling point of view, they can be justified as follows:  $n^0$  denoting the initial density, it is reasonable to consider that a bounded range of phenotypic traits is initially represented; the hypothesis on  $r$  at  $+\infty$  is made in order to prevent an unlikely proliferation of individuals with more and more extreme ( $\|x\| \rightarrow +\infty$ ) phenotypic traits.

Under the above hypotheses, we can prove that there exists a unique solution  $n \in C(\mathbb{R}_+, L^1(\mathbb{R}^d))$  for this Cauchy problem by coupling the well-known method of characteristics for the advection equation [51] with the method applied in [131] for the case  $f \equiv 0$ . We do not elaborate further here on the issue of existence and uniqueness, that is addressed in a more general framework in Chapter 5.

Since we are concerned with the long-time behaviour of the PDE (4.1) and we expect to obtain convergence either to Dirac masses or to regular functions, the space of Radon measures is a natural setting. We start with a few usual reminders.

### 4.2.1 The space of Radon measures

We recall that the space of **finite Radon measures** can be identified with the topological dual space of  $C_c(\mathbb{R}^d)$ , *i.e.* the space of continuous functions on  $\mathbb{R}^d$  with a compact support. Thus, we say that a sequence of finite Radon measures  $(\mu_k)_{k \in \mathbb{N}}$  **weakly converges** to a finite Radon measure  $\mu$  (denoted  $\mu_k \rightharpoonup \mu$ ) if

$$\forall \varphi \in C_c(\mathbb{R}^d), \quad \int_{\mathbb{R}^d} \varphi(x) d\mu_k(x) \xrightarrow{k \rightarrow +\infty} \int_{\mathbb{R}^d} \varphi(x) d\mu(x).$$

In this chapter, we will be confronted mainly with convergence to Dirac masses or to  $L^1$  functions. It is clear that the convergence in  $L^1$  to a certain function implies the weak convergence to this function. The following standard lemma provides a sufficient condition to prove the weak convergence to a single Dirac mass. For completeness, we provide a proof.

**Lemma 5.** *Let  $u : \mathbb{R}_+ \times \mathbb{R}^d \rightarrow \mathbb{R}$  be a non-negative mapping such that  $u(t, \cdot) \in L^1(\mathbb{R}^d)$  for all  $t \geq 0$ , and  $u(t, \cdot)$  is compactly supported, uniformly in  $t \geq 0$ . We assume that there exists  $\bar{x} \in \mathbb{R}^d$ , such that for all compact set  $K_{\bar{x}}$  which does not contain  $\bar{x}$ ,  $\int_{K_{\bar{x}}} u(t, x) dx \xrightarrow{t \rightarrow +\infty} 0$ , and that there exists  $V_{\bar{x}}$  a compact neighbourhood of  $\bar{x}$  and  $C \in \mathbb{R}$  such that  $\int_{V_{\bar{x}}} u(t, x) dx \xrightarrow{t \rightarrow +\infty} C$ . Then,  $u(t, \cdot) \xrightarrow{t \rightarrow +\infty} C \delta_{\bar{x}}$ .*

*Proof.* Let  $\varphi \in \mathcal{C}_c(\mathbb{R}^d)$ , and let  $K$  be a compact set such that, for all  $t \geq 0$ ,  $\text{supp}(u(t, \cdot)) \cup V_{\bar{x}} \subset K$ . Then,

$$\begin{aligned} \left| \int_{\mathbb{R}^d} \varphi(x)u(t, x)dx - C\varphi(\bar{x}) \right| &= \left| \int_K \varphi(x)u(t, x)dx - \int_K \varphi(\bar{x})u(t, x)dx \right. \\ &\quad \left. + \int_K \varphi(\bar{x})u(t, x)dx - C\varphi(\bar{x}) \right| \\ &\leq \int_K |\varphi(x) - \varphi(\bar{x})|u(t, x)dx + |\varphi(\bar{x})| \left| \int_K u(t, x)dx - C \right|. \end{aligned}$$

The second term tends to 0 since  $K$  contains  $V_{\bar{x}}$ . It remains to prove that  $t \mapsto \int_{\mathbb{R}^d} |\varphi(x) - \varphi(\bar{x})|u(t, x)dx$  converges to zero. Let  $\varepsilon > 0$  be given. Since  $\varphi$  is continuous, there exists  $B_{\bar{x}}$  a neighbourhood of  $\bar{x}$ , which can be chosen as a subset of  $V_{\bar{x}}$ , such that  $|\varphi(x) - \varphi(\bar{x})| \leq \varepsilon$ , for all  $x \in B_{\bar{x}}$ . Thus, for all  $t \geq 0$ ,

$$\begin{aligned} \int_K |\varphi(x) - \varphi(\bar{x})|u(t, x)dx &= \int_{K \setminus B_{\bar{x}}} |\varphi(x) - \varphi(\bar{x})|u(t, x)dx + \int_{B_{\bar{x}}} |\varphi(x) - \varphi(\bar{x})|u(t, x)dx \\ &\leq 2\|\varphi\|_{\infty} \int_{K \setminus B_{\bar{x}}} u(t, x)dx + \varepsilon \int_{B_{\bar{x}}} u(t, x)dx. \end{aligned}$$

This concludes the proof, since  $t \mapsto \int_{K \setminus V_{\bar{x}}} u(t, x)dx$  converges to zero and for any  $t$  large enough,  $\int_{B_{\bar{x}}} u(t, x)dx \leq \int_{V_{\bar{x}}} u(t, x)dx \leq C + \varepsilon$ .  $\square$

#### 4.2.2 General statement regarding the characteristics curves

We are led to consider the **characteristics curves** associated with the advection term. In this section, we introduce some notations and state some classical results from ODE theory, that will prove to be useful later on.

Since  $f$  is assumed to be Lipschitz-continuous, the global Cauchy-Lipschitz theorem ensures the global existence on  $\mathbb{R}_+$  and the uniqueness of the characteristic curves related to  $f$  defined for all  $y \in \mathbb{R}^d$  as the solution to the ODE

$$\begin{cases} \dot{X}(t, y) = f(X(t, y)) & t \geq 0 \\ X(0, y) = y \end{cases} . \quad (4.3)$$

It is well-known that for all  $t \geq 0$ ,  $y \mapsto X(t, y)$  is a  $C^1$ -diffeomorphism between  $\mathbb{R}^d$  and itself [51], and that the inverse function of  $X(t, \cdot)$ , that we denote  $x \mapsto Y(t, x)$ , is the unique solution of

$$\begin{cases} \dot{Y}(t, x) = -f(Y(t, x)) & t \geq 0 \\ Y(0, x) = x \end{cases} . \quad (4.4)$$

Moreover, **Liouville's formula** states that for all  $t \geq 0$  and  $y \in \mathbb{R}^d$ ,

$$\det(\text{Jac}_y X(t, y)) = e^{\int_0^t \nabla \cdot f(X(s, y))ds} . \quad (4.5)$$

It follows from the uniqueness of solutions to (4.3) that for all  $0 \leq s \leq t$ ,

$$X(s, Y(t, x)) = Y(t - s, x) . \quad (4.6)$$

**Specific results in  $\mathbb{R}$ .** Let us note that the behaviour of the characteristic curves is particularly simple in  $\mathbb{R}$ . Indeed, an elementary ODE analysis shows that for all  $x, y \in \mathbb{R}$ ,

$t \mapsto X(t, y)$  and  $t \mapsto Y(t, x)$  are monotonic functions. This implies that these characteristic curves either converge to a root of  $f$ , or go to  $\pm\infty$  as  $t \rightarrow +\infty$ . More precisely, if  $f$  has a finite number of roots, then for all  $y \in \mathbb{R}$  such that  $f(y) > 0$ ,  $t \mapsto X(t, y)$  converges to the closest root of  $f$  which is greater than  $y$ , or to  $+\infty$  if  $y$  is greater than the greatest root of  $f$ . Similarly, for all  $y \in \mathbb{R}$  such that  $f(y) < 0$ ,  $t \mapsto X(t, y)$  converges to the closest root of  $f$  which is lesser  $y$ , and to  $-\infty$  if  $y$  is lesser the smallest root of  $f$ .

Moreover, if each of these roots are hyperbolic equilibrium points for the ODE  $\dot{x} = f(x)$ , *i.e.* if  $f'(\bar{x}) \neq 0$  for all  $\bar{x}$  root of  $f$ , then a given root of  $f$  is either asymptotically unstable (*i.e.*  $f'(\bar{x}) > 0$ ), which implies that its basin of attraction is limited to itself, or asymptotically stable (*i.e.*  $f'(\bar{x}) < 0$ ), which implies that its basin of attraction in an open interval containing  $\bar{x}$ .

Lastly, let us recall that under these hypotheses, the convergence to an asymptotically stable point happens with an exponential speed, which means that for all  $y \in \mathbb{R}$ ,  $\bar{x}$  root of  $f$ ,

$$X(t, y) \xrightarrow[t \rightarrow +\infty]{} \bar{x} \quad \Rightarrow \quad \exists \delta_y > 0 : X(t, y) - \bar{x} = \underset{t \rightarrow +\infty}{O} (e^{-\delta_y t})$$

Since the reverse characteristic curves satisfy (4.4), the same results hold for  $Y(t, x)$ , provided that we replace  $f$  by  $-f$ . In brief, the asymptotically stable equilibria become unstable for the reverse ODE, and vice versa, and if  $t \mapsto X(t, y)$  is increasing (respectively decreasing), then  $t \mapsto Y(t, x)$  is decreasing (respectively increasing).

### 4.3 Resolution method

The method of resolution to determine the asymptotic behaviour of  $n$  that we propose here is based on the following two propositions, which are developed in the following two subsections, respectively:

1. For all  $t \geq 0$ ,  $x \in \mathbb{R}^d$ , we can express  $n(t, x)$  as a function which only depends on  $t, x$ , on the functions  $n^0$ ,  $f$  and  $r$ , on the inverse characteristic curves  $Y(t, x)$ , and on the population size  $\rho$ . Therefore, knowing the limit of  $Y(t, x)$  and  $\rho(t)$  as  $t$  goes to  $+\infty$  is enough to understand the long-time behaviour of  $n$ .
2. The population size  $\rho$  is the solution of a non-autonomous ODE, and its long-time behaviour may be inferred from the limit of some parameter-dependent integrals.

Combining these two propositions allows us to reduce the study of the asymptotic behaviour of  $n$  to that of parameter-dependent integrals.

#### 4.3.1 Semi-explicit expression of the solution

According to the definition of the characteristic curves (4.3), for all  $t \geq 0$  and all  $y \in \mathbb{R}^d$ ,

$$\frac{d}{dt} n(t, X(t, y)) = \left( r(X(t, y)) - \nabla \cdot f(X(t, y)) - \rho(t) \right) n(t, X(t, y)),$$

*i.e.*

$$n(t, X(t, y)) = e^{\int_0^t (r(X(s, y)) - \nabla \cdot f(X(s, y)) - \rho(s)) ds} n^0(y).$$

Replacing  $y$  by  $Y(t, x)$  in this last expression, we get a **semi-explicit** expression for  $n$ , which is expressed as a function of  $t, x$  and  $\rho$ :

$$\begin{aligned} n(t, x) &= n^0(Y(t, x)) e^{\int_0^t ((r - \nabla \cdot f)(X(s, Y(t, x))) - \rho(s)) ds} \\ &= n^0(Y(t, x)) e^{\int_0^t ((r - \nabla \cdot f)(Y(s, x)) - \rho(s)) ds}, \end{aligned} \tag{4.7}$$

The second equality holds according to equality (4.6) and the change of variable  $s' = t - s$ .

Beyond the non-negativity of  $n$ , this semi-explicit expression shows that determining the asymptotic behaviour of  $\rho$  and  $Y$  is enough to uncover that of  $n$ . In the following section, we show that  $\rho$  is the solution of a non-autonomous ODE, and that its asymptotic behaviour is related to that of parameter-dependent integrals.

This expression also provides exhaustive information about the support of  $n(t, \cdot)$ : indeed, it ensures that for all  $t \geq 0$ ,

$$\text{supp}(n(t, \cdot)) = \text{supp}(n^0 \circ Y(t, \cdot)) = X(t, \text{supp}(n^0)). \quad (4.8)$$

Since  $n^0$  is assumed to have a compact support, then so does  $n(t, \cdot)$  for any  $t \geq 0$ .

We recall that a set  $\mathcal{E} \subset \mathbb{R}^d$  is said to be **positively invariant** for the ODE  $\dot{x} = f(x)$  if for all  $t \geq 0$ ,  $X(t, \mathcal{E}) \subset \mathcal{E}$ .

With this definition in mind, it becomes clear, according to (4.8), that if  $\text{supp}(n^0)$  is positively invariant for the ODE  $\dot{x} = f(x)$ , then  $\text{supp}(n(t, \cdot)) \subset \text{supp}(n^0)$ , for all  $t \geq 0$ , and, more generally, that if there exists  $\mathcal{E} \subset \mathbb{R}^d$  a set which is positively invariant for this ODE such that  $\text{supp}(n^0) \subset \mathcal{E}$ , then  $\text{supp}(n(t, \cdot)) \subset \mathcal{E}$ , for all  $t \geq 0$ . Hence, even if PDE (4.1) is defined for all  $x \in \mathbb{R}^d$ , if the support of  $n^0$  is included in a compact subset of  $\mathbb{R}^d$  which is positively invariant, then everything happens as if we were working in this compact set. In particular, the functions  $f$  and  $r$  do not need to be defined outside this set.

### 4.3.2 ODE satisfied by the population size

Let us start with a basic lemma which ensures that the population size  $\rho$  does not blow up as  $t$  tends to  $+\infty$ .

**Lemma 6** Bounds on  $\rho$ . *Let  $\rho$  be defined as in (4.1). Then for all  $t \geq 0$ ,  $\rho(t) \leq \max(\|r\|_\infty, \rho(0))$ .*

*Proof.* According to (4.8), since,  $n^0$  is assumed to have a compact support,  $n(t, \cdot)$  has a compact support for all  $t \geq 0$ . Hence, when integrating the first line of (4.1), the advection term vanishes, and we get

$$\dot{\rho}(t) = \int_{\mathbb{R}^d} (r(x) - n(t, x))n(t, x)dx \leq (\|r\|_\infty - \rho(t))\rho(t).$$

In other words,  $\rho$  is a sub-solution of the logistic ODE  $\dot{u} = (\|r\|_\infty - u)u$ , which proves the result.  $\square$

In the remainder of this section, we show that  $\rho$  is in fact the solution to a non-autonomous logistic equation, which can be written in different forms. In order to lighten the future expressions, we now denote

$$\tilde{r} := r - \nabla \cdot f.$$

Let  $\mathcal{E} \subset \mathbb{R}^d$  be any measurable subset of  $\mathbb{R}^d$ , and let us denote

$$\rho_{\mathcal{E}}(t) := \int_{\mathcal{E}} n(t, x)dx,$$

which is well-defined and bounded, according to Lemma 4.3.2. By integrating the semi-explicit expression (4.7) of  $n$  over  $\mathcal{E}$ , we obtain the equality

$$\rho_{\mathcal{E}}(t) = S_{\mathcal{E}}(t)e^{-\int_0^t \rho(s)ds}, \quad (4.9)$$



where

$$S_{\mathcal{E}}(t) := \int_{\mathcal{E}} n^0(Y(t, x)) e^{\int_0^t \tilde{r}(Y(s, x)) ds} dx$$

is a function which only depends on the parameters  $f, r$  and  $n^0$ . This function is well-defined, and differentiable, thanks to our regularity assumptions, and since for all  $t \geq 0$   $n^0(Y(t, \cdot))$  has compact support. Thus, under the hypothesis that for all  $t \geq 0$ ,  $S_{\mathcal{E}}(t) > 0$ , we obtain

$$\ln(\rho_{\mathcal{E}}(t)) = \ln(S_{\mathcal{E}}(t)) - \int_0^t \rho(s) ds,$$

and finally, by differentiating and multiplying by  $\rho_{\mathcal{E}}$  on both sides,

$$\dot{\rho}_{\mathcal{E}}(t) = \left( \frac{\dot{S}_{\mathcal{E}}(t)}{S_{\mathcal{E}}(t)} - \rho(t) \right) \rho_{\mathcal{E}}(t). \quad (4.10)$$

At this stage, one might be tempted to choose  $\mathcal{E} = \mathbb{R}^d$  to obtain, denoting  $S := S_{\mathbb{R}^d}(t)$ ,

$$\dot{\rho}(t) = \left( \frac{\dot{S}(t)}{S(t)} - \rho(t) \right) \rho(t). \quad (4.11)$$

This proves that  $\rho$  is the solution to a non-autonomous logistic equation, and the study of such equations [70] proves that if the time-dependant carrying capacity  $t \mapsto \frac{\dot{S}(t)}{S(t)}$  converges, then  $\rho$  converges to the same limit. Unfortunately, computing the limit of  $t \mapsto \frac{\dot{S}(t)}{S(t)}$  is intricate (except in very specific cases). This brings us to introducing a more general framework, which involves simpler functions whose limit can be computed (at least in the case  $x \in \mathbb{R}$ ). The idea is to partition the space  $\mathbb{R}^d$  into several well-chosen subsets, and to consider the size of the population on each of these sets. As seen above, to obtain equations of the type (4.10), we must be cautious when choosing these subsets in order for the corresponding functions  $S_{\mathcal{E}}$  to be positive. All this leads us the following proposition:

**Proposition 16.** *Let  $\mathcal{U} \subset \mathbb{R}^d$  be a set such that*

$$X(\mathbb{R}_+ \times \text{supp}(n^0)) \subset \mathcal{U} \quad (4.12)$$

and let  $(\mathcal{O}_i)_{i \in \{1, \dots, N\}}$  be a finite family of open subsets of  $\mathcal{U}$  such that

$$(i) \quad \forall i \neq j, \quad \mathcal{O}_i \cap \mathcal{O}_j = \emptyset.$$

$$(ii) \quad \nu \left( \mathcal{U} \setminus \bigcup_{i=1}^N \mathcal{O}_i \right) = 0, \text{ where } \nu \text{ denotes the Lebesgue measure.}$$

$$(iii) \quad \forall i \in \{1, \dots, N\}, \quad \forall t \geq 0, \quad X(t, \text{supp}(n^0)) \cap \mathcal{O}_i \neq \emptyset.$$

Then, by denoting for all  $i \in \{1, \dots, N\}$

$$\rho_i(t) := \int_{\mathcal{O}_i} n(t, x) dx, \quad (4.13)$$

$$S_i(t) := \int_{\mathcal{O}_i} n^0(Y(t, x)) e^{\int_0^t \tilde{r}(Y(s, x)) ds} dx, \quad (4.14)$$

$$R_i(t) := \frac{\dot{S}_i(t)}{S_i(t)}, \quad (4.15)$$

the following equation holds:

$$\begin{cases} \dot{\rho}_i(t) = (R_i(t) - \rho(t)) \rho_i(t) & \forall t \geq 0, \quad \forall i \in \{1, \dots, N\} \\ \rho(t) = \sum_{i=0}^N \rho_i(t) & \forall t \geq 0 \\ \rho_i(0) > 0 & \forall i \in \{1, \dots, N\} \end{cases}. \quad (4.16)$$

**Remark.** Note that a sufficient condition for the third condition (iii) to hold is the following: for any  $i \in \{1, \dots, N\}$ , there exists  $x_i$  in the closure of  $\mathcal{O}_i$  such that  $f(x_i) = 0$  and  $n^0(x_i) > 0$ .

*Proof.* As a consequence of the discussion at the beginning of this section, it is enough to prove that

1. For all  $i \in \{1, \dots, N\}$  and all  $t \geq 0$ ,  $S_i(t) > 0$
2. For all  $t \geq 0$ ,  $\rho(t) = \sum_{i=1}^N \rho_i(t)$ .

First, notice that hypothesis (iii) is equivalent to  $\text{supp}(n^0) \cap Y(t, \mathcal{O}_i) \neq \emptyset$  for all  $i \in \{1, \dots, N\}$  and all  $t \geq 0$ . Moreover,  $\mathcal{O}_i$  is an open set, which ensures, thanks to the continuity of  $n^0$ , that  $\{x \in \mathcal{O}_i : n^0(Y(t, x)) > 0\}$  has a positive measure for all  $t \geq 0$ . This proves the first point by definition of  $S_i$ . Since  $\rho_i(0) = S_i(0)$ , we also infer  $\rho_i(0) > 0$ .

The second point is due to hypothesis (4.12): Indeed, for any  $t \geq 0$ , according to the semi-explicit expression of  $n$  provided by (4.7),  $n(t, x) = 0$  if  $Y(t, x) \notin \text{supp}(n^0)$  *i.e.* if  $x \notin X(t, \text{supp}(n^0))$ , which ensures that

$$\rho(t) = \int_{\mathbb{R}^d} n(t, x) dx = \int_{\mathcal{U}} n(t, x) dx.$$

The first two hypotheses satisfied by the sets  $\mathcal{O}_i$  ensure that  $\rho(t) = \sum_{i=1}^N \rho_i(t)$ .

**Proof of the remark:** Let  $x_i$  be a root of  $f$ . A classical ODE result ensures that for all  $t \geq 0$ ,  $x \in \mathbb{R}^d$ ,  $\|Y(t, x) - x_i\| \leq e^{Lt} \|x - x_i\|$ , with  $L > 0$  the Lipschitz constant of  $f$ . Since  $n^0(x_i) > 0$  and  $n^0$  is continuous, there exists  $\varepsilon > 0$  such that  $B(x_i, \varepsilon) \subset \text{supp}(n^0)$ . Let  $t \geq 0$ ,  $x \in \mathcal{O}_i \cap B(x_i, \varepsilon e^{-Lt}/2)$  (such a point does exist, by definition of the closure). Then,  $Y(t, x) \in B(x_i, \varepsilon) \subset \text{supp}(n^0)$ , which ensures that  $x \in X(t, \text{supp}(n^0))$ , and thus concludes the proof.  $\square$

In the one-dimensional case, assuming that  $f$  has a finite number of roots, an efficient choice for the sets  $\mathcal{O}_i$  is to take the segments between the roots of  $f$  which intersect the support of  $n^0$ , as the following result shows.

**Lemma 7.** *Let  $x \in \mathbb{R}$  and assume that  $f : \mathbb{R} \rightarrow \mathbb{R}$  has a finite number of roots, that we denote  $x_1 < x_2 < \dots < x_N$ . Let us denote*

$$\mathcal{O}_0 := (-\infty, x_1), \quad \mathcal{O}_i := (x_i, x_{i+1}), \quad i \in \{1, \dots, N-1\}, \quad \mathcal{O}_N := (x_N, +\infty),$$

and, among these segments, let us consider  $\mathcal{O}_{i_1}, \dots, \mathcal{O}_{i_M}$  those which have a non-empty intersection with  $\text{supp}(n^0)$ . Then, the set  $\mathcal{U} := \bigcup_{1 \leq j \leq M} \mathcal{O}_{i_j}$  and the family of sets  $(\mathcal{O}_{i_j})_{1 \leq j \leq M}$  satisfy the hypotheses of Proposition 16.

*Proof.* By applying the results stated at the end of Section 4.2.2, we note that for all  $i \in \{1, \dots, N\}$ ,  $\mathcal{O}_i$  is positively invariant for the ODE  $\dot{x} = f(x)$ . Thus, for all  $y \in \text{supp}(n^0) \subset \mathcal{U}$ ,  $t \geq 0$ ,  $X(t, y) \in \mathcal{U}$ , which ensures that  $X(\mathbb{R}_+ \times \text{supp}(n^0)) \subset \mathcal{U}$ . Moreover, the same results show that for all  $j \in \{1, \dots, M\}$ ,  $X(t, \text{supp}(n^0) \cap \mathcal{O}_{i_j}) \subset \mathcal{O}_{i_j}$ , and thus that  $X(t, \text{supp}(n^0)) \cap \mathcal{O}_{i_j} \neq \emptyset$ . The other two points are automatically satisfied, thanks to the definition of  $\mathcal{U}$  and the sets  $\mathcal{O}_i$ .  $\square$

Proposition 16 shows us that  $\rho$  satisfies ODE (4.16). Our next result shows that the long-time behaviour of this ODE depends on the long-time behaviour of the functions  $R_i$ . In particular, it states that if all the functions  $R_i$  converge, then  $\rho$  converges to the maximum of their limit. Before stating the result, we introduce some notations.

**Notation.** For any function  $g : \mathbb{R}_+ \rightarrow \mathbb{R}$ , we denote:

$$\underline{g} := \liminf_{t \rightarrow +\infty} g(t) \quad \text{and} \quad \bar{g} := \limsup_{t \rightarrow +\infty} g(t),$$

and we say that  $g$  converges to  $l \in \mathbb{R}$  **with an exponential speed** if there exist  $\delta > 0$  such that

$$g(t) - l = O_{t \rightarrow +\infty} \left( e^{-\delta t} \right).$$

**Proposition 17.** The coupled system of ODEs (4.16) has the following properties:

- (i) For all  $i \in \{1, \dots, N\}$  and all  $t \geq 0$ ,  $\rho_i(t) > 0$ .
- (ii)  $\underline{\rho} \geq \min_{1 \leq i \leq N} (\underline{R}_i)$  and  $\bar{\rho} \leq \max_{1 \leq i \leq N} (\bar{R}_i)$ .
- (iii) Let  $j \in \{1, \dots, N\}$ . If there exists  $i \in \{1, \dots, N\}$  such that  $\bar{R}_j < \underline{R}_i$ , then  $\rho_j(t) \xrightarrow[t \rightarrow +\infty]{} 0$ .
- (iv) Let us assume that there exists  $l \in \mathbb{R}_+ \cup \{+\infty\}$ , and a non empty set  $I \subset \{1, \dots, N\}$  (where potentially  $I = \{1, \dots, N\}$ ) such that for all  $i \in I$ ,  $R_i(t) \xrightarrow[t \rightarrow +\infty]{} l$ , and  $\bar{R}_j < l$  for all  $j \notin I$ . Then,  $\rho(t) \xrightarrow[t \rightarrow +\infty]{} l$ .
- (v) Under the hypotheses of (iv), if moreover  $0 < l < +\infty$  and for all  $i \in I$  the function  $R_i$  converges to  $l$  with an exponential speed, then  $\rho$  converges to  $l$  with an exponential speed.

*Proof.* (i) According to the first line of ODE (4.16),  $\rho_i(t) = e^{\int_0^t R_i(s) - \rho(s) ds} \rho_i(0)$ , which is positive according to the third line.

- (ii) If  $\min_{1 \leq i \leq N} (\underline{R}_i) = 0$ , there is nothing to prove: we assume  $\min_{1 \leq i \leq N} (\underline{R}_i) > 0$  and let  $m < \min_{1 \leq i \leq N} (\underline{R}_i)$ . There exists  $T_m \geq 0$  such that for all  $t \geq T_m$ , and all  $i \in \{1, \dots, N\}$ ,  $R_i(t) \geq m$ . Thus

$$\dot{\rho}(t) = \sum_{i=1}^N \dot{\rho}_i(t) = \sum_{i=1}^N (R_i(t) - \rho(t)) \rho_i(t) \geq (m - \rho(t)) \rho(t),$$

which means that  $\rho$  is a super-solution of a logistic equation which converges to  $m$ , and thus that  $\underline{\rho} \geq m$ . Since this inequality holds for any  $m < \min_{1 \leq i \leq N} (\underline{R}_i)$  it proves that  $\underline{\rho} \geq \min_{1 \leq i \leq N} (\underline{R}_i)$ . By proceeding in the same way with the limit superior, we get the second inequality.

- (iii) Let  $i, j \in \{1, \dots, N\}$  such that  $\bar{R}_j < \bar{R}_i$ . The latter inequality is written with the convention that if  $\bar{R}_i = +\infty$ , then  $\bar{R}_j \in \mathbb{R}$ . Using the first point,  $\rho_j, \rho_i > 0$  on  $\mathbb{R}_+$ . We can compute

$$\frac{d}{dt} \ln \left( \frac{\rho_i(t)}{\rho_j(t)} \right) = R_i(t) - R_j(t) > \varepsilon,$$

for a certain  $\varepsilon > 0$  and  $t$  large enough. Thus,  $\rho(t) \geq \rho_i(t) \geq C e^{\varepsilon t} \rho_j(t)$ , for a certain constant  $C > 0$ , which yields

$$\dot{\rho}_j(t) \leq \left( \sup_{t>0} R_j(t) - C e^{\varepsilon t} \rho_j(t) \right) \rho_j(t),$$

with  $\sup_{t>0} R_j(t) < +\infty$  by hypothesis, and thus  $\rho_j$  goes to zero as  $t$  goes to  $+\infty$ .

- (iv) Let us denote  $\rho_J := \sum_{j \notin I} \rho_j$ . (This first step is not necessary in the case  $I = \{1, \dots, N\}$ ).

According to the previous property,  $\rho_J$  converges to zero. By denoting  $\tilde{R}_i := R_i - \rho_J$ , we can thus rewrite system (4.16) as:

$$\begin{cases} \dot{\rho}_i(t) = \left( \tilde{R}_i(t) - \rho_I(t) \right) \rho_i(t) & \forall t \geq 0, \quad \forall i \in I \\ \rho_I(t) = \sum_{i \in I} \rho_i(t) & \forall t \geq 0 \\ \rho_i(0) > 0 & \forall i \in \{1, \dots, N\} \end{cases}.$$

Applying Property (ii) to this new system proves the desired result, since

$$\min_{i \in I} \left( \tilde{R}_i \right) = \max_{i \in I} \left( \tilde{R}_i \right) = l.$$

- (v) Let  $l \in (0, +\infty)$ . According to the previous point,  $\rho$  is bounded by two positive constants (and so is  $\rho_I$ ), that we denote  $\rho^m < \rho^M$ . Using the same argument as in the proof of the third point, one proves that for all  $j \notin I$ , there exists  $\varepsilon > 0$  such that  $\rho_j(t) \leq C e^{-\varepsilon t} \rho^M$ , and thus that  $\rho_J$  converges to 0 with an exponential speed. Thus, it remains to prove that the convergence of  $\rho_I$  to  $l$  also occurs with an exponential speed. By hypothesis, there exists  $C, \delta > 0$  such that for all  $t \geq 0$ ,  $\sum_{i \in I} |\tilde{R}_i(t) - l| \leq C e^{-\delta t}$ . Thus, by denoting  $C' := C \|\rho_I(\cdot) - l\|_\infty \rho^M$ , we find

$$\begin{aligned} \frac{d}{dt} \frac{1}{2} (\rho_I(t) - l)^2 &= (\rho_I(t) - l) \sum_{i \in I} ((\tilde{R}_i(t) - l) - (\rho_I(t) - l)) \rho_i(t) \\ &\leq C' e^{-\delta t} - \rho^m (\rho_I(t) - l)^2, \end{aligned}$$

which concludes the proof, according to Grönwall's lemma.  $\square$

## 4.4 Results in the one-dimensional case

### 4.4.1 Asymptotic behaviour of the carrying capacities

As evidenced by the previous section and in particular by Proposition 17, the long-time behaviour of  $\rho$  is completely determined by that of the functions  $R_i$ , which we call **carrying capacities** by analogy with the logistic equation. As their definition suggests, computing the limit of these functions is a delicate issue: this section is dedicated to these

computations. The multidimensional case seems out of reach with this method, because, as we shall see, we use a change of variable that requires to be working in 1D.

In order to simplify the notations, we will now denote  $R$  instead of  $R_{\mathcal{E}}$  or  $R_i$ , when there is no ambiguity as to which sets we are working with. We are thus interested in the asymptotic behaviour of the function

$$R(t) = \frac{\dot{S}(t)}{S(t)}, \quad \text{with} \quad S(t) = \int_{\mathcal{E}} n^0(Y(t, x)) e^{\int_0^t \tilde{r}(Y(s, x)) ds} dx, \quad (4.17)$$

where  $\mathcal{E} \subset \mathbb{R}^d$  is an open set which satisfies  $\text{supp}(n^0) \cap Y(t, \mathcal{E}) \neq \emptyset$  for all  $t \geq 0$ .

First, let us note that for all  $l \in \mathbb{R}$ ,

$$R(t) - l = \frac{\frac{d}{dt}(S(t)e^{-lt})}{S(t)e^{-lt}}. \quad (4.18)$$

Thus, in order to prove that  $R$  converges to  $l \in \mathbb{R}$  with an exponential speed, it is enough to prove that:

(a)  $\liminf_{t \rightarrow +\infty} S(t)e^{-lt} > 0$ .

(b)  $t \mapsto e^{\delta t} \frac{d}{dt}(S(t)e^{-lt})$  is bounded for a certain  $\delta > 0$ .

Indeed, we immediately deduce from (4.18), and the fact that  $S$  is positive, according to its definition (4.14), that these two hypotheses imply that for any  $\delta' \in (0, \delta)$ ,

$$R(t) - l = O_{t \rightarrow +\infty} \left( e^{-\delta' t} \right).$$

### Integral formulae for the carrying capacities

This section aims at listing several alternative formulae of  $S$ . In the following section, we will use one or the other, depending on the studied case.

We recall that  $S$  is defined as

$$S(t) = \int_{\mathcal{E}} n^0(Y(t, x)) e^{\int_0^t \tilde{r}(Y(s, x)) ds} dx, \quad (4.19)$$

with  $\tilde{r} := r - \nabla \cdot f$ .

As seen in the first section, for any  $t \geq 0$ ,  $x \mapsto Y(t, x)$  is a  $C^1$ -diffeomorphism from  $\mathcal{E}$  to  $Y(t, \mathcal{E})$ . Thus, the change of variable  $y = Y(t, x)$ , and Liouville's formula which ensures that  $|\det(\text{Jac}(Y(t, x)))| = e^{\int_0^t -\nabla \cdot f(Y(s, x)) ds}$  provide a second expression for  $S$ , namely

$$S(t) = \int_{Y(t, \mathcal{E})} n^0(y) e^{\int_0^t r(X(s, y)) ds} dy. \quad (4.20)$$

Moreover, **in the one-dimensional case**  $x \in \mathbb{R}$ , if  $\mathcal{E}$  is an interval on which  $f \neq 0$ , then for all  $y \in \mathcal{E}$ ,  $t \mapsto X(t, y)$  is also a  $C^1$ -diffeomorphism from  $(0, t)$  to  $(y, X(t, y))$  or  $(X(t, y), y)$ . This allows us to make the change of variable  $s' = Y(s, x)$  and  $s' = X(s, y)$  in the two expressions for  $S$ , thereby obtaining two new formulations

$$S(t) = \int_{\mathcal{E}} n^0(Y(t, x)) e^{\int_{Y(t, x)}^x \frac{\tilde{r}(s)}{f(s)} ds} dx = \int_{Y(t, \mathcal{E})} n^0(y) e^{\int_y^{X(t, y)} \frac{r(s)}{f(s)} ds} dy, \quad (4.21)$$

and, in the same way, for all  $l \in \mathbb{R}$ ,

$$S(t)e^{-lt} = \int_{\mathcal{E}} n^0(Y(t, x)) e^{\int_{Y(t, x)}^x \frac{\tilde{r}(s)-l}{f(s)} ds} dx = \int_{Y(t, \mathcal{E})} n^0(y) e^{\int_y^{X(t, y)} \frac{r(s)-l}{f(s)} ds} dy. \quad (4.22)$$

Likewise, by differentiating expressions (4.19) and (4.20), we are led to several formulae for  $\frac{d}{dt}(S(t)e^{-lt})$ , namely

$$\frac{d}{dt}(S(t)e^{-lt}) = \int_{\mathcal{E}} m(Y(t,x)) e^{\int_0^t \tilde{r}(Y(s,x)) - l ds} dx = \int_{\mathcal{E}} m(y) e^{\int_0^t r(X(s,y)) - l ds} dy, \quad (4.23)$$

with

$$m(y) := n^0(y)(\tilde{r}(y) - l) - f(y)n^{0'}(y). \quad (4.24)$$

**In the one-dimensional case**  $x \in \mathbb{R}$ , assuming that  $\mathcal{E}$  is an interval in which  $f \neq 0$ , we get the additional expressions

$$\frac{d}{dt}(S(t)e^{-lt}) = \int_{\mathcal{E}} m(Y(t,x)) e^{\int_{Y(t,x)}^x \frac{\tilde{r}(s)-l}{f(s)} ds} dx = \int_{\mathcal{E}} m(y) e^{\int_y^{X(t,y)} \frac{r(s)-l}{f(s)} ds} dy. \quad (4.25)$$

Lastly, in the particular one-dimensional case where  $\mathcal{E}$  is an interval such that  $Y(t, \mathcal{E}) = \mathcal{E}$  for all  $t \geq 0$ , and  $f \neq 0$  on  $\mathcal{E}$ , (which is the case if  $\mathcal{E}$  is an interval delimited by two consecutive roots of  $f$ ) one can differentiate (4.20) to get

$$\frac{d}{dt}(S(t)e^{-lt}) = \int_{\mathcal{E}} n^0(y)(r(X(t,y)) - l) e^{\int_0^t r(X(s,y)) - l ds} dy \quad (4.26)$$

and the second expression of (4.22) to get

$$\frac{d}{dt}(S(t)e^{-lt}) = \int_{\mathcal{E}} n^0(y)(r(X(t,y)) - l) e^{\int_y^{X(t,y)} \frac{r(s)-l}{f(s)} ds} dy \quad (4.27)$$

$$= \int_{\mathcal{E}} n^0(Y(t,x))(r(x) - l) e^{\int_{Y(t,x)}^x \frac{\tilde{r}(s)-l}{f(s)} ds} dx. \quad (4.28)$$

### An important estimate

The lemma stated in this section will be crucial in computing limits of the relevant parameter-dependent integrals in the next section.

**Notation.** Let  $x_0 \in \mathbb{R} \cup \{\pm\infty\}$ , and  $h$  and  $g$  be two functions defined in the neighbourhood of  $x_0$ . If there exist  $C_1, C_2 > 0$  such that

$$C_1|g(x)| \leq |h(x)| \leq C_2|g(x)|$$

for any  $x$  close enough to  $x_0$ , we write

$$h(x) = \Theta_{x \rightarrow x_0}(g(x)).$$

**Remark.** According to the definition of  $\Theta$ , it is clear that for any  $x_0 \in \mathbb{R} \cup \{\pm\infty\}$ ,  $g, h$  defined in the neighbourhood of  $x_0$ ,  $f$  such that  $h(x) = \Theta_{x \rightarrow x_0}(g(x))$ ,  $h$  is integrable near  $x_0$  if and only if  $g$  is integrable near  $x_0$ .

**Lemma 8.** Let  $x_0, y \in \mathbb{R}$ , with  $x_0 < y$ , and let  $\beta \in C^2([x_0, y])$  such that  $\beta(y) = 0$ ,  $\beta'(y) \neq 0$  and  $\beta \neq 0$  on  $[x_0, y)$ , and  $\alpha \in C^1([x_0, y])$ . Then,

$$e^{\int_{x_0}^x \frac{\alpha(s)}{\beta(s)} ds} = \Theta_{x \rightarrow y^-} \left( |y - x|^{\frac{\alpha(y)}{\beta'(y)}} \right).$$

Similarly, if  $x_0 > y$ ,  $\beta \in C^2([y, x_0])$  is such that  $\beta(y) = 0$ ,  $\beta'(y) \neq 0$  and  $\beta \neq 0$  on  $[y, x_0)$ , and  $\alpha \in C^1([y, x_0])$ , then

$$e^{\int_{x_0}^x \frac{\alpha(s)}{\beta(s)} ds} = \Theta_{x \rightarrow y^+} \left( |y - x|^{\frac{\alpha(y)}{\beta'(y)}} \right).$$

*Proof.* we show only the first equivalence: the proof of the second is analogous. According to the regularity of  $\alpha$  and  $\beta$ , the integral  $\int_{x_0}^x \frac{\alpha(s)}{\beta(s)}$  is well-defined for any  $x \in (x_0, y)$ , and for all  $s \in (x_0, y)$ ,

$$\alpha(s) = \alpha(y) + O(s - y), \quad \text{and} \quad \beta(s) = (s - y)\beta'(y) + O((s - y)^2).$$

Thus,

$$\frac{\alpha(s)}{\beta(s)} - \frac{\alpha(y)}{(s - y)\beta'(y)} = \frac{\alpha(s)(s - y)\beta'(y) - \alpha(y)\beta(s)}{\beta(s)(s - y)\beta'(y)} = \frac{O(s - y)^2}{\beta'(y)^2(s - y)^2 + O((s - y)^3)} = O(1).$$

Hence,

$$e^{\int_{x_0}^x \frac{\alpha(s)}{\beta(s)} ds} = e^{\int_{x_0}^x \frac{\alpha(y)}{\beta'(y)} \frac{1}{s - y} + O(1) ds} = e^{O(1)} |y - x|^{\frac{\alpha(y)}{\beta'(y)}},$$

which proves the result of this lemma.  $\square$

### Asymptotic behaviour of the carrying capacity in one dimension

We here focus on the one-dimensional case. We recall that we assume that  $n^0 \in \mathcal{C}_c(\mathbb{R})$ ,  $f \in C^2(\mathbb{R}) \cap \text{Lip}(\mathbb{R})$ ,  $r \in C^1(\mathbb{R}) \cap L^1(\mathbb{R})$ , and that  $r(x)$  goes to 0 as  $x$  goes to  $\pm\infty$ . In this section, we further assume that  $f \in \text{BV}(\mathbb{R})$ , *i.e.*  $f' \in L^1(\mathbb{R})$ , and that  $f$  converges to a non-zero limit at  $\pm\infty$ .

In order to apply Proposition 17 (as explained in Lemma 7), the most insightful division is to consider each segment between the roots of  $f$ . Hence, we must first compute the limit of the function  $R$  when the chosen set  $\mathcal{E}$  is such a segment.

To be more precise, we must therefore distinguish between several cases, depending on whether the considered interval is bounded (delimited by two consecutive roots of  $f$ ) or not (delimited by the smallest or the greatest root of  $f$ ), and the sign of the derivative at these boundary roots.

In fact, when  $n^0$  vanishes at a given root  $a$ , the limit may depend on *how fast*  $n^0$  vanishes, *i.e.* on the value  $\alpha > 0$  such that  $n^0(y)$  vanishes like  $(y - a)^\alpha$ . For our method of proof to accommodate this case, we will need to make a slightly stronger assumption involving the derivative of  $n^0$ .

We will see in the next section that a slight change in the limit of  $R$  may have a drastic impact on the long-time behaviour of  $n$ . We also deal with cases where  $f$  does not have any root (which ensures, as one might expect, that  $R$  converges to 0), and the case where  $f$  is zero on a whole interval. Hence, this result can be seen as a generalisation of the one stated in [106].

**Proposition 18.** *In each case, we assume that  $\mathcal{E} \cap \text{supp}(n^0) \neq \emptyset$ .*

- (i) *If  $\mathcal{E} = (a, +\infty)$ ,  $f < 0$  on  $\mathcal{E}$ ,  $f(a) = 0$  and  $f'(a) < 0$ , then  $R$  converges to  $r(a)$ .*
- (ii) *If  $\mathcal{E} = (-\infty, a)$ ,  $f > 0$  on  $\mathcal{E}$ ,  $f(a) = 0$  and  $f'(a) < 0$ , then  $R$  converges to  $r(a)$ .*
- (iii) *If  $\mathcal{E} = (a, +\infty)$ ,  $f > 0$  on  $\mathcal{E}$ ,  $f(a) = 0$ ,  $f'(a) > 0$ , then*
  - *If  $n^0(a) > 0$ , then*
    - *If  $r(a) - f'(a) > 0$ , then  $R$  converges to  $r(a) - f'(a)$ .*
    - *If  $r(a) - f'(a) < 0$ , then  $R$  converges to 0.*
  - *If  $n^0(a) = 0$ , and if there exist  $C, \alpha > 0$  such that  $n^0(y) = C\alpha(y - a)^{\alpha-1} + O_{y \rightarrow a^+}((y - a)^\alpha)$ , then*
    - *If  $r(a) - (1 + \alpha)f'(a) > 0$ , then  $R$  converges to  $r(a) - (1 + \alpha)f'(a)$ .*

- If  $r(a) - (1 + \alpha)f'(a) < 0$ , then  $R$  converges to 0.
  - If  $n^0(a) = 0$ , and if there exists  $\varepsilon > 0$  such that  $n^0(\cdot) = 0$  on  $[a, a + \varepsilon]$ , then  $R$  converges to 0.
- (iv) If  $\mathcal{E} = (-\infty, a)$ ,  $f < 0$  on  $\mathcal{E}$ ,  $f(a) = 0$ ,  $f'(a) > 0$ , then
- If  $n^0(a) > 0$ , then
    - If  $r(a) - f'(a) > 0$ , then  $R$  converges to  $r(a) - f'(a)$ .
    - If  $r(a) - f'(a) < 0$ , then  $R$  converges to 0.
  - If  $n^0(a) = 0$ , and if there exist  $C, \alpha > 0$  such that  $n^0(y) = -C\alpha(a - y)^{\alpha-1} + O((a - y)^\alpha)$ ,  $y \rightarrow a^-$ , then
    - If  $r(a) - (1 + \alpha)f'(a) > 0$ , then  $R$  converges to  $r(a) - (1 + \alpha)f'(a)$ .
    - If  $r(a) - (1 + \alpha)f'(a) < 0$ , then  $R$  converges to 0.
  - If  $n^0(a) = 0$ , and if there exists  $\varepsilon > 0$  such that  $n^0(\cdot) = 0$  on  $[a - \varepsilon, a]$ , then  $R$  converges to 0.
- (v) If  $\mathcal{E} = (a, b)$ ,  $f > 0$  on  $(a, b)$ ,  $f(a) = f(b) = 0$ ,  $f'(a) > 0$ ,  $f'(b) < 0$ , then
- If  $n^0(a) > 0$ , then
    - If  $r(b) > r(a) - f'(a)$ , then  $R$  converges to  $r(b)$ .
    - If  $r(b) < r(a) - f'(a)$ , then  $R$  converges to  $r(a) - f'(a)$ .
  - If  $n^0(a) = 0$ , and if there exist  $C, \alpha > 0$  such that  $n^0(y) = C\alpha(y - a)^{\alpha-1} + O((y - a)^\alpha)$ ,  $y \rightarrow a^+$ , then
    - If  $r(b) > r(a) - (\alpha + 1)f'(a)$ , then  $R$  converges to  $r(b)$ .
    - If  $r(b) < r(a) - (\alpha + 1)f'(a)$ , then  $R$  converges to  $r(a) - (\alpha + 1)f'(a)$ .
  - If  $n^0(a) = 0$ , and if there exists  $\varepsilon > 0$  such that  $n^0(\cdot) = 0$  on  $[a, a + \varepsilon]$ , then  $R$  converges to  $r(b)$ .
- (vi) If  $\mathcal{E} = (a, b)$ ,  $f < 0$  on  $(a, b)$ ,  $f(a) = f(b) = 0$ ,  $f'(a) < 0$ ,  $f'(b) > 0$ , then
- If  $n^0(b) > 0$ , then
    - If  $r(a) > r(b) - f'(b)$ , then  $R$  converges to  $r(a)$ .
    - If  $r(a) < r(b) - f'(b)$ , then  $R$  converges to  $r(b) - f'(b)$ .
  - If  $n^0(b) = 0$ , and if there exist  $C, \alpha > 0$  such that  $n^0(y) = -C\alpha(b - y)^{\alpha-1} + O((b - y)^\alpha)$ ,  $y \rightarrow b^-$ , then
    - If  $r(a) > r(b) - (\alpha + 1)f'(b)$ , then  $R$  converges to  $r(a)$ .
    - If  $r(a) < r(b) - (\alpha + 1)f'(b)$ , then  $R$  converges to  $r(b) - (\alpha + 1)f'(b)$ .
  - If  $n^0(b) = 0$ , and if there exists  $\varepsilon > 0$  such that  $n^0(\cdot) = 0$  on  $[b - \varepsilon, b]$ , then  $R$  converges to  $r(a)$ .
- (vii) If  $\mathcal{E} = \mathbb{R}$ , and  $f > 0$  on  $\mathbb{R}$ , then  $R$  converges to 0.
- (viii) If  $\mathcal{E} = \mathbb{R}$ , and  $f < 0$  on  $\mathbb{R}$ , then  $R$  converges to 0.
- (ix) If  $\mathcal{E}$  is a interval in which  $f \equiv 0$ , and  $n^0 > 0$ , and  $\arg \max_{\mathcal{E}} r = \{x_1, \dots, x_p\} \subset \mathcal{E}$ , with

$$r'(x_i) = 0, \quad r''(x_i) < 0 \quad \text{for all } i \in \{1, \dots, p\},$$

then,  $R$  converges to  $\bar{r} := \max_{x \in \mathcal{E}} r(x)$ .



Moreover, except in this last case,  $R$  converges with an exponential speed whenever it does not converge to 0.

*Proof.* As explained at the beginning of this section, whenever we show that  $R$  converges with an exponential speed, we must prove successively that

$$(a) \liminf_{t \rightarrow +\infty} S(t)e^{-lt} > 0$$

$$(b) t \mapsto e^{\delta t} \frac{d}{dt} (S(t)e^{-lt}) \text{ is bounded for a certain } \delta > 0,$$

where  $l$  is the expected limit. By Fatou's lemma, the point (a) can be proven by showing that the integrand involved in the expression of  $S$  (which depends on the chosen formula) converges pointwise to a non-negative function which is positive on a set of positive Lebesgue measure. Depending on the case, we will use different expressions for  $S$  and  $S'$  among those determined in Section 4.4.1. In order to lighten the proof, we assume without loss of generality that  $a = 0$  and  $b = 1$ , and we denote

$$\tilde{r} = r - f' \quad \text{and} \quad \tilde{r}_\alpha := \tilde{r} - \alpha f' = r - (\alpha + 1)f' \quad \text{for } \alpha \in \mathbb{R}.$$

Moreover since the cases (ii), (iv), (vi) and (viii) are symmetric to the cases (i), (iii), (v) and (vii) respectively, we omit their proof.

(i) Note that, according to the hypotheses satisfied by  $f$ , for all  $y \in (0, +\infty)$ ,  $t \mapsto X(t, y)$  converges to 0.

(a) According to (4.20),

$$S(t)e^{-r(0)t} = \int_0^{+\infty} n^0(y) e^{\int_0^t r(X(s,y)) - r(0) ds} dy = \int_0^M n^0(y) e^{\int_0^t r(X(s,y)) - r(0) ds} dy,$$

for a certain  $M > 0$ , since  $n^0$  has a compact support. Since  $f'(0) < 0$ , there exist  $C, \delta > 0$  such that  $X(t, y) \leq Ce^{-\delta t}$  for all  $y \in [0, M]$ ,  $t \geq 0$ . This proves that for all  $y \in [0, M]$ ,  $s \mapsto r(X(s, y)) - r(0)$  is integrable on  $(0, +\infty)$ , and thus that  $y \mapsto n^0(y) e^{\int_0^{+\infty} r(X(s,y)) - r(0) ds}$  is well-defined on  $[0, M]$ . Since this function is positive on a sub-interval of  $[0, M]$ , its integral on this segment is positive. Moreover,  $t \mapsto n^0(y) e^{\int_0^t r(X(s,y)) - r(0) ds}$  converges pointwise to this function.

(b) As seen in the first point, there exist  $C, \delta > 0$  such that for all  $y \in [0, M]$  and all  $t \geq 0$ ,  $0 \leq X(t, y) \leq Ce^{-\delta t}$ . Thus, using expression (4.26), and the mean value theorem,

$$\begin{aligned} \left| e^{\delta t} \frac{d}{dt} (S(t)e^{-r(0)t}) \right| &= e^{\delta t} \left| \int_0^M n^0(y) (r(X(t, y)) - r(0)) e^{\int_0^t r(X(s,y)) - r(0) ds} dy \right| \\ &\leq 2 \|n^0\|_\infty \|r\|_{L^\infty(0, M)} C \int_0^M e^{\int_0^t |r(X(s,y)) - r(0)| ds} dy \\ &\leq \|n^0\|_\infty \|r'\|_{L^\infty(0, M)} C M e^{\int_0^t C \|r'\|_{L^\infty(0, M)} e^{-\delta s} ds} \end{aligned}$$

which is bounded.

(iii) Note that, according to the hypothesis on  $f$ , for all  $x, y \in (0, +\infty)$ ,  $t \mapsto X(t, y)$  is increasing and goes to  $+\infty$ , and  $t \mapsto Y(t, x)$  is decreasing and converges to 0.

• Let us assume that  $n^0(0) > 0$ . We distinguish two cases:

– **Case**  $r(0) - f'(0) > 0$ :

(a) According to (4.22),

$$S(t)e^{-\tilde{r}(0)t} = \int_{\mathcal{E}} n^0(Y(t, x)) e^{\int_{Y(t, x)}^x \frac{\tilde{r}(s) - \tilde{r}(0)}{f(s)} ds} dx.$$

For all  $x \in (0, +\infty)$ ,  $n^0(Y(t, x)) e^{\int_{Y(t, x)}^x \frac{\tilde{r}(s) - \tilde{r}(0)}{f(s)} ds} \xrightarrow{t \rightarrow +\infty} n^0(0) e^{\int_0^x \frac{\tilde{r}(s) - \tilde{r}(0)}{f(s)} ds}$ ,

which is well defined since  $s \mapsto \frac{\tilde{r}(s) - \tilde{r}(0)}{f(s)}$  is continuous on  $[0, x)$ , thanks to the regularity of  $r$  and  $f$ , and positive, since  $n^0(0) > 0$  by hypothesis.

- (b) Let  $\delta \in (0, \min(\tilde{r}(0), f'(0)))$ . Since  $\delta - \tilde{r}(0) < 0$ ,  $r$  goes to 0 at  $+\infty$  and  $f$  is positive, we can find  $M \geq 0$  such that  $\frac{r(s) - \tilde{r}(0) + \delta}{f(s)} \leq 0$  for all  $s \in [M, +\infty)$ , and  $\text{supp}(n^0) \cap \mathcal{E} \subset [0, M]$ . Thus, for all  $t \geq 0$ , and all  $y \in (0, M)$ ,

$$\int_y^{X(t,y)} \frac{r(s) - \tilde{r}(0) + \delta}{f(s)} ds \leq \int_y^M \left| \frac{r(s) - \tilde{r}(0) + \delta}{f(s)} \right| ds.$$

According to (4.25),

$$e^{\delta t} \frac{d}{dt} \left( S(t) e^{-\tilde{r}(0)t} \right) = \int_0^{+\infty} m(y) e^{\int_y^{X(t,y)} \frac{r(s) - \tilde{r}(0) + \delta}{f(s)} ds} dy.$$

Thus, since  $\text{supp}(m) \cap \mathcal{E} = \text{supp}(n^0) \cap \mathcal{E} \subset [0, M]$ , and by the previous inequality,

$$\left| e^{\delta t} \frac{d}{dt} \left( S(t) e^{-\tilde{r}(0)t} \right) \right| \leq \int_0^M |m(y)| e^{\int_y^M \frac{|r(s) - \tilde{r}(0) + \delta|}{f(s)} ds} dy.$$

Since  $m(y) = n^0(y)(\tilde{r}(y) - \tilde{r}(0)) - f(y)n^{0'}(y)$ ,  $|m(y)| = O_{y \rightarrow 0}(y)$ . Moreover, since  $|r(0) - \tilde{r}(0) + \delta| = f'(0) + \delta$ , Lemme 8 yields  $e^{\int_y^M \frac{|r(s) - \tilde{r}(0) + \delta|}{f(s)} ds} = O_{y \rightarrow 0}(y^{-1 - \delta/f'(0)})$ .

Therefore,

$$|m(y)| e^{\int_y^M \frac{|r(s) - \tilde{r}(0) + \delta|}{f(s)} ds} = O_{y \rightarrow 0}(y^{-\delta/f'(0)}),$$

and is thus integrable since  $\delta < f'(0)$ .

- **Case  $r(0) - f'(0) < 0$ :** in this case, we do not show that convergence occurs with an exponential speed. Thus, we do not prove the two points as before, but simply that

$$\limsup_{t \rightarrow +\infty} S(t) > 0 \quad \text{and} \quad \lim_{t \rightarrow +\infty} S'(t) = 0,$$

which will imply, by definition of  $R$  (4.17), that  $R$  converges to 0. According to (4.22),

$$S(t) = \int_0^{+\infty} n^0(y) e^{\int_y^{X(t,y)} \frac{r(s)}{f(s)} ds} dy.$$

By hypothesis,  $f$  converges to a positive limit. Thus, for all  $y > 0$ , there exist  $\varepsilon_y > 0$  such that  $f(s) > \varepsilon_y$ , for all  $s \geq y$ . Thus, for all  $y > 0$ ,  $\int_y^{+\infty} \frac{r(s)}{f(s)} ds \leq \frac{1}{\varepsilon_y} \|r\|_{L^1} < +\infty$ . This implies that  $y \mapsto n^0(y) e^{\int_y^{+\infty} \frac{r(s)}{f(s)} ds}$  is well defined on  $\mathbb{R}_+$ . Moreover, this function is positive at any  $y$  such that  $n^0(y) > 0$ , hence its integral is positive. Finally,  $t \mapsto n^0(y) e^{\int_y^{X(t,y)} \frac{r(s)}{f(s)} ds}$  converges to this function pointwise. Owing to (4.26) (with  $l = 0$ ),

$$S'(t) = \int_{\text{supp}(n^0)} n^0(y) r(X(t,y)) e^{\int_y^{X(t,y)} \frac{r(s)}{f(s)} ds} dy.$$

By hypothesis, there exist  $\varepsilon, M > 0$  such that  $f(s) \geq \varepsilon$  for all  $s \geq M$ . Thus, for all  $y > 0$ ,

$$\int_y^{X(t,y)} \frac{r(s)}{f(s)} ds \leq \int_y^{+\infty} \frac{r(s)}{f(s)} ds \leq \underbrace{\int_M^{+\infty} \frac{r(s)}{f(s)} ds}_{\leq \frac{\|r\|_{L^1}}{\varepsilon}} + \int_y^M \frac{r(s)}{f(s)} ds \mathbb{1}_{(0,M)}(y).$$

Since  $r \in L^1(\mathbb{R}_+)$ , by hypothesis, this proves that there exists a constant  $K > 0$  such that for all  $t \geq 0$ ,  $y > 0$ ,

$$\left| n^0(y) r(X(t,y)) e^{\int_y^{X(t,y)} \frac{r(s)}{f(s)} ds} \right| \leq \|n^0\|_{\infty} \|r\|_{\infty} e^K e^{\int_y^M \frac{r(s)}{f(s)} ds} \mathbb{1}_{(0,M)}(y).$$

By virtue of Lemma 8, this last quantity is integrable, since

$$e^{\int_y^M \frac{r(s)}{f(s)} ds} = O\left(y^{-\frac{r(0)}{f'(0)}}\right),$$

with  $r(0) < f'(0)$ , by hypothesis. Moreover, since  $t \mapsto r(X(t, y))$  converges to 0 as  $t$  goes to  $+\infty$  for any  $y > 0$ ,  $n^0(y) r(X(t, y)) e^{\int_y^{X(t, y)} \frac{r(s)}{f(s)} ds} dy$  converges to 0 pointwise. According to the dominated convergence theorem,  $S'$  thus converges to 0.

- Let us assume that  $n^0(a) = 0$ , and that the hypothesis of the theorem regarding  $n^{0'}$  holds. We follow exactly the same steps and use the same formulae as in the case ' $n^0(0) > 0$ ', by adapting the computations. We distinguish again two cases.

– **Case  $\tilde{r}_\alpha(0) > 0$ :**

(a) According to (4.22),

$$S(t)e^{-\tilde{r}_\alpha(0)t} = \int_{\mathcal{E}} n^0(Y(t, x)) e^{\int_{Y(t, x)}^x \frac{\tilde{r}(s) - \tilde{r}(0)}{f(s)} ds} dx.$$

For all  $x \in (0, +\infty)$ ,

$$\begin{aligned} n^0(Y(t, x)) e^{\int_{Y(t, x)}^x \frac{\tilde{r}(s) - \tilde{r}_\alpha(0)}{f(s)} ds} &= \frac{n^0(Y(t, x))}{Y(t, x)^\alpha} e^{\int_{Y(t, x)}^x \frac{\tilde{r}(s) - \tilde{r}(0)}{f(s)} ds} \\ &\times Y(t, x)^\alpha e^{\int_{Y(t, x)}^x \frac{\alpha f'(0)}{f(s)} ds}. \end{aligned}$$

Let  $x > 0$ . On the one hand,

$$\frac{n^0(Y(t, x))}{Y(t, x)^\alpha} e^{\int_{Y(t, x)}^x \frac{\tilde{r}(s) - \tilde{r}(0)}{f(s)} ds} \xrightarrow{t \rightarrow +\infty} C e^{\int_0^x \frac{\tilde{r}(s) - \tilde{r}(0)}{f(s)} ds}$$

which is well defined since  $s \mapsto \frac{\tilde{r}(s) - \tilde{r}(0)}{f(s)}$  is continuous on  $[0, x)$ , according to the regularity assumptions on  $r$  and  $f$ , and positive. On the other hand, by rewriting

$$\begin{aligned} Y(t, x)^\alpha e^{\int_{Y(t, x)}^x \frac{-\alpha f'(0)}{f(s)} ds} &= e^{\alpha(\ln(Y(t, x)) - \ln(x))} x^\alpha e^{\int_{Y(t, x)}^x \frac{\alpha f'(0)}{f(s)} ds} \\ &= x^\alpha e^{\int_{Y(t, x)}^x \frac{\alpha f'(0)}{f(s)} - \frac{\alpha}{s} ds}, \end{aligned}$$

and by noting that  $s \mapsto \frac{\alpha f'(0)}{f(s)} - \frac{\alpha}{s}$  is continuous at 0, since

$$\frac{\alpha f'(0)}{f(s)} - \frac{\alpha}{s} = \frac{\alpha f'(0)s - \alpha f(s)}{sf(s)} \quad (4.29)$$

$$= \frac{\alpha f'(0)s - \alpha f'(0)s + f''(0)/2s^2 + o(s^2)}{f'(0)s^2 + o(s^2)} \xrightarrow{s \rightarrow 0} -\frac{\alpha f''(0)}{2f'(0)} \quad (4.30)$$

we show that

$$Y(t, x)^\alpha e^{\int_{Y(t, x)}^x \frac{-\alpha f'(0)}{f(s)} ds} \xrightarrow{t \rightarrow +\infty} x^\alpha e^{\int_0^x \frac{\alpha f'(0)}{f(s)} - \frac{\alpha}{s} ds},$$

which is also well defined, and positive.

- (b) Let  $\delta \in (0, \min(\tilde{r}_\alpha(0), f'(0)))$ . Since  $\delta - \tilde{r}_\alpha(0) < 0$ ,  $r$  goes to 0 at  $+\infty$  and  $f$  is positive, we can find  $M \geq 0$  such that  $\frac{r(s) - \tilde{r}_\alpha(0) + \delta}{f(s)} \leq 0$  for all  $s \in [M, +\infty)$ , and  $\text{supp}(n^0) \subset [0, M]$ . Thus, for all  $t \geq 0$ , and all  $y \in (0, M)$ ,

$$\int_y^{X(t, y)} \frac{r(s) - \tilde{r}_\alpha(0) + \delta}{f(s)} ds \leq \int_y^M \left| \frac{r(s) - \tilde{r}_\alpha(0) + \delta}{f(s)} \right| ds.$$

According to (4.25),

$$e^{\delta t} \frac{d}{dt} \left( S(t) e^{-\tilde{r}_\alpha(0)t} \right) = \int_0^{+\infty} m(y) e^{\int_y^{X(t,y)} \frac{r(s) - \tilde{r}_\alpha(0) + \delta}{f(s)} ds} dy.$$

Thus, since  $\text{supp}(m) \cap \mathcal{E} = \text{supp}(n^0) \cap \mathcal{E} \subset [0, M]$ , and thanks to the previous inequality,

$$\left| e^{\delta t} \frac{d}{dt} \left( S(t) e^{-\tilde{r}_\alpha(0)t} \right) \right| \leq \int_0^M |m(y)| e^{\int_y^M \frac{|r(s) - \tilde{r}_\alpha(0) + \delta|}{f(s)} ds} dy.$$

Let us prove that this integral is bounded. First, let us note that

$$m(y) = n^0(y)(\tilde{r}(y) - \tilde{r}_\alpha(0)) - f(y)n^{0'}(y) = O_{y \rightarrow 0^+}(y^{\alpha+1})$$

Indeed, since  $n^0(y) = Cy^\alpha + O_{y \rightarrow 0^+}(y^{\alpha+1})$  and  $n^{0'}(y) = C\alpha y^{\alpha-1} + O_{y \rightarrow 0^+}(y^\alpha)$ ,

$$\begin{aligned} \frac{|m(y)|}{y^{\alpha+1}} &\leq \frac{n^0(y)}{y^\alpha} \frac{|\tilde{r}(y) - \tilde{r}(0)|}{y} + \frac{|\alpha f'(0)n^0(y) - f(y)n^{0'}(y)|}{y^{\alpha+1}} \\ &\leq \frac{n^0(y)}{y^\alpha} \|\tilde{r}'\|_\infty + \frac{|C\alpha f'(0)y^\alpha - C\alpha f'(0)y^\alpha + O(y^{\alpha+1})|}{y^{\alpha+1}} \\ &= O_{y \rightarrow 0^+}(1). \end{aligned}$$

Moreover, according to Lemma 8, since  $|r(0) - \tilde{r}_\alpha(0) + \delta| = (\alpha + 1)f'(0) + \delta$ ,

$$e^{\int_y^M \frac{|r(s) - \tilde{r}_\alpha(0) + \delta|}{f(s)} ds} = O_{y \rightarrow 0^+}(y^{-\alpha-1-\delta/f'(0)}).$$

Therefore,

$$|m(y)| e^{\int_y^M \frac{|r(s) - \tilde{r}_\alpha(0) + \delta|}{f(s)} ds} = O_{y \rightarrow 0^+}(y^{-\delta/f'(0)}),$$

and is thus integrable since  $\delta < f'(0)$ .

– **Case  $\tilde{r}_\alpha(0) < 0$ :** again, we just prove that

$$\limsup_{t \rightarrow +\infty} S(t) > 0 \quad \text{and} \quad \lim_{t \rightarrow +\infty} S'(t) = 0.$$

According to (4.22),

$$S(t) = \int_0^{+\infty} n^0(y) e^{\int_y^{X(t,y)} \frac{r(s)}{f(s)} ds} dy.$$

By hypothesis,  $f$  converges to a positive limit. Thus, for all  $y > 0$ , there exist  $\varepsilon_y > 0$  such that  $f(s) > \varepsilon_y$ , for all  $s \geq y$ . Hence, for all  $y > 0$ ,  $\int_y^{+\infty} \frac{r(s)}{f(s)} ds \leq \frac{1}{\varepsilon_y} \|r\|_{L^1} < +\infty$ . This ensures that  $y \mapsto n^0(y) e^{\int_y^{+\infty} \frac{r(s)}{f(s)} ds}$  is well defined on  $\mathbb{R}_+$ . Moreover, this function is positive for every  $y$  such that  $n^0(y) > 0$ , which ensures that its integral is positive, and  $t \mapsto n^0(y) e^{\int_y^{X(t,y)} \frac{r(s)}{f(s)} ds}$  converges to this function pointwise. According to (4.26), (with  $l = 0$ ),

$$S'(t) = \int_{\text{supp}(n^0)} n^0(y) r(X(t,y)) e^{\int_y^{X(t,y)} \frac{r(s)}{f(s)} ds} dy.$$

By hypothesis, there exist  $\varepsilon, M > 0$  such that  $f(s) \geq \varepsilon$  for all  $s \geq M$ . Thus, for all  $y > 0$ ,

$$\int_y^{X(t,y)} \frac{r(s)}{f(s)} ds \leq \int_y^{+\infty} \frac{r(s)}{f(s)} ds \leq \underbrace{\int_M^{+\infty} \frac{r(s)}{f(s)} ds}_{\leq \frac{\|r\|_{L^1}}{\varepsilon}} + \int_y^M \frac{r(s)}{f(s)} ds \mathbb{1}_{(0,M)}(y).$$

Since  $r \in L^1(\mathbb{R}_+)$ , this proves that there exist a constant  $K > 0$  such that for all  $t \geq 0, y > 0$ ,

$$\left| n^0(y) r(X(t, y)) e^{\int_y^{X(t, y)} \frac{r(s)}{f(s)} ds} \right| \leq \|r\|_\infty e^K n^0(y) e^{\int_y^M \frac{r(s)}{f(s)} ds} \mathbb{1}_{(0, M)}(y).$$

By hypothesis, and according to Lemma 8,

$$n^0(y) = O_{y \rightarrow 0^+}(y^\alpha) \quad \text{and} \quad e^{\int_y^M \frac{r(s)}{f(s)} ds} = O_{y \rightarrow 0^+}\left(y^{-\frac{r(0)}{f'(0)}}\right).$$

Thus,

$$n^0(y) e^{\int_y^M \frac{r(s)}{f(s)} ds} = O_{y \rightarrow 0^+}\left(y^{\alpha - \frac{r(0)}{f'(0)}}\right),$$

with  $\alpha - \frac{r(0)}{f'(0)} > -1$ . Moreover, since  $t \mapsto r(X(t, y))$  converges to 0 as  $t$  goes to  $+\infty$  for any  $y > 0$ ,  $n^0(y) r(X(t, y)) e^{\int_y^{X(t, y)} \frac{r(s)}{f(s)} ds} dy$  converges to 0 pointwise. By the dominated convergence theorem,  $S'$  thus converges to 0.

- We can prove this point exactly as we treat the case  $f > 0$  on  $\mathbb{R}$ . We therefore leave it to the reader and refer to the proof of (vii).
- (v) Let us note that, for any  $x, y \in (0, 1)$ ,  $t \mapsto X(t, y)$  is increasing and converges to 1, and  $t \mapsto Y(t, x)$  is decreasing and converges to 0.
  - Let us assume that  $n^0(a) > 0$ . We distinguish again between two cases:
    - **Case  $r(1) > \tilde{r}(0)$ :**
      - (a) Let us use the second expression (4.22) for  $S$ , *i.e.*

$$S(t) e^{-r(1)t} = \int_0^1 e^{\int_y^{X(t, y)} \frac{r(s) - r(1)}{f(s)} ds} dy.$$

For all  $y \in (0, 1)$ ,  $n^0(y) e^{\int_y^{X(t, y)} \frac{r(s) - r(1)}{f(s)} ds} \xrightarrow[t \rightarrow +\infty]{} n^0(y) e^{\int_y^1 \frac{r(s) - r(1)}{f(s)} ds}$ , which is well-defined for all  $y \in (0, 1)$ , since  $s \mapsto \frac{r(s) - r(1)}{f(s)}$  is continuous on  $(0, 1]$ , and positive on a set of non-zero measure, since it is positive where  $n^0$  is positive.

- (b) Let  $\delta \in (0, \min(r(1) - \tilde{r}(0), -f'(1)))$ . Since  $\tilde{r}(0) - r(1) + \delta < 0$ , there exists  $m \in (0, 1)$  such that  $\tilde{r}(s) - r(1) + \delta$  for all  $s \in (0, m]$ . Thus, for all  $x \in (0, 1)$ ,  $t \geq 0$ ,

$$\int_{Y(t, x)}^x \frac{\tilde{r}(s) - r(1) + \delta}{f(s)} ds \leq \int_m^x \frac{|\tilde{r}(s) - r(1) + \delta|}{f(s)} ds \mathbb{1}_{(m, 1)}(x).$$

Thus, using expression (4.28),

$$\begin{aligned} e^{\delta t} \frac{d}{dt} \left( S(t) e^{-r(1)t} \right) &= \int_0^1 n^0(Y(t, x)) (r(x) - r(1)) e^{\int_{Y(t, x)}^x \frac{\tilde{r}(s) - r(1) + \delta}{f(s)} ds} dx \\ &\leq \|n^0\| \int_0^1 |r(x) - r(1)| e^{\int_m^x \frac{|\tilde{r}(s) - r(1) + \delta|}{f(s)} ds} \mathbb{1}_{(m, 1)}(x) dx \\ &< +\infty. \end{aligned}$$

This last integral is finite since  $|\tilde{r}(1) - r(1) + \delta| = -f'(1) + \delta$ , and thus  $e^{\int_m^x \frac{|\tilde{r}(s) - r(1) + \delta|}{f(s)} ds} = O_{x \rightarrow 1}\left(|x - 1|^{\frac{\delta}{f'(1)} - 1}\right)$ , by Lemma 8)  $|r(x) - r(1)| = O_{x \rightarrow 1}|x - 1|$ , and  $\frac{\delta}{f'(1)} > -1$  by hypothesis.

- **Case  $\tilde{r}(0) > r(1)$ :**
  - (a) Using (4.22), we find

$$S(t) e^{-\tilde{r}(0)t} = \int_0^1 n^0(Y(t, x)) e^{\int_{Y(t, x)}^x \frac{\tilde{r}(s) - \tilde{r}(0)}{f(s)} ds} dx.$$

For all  $x \in (0, 1)$ ,  $n^0(Y(t, x))e^{\int_Y(t, x) \frac{\tilde{r}(s) - \tilde{r}(0)}{f(s)} ds} \xrightarrow[t \rightarrow +\infty]{} n^0(0)e^{\int_0^x \frac{\tilde{r}(s) - \tilde{r}(0)}{f(s)} ds}$ , which is well-defined since  $s \mapsto \frac{\tilde{r}(s) - \tilde{r}(0)}{f(s)}$  is continuous on  $[0, 1]$ , and positive by hypothesis on  $n^0$ .

- (b) Let  $\delta \in (0, \min(\tilde{r}(0) - r(1), f'(0)))$ . Since  $r(1) - \tilde{r}(0) + \delta < 0$ , there exists  $M \in (0, 1)$  such that  $r(s) - \tilde{r}(0) + \delta < 0$  for all  $s \geq M$ . Thus, for all  $y \in (0, 1)$ ,  $t \geq 0$ ,  $\int_y^{X(t, y)} \frac{r(s) - \tilde{r}(0) + \delta}{f(s)} ds \leq \int_y^M \frac{|r(s) - \tilde{r}(0) + \delta|}{f(s)} ds \mathbb{1}_{(0, M)}(y)$ . Thus,

$$\begin{aligned} \left| e^{\delta t} \frac{d}{dt} \left( S(t) e^{-\tilde{r}(0)t} \right) \right| &= \left| \int_0^1 m(y) e^{\int_y^{X(t, y)} \frac{r(s) - \tilde{r}(0) + \delta}{f(s)} ds} dy \right| \\ &\leq \int_0^1 |m(y)| e^{\int_y^M \frac{|r(s) - \tilde{r}(0) + \delta|}{f(s)} ds} \mathbb{1}_{(0, M)}(y) dy, \end{aligned}$$

which is a finite integral, since  $|r(0) - \tilde{r}(0) + \delta| = f'(0) + \delta$ , and thus

$$e^{\int_y^M \frac{|r(s) - \tilde{r}(0) + \delta|}{f(s)} ds} = O_{y \rightarrow 0} \left( y^{-\frac{\delta}{f'(0)} - 1} \right)$$

(by Lemma 8),

$$m(y) = n^0(y) (\tilde{r}(y) - \tilde{r}(0)) - f(y) n^{0'}(y) = O_{y \rightarrow 0}(y),$$

and  $\frac{\delta}{f'(0)} < 1$  thanks to our choice for  $\delta$ .

- Let us assume that  $n^0(a) = 0$ , and that the hypothesis on  $n^{0'}$  of the theorem holds. As usual, we distinguish two cases.

– **Case**  $r(1) > \tilde{r}_\alpha(0)$ :

- (a) This first point is exactly the same as in the case  $n^0 > 0$ . Let us use the second expression (4.22) for  $S$ , *i.e.*

$$S(t) e^{-r(1)t} = \int_0^1 e^{\int_y^{X(t, y)} \frac{r(s) - r(1)}{f(s)} ds} dy.$$

For all  $y \in (0, 1)$ ,  $n^0(y) e^{\int_y^{X(t, y)} \frac{r(s) - r(1)}{f(s)} ds} \xrightarrow[t \rightarrow +\infty]{} n^0(y) e^{\int_y^1 \frac{r(s) - r(1)}{f(s)} ds}$ , which is well-defined for all  $y \in (0, 1)$ , since  $s \mapsto \frac{r(s) - r(1)}{f(s)}$  is continuous on  $(0, 1]$ , and positive on a set of measure non-zero, since it is positive where  $n^0$  is positive.

- (b) Let  $\delta \in (0, \min(r(1) - \tilde{r}_\alpha(0), -f'(1)))$ . First, let us note that we can rewrite

$$Y(t, x)^\alpha = e^{\ln(Y(t, x)) - \ln(x)} x^\alpha = x^\alpha e^{\int_Y(t, x) \frac{\varphi(s)}{f(s)} ds}.$$

Thus, by using expression (4.28), we get

$$\begin{aligned} e^{\delta t} \frac{d}{dt} \left( S(t) e^{-r(1)t} \right) &= \int_0^1 n^0(Y(t, x)) (r(x) - r(1)) e^{\int_Y(t, x) \frac{\tilde{r}(s) - r(1) + \delta}{f(s)} ds} dx \\ &= \int_0^1 \frac{n^0(Y(t, x))}{Y(t, x)^\alpha} x^\alpha (r(x) - r(1)) e^{\int_Y(t, x) \frac{\varphi(s)}{f(s)} ds} dx, \end{aligned}$$

with

$$\varphi(s) := \tilde{r}(s) - r(1) + \delta - \frac{\alpha f(s)}{s}.$$

By hypothesis on  $n^0$ ,  $f$  and  $r$ ,  $\tilde{n}^0 : y \mapsto \frac{n^0(y)}{y^\alpha}$  and  $\varphi$  are both continuous on  $[0, 1]$ . Moreover, since  $\varphi(0) = \tilde{r}_\alpha(0) - r(1) + \delta < 0$ , there exists  $\varepsilon \in (0, 1)$  such that  $\varphi(s) < 0$  for all  $s \in [0, \varepsilon]$ . Thus,

$$\left| e^{\delta t} \frac{d}{dt} \left( S(t) e^{-r(1)t} \right) \right| \leq \|\tilde{n}^0\|_\infty \int_0^1 |r(x) - r(1)| x^\alpha e^{\int_\varepsilon^x \frac{|\varphi(s)|}{f(s)} ds} \mathbb{1}_{(\varepsilon, 1)}(x) dx.$$

since  $|\varphi(1)| = \delta - f'(1)$ , Lemma 8 yields

$$e^{\int_\varepsilon^x \frac{|\varphi(s)|}{f(s)} ds} = O\left(|x-1|^{\frac{\delta}{f'(1)}-1}\right).$$

Since  $|r(x) - r(1)| = O_{x \rightarrow 1}(x)$  and  $\frac{\delta}{f'(1)} > -1$  (by hypothesis on  $\delta$ ), this proves that this last integral is bounded.

– **Case  $\tilde{r}_\alpha(0) > r(1)$ :**

(a) According to (4.22),

$$\begin{aligned} S(t)e^{-\tilde{r}_\alpha(0)t} &= \int_0^1 n^0(Y(t,x)) e^{\int_{Y(t,x)}^x \frac{\tilde{r}(s) - \tilde{r}_\alpha(0)}{f(s)} ds} dx \\ &= \int_0^1 \frac{n^0(Y(t,x))}{Y(t,x)^\alpha} Y(t,x)^\alpha e^{\int_{Y(t,x)}^x \frac{\tilde{r}(s) - \tilde{r}_\alpha(0)}{f(s)} ds} dx. \end{aligned}$$

By rewriting  $Y(t,x)^\alpha = x^\alpha e^{-\int_{Y(t,x)}^x \frac{\alpha}{s} ds}$ , we get

$$S(t)e^{-\tilde{r}_\alpha(0)t} = \int_0^1 \frac{n^0(Y(t,x))}{Y(t,x)^\alpha} x^\alpha e^{\int_{Y(t,x)}^x \frac{\tilde{r}(s) - \tilde{r}_\alpha(0)}{f(s)} - \frac{\alpha}{s} ds} dx.$$

Since  $\frac{n^0(Y(t,x))}{Y(t,x)^\alpha} x^\alpha e^{\int_{Y(t,x)}^x \frac{\tilde{r}(s) - \tilde{r}_\alpha(0)}{f(s)} - \frac{\alpha}{s} ds}$  converges pointwise to

$C x^\alpha e^{\int_0^x \frac{\tilde{r}(s) - \tilde{r}_\alpha(0)}{f(s)} - \frac{\alpha}{s} ds}$ , which is well-defined, since  $s \mapsto \frac{\tilde{r}(s) - \tilde{r}_\alpha(0)}{f(s)} - \frac{\alpha}{s}$  is continuous at 0 and positive, we are done.

(b) Let  $\delta \in (0, \min(\tilde{r}_\alpha(0) - r(1), f'(0)))$ . Since  $r(1) - \tilde{r}_\alpha(0) + \delta < 0$ , there exists  $M \in (0, 1)$  such that  $r(s) - \tilde{r}_\alpha(0) + \delta < 0$  for all  $s \geq M$ . Thus, for all  $y \in (0, 1)$ ,  $t \geq 0$ ,

$$\int_y^{X(t,y)} \frac{r(s) - \tilde{r}_\alpha(0) + \delta}{f(s)} ds \leq \int_y^M \frac{|r(s) - \tilde{r}_\alpha(0) + \delta|}{f(s)} ds \mathbb{1}_{(0,M)}(y).$$

Hence, using expression (4.25), we get

$$\begin{aligned} \left| e^{\delta t} \frac{d}{dt} \left( S(t) e^{-\tilde{r}_\alpha(0)t} \right) \right| &= \left| \int_0^1 m(y) e^{\int_y^{X(t,y)} \frac{r(s) - \tilde{r}_\alpha(0) + \delta}{f(s)} ds} \right| \\ &\leq \int_0^1 |m(y)| e^{\int_y^M \frac{|r(s) - \tilde{r}_\alpha(0) + \delta|}{f(s)} ds} \mathbb{1}_{(0,M)}(y) dy, \end{aligned}$$

which is a finite integral, since  $|r(0) - \tilde{r}_\alpha(0) + \delta| = (1 + \alpha)f'(0) + \delta$ . Lemma 8 leads to

$$e^{\int_y^b \frac{|r(s) - \tilde{r}_\alpha(0) + \delta|}{f(s)} ds} = O\left(y^{-\frac{\delta}{f'(0)} - \alpha - 1}\right).$$

The integrability follows from

$$m(y) = n^0(y) (\tilde{r}(y) - \tilde{r}_\alpha(0)) - f(y) n^{0'}(y) = O_{y \rightarrow 0}(y^{\alpha+1})$$

(as seen previously), and  $\frac{\delta}{f'(0)} < 1$  thanks to our choice for  $\delta$ .

- We prove this case with exactly the same arguments that for the case of a unique root which is asymptotically unstable. We therefore apply the proof of (i).

(vii) In this case, since  $f > 0$ ,  $X(t,y) \xrightarrow[t \rightarrow +\infty]{} +\infty$ , for all  $y \in \mathbb{R}$ . Let us prove that

$$\liminf_{t \rightarrow +\infty} S(t) > 0 \quad \text{and} \quad \lim_{t \rightarrow +\infty} S(t) = 0.$$

According to (4.21),

$$S(t) = \int_{\text{supp}(n^0)} n^0(y) e^{\int_y^{X(t,y)} \frac{r(s)}{f(s)} ds} dy.$$

The integrand  $n^0(y)e^{\int_y^{X(t,y)} \frac{r(s)}{f(s)} ds}$  converges pointwise to  $n^0(y)e^{\int_y^{+\infty} \frac{r(s)}{f(s)} ds}$ , which is well defined (with values in  $[0, +\infty]$ ), and positive for all  $y \in \text{supp}(n^0)$ , since  $\frac{r}{f}$  is positive. According to (4.28),

$$S'(t) = \int_{\text{supp}(n^0)} n^0(y)r(X(t,y))e^{\int_y^{X(t,y)} \frac{r(s)}{f(s)} ds} dy.$$

Since  $f$  is continuous, positive, and converges to positive constants at  $\pm\infty$ ,  $\varepsilon := \min_{s \in \mathbb{R}} f(s) > 0$ .

Thus, for all  $y \in \mathbb{R}, t \geq 0$ ,

$$\left| n^0(y)r(X(t,y))e^{\int_y^{X(t,y)} \frac{r(s)}{f(s)} ds} \right| \leq \|n^0\|_\infty \|r\|_\infty e^{\frac{\|r\|_{L^1}}{\varepsilon}} < +\infty.$$

Combined with the fact that  $r(X(t,y))$  converges to 0 as  $t$  goes to  $+\infty$  pointwise, we deduce that  $S'$  converges to 0 by the dominated convergence theorem.

(ix) Since  $f \equiv 0$  on  $\mathcal{E}$ ,  $Y(t,x) = x$  for all  $(t,x) \in \mathbb{R}_+ \times \mathcal{E}$ . Thus, according to formula (4.20),

$$S(t) = \int_{\mathcal{E}} n^0(x)e^{r(x)t} dx \quad \text{and} \quad S'(t) = \int_{\mathcal{E}} n^0(x)r(x)e^{r(x)t} dx.$$

By Laplace's formula (see [162]),

$$S(t) \underset{t \rightarrow +\infty}{\sim} \sqrt{2\pi} \left( \sum_{i=1}^p \frac{n^0(x_i)}{\sqrt{|r''(x_i)|}} \right) \frac{e^{\bar{r}t}}{\sqrt{t}}$$

and

$$S'(t) \underset{t \rightarrow +\infty}{\sim} \sqrt{2\pi} \left( \sum_{i=1}^p \frac{n^0(x_i)r(x_i)}{\sqrt{|r''(x_i)|}} \right) \frac{e^{\bar{r}t}}{\sqrt{t}} = \sqrt{2\pi} \bar{r} \left( \sum_{i=1}^p \frac{n^0(x_i)}{\sqrt{|r''(x_i)|}} \right) \frac{e^{\bar{r}t}}{\sqrt{t}} \underset{t \rightarrow +\infty}{\sim} \bar{r} S(t).$$

Thus,  $R(t) = \frac{S'(t)}{S(t)} \xrightarrow{t \rightarrow +\infty} \bar{r}$ .

□

## 4.4.2 Applications

**Summary of the method.** The method that we propose in order to study the asymptotic behaviour of PDE (4.1) can be summarised by the following three steps:

1. Choose an appropriate family of set  $(\mathcal{O}_i)$  which satisfies the assumptions of Proposition 4.12, and such that we can compute the asymptotic behaviour of the functions  $R_i$ : a good choice when  $f$  has a finite number of roots is to take the interval between the roots, as suggested in Lemma 7.
2. Use Proposition 17 in order to determine the limit of  $\rho$ , and its speed of convergence when possible.
3. Use the semi-explicit expression of  $n$  provided by equation (4.7), and eventually Proposition 5 to deduce the asymptotic behaviour of  $n$ .

In each of the following subsections, we apply the three points detailed in this summary to study the asymptotic behaviour of  $n$  in different cases.

**Remark regarding the regularity of parameter functions.** As in subsection 4.4.1, we make the further assumptions that  $f \in \text{BV}(\mathbb{R})$ , and that  $f$  converges to a non-zero limit at  $\pm\infty$ . Moreover, we easily check that all the results of this previous section remain true if we assume that  $n^0$  is  $\mathcal{C}^1$  on each interval between the roots of  $f$ , and not necessarily on the whole of  $\mathbb{R}$ . As far as  $f$  is concerned, it is enough to assume that it is globally Lipschitz, and  $\mathcal{C}^2$  only on a neighbourhood of its roots. It will sometimes be advisable to make these two additional assumptions: we will indicate this at the beginning of each statement whenever this is the case.



### Case of a unique stable equilibrium

We start by assuming that  $f$  has a unique root (denoted  $a$ ), which is asymptotically stable for the ODE  $\dot{u} = f(u)$ . In this case, solutions converge to a weighted Dirac mass at  $a$ , regardless of the functions  $r$  and  $n^0$ . The weight in front of the Dirac mass is determined by the value of  $r$  at  $a$ . Note that this result can be generalised to higher dimensions, see Proposition 25.

**Proposition 19.** *Let us assume that  $f$  has a unique root (denoted  $a$ ), and that  $f'(a) < 0$ . Then,  $\rho$  converges to  $r(a)$  and  $n(t, \cdot) \xrightarrow[t \rightarrow +\infty]{} r(a)\delta_a$ .*

*Proof.* We apply the three points detailed in the summary:

1. Let us denote  $\mathcal{O}_1 := (-\infty, a)$ ,  $\mathcal{O}_2 := (a, +\infty)$ , which satisfy the assumptions of Proposition 16, by Lemma 7. By proposition 18,  $R_1$  and  $R_2$  both converge to  $r(a)$  (with an exponential speed).
2. By Proposition 17,  $\rho$  converges to  $r(a)$  with an exponential speed.
3. According to to the semi-explicit expression (4.7),  $n(t, x) = n^0(Y(t, x))e^{\int_0^t \tilde{r}(Y(s, x)) - \rho(s) ds}$ . Let  $\delta > 0$ . Since  $\|Y(t, x)\| \xrightarrow[t \rightarrow +\infty]{} +\infty$  for all  $x \in \mathbb{R}^d \setminus \{a\}$ , and  $n^0$  has a compact support, there exists  $T_0$  such that  $n(t, x) = 0$  for all  $t \geq T_0$ ,  $x \in \mathbb{R}^d \setminus [a - \delta, a + \delta]$ . Since  $\rho(t) = \int_{\text{supp}(n^0)} n(t, x) dx$  converges to  $r(a)$ , Propositions 5 allows us to conclude that  $n(t, \cdot) \xrightarrow[t \rightarrow +\infty]{} r(a)\delta_a$ .

□

### Case of a unique unstable equilibrium

We now assume that  $f$  has a unique root (denoted  $a$ ) which is asymptotically unstable for the ODE  $\dot{u} = f(u)$ . Under these hypotheses, the growth term can counterbalance the advection term: there exist two regimes of convergence, depending on how  $r(a)$  and  $f'(a)$  compare.

**Proposition 20.** *Let us assume that  $f$  has a unique root (denoted  $a$ ), and that  $f'(a) > 0$ , Then:*

- If  $r(a) < f'(a)$ , then  $\rho(t) \xrightarrow[t \rightarrow +\infty]{} 0$  and  $n(t, \cdot) \xrightarrow[t \rightarrow +\infty]{} 0$  in  $L^1(\mathbb{R})$ .
- If  $r(a) > f'(a)$ , and  $n^0(a) > 0$ , then  $\rho(t) \xrightarrow[t \rightarrow +\infty]{} r(a) - f'(a)$ , and  $n(t, \cdot) \xrightarrow[t \rightarrow +\infty]{} \bar{n}$  in  $L^1(\mathbb{R})$ , where

$$\bar{n}(x) := C e^{\int_a^x \frac{\tilde{r}(s) - \tilde{r}(a)}{f(s)} ds},$$

with  $\tilde{r} = r - f'$  and  $C$  such that  $\int_{\mathbb{R}} \bar{n}(x) dx = r(a) - f'(a)$ .

*Proof.* We apply the three points detailed in the summary:

- Let us assume that  $r(a) < f'(a)$ :
  1. Let us denote  $\mathcal{O}_1 := (-\infty, a)$ ,  $\mathcal{O}_2 := (a, +\infty)$ , which satisfy the assumptions of Proposition 16, by Lemma 7. Proposition 18 shows that  $R_1$  and  $R_2$  both converge to 0.
  2. By Proposition 17,  $\rho$  converges to 0.

3. We immediately deduce from the previous point that  $n(t, \cdot) \xrightarrow[t \rightarrow +\infty]{} 0$  in  $L^1(\mathbb{R})$ , by definition of  $\rho$ .

• Let us assume that  $r(a) > f'(a)$ :

1. With the same choice for  $\mathcal{O}_1$  and  $\mathcal{O}_2$ , Proposition 18 shows that  $R_1$  and  $R_2$  both converge to  $r(a) - f'(a)$ .
2. By Proposition 17,  $\rho$  converges to  $\tilde{r}(a)$  with an exponential speed.
3. By the semi-explicit expression (4.7),

$$\begin{aligned} n(t, x) &= n^0(Y(t, x)) e^{\int_0^t \tilde{r}(Y(s, x)) - \rho(s) ds} \\ &= n^0(Y(t, x)) e^{\int_0^t \tilde{r}(Y(s, x)) - \tilde{r}(a) ds} e^{\int_0^t \tilde{r}(a) - \rho(s) ds} \\ &= n^0(Y(t, x)) e^{\int_{Y(t, x)}^x \frac{\tilde{r}(s) - \tilde{r}(a)}{f(s)} ds} e^{\int_0^t \tilde{r}(a) - \rho(s) ds} \end{aligned}$$

(we use the change of variable  $s' = Y(s, x)$  in the first integral to get this last expression). Thus,  $n(t, \cdot)$  converges pointwise to

$$x \mapsto n^0(a) e^{\int_a^x \frac{\tilde{r}(s) - \tilde{r}(a)}{f(s)} ds} e^{\int_0^{+\infty} \tilde{r}(a) - \rho(s) ds},$$

which is well-defined, since  $\rho$  converges to  $\tilde{r}(a)$  with an exponential speed,  $f > 0$  on  $(a, +\infty)$  and  $s \mapsto \frac{\tilde{r}(s) - \tilde{r}(a)}{f(s)}$  is continuous at  $a$ .

Moreover, since  $r(x) \xrightarrow[x \rightarrow +\infty]{} 0$ , and  $f$  converges to a positive limit, there exist  $M, d > 0$  such that  $\frac{r(s) - \tilde{r}(a)}{f(s)} < -d$  for all  $s \geq M$ . Thus, for all  $t \geq 0$ ,  $x \in (a, +\infty)$ ,

$$\begin{aligned} \int_{Y(t, x)}^x \frac{\tilde{r}(s) - \tilde{r}(a)}{f(s)} ds &\leq \underbrace{\int_a^M \frac{|\tilde{r}(s) - \tilde{r}(a)|}{f(s)} ds}_{:=C_1} + \int_M^x \frac{\tilde{r}(s) - \tilde{r}(a)}{f(s)} ds \mathbb{1}_{(M, +\infty)}(x) \\ &\leq C_1 + \int_M^x \underbrace{\frac{r(s) - \tilde{r}(a)}{f(s)}}_{\leq -d} ds \mathbb{1}_{(M, +\infty)}(x) \\ &\quad + \int_M^x \frac{f'(s)}{f(s)} ds \mathbb{1}_{(M, +\infty)}(x) \\ &\leq C_1 - d(x - M) \mathbb{1}_{(M, +\infty)} + \underbrace{\int_M^{+\infty} \frac{|f'(s)|}{f(s)} ds}_{:=C_2}, \end{aligned}$$

with  $C_1, C_2 < +\infty$ , by the regularity of  $\tilde{r}$ ,  $f \in BV(\mathbb{R})$ , and the fact that  $f$  converges to a positive constant at infinity.

By proceeding in the same way for all  $x \leq a$ , we show that for all  $x \in \mathbb{R}$ ,  $t \geq 0$ ,

$$n(t, x) \leq C e^{-d|x|}$$

for some constants  $C, d > 0$ , which ensures, according to the dominated convergence theorem, that  $t \mapsto n(t, \cdot)$  converges to  $x \mapsto n^0(a) e^{\int_a^x \frac{tr(s) - \tilde{r}(a)}{f(s)} ds} e^{\int_0^{+\infty} \tilde{r}(a) - \rho(s) ds}$  in  $L^1(\mathbb{R})$ .

□

## Two equilibria

In this section we assume that  $f$  has exactly two roots,  $a < b$ , which satisfy  $f'(a) > 0$  and  $f'(b) < 0$  (hence  $f > 0$  on  $(a, b)$ ). The case  $f'(a) < 0$ ,  $f'(b) > 0$ ,  $f < 0$  on  $(a, b)$  is similar. Depending on the functions  $r$  and  $n^0$ ,  $n$  will either converge to a function in  $L^1$ , or converge to a Dirac mass at  $b$ . We split this result into two propositions: the first one assumes that the support of  $n^0$  crosses  $a$ , which means that  $n^0 > 0$  in a neighbourhood of  $a$ . The second one assumes that  $\text{supp}(n^0) \subset [a, +\infty)$ , and we consider the case where  $n^0(a) = 0$ , which leads to other regular functions being reached.

**Proposition 21.** *Let us assume that  $f$  has exactly two roots,  $a < b$ , which satisfy  $f'(a) > 0$ ,  $f'(b) < 0$ , and that  $n^0(a) > 0$ . Then:*

- If  $r(b) > r(a) - f'(a)$ , then  $\rho(t) \xrightarrow{t \rightarrow +\infty} r(b)$  and  $n(t, \cdot) \xrightarrow{t \rightarrow +\infty} r(b)\delta_b$ .
- If  $r(b) < r(a) - f'(a)$ , then  $\rho(t) \xrightarrow{t \rightarrow +\infty} r(a) - f'(a)$ , and  $n(t, \cdot) \xrightarrow{t \rightarrow +\infty} \bar{n}$  in  $L^1(\mathbb{R})$ , where

$$\bar{n}(x) := D e^{\int_a^x \frac{\tilde{r}(s) - \tilde{r}(a)}{f(s)} ds} \mathbb{1}_{(-\infty, b)},$$

with  $\tilde{r} = r - f'$ , and  $D > 0$  is such that  $\int_{\mathbb{R}} \bar{n}(x) dx = r(a) - f'(a)$ .

*Proof.* Note that since  $n^0$  is assumed to be continuous on  $\mathbb{R}$ ,  $n^0 > 0$  on a neighbourhood of  $a$ .

- Let us assume that  $r(b) > r(a) - f'(a)$ . We again follow the three points of the method outlined in the beginning of the subsection.
  1. Let us denote  $\mathcal{O}_1 = (-\infty, a)$ ,  $\mathcal{O}_2 = (a, b)$ ,  $\mathcal{O}_3 = (b, +\infty)$ . One easily checks that these sets satisfy the hypotheses of Proposition 16, thanks to Lemma 7. According to Proposition 18,  $R_1$  converges to  $\max(0, \tilde{r}(a)) < r(b)$  and  $R_2$  and  $R_3$  both converge to  $r(b)$  with an exponential speed.
  2. By Proposition 17,  $\rho$  converges to  $r(b)$  with an exponential speed, and  $\rho_1(t) = \int_{-\infty}^a n(t, x) dx$  converges to 0.
  3. Let  $x \in (a, b)$ . Using (4.7), we find  $n(t, x) = n^0(Y(t, x)) e^{\int_0^t \tilde{r}(Y(s, x)) - \rho(s) ds}$ , for all  $t \geq 0$ . Let  $K \subset (a, b)$  be a compact set,  $\delta \in (0, \frac{1}{2}(r(b) - \tilde{r}(a)))$ , and let us denote  $d := r(b) - \tilde{r}(a) - 2\delta > 0$ . Since  $\rho$  converges to  $r(b)$ , and  $(Y(s, x))_{s \geq 0}$  converges to  $a$  uniformly on  $K$ , there exists  $T_0$  such that for all  $s \geq T_0$  and all  $x \in K$ ,

$$\rho(s) \geq r(b) - \delta \quad \text{and} \quad \tilde{r}(Y(s, x)) \leq \tilde{r}(a) + \delta,$$

Thus,

$$\int_K n(t, x) dx \leq \|n^0\|_{\infty} \int_K e^{\int_0^{T_0} \tilde{r}(Y(s, x)) - \rho(s) ds} dx e^{-d(t-T_0)} \xrightarrow{t \rightarrow +\infty} 0.$$

Let  $K'$  be a compact subset of  $(b, +\infty)$ . Since  $n^0$  has compact support, there exists  $T_0$  such that  $n^0(Y(t, x)) = 0$  for all  $t \geq T_0$ ,  $x \in K'$ . Thus,  $t \mapsto \int_{K'} n(t, x) dx$  converges to 0. By Proposition 5,  $n(t, \cdot) \xrightarrow{t \rightarrow +\infty} r(b)\delta_b$ .

- Let us now assume that  $r(b) < r(a) - f'(a)$ .
  1. With the same choice for  $\mathcal{O}_1, \mathcal{O}_2$  and  $\mathcal{O}_3$ , Proposition 18 shows that  $R_1$  and  $R_2$  converge to  $\tilde{r}(a)$ , and that  $R_3$  converges to  $r(b) < \tilde{r}(a)$ .

2. We then apply Proposition 17 to infer that  $\rho$  converges to  $\tilde{r}(a)$  with an exponential speed, and  $\rho_3(t) = \int_b^{+\infty} n(t, x) dx$  converges to 0.
3. Let  $x \in (-\infty, b)$ ,  $t \geq 0$ . By the semi-explicit expression (4.7),

$$\begin{aligned} n(t, x) &= n^0(Y(t, x)) e^{\int_0^t \tilde{r}(Y(s, x)) - \rho(s) ds} \\ &= n^0(Y(t, x)) e^{\int_{Y(t, x)}^x \frac{\tilde{r}(s) - \tilde{r}(a)}{f(s)} ds} e^{\int_0^t \tilde{r}(a) - \rho(s) ds}, \end{aligned}$$

where we used the change of variable ‘ $s' = Y(t, x)$ ’. The latter function converges pointwise to

$$n^0(a) e^{\int_a^x \frac{\tilde{r}(s) - \tilde{r}(a)}{f(s)} ds} e^{\int_0^{+\infty} \tilde{r}(a) - \rho(s) ds}.$$

As for the case of a unique unstable equilibrium (proof of Proposition 20) one can find  $C, d \geq 0$  such that  $n(t, x) \leq C e^{-d|x|}$  for all  $x \leq a$ . Moreover, for all  $x \in (a, b)$ ,

$$n(t, x) \leq \|n^0\|_{\infty} e^{\int_0^{+\infty} |\tilde{r}(a) - \rho(s)| ds} e^{\int_a^x \frac{\tilde{r}(s) - \tilde{r}(a)}{f(s)} ds} \mathbb{1}_{\{\tilde{r}(s) > \tilde{r}(a)\}}(s) ds,$$

which provides an  $L^1$ -domination, since  $e^{\int_a^x \frac{\tilde{r}(s) - \tilde{r}(a)}{f(s)} ds} = O_{x \rightarrow +\infty} \left( |x - b|^{\frac{\tilde{r}(a) - r(b)}{-f'(b)} - 1} \right)$ , thanks to Lemma 8. By the dominated convergence theorem, combined with the fact that  $\rho$  converges to  $\tilde{r}(a)$  and  $\rho_3$  converges to 0, this ensures that  $n(t, \cdot)$  converges to the expected limit.

□

In the following proposition, we assume that  $n^0$  is  $\mathcal{C}^1$  on  $(a, b)$  and on  $(b, +\infty)$ , and not necessarily on the whole of  $\mathbb{R}$ .

**Proposition 22.** *Let us assume that  $f$  has exactly two roots,  $a < b$ , which satisfy  $f'(a) > 0$ ,  $f'(b) < 0$ , and that  $\text{supp}(n^0) \subset [a, +\infty)$ . We distinguish between several cases:*

- If  $n^0(a) > 0$ , then
  - If  $r(b) > r(a) - f'(a)$ , then  $\rho(t) \xrightarrow{t \rightarrow +\infty} r(b)$ , and  $n(t, \cdot) \xrightarrow{t \rightarrow +\infty} r(b) \delta_b$ .
  - If  $r(b) < r(a) - f'(a)$ , then  $\rho(t) \xrightarrow{t \rightarrow +\infty} r(a) - f'(a)$ , and  $n(t, \cdot) \xrightarrow{t \rightarrow +\infty} \bar{n}_0$  in  $L^1(\mathbb{R})$ , where

$$\bar{n}_0(x) := D_0 e^{\int_a^x \frac{\tilde{r}(s) - \tilde{r}(a)}{f(s)} ds} \mathbb{1}_{(a, b)},$$

with  $\tilde{r} = r - f'$ , and  $D_0 > 0$  is such that  $\int_{\mathbb{R}} \bar{n}_0(x) dx = r(a) - f'(a)$ .

- If  $n^0(a) = 0$ , and if there exist  $C, \alpha > 0$  such that  $n^0(y) = C \alpha (y - a)^{\alpha - 1} + O_{y \rightarrow a^+}((y - a)^\alpha)$ , then
  - If  $r(b) > r(a) - (1 + \alpha)f'(a)$ , then  $\rho(t) \xrightarrow{t \rightarrow +\infty} r(b)$ , and  $n(t, \cdot) \xrightarrow{t \rightarrow +\infty} r(b) \delta_b$ .
  - If  $r(b) < r(a) - (1 + \alpha)f'(a)$ , then  $\rho(t) \xrightarrow{t \rightarrow +\infty} r(a) - (1 + \alpha)f'(a)$ , and  $n(t, \cdot) \xrightarrow{t \rightarrow +\infty} \bar{n}_\alpha$  in  $L^1(\mathbb{R})$ , where

$$\bar{n}_\alpha(x) := D_\alpha (x - a)^\alpha e^{\int_a^x \frac{\tilde{r}(s) - \tilde{r}_\alpha(a)}{f(s)} - \frac{\alpha}{s - a} ds} \mathbb{1}_{(a, b)},$$

where  $\tilde{r} = r - f'$ ,  $\tilde{r}_\alpha = r - (1 + \alpha)f'$ , and  $D_\alpha > 0$  is such that  $\int_{\mathbb{R}} \bar{n}_\alpha(x) dx = r(a) - (1 + \alpha)f'(a)$ .

- If  $n^0(a) = 0$ , and if there exists  $\varepsilon > 0$  such that  $n^0(y) = 0$  for all  $y \in [a, a + \varepsilon]$ , then  $\rho(t) \xrightarrow{t \rightarrow +\infty} r(b)$ , and  $n(t, \cdot) \xrightarrow{t \rightarrow +\infty} r(b)\delta_b$ .

*Proof.* Since the proof of this proposition is very similar to the one of Proposition 21, we do not write it in full detail, but we simply underline the points that must be adapted.

- In the case where  $n^0(a) > 0$ , and  $\text{supp}(n^0) \subset [a, +\infty)$ , the proof is the same, but by considering only the two sets  $\mathcal{O}_2 = (a, b)$ , and  $\mathcal{O}_3 = (b, +\infty)$ , and not  $\mathcal{O}_1 = (-\infty, a)$ . We easily check using Lemma 7 that  $(\mathcal{O}_2, \mathcal{O}_3)$  satisfy the hypotheses of Lemma 7, since  $\text{supp}(n^0) \cap \mathcal{O}_1 = \emptyset$
- The case where  $n^0(a) = 0$  and the hypothesis on  $n^0$  holds is quite similar, except that Proposition 18 now shows that  $R_2$  converges to  $\max(\tilde{r}_\alpha(a), r(b))$ , with an exponential speed (if  $\tilde{r}_\alpha(a) \neq r(b)$ ). Thus, we treat the case  $r(b) > \tilde{r}_\alpha(a)$  in exactly the same way; the case  $r(b) < r(a) - (1 + \alpha)$  is a little more intricate: recalling that for all  $x \in (a, b)$ ,  $t \geq 0$ ,

$$n(t, x) = n^0(Y(t, x)) e^{\int_{Y(t, x)}^x \frac{\tilde{r}(s) - \tilde{r}_\alpha(a)}{f(s)} ds} e^{\int_0^t \tilde{r}_\alpha(a) - \rho(s) ds}$$

and

$$(Y(t, x) - a)^\alpha = (x - a)^\alpha e^{-\int_{Y(t, x)}^x \frac{\alpha}{s-a} ds},$$

one notes that

$$n(t, x) = \frac{n^0(Y(t, x))}{(Y(t, x) - a)^\alpha} e^{\int_0^t \tilde{r}_\alpha(a) - \rho(s) ds} (x - a)^\alpha e^{\int_{Y(t, x)}^x \frac{\tilde{r}(s) - \tilde{r}_\alpha(a)}{f(s)} - \frac{\alpha}{s-a} ds}$$

converges pointwise to

$$C e^{\int_0^{+\infty} \tilde{r}_\alpha(a) - \rho(s) ds} (x - a)^\alpha e^{\int_a^x \frac{\tilde{r}(s) - \tilde{r}_\alpha(a)}{f(s)} - \frac{\alpha}{s-a} ds},$$

which is well-defined, since  $\rho$  converges to  $\tilde{r}_\alpha(a)$  with an exponential speed, and  $s \mapsto \frac{\tilde{r}(s) - \tilde{r}_\alpha(s)}{f(s)} - \frac{\alpha}{s-a}$  is continuous on  $a$ , as seen in the proof of Proposition 18. Moreover,

$$x \mapsto \left\| \frac{n^0(\cdot)}{(\cdot - a)^\alpha} \right\|_\infty e^{\int_0^{+\infty} |\tilde{r}_\alpha(s) - \rho(s)| ds} e^{\int_a^x \frac{\varphi(s)}{f(s)} \mathbb{1}_{\{\varphi(s) \geq 0\}} ds},$$

with  $\varphi(s) = \tilde{r}(s) - \tilde{r}_\alpha(s) - \alpha \frac{f(s)}{s-a}$  is clearly a domination of  $n$ , and is in  $L^1$ , since  $\varphi(b) = r(b) - \tilde{r}_\alpha(a) - f'(b)$ , which implies by Lemma 8 that  $e^{\int_a^x \frac{\varphi(s)}{f(s)} ds} = O\left(\frac{r(b) - \tilde{r}_\alpha(a)}{f'(b)} - 1\right)$ , with  $\frac{r(b) - \tilde{r}_\alpha(a)}{f'(b)} > 0$ .

- This last point is the simplest, and is in fact analogous to the case of a single equilibrium point. According to Proposition 18,  $R_2$  and  $R_3$  converge to  $r(b)$ : we deduce the result by following the steps of Proposition 19.

□

**Remark.** Since Proposition 22 provides a completely explicit expression for the limit functions  $\bar{n}_\alpha$ ,  $\alpha \geq 0$ , one can easily determine their asymptotic behaviour at the boundary of the segment  $(a, b)$ . Since for all  $\alpha > 0$ ,  $x \in (a, b)$ ,

$$\bar{n}_\alpha(x) = D_\alpha (x - a)^\alpha e^{\int_a^x \frac{\tilde{r}(s) - \tilde{r}_\alpha(s)}{f(s)} - \frac{\alpha}{s-a} ds},$$

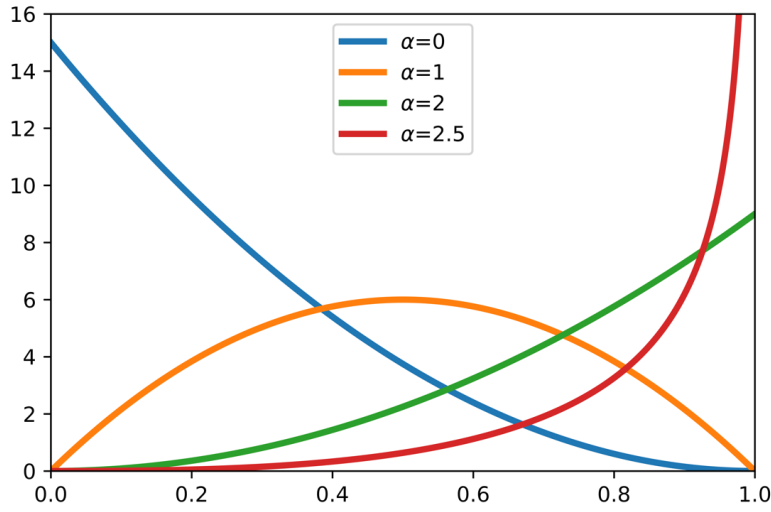


Figure 4.2: Continuous limit functions  $\bar{n}_\alpha$ , for different values of  $\alpha > 0$ , as defined in Proposition 22. In this example, we have chosen  $f(x) = x(1-x)$ , and  $r(x) = b - ax$  (with  $b = 6, a = 4$ ). With this choice, we easily compute that, for all  $\alpha \in [0, a-1)$ , and all  $x \in (0, 1)$   $\bar{n}_\alpha(x) = D_\alpha x^\alpha (1-x)^{a-\alpha-2}$ , for the appropriate constant  $D_\alpha$ . This illustrates the variety of limit functions that can be reached depending on the initial condition, as detailed in Proposition 22.

and  $s \mapsto \frac{\tilde{r}(s) - \tilde{r}_\alpha(a)}{f(s)} - \frac{\alpha}{s}$  is continuous on  $[a, b)$ , it is clear that  $\bar{n}_\alpha(x) = \Theta_{x \rightarrow a^+}((x-a)^\alpha)$ . In particular,  $\bar{n}_\alpha$  can be extended by continuity at 0, with  $\bar{n}_\alpha(a) = 0$  if  $\alpha > 0$ , and  $\bar{n}_0(a) \in (0, +\infty)$ .

Moreover, since  $\tilde{r}(b) - \tilde{r}_\alpha(a) - \frac{\alpha f(b)}{b-a} = r(b) - \tilde{r}_\alpha(a) - f(b)$ , Lemma 8 ensures that  $\bar{n}_\alpha(x) = \Theta_{x \rightarrow b^-} \left( (b-x)^{\frac{r(b) - \tilde{r}_\alpha(a)}{f'(b)} - 1} \right)$ . In particular

$$\lim_{x \rightarrow b^-} \bar{n}_\alpha(x) = \begin{cases} 0 & \text{if } \tilde{r}(b) < \tilde{r}_\alpha(a) \\ l > 0 & \text{if } \tilde{r}(b) = \tilde{r}_\alpha(a) \\ +\infty & \text{if } \tilde{r}(b) > \tilde{r}_\alpha(a) \end{cases}.$$

These different cases are illustrated by Figure 4.2.

The case where  $f$  has two roots  $a < b$  with  $f'(a) < 0$  and  $f'(b) > 0$  is symmetric to the cases here, and thus lead to the same results, by switching  $a$  and  $b$  in the Propositions.

### 4.4.3 More than two equilibria

In this subsection, we deal with the cases where  $f$  has more than two equilibria. As evidenced by the previous result, listing all possible scenarios when there are two roots already is cumbersome: this is why we will not do so in a more general case, and will focus on the case where  $n^0$  is positive on the neighbourhood of the unstable equilibrium points. The other cases can of course be treated as seen above, keeping in mind that this changes the value of the limits reached by the  $R$  functions.

**Proposition 23.** *Let us assume that  $f$  has a finite number of roots, which are all hyperbolic equilibrium points for the ODE  $\dot{u} = f(u)$ , i.e.  $f'$  has a sign at each root of  $f$ , and let us denote  $x_u^1, \dots, x_u^p$  the asymptotically unstable equilibria, and  $x_s^1, \dots, x_s^1$  the asymptotically stable one. Moreover, let us denote  $M_u := \max\{r(x_u^1) - f'(x_u^1), \dots, r(x_u^p) - f'(x_u^p)\}$ , and  $M_s := \max\{r(x_s^1), \dots, r(x_s^m)\}$ , and let us assume that these two maxima are both reached at a unique point. Lastly, let us assume that  $n^0(x_u^i) > 0$  for all  $i \in \{1, \dots, p\}$ .*

- If  $M_s > M_u$ , then  $\rho(t) \xrightarrow{t \rightarrow +\infty} M_s$ , and  $n(t, \cdot) \xrightarrow{t \rightarrow +\infty} M_s \delta_{x_s^i}$ , with  $x_s^i$  the unique stable equilibria such that  $M_s = r(x_s^i)$ .
- If  $M_s < M_u$ , then  $\rho(t) \xrightarrow{t \rightarrow +\infty} M_u$ , and  $n(t, \cdot) \xrightarrow{t \rightarrow +\infty} \bar{n}_{i^*}$  in  $L^1$ , where

$$\bar{n}_{i^*}(x) = C_{i^*} e^{\int_{x_u^{i^*}}^x \frac{\tilde{r}(s) - \tilde{r}(x_u^{i^*})}{f(s)} ds} \mathbb{1}_{I_{i^*}}(x),$$

with  $\tilde{r} = r - f'$ ,  $i^*$  the unique integer of  $\{1, \dots, p\}$  such that  $\tilde{r}(x_u^{i^*}) = M_u$ ,  $I_{i^*}$  the open interval delimited by the two stable equilibria which enclose  $x_u^{i^*}$  (or  $-\infty$  or  $+\infty$  if  $x_u^{i^*}$  is the smallest or the greatest root of  $f$ ), and  $C_{i^*}$  a positive constant such that  $\int_{I_{i^*}} \bar{n}_{i^*}(x) dx = M_u$ .

*Proof.* The proof of this proposition is in similar to that of Proposition 21: we denote  $\mathcal{O}_0, \dots, \mathcal{O}_{p+m}$ , the intervals between each roots of  $f$ , which satisfy the hypotheses of Proposition 16, according to Lemma 7, and, using Proposition 18, we are able to compute the limit of the function  $R_i$ , for all  $i \in \{0, \dots, p+m\}$ , and thus determine the long-time behaviour of  $\rho$  by Proposition 17. We conclude by using the semi-explicit expression (4.7) for  $n$ .  $\square$

This method also allows to deal with the case where  $f \equiv 0$  on a whole segment: we do not return to the case  $f \equiv 0$  on  $\mathbb{R}$ , which has already been studied in [131] and [106], but we consider the case where  $f \equiv 0$  on an interval, and then becomes positive.

To make the assumption of the following proposition possible, we assume that  $f$  is  $\mathcal{C}^2$  on  $(-\infty, a)$  and on  $(a, +\infty)$ , but not necessarily on the whole of  $\mathbb{R}$ .

**Proposition 24.** *Let us assume that there exists  $a \in \mathbb{R}$  such that  $f \equiv 0$  on  $(-\infty, a)$ ,  $f > 0$  on  $(a, +\infty)$ ,  $f'(a^+) > 0$ , and that  $\text{supp}(n^0) = [s^-, s^+]$ , with  $s^- < a < s^+$ . Then,*

- If there exists a unique  $x_M \in (s^-, a)$  such that  $r(x_M) = \max_{x \in [s^-, a]} r(x)$ , and  $f''(x_M) < 0$ , then  $\rho$  converges to  $r(x_M)$ , and  $n(t, \cdot) \xrightarrow{t \rightarrow +\infty} r(x_M) \delta_{x_M}$ .
- If  $r|_{[s^-, a]}$  reaches its maximum at  $a$  (and only at  $a$ ), then  $\rho$  converges to  $r(a)$ , and  $n(t, \cdot) \xrightarrow{t \rightarrow +\infty} r(a) \delta_a$ .

*Proof.*

1. Let us denote  $\mathcal{O}_1 := (s^-, a)$ ,  $\mathcal{O}_2 := (a, +\infty)$ , which satisfy the hypothesis of Proposition 16, according to Lemma 7. By Proposition 18,  $R_1$  converges to  $r(x_M)$  and  $R_2$  converges to  $r(x_M) - f'(x_M)$ .
2. From Proposition 17,  $\rho$  and  $\rho_1$  converge to  $r(x_M)$ , and  $\rho_2$  converges to 0.
3. Let  $K \subset [s^-, a]$  be a compact set that does not contain  $x_M$ . Thanks to the semi-explicit expression (4.7), and using the fact that  $f'(x) = 0$  and  $Y(t, x) = x$  for all  $x \in K$  and all  $t \geq 0$ ,

$$n(t, x) = n^0(x) e^{\int_0^t r(x) - \rho(s) ds} \leq n^0(x) e^{\int_0^t r_K - \rho(s) ds},$$

with  $r_K := \max_{x \in K} r(x) < r(x_M)$ .

Thus,

$$\int_K n(t, x) dx \leq \int_K n^0(x) dx e^{\int_0^t r_M - \rho(s) ds},$$

which converges to 0, since  $r_M - \rho(s)$  is negative for any  $s$  large enough. This proves the result thanks to Proposition 5.

- 1. Here we have to make a slightly more subtle choice of subsets than usual. Let  $\varepsilon > 0$ , and let us denote  $\mathcal{O}_1^\varepsilon := (s^-, a - 2\varepsilon)$ ,  $\mathcal{O}_2^\varepsilon := (a - 2\varepsilon, a - \varepsilon)$ ,  $\mathcal{O}_3^\varepsilon := (a - \varepsilon, a)$ ,  $\mathcal{O}_4 := (a, +\infty)$ . We easily check that these four sets satisfy the hypotheses of Proposition 16. Moreover, since  $f \equiv 0$  on  $[s^-, a]$  for all  $i \in \{1, 2, 3\}$ ,

$$R_i^\varepsilon(t) = \frac{\int_{\mathcal{O}_i^\varepsilon} r(x) e^{r(x)t} dx}{\int_{\mathcal{O}_i^\varepsilon} e^{r(x)t} dx}.$$

Thus, for all  $t \geq 0$ ,  $i \in \{1, 2, 3\}$ ,

$$\min_{x \in \mathcal{O}_i^\varepsilon} r(x) \leq R_i^\varepsilon(t) \leq \max_{x \in \mathcal{O}_i^\varepsilon} r(x).$$

In particular,

$$\bar{R}_1^\varepsilon \leq \max_{x \in [s^-, a - 2\varepsilon]} r(x) \quad \text{and} \quad \underline{R}_3^\varepsilon \geq \min_{x \in [a - \varepsilon, a]} r(x).$$

Finally, Proposition 18 shows that  $R_4$  converges to  $r(a) - f'(a^+)$ .

- 2. Since  $r$  reaches its unique maximum at  $a$ , for any  $\varepsilon$  small enough, we get

$$\underline{R}_3^\varepsilon > \bar{R}_1^\varepsilon \quad \text{and} \quad \underline{R}_3^\varepsilon > \lim_{t \rightarrow +\infty} R_4(t).$$

Thus, according to Proposition 17,  $\rho_1^\varepsilon$  and  $\rho_4$  converge to 0, for all  $\varepsilon > 0$ . The choice of  $\varepsilon$  being arbitrary, it also proves that  $\rho_2^\varepsilon$  converges to 0. Thus,  $\bar{\rho} = \bar{\rho}_3^\varepsilon$ , and  $\underline{\rho} = \underline{\rho}_3^\varepsilon$ , for all  $\varepsilon > 0$ . Since for all  $t \geq 0$

$$\min_{x \in [a - \varepsilon, a]} r(x) \leq R_3^\varepsilon(t) \leq r(a),$$

we prove that  $\rho$  converges to  $r(a)$  by making  $\varepsilon$  tend to 0.

- 3. We have proved that  $t \mapsto \int_{s^-}^a n(t, x) dx$  converges to  $r(a)$ , that  $t \mapsto \int_a^{+\infty} n(t, x) dx$  converges to 0 and that for all  $\varepsilon > 0$ ,  $\int_{s^-}^{a - \varepsilon} n(t, x) dx$  converges to 0. The hypotheses of Proposition 5 are therefore met, which concludes the proof.  $\square$

Note that the methods of Propositions 23 and 24 can be coupled to treat more complex cases, where, for example,  $f \equiv 0$  on several disjoint segments.

## 4.5 Some results in higher dimensions

As seen in the previous sections, our entire method is based on the computation of the limit of the  $R_i$  functions defined in Section 4.3. Unfortunately, these computations seem out of reach in the multidimensional case  $\mathbb{R}^d$ ,  $d \geq 2$ .

In this section, we nevertheless address the question of the possible convergence to smooth or singular measures in higher dimensions in some specific simple cases. We first



analyse how the solution support evolves over time. This allows us to conclude that that the solution converges to a Dirac mass in the case of a unique equilibrium which is asymptotically stable for the ODE  $\dot{u} = f(u)$ , and provide hypotheses under which the solution cannot converge to a smooth function. We then characterise which stationary measures may or may not be limits for solutions of (4.1), before providing a criterion ensuring the existence of continuous stationary solutions.

### 4.5.1 Limit support

**Definition 4** Limit support. We define the **limit support** of  $n$  as:

$$\sigma_\infty = \bigcap_{t \geq 0} \overline{\bigcup_{s \geq t} \text{supp}(n(s, \cdot))}.$$

Recalling the semi-explicit expression (4.7),

$$n(t, x) = n^0(Y(t, x)) e^{\int_0^t (r - \nabla \cdot f)(Y(s, x)) - \rho(s) ds},$$

and that for all  $t \geq 0$ ,  $\text{supp}(n^0(Y(t, \cdot))) = X(t, \text{supp}(n^0))$ , we get

$$\sigma_\infty = \bigcap_{t \geq 0} \overline{\bigcup_{s \geq t} \text{supp}(n^0(Y(s, \cdot)))} = \bigcap_{t \geq 0} \overline{\bigcup_{s \geq t} X(s, \text{supp}(n^0))}. \quad (4.31)$$

In the cases where we are able to determine the latter set, we gather information about possible limits for  $n$ .

**Lemma 9.** *If the limit support of  $n$  is of measure zero, then  $n$  does not converge (weakly) to a non-zero function in  $L^1(\mathbb{R}^d)$ .*

*Proof.* Let us argue by contradiction. By denoting  $\nu$  the Lebesgue measure, let us assume that  $\nu(\sigma_\infty) = 0$ , and that  $n$  converges weakly to  $\bar{n} \in L^1(\mathbb{R}^d)$ ,  $\bar{n} \neq 0$ . Since  $\limsup_{t \rightarrow +\infty} \text{supp}(n(t, \cdot)) = \bigcap_{t \geq 0} \overline{\bigcup_{s \geq t} \text{supp}(n(s, \cdot))} \subset \sigma_\infty$ ,

$$\limsup_{t \rightarrow +\infty} \nu(\text{supp}(n(t, \cdot))) \leq \nu\left(\limsup_{t \rightarrow +\infty} \text{supp}(n(t, \cdot))\right) \leq \nu(\sigma_\infty) = 0,$$

which contradicts the initial hypothesis.  $\square$

**Proposition 25.** *Let us assume that  $f$  has a unique root, denoted  $\bar{x}$ , which is **globally asymptotically stable** for the ODE  $\dot{u} = f(u)$  over  $\mathbb{R}^d$ , and that the set  $\bigcup_{t \geq 0} X(t, \text{supp}(n^0))$  is bounded. Then,  $n(t, \cdot) \xrightarrow{t \rightarrow +\infty} r(\bar{x})\delta_{\bar{x}}$ , and  $\rho$  converges to  $r(\bar{x})$ .*

*Proof.* Since the support of  $n^0$  is compact, and  $\bar{x}$  is globally asymptotically stable, we easily check, according to (4.31), that  $\sigma_\infty = \{\bar{x}\}$ . By Lemma 5, it is hence enough to prove that  $\rho$  converges to  $r(\bar{x})$ . As seen in the proof of Lemma 4.3.2,  $\rho$  satisfies, for all  $t \geq 0$ ,

$$\dot{\rho}(t) = \int_{\mathbb{R}^d} (r(x) - \rho(t))n(t, x)dx = \int_{\text{supp}(n(t, \cdot))} (r(x) - \rho(t))n(t, x)dx.$$

Let  $\varepsilon > 0$ . Since  $\sigma_\infty = \{\bar{x}\}$  is the intersection of compact decreasing sets, there exists  $T_\varepsilon > 0$  such that, for all  $t \geq T_\varepsilon$ ,  $\text{supp}(n(t, \cdot)) \subset B(\bar{x}, \varepsilon)$ . Thus, by denoting

$$r_m^\varepsilon := \min_{x \in B(\bar{x}, \varepsilon)} r(x) \quad \text{and} \quad r_M^\varepsilon := \max_{x \in B(\bar{x}, \varepsilon)} r(x),$$

we get, for all  $t \geq T_\varepsilon$ ,

$$(r_m^\varepsilon - \rho(t))\rho(t) \leq \dot{\rho}(t) \leq (r_M^\varepsilon - \rho(t))\rho(t),$$

which ensures that

$$\liminf_{t \rightarrow +\infty} \rho(t) \geq r_m^\varepsilon \quad \text{and} \quad \limsup_{t \rightarrow +\infty} \rho(t) \leq r_M^\varepsilon.$$

Since these inequalities hold for any  $\varepsilon > 0$ , and  $r_m^\varepsilon$  and  $r_M^\varepsilon$  both converge to  $r(\bar{x})$  when  $\varepsilon$  goes to 0, it concludes the proof.  $\square$

Because of the diversity of possible behaviours of ODE systems, it is difficult to compute the limit support for a given  $f$ , unless very strong assumptions are made about the ODE  $\dot{u} = f(u)$ . This is what we do in the following proposition, motivated by a family of ODE systems commonly used in systems biology.

We say that the two-dimensional system

$$\begin{cases} \dot{x}_1 = f_1(x_1, x_2) \\ \dot{x}_2 = f_2(x_1, x_2) \end{cases} \quad (4.32)$$

is **competitive** if  $\partial_2 f_1 \leq 0$  and  $\partial_1 f_2 \leq 0$ , and **cooperative** if  $\partial_2 f_1 \geq 0$  and  $\partial_1 f_2 \geq 0$ . For instance, such systems are commonly used to model the interactions between two proteins in the context of cell differentiation [60, 67, 91, 150], and are known to have an interesting property: trajectories either go to  $+\infty$ , or converge [77], *i.e.* for all  $x \in \mathbb{R}^d$ ,

$$\|Y(t, x)\| \xrightarrow[t \rightarrow +\infty]{} +\infty \quad \Rightarrow \quad t \mapsto Y(t, x) \text{ converges.} \quad (4.33)$$

Note that if the ODE (4.32) is competitive (or cooperative), then the reverse ODE  $\dot{u} = -f(u)$  is cooperative (or competitive). This motivates the hypothesis of the following proposition. Before giving its statement, we recall that if  $\bar{x}$  is a root of  $f$ ,  $\bar{x}$  is called a **hyperbolic equilibrium** if all the eigenvalues of  $\text{Jac } f(\bar{x})$  have a non-zero real part, and is called a **repellor** if all these eigenvalues have a positive real part. Lastly, we recall that the **unstable set** of  $\bar{x}$  is defined by  $\{x \in \mathbb{R}^d : Y(t, x) \xrightarrow[t \rightarrow +\infty]{} \bar{x}\}$ .

**Proposition 26.** *Let us assume that  $f$  has a finite number of roots, and is such that identity (4.33) holds. Then, the limit support of  $n$  is included in the closure of the union of the unstable sets of the roots of  $f$ , *i.e.* by denoting  $\bar{x}_1, \dots, \bar{x}_N$  the roots of  $f$ ,*

$$\sigma_\infty \subset \overline{\bigcup_{1 \leq i \leq N} \left\{ x \in \mathbb{R}^d : Y(t, x) \xrightarrow[t \rightarrow +\infty]{} \bar{x}_i \right\}}.$$

Moreover, if all the roots of  $f$  are hyperbolic points, and if none of them is a repellor, then the limit support of  $n$  is of measure 0. In particular,  $n$  does not converge (weakly) to a function in  $L^1$ .

*Proof.* The inclusion is clear: by hypothesis for all  $x \in \mathbb{R}^d$  such that  $t \mapsto Y(t, x)$  does not converge,  $t \mapsto \|Y(t, x)\|$  goes to  $+\infty$ , and since the support of  $n^0$  is bounded, the points of the limit support are necessary in the unstable set of one of the equilibria. The second part of the proposition is a consequence of the stable manifold theorem [130], which ensures that the unstable set of an equilibrium which is not a repellor is a smooth manifold of dimension at most  $d - 1$ , hence a set of measure zero. We conclude with Lemma 9.  $\square$

### 4.5.2 Stationary solutions

In this subsection, we define the stationary solution in the weak sense, which allows to include measures. As seen in the previous section, under appropriate hypotheses on  $f$ , the presence of a repeller is necessary to hope for solutions which converge to smooth functions. In this section, we prove that, under appropriate hypotheses, the presence of a repeller ensures the existence of smooth stationary solutions.

**Definition 5** Weak stationary solution. Let  $\mu$  be a finite positive Radon measure. We say that  $\mu$  is a **weak stationary solution** of equation (4.1) if it satisfies

$$\forall \varphi \in C_c^1(\mathbb{R}^d), \quad \int_{\mathbb{R}^d} \left( f(x) \cdot \nabla \varphi(x) + (r(x) - \mu(\mathbb{R}^d)) \varphi(x) \right) d\mu(x) = 0. \quad (4.34)$$

**Remark.** If  $\bar{x}$  is a root of  $f$ , let us note that  $r(\bar{x})\delta_{\bar{x}}$  is a weak stationary solution of (4.1).

The following proposition shows, as we might expect, that convergent solutions of (4.1) (in the weak sense) necessarily converge to a weak stationary solution.

**Proposition 27.** *Let us assume that  $r \in C_0(\mathbb{R}^d)$ , and let  $n(t, \cdot)$  be a solution of (4.1) which converges in the weak sense in the space of Radon measure. Then its limit is a weak stationary solution of equation (4.1).*

*Proof.* We let  $\mu$  be the limit of  $n(t, \cdot)$ . Let us first prove that, under these conditions,  $\rho(t) = \int_{\mathbb{R}^d} n(t, x) dx$  converges when  $t$  goes to  $+\infty$ .

Let us denote  $\psi(t) := \int_{\mathbb{R}^d} r(x) n(t, x) dx$ , which is non-negative, according to the non-negativity of  $r$  and  $n$ , and converges to  $\bar{\psi} := \int_{\mathbb{R}^d} r(x) d\mu(x)$ , by definition of the weak convergence, and since  $r \in C_0(\mathbb{R}^d)$ . Let us assume that  $\bar{\psi} > 0$ . Let  $\varepsilon \in (0, \bar{\psi})$ . Since  $\psi$  converges to  $\bar{\psi}$ , and since  $\rho$  satisfies the ODE

$$\dot{\rho}(t) = \psi(t) - \rho(t)^2,$$

there exists  $T_\varepsilon > 0$  such that for all  $t \geq T_\varepsilon$ ,

$$\bar{\psi} - \varepsilon - \rho(t)^2 \leq \dot{\rho}(t) \leq \bar{\psi} + \varepsilon - \rho(t)^2.$$

In other words,  $\rho$  is a super-solution of  $\dot{u} = \bar{\psi} - \varepsilon - u^2$ , and a sub-solution of  $\dot{u} = \bar{\psi} + \varepsilon - u^2$ . Since the solutions of these equations converge to  $\sqrt{\bar{\psi} - \varepsilon}$  and  $\sqrt{\bar{\psi} + \varepsilon}$  respectively,

$$\liminf_{t \rightarrow +\infty} \rho(t) \geq \sqrt{\bar{\psi} - \varepsilon} \quad \text{and} \quad \limsup_{t \rightarrow +\infty} \rho(t) \leq \sqrt{\bar{\psi} + \varepsilon}.$$

Since these inequalities hold for any  $\varepsilon \in (0, \bar{\psi})$ , it proves that  $\rho$  indeed converges to  $\sqrt{\bar{\psi}}$ .

If  $\bar{\psi} = 0$ , we prove that  $\limsup \rho \leq 0$  with the same method, and the non-negativity of  $\psi$  ensures that  $\liminf \rho \geq 0$ .

Let  $\varphi \in C_c^1(\mathbb{R}^d)$ , and let us denote  $\bar{\rho} := \lim_{t \rightarrow +\infty} \rho(t)$ . We recall that if a differentiable function converges, then its derivative is either divergent or converges to 0. Hence, since  $t \mapsto \int_{\mathbb{R}^d} \varphi(x) n(t, x) dx$  converges (by hypothesis), and

$$\begin{aligned} \frac{d}{dt} \int_{\mathbb{R}^d} \varphi(x) n(t, x) dx &= - \int_{\mathbb{R}^d} \nabla \cdot (f(x) n(t, x)) \varphi(x) dx - \int_{\mathbb{R}^d} (r(x) - \rho(t)) \varphi(x) n(t, x) dx \\ &= + \int_{\mathbb{R}^d} f(x) \cdot \nabla \varphi(x) n(t, x) dx + \int_{\mathbb{R}^d} (r(x) - \rho(t)) \varphi(x) n(t, x) dx \\ &\xrightarrow{t \rightarrow +\infty} \int_{\mathbb{R}^d} f(x) \cdot \nabla \varphi(x) + (r(x) - \bar{\rho}) \varphi(x) d\mu(x), \end{aligned}$$

the equality

$$\int_{\mathbb{R}^d} f(x) \cdot \nabla \varphi(x) + (r(x) - \bar{\rho}) \varphi(x) d\mu(x) = 0 \quad (4.35)$$

holds for any  $\varphi \in \mathcal{C}_c^1(\mathbb{R}^d)$ .

It remains to prove that  $\mu(\mathbb{R}^d) = \bar{\rho}$ . If  $\bar{\rho} = 0$ , then the non-negativity of  $n$  and the definition of  $\rho$  lead to  $\mu = 0$ . Let us now assume that  $\bar{\rho} > 0$ , and let  $\varepsilon > 0$ . Since  $\mu$  is a finite measure,  $r \in \mathcal{C}^0(\mathbb{R}^d)$ , and owing to the definition of  $\psi$  and  $\bar{\rho}$ , there exists  $K \subset \mathbb{R}^d$  a compact set such that

- $\mu(K) \geq \mu(\mathbb{R}^d) - \varepsilon$
- $\int_K r(x) d\mu(x) \geq \int_{\mathbb{R}^d} r(x) d\mu(x) - \varepsilon = \bar{\rho}^2 - \varepsilon$ .

Let  $\varphi_K \in \mathcal{C}_c^1(\mathbb{R}^d)$  such that  $\varphi_K \equiv 1$  on  $K$ ,  $0 \leq \varphi \leq 1$  on  $\mathbb{R}^d$ . Since  $\nabla \varphi_K \equiv 0$  on  $K$ ,

$$\left| \int_{\mathbb{R}^d} f(x) \cdot \nabla \varphi_K(x) d\mu(x) \right| \leq \|f \cdot \nabla \varphi_K\|_{\infty} \mu(\mathbb{R}^d \setminus K) \leq \varepsilon \|f \cdot \nabla \varphi_K\|_{\infty}.$$

Moreover, according to the choice of  $\varphi_K$

$$\int_{\mathbb{R}^d} r(x) \varphi_K(x) d\mu(x) \in [\bar{\rho}^2 - \varepsilon, \bar{\rho}^2], \quad \text{and} \quad \int_{\mathbb{R}^d} \varphi_K(x) d\mu(x) \in [\mu(\mathbb{R}^d) - \varepsilon, \mu(\mathbb{R}^d)].$$

Hence, injecting these inequalities in (4.35), we obtain

$$-C\varepsilon \leq \bar{\rho}(\bar{\rho} - \mu(\mathbb{R}^d)) \leq C\varepsilon$$

for some  $C \geq 0$ . Since this equality holds for any  $\varepsilon$ , and  $\bar{\rho}$  is positive, it proves that  $\mu(\mathbb{R}^d) = \bar{\rho}$ .  $\square$

Weak stationary solutions which are smooth enough (at least in  $\mathcal{C}^1(\mathbb{R}^d)$ ) are in fact stationary solutions in the strong sense, as defined in the following lemma, and can be further characterised.

**Lemma 10.** *Let  $\bar{n} \in \mathcal{C}^1(\mathbb{R}^d)$ . Then,  $\bar{n}$  is a weak stationary solution of (4.1) if and only if for all  $t \geq 0$ ,  $y \in \mathbb{R}^d$ ,*

$$\begin{cases} \bar{n}(X(t, y)) = e^{\int_0^t \tilde{r}(X(s, y)) - \bar{\rho} ds} \bar{n}(y) \\ \int_{\mathbb{R}^d} \bar{n}(x) dx = \bar{\rho} \end{cases}.$$

*Proof.* First, let us note that, since  $\bar{n} \in \mathcal{C}^1(\mathbb{R}^d)$ , one can integrate by parts in the expression (4.34) in order to prove that  $\bar{n}$  is a weak stationary solution if and only if for any  $\varphi \in \mathcal{C}_c^1(\mathbb{R}^d)$ ,

$$\int_{\mathbb{R}^d} (-\nabla \cdot (f(x) \bar{n}(x)) + (r(x) - \bar{\rho}) \bar{n}(x)) \varphi(x) dx = 0,$$

with  $\bar{\rho} = \int_{\mathbb{R}^d} \bar{n}(x) dx$ , which means that  $\bar{n}$  is a weak stationary solution if and only if it is a stationary solution in the strong sense, *i.e.*

$$-\nabla \cdot (f(x) \bar{n}(x)) + (r(x) - \bar{\rho}) \bar{n}(x) = 0 \quad \text{for all } x \in \mathbb{R}^d.$$

The result follows, since for any  $y \in \mathbb{R}^d$

$$\begin{aligned} & \frac{d}{dt} \left( \bar{n}(X(t, y)) e^{-\int_0^t \tilde{r}(X(s, y)) - \bar{\rho} ds} \right) \\ &= \left( f(X(t, y)) \cdot \nabla \bar{n}(X(t, y)) - (\tilde{r}(X(t, y)) - \bar{\rho}) \bar{n}(X(t, y)) \right) e^{-\int_0^t \tilde{r}(X(s, y)) - \bar{\rho} ds} \\ &= - \left( -\nabla \cdot (f(X(t, y)) \bar{n}(X(t, y))) + (r(X(t, y)) - \bar{\rho}) \bar{n}(X(t, y)) \right) e^{-\int_0^t \tilde{r}(X(s, y)) - \bar{\rho} ds} = 0. \end{aligned}$$

□

Lemma 10 allows us to conclude that in the case where the ODE  $\dot{u} = f(u)$  has a repeller with a bounded unstable set, there exists a smooth stationary solution for (4.1).

**Corollary 28.** *Let  $x_u \in \mathbb{R}^d$  be a repeller point for the ODE  $\dot{x} = f(x)$ , and let us assume that*

$$\bar{n}(x) := \frac{\tilde{r}(x_u)}{\alpha} e^{\int_0^{+\infty} \tilde{r}(Y(s, x)) - \tilde{r}(x_u) ds} \mathbb{1}_B(x)$$

*is well-defined, and that  $\bar{n} \in \mathcal{C}^1(B) \cap L^1(B)$ , where  $B = \{x \in \mathbb{R}^d : Y(t, x) \xrightarrow[t \rightarrow +\infty]{} x_u\}$  is the unstable set of  $x_u$ . Then,  $\bar{n}$  is a  $\mathcal{C}^1$  stationary solution.*

*Proof.* For all  $y \in \mathbb{R}$ ,  $t \geq 0$ ,

$$\bar{n}(X(t, y)) = \frac{\tilde{r}(x_u)}{\alpha} e^{\int_0^{+\infty} \tilde{r}(X(t-s, x)) - \tilde{r}(x_u) ds} = \frac{\tilde{r}(x_u)}{\alpha} e^{\int_{-\infty}^t \tilde{r}(X(s, y)) - \tilde{r}(x_u) ds},$$

with the change of variable  $s' = t - s$  and

$$\bar{n}(y) = \frac{\tilde{r}(x_u)}{\alpha} e^{\int_{-\infty}^0 \tilde{r}(X(s, y)) - \tilde{r}(x_u) ds},$$

with the change of variable  $s' = -s$ . Thus, the equality of Lemma 10 holds, which concludes the proof. □





## Chapter 5

# A particle method for an advection-selection-mutation equation

We study a non-local advection-selection-mutation problem deriving from adaptive dynamics models. We prove well-posedness for a large family of initial data. We then develop a particle method to approximate the solution of such problem by a regularised sum of weighted Dirac masses whose characteristics solve a suitably defined ODE system. We establish the convergence of the particle method over any finite time interval, providing an explicit rate of convergence. Finally, we investigate the asymptotic-preserving properties of the method in large times, providing sufficient conditions for it to hold true as well as examples and counter-examples. Finally, we illustrate the method in two cases taken from the literature.

This work has been done with Frank Ernesto Alvarez and has been submitted for publication.



## 5.1 Introduction

### Presentation of the model

The goal of this chapter is to develop a numerical method allowing to approximate the solutions of equations of the form

$$\begin{cases} \partial_t v(t, x) + \nabla_x \cdot (a(t, x, I_a v(t, x))v(t, x)) = R(t, x, I_g v(t, x))v(t, x) + \int_{\mathbb{R}^d} m(t, x, y, I_d v(t, x))v(t, y)dy, \\ v \in \mathcal{C}([0, T], L^1(\mathbb{R}^d)), \\ v(0, \cdot) = v^0(\cdot) \in W^{1,1}(\mathbb{R}^d), \end{cases} \quad (5.1)$$

where

$$(I_l u)(t, x) = \int_{\mathbb{R}^d} \psi_l(t, x, y)u(t, y)dy, \quad l = a, g, d$$

are non-local terms and  $a$ ,  $R$ ,  $m$  and  $\psi_l$  are smooth functions.

This general formulation aims to bring together a wide family of PDE models typically used in the field of *adaptive dynamics*. In this context,  $x$  represents a phenotypic trait (usually simply called ‘phenotype’ or ‘trait’) which is a characteristic inherent to individuals, and  $x \mapsto v(t, x)$  represents the density of the studied population at time  $t \geq 0$ . One purpose of adaptive dynamics is to understand the combined effect of selection and mutations (which are usually assumed to be rare and small [49, 62, 117]) on living populations [18]. The literature concerning phenotype-structured equations is abundant [10, 42, 50, 86, 107, 131, 132]. The model proposed in this chapter (which includes, among others, the equations studied in [13, 21, 48, 59, 106]) takes into account

- Selection and growth, via the term ‘ $R(t, x, I_g v(t, x))v(t, x)$ ’, where  $R$  can be interpreted as the instantaneous growth rate, which depends on the trait  $x$  and the whole population.
- Mutation, via the term ‘ $\int_{\mathbb{R}^d} m(t, x, y, I_d v(t, x))v(t, y)dy$ ’, where the function  $m$  can be seen as the probability density for a cell of trait  $y$  to mutate into a cell of trait  $x$ .
- Advection, via the term ‘ $\nabla_x \cdot (a(t, x, I_a v(t, x))v(t, x))$ ’. This term models how the environment drives the individuals towards specific regions, as opposed to random mutations. Among others, this term can be used in order to model a cell differentiation phenomenon.

Mutations can also be modelled through a second order differential operator such as in [32] and [4]. Laplacian-like terms can be approximated by integral operators, as shown in [46], which means that, after choosing an appropriate integral approximation, our analysis could be extended to deal with second order equations. The other two non local-terms ( $I_a$  and  $I_g$ ) allow to take into account the influence of the environment, created by the whole population, over the behaviour of the individuals [59, 164], or competition between individuals [131].

The long time behaviour of models considering only one phenomenon among selection, advection and mutation is well-known: broadly speaking, it has been shown that considering selection alone, or advection alone, leads to concentration phenomena (towards a finite number of traits) [51, 106, 131], meaning that the density converges to a sum of Dirac masses, while mutations by themselves have a smoothing effect [69]. Nevertheless, the combined effects of these terms remains unclear, and may lead to different and

non-intuitive behaviours. As an example, considering both selection and advection can lead to convergence either to a Dirac mass or to a continuous function (see Chapter 4, cooney2022long), and considering mutation and selection leads either to convergence to a non-smooth measure or to a continuous function [13]. Note that this model also includes the equation studied in [9], which was also approximated with a particle method.

Upon establishing the well-posedness of (5.1), this chapter is concerned with the derivation of a particle method inspired by [46], the analysis of its convergence and asymptotic-preserving properties. However, we must emphasise the two main novelties with respect to that work: First, the use of non local terms, which as we will show, poses technical difficulties and affects the existence of smooth solutions in certain cases. Secondly, the study of the asymptotic preserving property, which guarantees that, under certain hypotheses, the long time behaviour of the solution is conserved. As will be seen, these equations are naturally posed in the space of Radon measures, making particle methods a natural tool to approximate them. Compared to finite volume or finite element methods, they are more easily implemented. Furthermore, a change in model leads to very few changes in the corresponding code, a clear advantage over other methods.

## Particle method

Particle methods use ODE resolution in order to approximate the solution of PDEs. This makes them particularly easy to implement, as they only require a classical ODE solver. The main idea is to seek a sum of weighted Dirac masses, called particle solution, which is denoted

$$v^N(t) = \sum_{i=1}^N \alpha_i(t) \delta_{x_i(t)}, \quad (5.2)$$

where the weights  $\alpha_i$  and the points  $x_i$  are solutions of a suitable ODE system.

In order to recover a smooth function close to the solution of the studied PDE, the particle solution needs to be regularised: this is usually done by means of a convolution with a so-called ‘cut-off function’  $\varphi_\varepsilon$  which must satisfy some specific properties. We denote this regular solution

$$v_\varepsilon^N(t, x) = \sum_{i=1}^N \alpha_i(t) \varphi_\varepsilon(x - x_i(t)), \quad (5.3)$$

where the scaling parameter  $\varepsilon$  is a function of  $N$ .

This method is especially adapted for the linear advection equation ‘ $\partial_t v(t, x) + \nabla \cdot (a(x)v(t, x)) = 0$ ’, but has been generalised to many other kinds of equations which mostly come from physics [81], such as diffusion equations [22, 47, 100, 101, 140], advection-diffusion equations [74, 111], convection-diffusion equations [46], the Navier-Stokes equation [33, 41] or the Vlasov-Poisson equation [40, 136].

We apply the particle method by following its main three steps, as described in [27]:

- (i) **Particle approximation of the initial data.** This first step consists of approaching the initial condition of  $v^0$  with a sum of weighted Dirac masses, *i.e.* choosing  $N \in \mathbb{N}$ ,  $x_1^0, \dots, x_N^0 \in \mathbb{R}^d$ ,  $\alpha_1^0, \dots, \alpha_N^0 \in \mathbb{R}$  such that

$$v_0^N := \sum_{i=1}^N \alpha_i^0 \delta_{x_i^0} \sim v^0,$$

in the sense of Radon measures, which means that, for any  $\phi \in \mathcal{C}_c^0(\mathbb{R}^d)$ ,

$$\sum_{i=1}^N \alpha_i^0 \phi(x_i^0) \xrightarrow{N \rightarrow +\infty} \int_{\mathbb{R}^d} \phi(x) v^0(x) dx.$$

Assuming that  $v^0$  has a compact support, a canonical way of choosing these values is to choose a finite collection of subsets  $\Omega_i^0 \subset \text{supp}(v^0)$  satisfying

$$\Omega_i^0 \cap \Omega_j^0 = \emptyset, \text{ if } i \neq j, \text{ and } \bigcup_{i \in \{1, \dots, N\}} \Omega_i^0 = \text{supp}(v^0),$$

and to take, for any  $i \in \{1, \dots, N\}$

$$x_i^0 \in \Omega_i^0, \quad w_i^0 = |\Omega_i^0|, \quad \nu_i^0 = v^0(x_i^0) \quad \alpha_i = \nu_i w_i.$$

- (ii) **Time-evolution of the particles.** By using a weak formulation of the PDE, we determine the ODE satisfied by the positions (denoted  $x_i$ ), the volumes (denoted  $w_i$ ) and the weights (denoted  $\nu_i$ ) associated to each of the  $N$  particles, with initial conditions  $(x_i^0, w_i^0, \nu_i^0)$  given at the previous step. The exact ODE, and the way the particles are correlated with each other depends on the complexity of the PDE. In the case where the advection term is local ( $a(t, x, I) = a(t, x)$ ), the positions and the volumes satisfy

$$\dot{x}_i(t) = a(t, x_i(t)), \quad \dot{w}_i(t) = \nabla_x \cdot (a(t, x_i(t)))w_i(t) \quad (5.4)$$

and the formula for the  $\nu_i$  depends on the selection and the mutation terms. The method described in the core of this chapter also allows to use this method in the case of a non-local advection, which modifies the ODE satisfied by the positions of the particles. Full formulas are given in Section 5.3.

Rewriting  $\alpha_i = \nu_i w_i$ , with the volumes  $w_i$  which satisfy (5.4) is required for the approximation of the different integral terms. Indeed, by Liouville's formula [73], for any  $f$  smooth enough,

$$\int_{\mathbb{R}^d} f(x) dx \sim \sum_{i=1}^N f(x_i(t))w_i(t).$$

Formally,  $v$  and  $v^N$  (as defined by (5.2)) are both solution of (5.1) in the weak sense, with  $v^N(0) \sim v^0$ , which implies that, assuming that the parameters of the PDE are smooth enough, for any given time  $T > 0$ ,  $v^N(T) \sim v(T, \cdot)$ .

- (iii) **Regularisation.** In order to transform the discrete measure  $v^N$  into a smooth function, we use a regularisation process based on convolution, which writes as a sum, shown in (5.3), since the convoluted measure is a sum of Dirac masses. The function

$$\varphi_\varepsilon := \frac{1}{\varepsilon^d} \varphi\left(\frac{\cdot}{\varepsilon}\right),$$

depending on a parameter  $\varepsilon > 0$ , is a scaling of the so-called *cut-off function*  $\varphi$ , which must satisfy some regularity and symmetry properties (which we specify in Section 5.4.1). The choice of  $\varepsilon$ , which intimately depends on the choice of  $N$ , is intricate: if  $\varepsilon$  is too large, then the solution is 'over-regularised', and the scheme loses its accuracy. Conversely, if  $\varepsilon$  is too small, then some of the particles will be neglected, and the scheme does not converge towards the solution. Choosing the optimal  $\varepsilon$  as a function of  $N$  is thus not a trivial question, and it is possible in some cases to optimise the convergence rate by improving this regularisation step [38, 83].

## Main results

**Well-posedness.** We first prove that problem (5.1) is well-posed, *i.e.* that for any family of parameters satisfying some regularity properties, defined in  $\mathcal{C}([0, +\infty), L^1(\mathbb{R}^d))$ . The proof heavily relies on the use of the characteristic curves  $X_u(t, y)$ , solution to the equation

$$\begin{cases} \dot{X}_u(t, y) = a(t, X_u(t, y), (I_a u)(t, X_u(t, y))) =: \mathcal{A}_u(t, X_u(t, y)), & t \in [0, T], \\ X_u(0, y) = y, \end{cases}$$

for  $y \in \mathbb{R}^d$  and  $u \in \mathcal{C}([0, T], L^1(\mathbb{R}^d))$ . The main difference with respect to the approach taken for similar problems, as in [46], is the need for continuity results for  $X_u(t, y)$ , not only with respect to the trait variable  $y$ , but with respect to  $u$  as well. The required results are stated in Section 5.2.1 and proved in Appendix C.1.

The well posedness of the problem for smooth initial data is proved in Section 5.2.2 using a standard fixed point argument. Moreover, in Section 5.2.3, we consider a more general family of initial conditions and we prove that the regularity of  $v(t, \cdot)$  is linked to that of  $v^0$ : more precisely, if  $v^0 \in W^{k, \infty}(\mathbb{R}^d)$  with compact support, then  $v \in \mathcal{C}([0, +\infty), W^{k-1, 1}(\mathbb{R}^d))$ . This result can be improved if the advection is local, *i.e.*  $a(t, x, I) = a(t, x)$ : In this case, for any initial condition in  $W^{k, 1}(\mathbb{R}^d)$ , the solution  $v$  is in  $\mathcal{C}([0, +\infty), W^{k, 1}(\mathbb{R}^d))$ .

**Particle method: definition and well-posedness.** In Section 5.3, we define the particle method corresponding to this PDE, by deriving the ODE system satisfied by the particles  $(x_i, w_i$  and  $\nu_i)$ . For the non-local case, the equation we obtain is a coupled system with infinitely many equations and unknowns. We generalise some classical results from the Cauchy-Lipschitz theory in order to deal with this problem, and prove that the ODE system is well posed.

**Convergence of the particle method.** Section 5.4 is structured as follows: in Section 5.4.1 we prove the following estimate, detailed in Theorem 14:

$$\|v - v_\varepsilon^h\|_{L^1(\mathbb{R}^d)} \leq C(\varepsilon^r + \left(\frac{h}{\varepsilon}\right)^\kappa + h^\kappa) \|v^0\|_{W^{\mu, 1}(\mathbb{R}^d)}, \quad \text{for all } 0 \leq t \leq T.$$

Nevertheless, in general this scheme is not asymptotic preserving. Therefore, in subsection 5.4.2, we show examples for which the scheme is asymptotic preserving, and others for which it is not. In general, the asymptotic behaviour of a solution is preserved when it converges to a sum of Dirac masses, and is not when it converges to a smooth solution.

## Perspectives and open problems

Although a loss of regularity appears to be taking place when advection is non local, we do not know whether such a loss of regularity does happen in certain cases. The construction of such an example or, on the contrary, the improvement of our results in order to prove that, in fact, no regularity is lost could be a first way to extend our work.

Another open problem is the optimisation of the order of convergence for the numerical solution, improving upon the order  $\frac{\kappa r}{\kappa + r}$  obtained in the present work. The approach from [38], where the local averages are viewed as point values of an approximation of the solution, and the regularisation of the solution at time  $t > 0$  is performed by interpolation

rather than convolution, could be a suitable choice.

Lastly, as mentioned before, another direction could be the extension of our results in order to deal with second order equations, as done in [46], where the Laplacian operator is approximated by an appropriate sequence of mutation kernels.

## 5.2 The problem

For  $T > 0$  and  $k \in \mathbb{N}$  we consider the functions

$$(t, x, I) \mapsto a(t, x, I) \in W^{1,\infty} \left( [0, T], (W^{k+1,\infty}(\mathbb{R}^{d+1}))^d \right), \quad (5.5)$$

$$(t, x, I) \mapsto R(t, x, I) \in \mathcal{C} \left( [0, T] \times \mathbb{R}_x^d, W_{loc}^{k+1,\infty}(\mathbb{R}_I) \right) \cap \mathcal{C} \left( [0, T] \times \mathbb{R}_I, W^{k+1,\infty}(\mathbb{R}_x^d) \right). \quad (5.6)$$

We consider as well  $(t, x, y, I) \mapsto m(t, x, y, I)$  such that  $m \geq 0$  and

$$\begin{aligned} m \in \mathcal{C} \left( [0, T] \times \mathbb{R}_x^d \times \mathbb{R}_y^d, W^{k+1,\infty}(\mathbb{R}_I) \right) \cap \mathcal{C} \left( [0, T] \times \mathbb{R}_x^d \times \mathbb{R}_I, L^\infty(\mathbb{R}_y^d) \right) \\ \cap \mathcal{C} \left( [0, T] \times \mathbb{R}_y^d \times \mathbb{R}_I, \mathcal{C}_c^k(\mathbb{R}_x^d) \right), \end{aligned}$$

which is globally Lipschitz with respect to the non local variables and uniformly compactly supported with respect to the  $x$  variable. That is, we suppose  $m$  to satisfy the following hypotheses:

- There exists  $\mu > 0$  such that

$$\sup_{t,x,y} \sum_{i=1}^k \sum_{|\alpha|=i} \sum_{j=1}^k |\partial_x^\alpha \partial_I^j m(t, x, y, I) - \partial_x^\alpha \partial_I^j m(t, x, y, J)| \leq \mu |I - J|. \quad (5.7)$$

- There exists a compact set  $\mathcal{K}$  such that the function

$$M(x) := \sup_{t,y,I} \sum_{i=1}^k \sum_{|\alpha|=i} \sum_{j=1}^k |\partial_x^\alpha \partial_I^j m(t, x, y, I)|, \quad (5.8)$$

satisfies

$$\text{supp } M(x) \subset \mathcal{K}. \quad (5.9)$$

Furthermore, we assume that

$$\|M\|_{L^\infty(\mathbb{R}^d)} \leq \overline{M} < \infty. \quad (5.10)$$

From a modelling point of view, assuming  $m$  to be uniformly compactly supported reflects the fact that some traits are realistically out of reach for a given population, and that, in general, mutations are rare and small.

We remark that hypotheses (5.9) and (5.10) imply that

$$\|M\|_{L^1(\mathbb{R}^d)} \leq |\mathcal{K}| \overline{M} =: K < \infty. \quad (5.11)$$

For all functions  $u \in \mathcal{C}([0, T], L^1(\mathbb{R}^d))$  we consider the linear mappings  $I_a$ ,  $I_g$  and  $I_d$  which satisfy, for all  $t \in [0, T]$ ,  $x \in \mathbb{R}^d$ ,

$$(I_a u)(t, x) := \int_{\mathbb{R}^d} \psi_a(t, x, y) u(t, y) dy, \quad (5.12)$$

$$(I_g u)(t, x) := \int_{\mathbb{R}^d} \psi_g(t, x, y) u(t, y) dy, \quad (5.13)$$

$$(I_d u)(t, x) := \int_{\mathbb{R}^d} \psi_d(t, x, y) u(t, y) dy, \quad (5.14)$$

where

$$\psi_a \in W^{1, \infty}([0, T] \times \mathbb{R}_x^d, L^\infty(\mathbb{R}_y^d)) \cap \mathcal{C}([0, T] \times \mathbb{R}_y^d, W^{k+1, \infty}(\mathbb{R}_x^d)), \quad (5.15)$$

$$0 < \underline{\psi}_g \leq \psi_g \in \mathcal{C}([0, T] \times \mathbb{R}_x^d, L^\infty(\mathbb{R}_y^d)) \cap \mathcal{C}([0, T] \times \mathbb{R}_y^d, W^{k+1, \infty}(\mathbb{R}_x^d)), \quad (5.16)$$

$$\psi_d \in \mathcal{C}([0, T] \times \mathbb{R}_x^d, L^\infty(\mathbb{R}_y^d)) \cap \mathcal{C}([0, T] \times \mathbb{R}_y^d, W^{k+1, \infty}(\mathbb{R}_x^d)), \quad (5.17)$$

for a certain  $\underline{\psi}_g > 0$ . We remark that  $I_a u, I_g u, I_d u \in \mathcal{C}([0, T], L^\infty(\mathbb{R}^d))$ . The functions  $\psi_a$  and  $\psi_d$  do not need to be positive, reflecting this way how different traits have different impacts (which are not always beneficial) on the environment, and ultimately on the population itself. On the other hand,  $\psi_g$  has to be bounded away from zero. This hypothesis reflects the fact that, at least for the growth term, all interactions between individuals are of the same type. These interactions may be interpreted as either strictly competitive or strictly cooperative. In particular, this means that only “very” non-local dependence with respect to  $u$  are allowed. This excludes partial densities, for instance on part of the traits. Lastly, we assume that there exist non-negative constants  $I^*$  and  $r^*$  such that, for all  $t \in [0, T]$ ,  $x \in \mathbb{R}^d$ , and  $I \geq I^*$ ,

$$R(t, x, I) + K < -r^*, \quad (5.18)$$

uniformly on  $t$  and  $x$ . It is somewhat natural to assume that  $R$  is negative for a large population: for example, if a carrying capacity is assumed to exist, and the population size is approaching such value, then the growth rate will inevitably drop to levels where no amount of mutations will be able to compensate for it.

For a given function  $v^0 \in L^1(\mathbb{R}^d)$ , we will study the existence and uniqueness of solution for the problem

$$\begin{cases} \partial_t v(t, x) + \nabla_x \cdot (a(t, x, I_a v(t, x)) v(t, x)) &= R(t, x, I_g v(t, x)) v(t, x) \\ &+ \int_{\mathbb{R}^d} m(t, x, y, I_d v(t, x)) v(t, y) dy, \\ v \in \mathcal{C}([0, T], L^1(\mathbb{R}^d)), \\ v(0, \cdot) = v^0(\cdot). \end{cases} \quad (5.19)$$

We will show that, under additional hypotheses, either over  $a$  or  $v^0$ , we can guarantee the well-posedness of this problem. In particular, we provide the results regarding the cases of local advection ( $\partial_I a = 0$ ) and non-local advection ( $\partial_I a \neq 0$ ). We will see that this distinction directly affects the set of initial data  $v^0$  for which the existence of solutions is guaranteed.

### 5.2.1 Some bounds over the characteristics

Consider  $a$  satisfying (5.5) and  $\psi_a$  satisfying (5.15) for some  $k \geq 1$ . For all  $y \in \mathbb{R}^d$  and  $u \in \mathcal{C}([0, T], L^1(\mathbb{R}^d))$  we define the characteristic curve  $t \mapsto X_u(t, y)$  as the unique solution the following ODE

$$\begin{cases} \dot{X}_u(t, y) = a(t, X_u(t, y), (I_a u)(t, X_u(t, y))) =: \mathcal{A}_u(t, X_u(t, y)), & t \in [0, T], \\ X_u(0, y) = y, \end{cases} \quad (5.20)$$

where  $(I_a u)(t, x)$  is defined in (5.12). Since the function  $\psi_a$  belongs to  $W^{1, \infty}([0, T] \times \mathbb{R}_x^d, L^\infty(\mathbb{R}_y^d))$ , then  $(I_a u)(t, x)$  belongs to  $W^{1, \infty}([0, T] \times \mathbb{R}_x^d) \subset \mathcal{C}^{0,1}([0, T] \times \mathbb{R}_x^d)$ . The regularity of  $a$  then implies that  $\mathcal{A}_u(t, x)$  is a Lipschitz function with respect to the  $x$  variable, uniformly with respect to  $t$ , guaranteeing this way the global existence of solution for (5.20).

For all  $u \in \mathcal{C}([0, T], L^1(\mathbb{R}^d))$ , we define the norms

$$\|u\|_1 := \|u\|_{L^1([0, T] \times \mathbb{R}^d)} = \int_0^T \int_{\mathbb{R}^d} |u(t, x)| dx dt \quad \text{and} \quad \|u\| := \sup_{t \in [0, T]} \|u(t, \cdot)\|_{L^1(\mathbb{R}^d)}.$$

We present some results involving the characteristics. The proofs for such results are given in Appendix C.1.

The first property we describe is the continuity of the family of characteristics with respect to the spatial variables and the function  $u$ .

**Lemma 11.** *Let  $\psi_a$  satisfy (5.15) and  $a$  satisfy (5.5) for  $k = 0$ . Consider  $y_1, y_2 \in \mathbb{R}^d$  and  $u_1, u_2 \in \mathcal{C}([0, T], L^1(\mathbb{R}^d))$ . Then there exists a positive constant  $C(T, \|u_1\|, \|u_2\|)$ , such that the solutions  $X_{u_1}$  and  $X_{u_2}$  of (5.20) satisfy for any  $t \in [0, T]$ ,*

$$\sum_{j=1}^d |X_{u_1}^j(t, y_1) - X_{u_2}^j(t, y_2)| \leq C(T, \|u_1\|, \|u_2\|) (|y_1 - y_2| + \|u_1 - u_2\|_1).$$

Secondly, we claim that the spatial derivatives of the characteristics remain bounded by a constant only depending on  $T$  and  $\|u\|$ . We also claim that the spatial derivatives are continuous with respect to the spatial variables and the function  $u$ .

**Lemma 12.** *Let  $a$  satisfy (5.5) and  $\psi_a$  satisfy (5.15) for some  $k \geq 1$ . Consider  $u \in \mathcal{C}([0, T], L^1(\mathbb{R}^d))$ . Then there exists a positive constant  $C(T, \|u\|)$ , such that the solution  $X_u(t, y)$  of (5.20) satisfies, for all  $t \in [0, T]$  and  $y \in \mathbb{R}^d$ ,*

$$\sum_{i=1}^k \sum_{|\alpha| \leq i} \sum_{j=1}^d |\partial_y^\alpha X_u^j(t, y)| \leq C(T, \|u\|). \quad (5.21)$$

Furthermore, for any two points  $y_1, y_2 \in \mathbb{R}^d$ , and any two functions  $u_1, u_2 \in \mathcal{C}([0, T], L^1(\mathbb{R}^d))$ , there exists a positive constant  $C_2(T, \|u_1\|, \|u_2\|)$ , such that the solutions  $X_{u_1}$  and  $X_{u_2}$  of (5.20) satisfy for any  $t \in [0, T]$

$$\sum_{i=0}^k \sum_{|\alpha| \leq i} \sum_{j=1}^d |\partial_y^\alpha X_{u_1}^j(t, y_1) - \partial_y^\alpha X_{u_2}^j(t, y_2)| \leq C(T, \|u_1\|, \|u_2\|) (|y_1 - y_2| + \|u_1 - u_2\|_1). \quad (5.22)$$

Since for all  $t \in [0, T]$ ,  $y \mapsto X_u(t, y)$  is a  $C^1$ -diffeomorphism from  $\mathbb{R}^d$  onto itself, we may define its inverse as the function satisfying  $X_u(t, X_u^{-1}(t, x)) = x$  for all  $(t, x) \in [0, T] \times \mathbb{R}^d$ . We have the following results for  $X_u^{-1}$ .

**Lemma 13.** *Let  $a$  satisfy (5.5) and  $\psi_a$  satisfy (5.15) for some  $k \geq 1$ . Consider  $u \in C([0, T], L^1(\mathbb{R}^d))$ . Then there exists a positive constant  $\tilde{C}(T, \|u\|)$ , such that the inverse of the solution  $X_u(t, y)$  of (5.20) satisfies, for all  $t \in [0, T]$ ,  $x \in \mathbb{R}^d$*

$$\sum_{i=1}^k \sum_{|\alpha| \leq i} \sum_{j=1}^d |\partial_x^\alpha (X_u^{-1})^j(t, x)| \leq \tilde{C}(T, \|u\|). \quad (5.23)$$

**Lemma 14.** *Let  $a$  satisfy (5.5) and  $\psi_a$  satisfy (5.15) for some  $k \geq 1$ . Consider any two functions  $u_1, u_2 \in C([0, T], L^1(\mathbb{R}^d))$ . Then there exists a positive constant  $\tilde{C}(T, \|u_1\|, \|u_2\|)$ , which satisfies  $\lim_{T \rightarrow 0} \tilde{C}(T, \|u_1\|, \|u_2\|) = 0$  and such that the inverses of the solutions  $X_{u_1}$  and  $X_{u_2}$  of (5.20) satisfy, for all  $t \in [0, T]$  and  $x \in \mathbb{R}^d$*

$$\sum_{i=0}^{k-1} \sum_{|\alpha| \leq i} \sum_{j=1}^d |(\partial_x^\alpha X_{u_1}^{-1})^j(t, x) - \partial_x^\alpha (X_{u_2}^{-1})^j(t, x)| \leq \tilde{C}(T, \|u_1\|, \|u_2\|) \|u_1 - u_2\|_1. \quad (5.24)$$

We remark that thanks to the relation  $\|u\|_1 \leq T\|u\|$ , the relations (5.22) and (5.24) also hold true when replacing  $\|u_1 - u_2\|_1$  by  $\|u_1 - u_2\|$ . Lastly, we give a result regarding the regularity of  $X_u(t, x)$  with respect to  $t$ .

**Lemma 15.** *Let  $a$  satisfy (5.5) and  $\psi_a$  satisfy (5.15) for some  $k \geq 1$ . Consider  $u \in C^1([0, T_1] \times \mathbb{R}^d)$  such that  $\sup_{t \in [0, T_1]} (\|u(t, \cdot)\|_{L^1(\mathbb{R}^d)} + \|\partial_t u(t, \cdot)\|_{L^1(\mathbb{R}^d)}) < +\infty$ . Then  $X_u(t, y) \in C^1([0, T_1], C^k(\mathbb{R}^d))$ . As a consequence,  $X_u^{-1}(t, x) \in C^1([0, T_1], C^k(\mathbb{R}^d))$ .*

*Proof.* Thanks to Lemma 12, we know that under these hypotheses, for all  $T < T_1$ ,  $X_u(t, y)$  exists and belongs to  $C^1([0, T], C^k(\mathbb{R}^d))$ . Consider  $0 < t_1, t_2 < T_1$ , then

$$\begin{aligned} \sum_{i=0}^k \sum_{|\alpha| \leq i} \sum_{j=1}^d |\partial_y^\alpha X_u^j(t_1, y) - \partial_y^\alpha X_u^j(t_2, y)| &= \sum_{i=0}^k \sum_{|\alpha| \leq i} \sum_{j=1}^d \left| \int_{t_1}^{t_2} \partial_y^\alpha \dot{X}_u^j(s, y) ds \right| \\ &= \sum_{i=0}^k \sum_{|\alpha| \leq i} \sum_{j=1}^d \left| \int_{t_1}^{t_2} \partial_y^\alpha a_j(s, X_u(s, y), (I_a u)(s, X_u(s, y))) ds \right|. \end{aligned}$$

Thanks to the regularity of  $a$  and  $\psi_a$ , the bounds given in Lemma 12 for the derivatives of  $X_u(t, y)$  and the uniform bound for  $\|u\|_{L^1(\mathbb{R}^d)}$  we conclude that there exists a positive constant such that

$$\sum_{i=0}^k \sum_{|\alpha| \leq i} \sum_{j=1}^d |\partial_y^\alpha X_u^j(t_1, y) - \partial_y^\alpha X_u^j(t_2, y)| \leq C|t_1 - t_2|.$$

Similarly, we have

$$\sum_{i=0}^k \sum_{|\alpha| \leq i} \sum_{j=1}^d |\partial_y^\alpha \dot{X}_u^j(t_1, y) - \partial_y^\alpha \dot{X}_u^j(t_2, y)| = \sum_{i=0}^k \sum_{|\alpha| \leq i} \sum_{j=1}^d |\partial_y^\alpha (\mathcal{A}_u)_j(t_1, y) - \partial_y^\alpha (\mathcal{A}_u)_j(t_2, y)|,$$

where

$$(\mathcal{A}_u)_j(t, y) := a_j(t, X_u(t, y), (I_a u)(t_1, X_u(t, y))).$$



Again, the regularity up to order  $k + 1$  of the involved coefficients allow us to conclude that

$$\sum_{i=0}^k \sum_{|\alpha| \leq i} \sum_{j=1}^d |\partial_y^\alpha \dot{X}_u^j(t_1, y) - \dot{X}_u^j(t_2, y)| \leq C|t_1 - t_2|.$$

We have shown that  $X_u(t, \cdot)$  is a Cauchy sequence in  $\mathcal{C}^k(\mathbb{R}^d)$  when  $t$  goes to  $T_M$ , therefore, there exists  $X^*(x) \in \mathcal{C}^k(\mathbb{R}^d)$  and  $Y^*(x) \in \mathcal{C}^k(\mathbb{R}^d)$  such that

$$\lim_{t \rightarrow T_M} \|X_u(t, \cdot) - X^*\|_{\mathcal{C}^k(\mathbb{R}^d)} + \|\dot{X}_u(t, \cdot) - Y^*\|_{\mathcal{C}^k(\mathbb{R}^d)},$$

which is the desired result.  $\square$

### 5.2.2 Existence of solution for smooth initial data

We first provide the proof of existence and uniqueness of solution for problem (5.19) when the initial condition  $v^0$  is a smooth enough function. We still assume hypotheses (5.5) through (5.18) hold. For a smooth initial condition  $v^0$  we denote by *classical solution* any function  $v \in \mathcal{C}^1([0, T] \times \mathbb{R}^d)$  which satisfies problem (5.19).

The following *a priori* estimate will allow us to guarantee the global existence of a classical solution given that such solution exists over a certain interval  $[0, T_1]$ .

**Lemma 16.** *Let  $v^0 \in \mathcal{C}_c^1(\mathbb{R}^d)$  and  $T_1 > 0$  be such that a classical solution  $v \in \mathcal{C}^1([0, T] \times \mathbb{R}^d)$  exists for problem (5.19) for all  $T < T_1$ , which is positive and has compact support with respect to the  $x$  variable. Then, such solution satisfies the estimate*

$$\sup_{t \in [0, T_1]} \|v(t, \cdot)\|_{L^1(\mathbb{R}^d)} \leq \max\left\{\|v^0\|_{L^1(\mathbb{R}^d)}, \frac{I^*}{\underline{\psi}_g}\right\}. \quad (5.25)$$

*Proof.* Let  $v$  be the aforementioned positive solution. Denoting  $\rho(t) := \|v(t, \cdot)\|_{L^1(\mathbb{R}^d)}$ , we see directly from equation (5.19) that

$$\dot{\rho}(t) = \int_{\mathbb{R}^d} \left( R(t, y, (I_g v)(t, y)) + \int_{\mathbb{R}^d} m(t, x, y, (I_a v)(t, x)) dx \right) v(t, y) dy.$$

If there exists  $t$  such that  $\rho(t) > \frac{I^*}{\underline{\psi}_g}$ , then  $(I_g v)(t, y) \geq \underline{\psi}_g \rho(t) > I^*$ , which allows us to use hypothesis (5.18) in order to conclude

$$\dot{\rho}(t) < -r^* \rho(t) < 0.$$

This way, we see that either  $\rho(t)$  is smaller than  $\frac{I^*}{\underline{\psi}_g}$  or  $\rho(t)$  is decreasing, which in turn implies the bound (5.25).  $\square$

A fixed point argument together with estimate (5.25) will allow us to conclude the existence of solution for problem (5.19) for smooth initial conditions.

**Theorem 10.** *Consider  $k \geq 1$  and  $T > 0$ . Under hypotheses (5.15) through (5.18), for all non-negative functions  $v^0 \in \mathcal{C}_c^k(\mathbb{R}^d)$ , there exists a unique non-negative classical solution  $v \in \mathcal{C}^1([0, T], \mathcal{C}_c^k(\mathbb{R}^d))$  to problem (5.19). Furthermore, such solution satisfies*

$$\sup_{t \in [0, T]} \|v(t, \cdot)\|_{L^1(\mathbb{R}^d)} \leq \max\left\{\|v^0\|_{L^1(\mathbb{R}^d)}, \frac{I^*}{\underline{\psi}_g}\right\}, \quad (5.26)$$

$$\sup_{t \in [0, T]} \|v(t, \cdot)\|_{W^{k,1}(\mathbb{R}^d)} \leq C_T \|v^0\|_{W^{k,1}(\mathbb{R}^d)}. \quad (5.27)$$

*Proof.* Consider  $v^0 \in \mathcal{C}_c^k(\mathbb{R}^d)$ . For  $t \geq 0$  and the function  $M$  introduced in (5.8) we define  $r_t := 2\|a\|_{L^\infty}t$ ,  $B_{r_t}$  the open ball centred at 0 and of radius  $r_t$ ,

$$O_t := \text{supp}(v^0) \cup \text{supp}(M) + B_{r_t},$$

and for  $\alpha > 1$  we define  $\bar{\rho}_\alpha := \max\{\alpha\|v^0\|_{L^1(\mathbb{R}^d)}, \frac{I^*}{\underline{\psi}_g}\}$ . We consider

$$u \in D_\alpha^T := \left\{ u \in \mathcal{C}([0, T] \times \mathbb{R}^d) : u \geq 0, \text{supp}(u(t, \cdot)) \subset O_t, \int_{\mathbb{R}^d} u(t, x) dx \leq \bar{\rho}_\alpha, \forall t \in [0, T] \right\}.$$

We denote as  $\Phi u$  the mapping defined by  $v = \Phi u$ , where  $v$  is the solution of

$$\begin{cases} \partial_t v(t, x) + \nabla_x \cdot (a(t, x, (I_a u)(t, x))v(t, x)) & -R(t, x, (I_g u)(t, x))v(t, x) \\ & = \int_{\mathbb{R}^d} m(t, x, y, (I_d u)(t, x))u(t, y) dy \\ v(0, \cdot) = v^0(\cdot). \end{cases}$$

Let us denote, for any  $y \in \mathbb{R}^d$ , as  $X_u(\cdot, y)$  the unique solution of

$$\begin{cases} \dot{X}_u(t, y) = a(t, X_u(t, y), (I_a u)(t, X_u(t, y))) =: \mathcal{A}_u(t, X_u(t, y)) & t \geq 0, \\ X(0, y) = y. \end{cases}$$

As stated before, for any  $t \in [0, T]$ ,  $x \mapsto X_u(t, x)$  is a  $\mathcal{C}^1$ -diffeomorphism, therefore, for all  $t \geq 0$  and all  $x \in \mathbb{R}^d$  there exists a unique  $y \in \mathbb{R}^d$  such that  $x = X_u(t, y)$ , which we denote  $y = X_u^{-1}(t, x)$ .

We see that,

$$\begin{aligned} \frac{d}{dt}v(t, X_u(t, y)) &= [R(t, X_u(t, y), (I_g u)(t, X_u(t, y))) - \text{div} \mathcal{A}_u(t, X_u(t, y))]v(t, X_u(t, y)) \\ &\quad + \int_{\mathbb{R}^d} m(t, X_u(t, y), z, (I_d u)(t, X_u(t, y)))u(t, z) dz, \end{aligned}$$

and thus, denoting

$$\mathcal{G}_u(t, y) := R(t, X_u(t, y), (I_g u)(t, X_u(t, y))) - \text{div} \mathcal{A}_u(t, X_u(t, y)),$$

we get

$$\begin{aligned} \Phi u(t, X_u(t, y)) &= v(t, X_u(t, y)) \\ &= v^0(y) \exp\left(\int_0^t \mathcal{G}_u(s, y) ds\right) \\ &\quad + \int_0^t \int_{\mathbb{R}^d} m(s, X_u(s, y), z, (I_d u)(s, X_u(s, y)))u(s, z) dz \exp\left(\int_s^t \mathcal{G}_u(\tau, y) d\tau\right) ds \end{aligned}$$

or, equivalently

$$\begin{aligned} \Phi u(t, x) &= v(t, x) \\ &= v^0(X_u^{-1}(t, x)) \exp\left(\int_0^t \mathcal{G}_u(s, X_u^{-1}(t, x)) ds\right) \\ &\quad + \int_0^t \int_{\mathbb{R}^d} m(s, X_u(s, X_u^{-1}(t, x)), z, (I_d u)(s, X_u(s, X_u^{-1}(t, x))))u(s, z) dz \\ &\quad \exp\left(\int_s^t \mathcal{G}_u(\tau, X_u^{-1}(t, x)) d\tau\right) ds. \end{aligned}$$

The solution  $v$  is thus non-negative, according to the non-negativity of  $v^0$ ,  $u$  and  $m$ . Thanks to Lemma 13,  $v$  belongs to  $\mathcal{C}^1([0, T], \mathcal{C}^\kappa(\mathbb{R}^d)) \subset \mathcal{C}([0, T] \times \mathbb{R}^d)$ . Furthermore, the

fact that for all  $t \in [0, T]$ ,  $|X(t, y) - y| \leq \|a\|_{L^\infty} t$  implies that  $\text{supp}(v(t, \cdot)) \subset O_t$ . Additionally, directly from its definition, we see that

$$\mathcal{G}_u(s, X_u^{-1}(t, x)) \leq \gamma := \|R\|_{L^\infty_{t,x,I}} + \|a\|_{W_x^{1,\infty} L_{t,I}^\infty} + \|a\|_{W_I^{1,\infty} L_{t,x}^\infty} \|\psi_a\|_{W_x^{1,\infty} L_y^\infty} \bar{\rho}_\alpha,$$

and consequently, for all  $u \in D_\alpha^T$ , we have

$$\begin{aligned} \|\Phi u(t, \cdot)\|_{L^1(\mathbb{R}^d)} &\leq e^{\gamma T} \left( \int_{\mathbb{R}^d} v^0(X_u^{-1}(t, x)) dx \right. \\ &\quad \left. + \int_0^t \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} m(s, X_u(s, X_u^{-1}(t, x)), z, (Idu)(s, X_u(s, X_u^{-1}(t, x)))) u(s, z) dz dx ds \right) \\ &\leq e^{\gamma T} \left( \int_{\mathbb{R}^d} v^0(X_u^{-1}(t, x)) dx + \int_0^t \int_{\mathbb{R}^d} M(s, X_u(s, X_u^{-1}(t, x))) dx \int_{\mathbb{R}^d} u(s, z) dz ds \right). \end{aligned}$$

Making the changes of variables  $y = X_u^{-1}(t, x)$  and  $y = X_u(s, X_u^{-1}(t, x))$  respectively on each of the integrals on the last expression, recalling that, according to Liouville's formula

$$\begin{aligned} |J_{X_u(t,y)}| &= e^{\int_0^t \text{div } \mathcal{A}_u(s, X_u(s, y)) ds}, \\ |J_{X_u(t, X_u^{-1}(s,y))}| &= e^{\int_0^t \text{div } \mathcal{A}_u(\tau, X_u(\tau, X_u^{-1}(s,y))) d\tau - \int_0^s \text{div } \mathcal{A}_u(\tau, y) d\tau}, \end{aligned}$$

and using the hypotheses over  $a$  and  $m$  we obtain that for all  $t \in [0, T]$ ,

$$\|\Phi u(t, \cdot)\|_{L^1(\mathbb{R}^d)} \leq e^{(\gamma+2\tilde{a})T} \left( \|v^0\|_{L^1(\mathbb{R}^d)} + KT \|u\|_{L^1(\mathbb{R}^d)} \right),$$

where  $\tilde{a} := \|a\|_{W_x^{1,\infty} L_{t,I}^\infty} + \|a\|_{W_I^{1,\infty} L_{t,x}^\infty} \|\psi_a\|_{W_x^{1,\infty} L_y^\infty} \bar{\rho}_\alpha$ . Finally, using the hypothesis over  $v^0$  and  $u$  we see that

$$\|\Phi u(t, \cdot)\|_{L^1(\mathbb{R}^d)} \leq e^{(\gamma+2\tilde{a})T} \left( \frac{1}{\alpha} + KT \right) \bar{\rho}_\alpha.$$

Thanks to the condition  $\alpha > 1$  there exists  $T_\alpha$  (only depending on  $\alpha$  and on the coefficients of the problem) such that  $\|\Phi u\|_{L^1(\mathbb{R}^d)} \leq \bar{\rho}_\alpha$  for all  $t \in [0, T_\alpha]$ . In other words,  $\Phi : D_\alpha^{T_\alpha} \rightarrow D_\alpha^{T_\alpha}$ .

We now claim that the mapping  $\Phi u$  is a contraction on some  $D_\alpha^{T_1}$ ,  $0 < T_1 \leq T_\alpha$ , with respect to the usual norm in  $\mathcal{C}([0, T] \times \mathbb{R}^d)$ . For two functions  $u_1, u_2 \in D_\alpha^{T_\alpha}$  and any  $t \in [0, T_\alpha]$ , thanks to Lemma 11 and Lemma 14, we have

$$\begin{aligned} |\Phi u_1 - \Phi u_2| &\leq C e^{\gamma t} |X_{u_1}^{-1}(t, x) - X_{u_2}^{-1}(t, x)| \\ &\leq C e^{\gamma t} \|u_1 - u_2\|_1 \\ &\leq C e^{\gamma t} O_t |t| \|u_1 - u_2\|_{\mathcal{C}([0,t] \times \mathbb{R}^d)}. \end{aligned}$$

Clearly, for  $t = T_1$  small enough,  $\Phi u$  is a contraction, and therefore, thanks to the Banach fixed point theorem, there exists a unique  $v \in D_\alpha^{T_1}$  such that  $\Phi v = v$ . Such  $v$  is a solution of (5.19) over  $[0, T_1]$ . Furthermore, directly from the relation  $v = \Phi v$  we see that  $v \in \mathcal{C}^1([0, T], \mathcal{C}_c^\kappa(\mathbb{R}^d))$ .

Let us now assume that there exists  $T_M$ , a finite maximal time such that a solution exists in  $B_\alpha^T$  for all  $T < T_M$ , and let  $v$  be such solution. Directly from Lemma 16,

the solution  $v$  satisfies estimate (5.25) over  $[0, T_M]$ . Furthermore, thanks to the relation  $v = \Phi v$ , we are able to show that

$$\sup_{[0, T_M]} (\|v(t, \cdot)\|_{L^1(\mathbb{R}^d)} + \|\dot{v}(t, \cdot)\|_{L^1(\mathbb{R}^d)}) < +\infty.$$

From Lemma 15 we get then that  $X_v^{-1}(t, x) \in \mathcal{C}^1([0, T_M], \mathcal{C}^k(\mathbb{R}^d))$ , and by composition of functions, so is  $v = \Phi v$ . We can then iterate the previous ideas using  $v(T_M) \in \mathcal{C}^k(\mathbb{R}^d)$  as a starting point in order to obtain the existence of solution over a certain interval  $[T_M, T_M + \delta)$ , contradicting this way the maximal character of  $T_M$ . Hence, there exists a classical solution of (5.19) for all  $t > 0$ .

The uniqueness on  $\mathcal{C}([0, T], \mathcal{C}_c^k(\mathbb{R}^d))$  comes from the fact that every other solution on  $D_\alpha^\infty$  will coincide with  $v$  at least over a small interval  $(0, t_0)$  and then, by continuity, the same would hold for all  $t$ .

To obtain the  $W^{k,1}(\mathbb{R}^d)$  estimates, we differentiate the relation  $v = \Phi v$  and notice that for all multi-index  $\beta$  such that  $|\beta| \leq k$

$$\begin{aligned} \partial_x^\beta v &= \exp\left(\int_0^t \mathcal{G}_v(s, X_v^{-1}(t, x)) ds\right) \sum_{|\gamma| \leq |\beta|} \partial_x^\gamma v^0(X_v^{-1}(t, x)) F_1^\gamma(t, x) \\ &\quad + \int_0^t \int_{\mathbb{R}^d} F_2(s, x, y) v(s, y) dy \exp\left(\int_s^t \mathcal{G}_v(\tau, x) d\tau\right) ds \end{aligned}$$

where the functions  $F_1^\gamma$  and  $F_2$  are combinations of sums and multiplications of the derivatives up to order  $|\beta|$  of  $X_v^{-1}$ ,  $R$ ,  $m$ ,  $I_g$  and  $I_d$ . Taking absolute values, integrating over  $\mathbb{R}^d$  and using the boundedness of all the involved coefficients we arrive at

$$\|v(t, \cdot)\|_{W^{k,1}(\mathbb{R}^d)} \leq C_T^1 \|v^0\|_{W^{k,1}(\mathbb{R}^d)} + C_T^2 \int_0^t \|v(s, \cdot)\|_{W^{k,1}(\mathbb{R}^d)} ds$$

and thanks to Grönwall's lemma we get (5.27).  $\square$

### 5.2.3 Existence of solution for more general initial data

Depending on whether  $\partial_I a = 0$  or  $\partial_I a \neq 0$ , we will have a different class of initial data for which we are able to guarantee existence of solution for problem (5.19). Furthermore, the regularity of such solution might also be affected.

We first prove that, when  $\partial_I a \neq 0$ , a solution exists (in a sense that will be defined below) for any initial condition  $v^0 \in W^{k,\infty}(\mathbb{R}^d)$ , with compact support. However, we do not prove that the regularity of the solution is preserved over time, even if we did not manage to highlight the existence of cases where a loss of regularity is observed. Secondly, we will show that, when  $\partial_I a = 0$ , not only the set of initial conditions for which we can claim existence of solution is more general ( $v^0 \in W^{k,1}(\mathbb{R}^d)$ ), but the regularity of such solution is preserved for all  $t > 0$ .

We introduce the definition of *weak solution* for problem (5.19). We say that  $v$  is a *weak solution* of problem (5.19) associated to  $v^0 \in L^p(\mathbb{R}^d)$  if

$$v \in L^\infty([0, T], L^p(\mathbb{R}^d)),$$

and it satisfies the equation in the following weak sense

$$\int_0^T \int_{\mathbb{R}^d} v L_v^* \varphi dx dt = \int_{\mathbb{R}^d} v^0 \varphi dx,$$

for any  $\varphi \in \mathcal{C}_c^1([0, T] \times \mathbb{R}^d)$ , where we define the operator  $L_v^*$  by

$$\begin{aligned} L_v^* \varphi(t, x) &= -\partial_t \varphi(t, x) - a(t, x, (I_a v)(t, x)) \cdot \nabla \varphi(t, x) - R(t, x, (I_g v)(t, x)) \varphi(t, x) \\ &\quad - \int_{\mathbb{R}^d} m(t, y, x, (I_d v)(t, y)) \varphi(t, y) dy, \end{aligned}$$

for all  $t \in [0, T]$ ,  $y \in \mathbb{R}^d$ . We remark that a *classical solution* is always a *weak solution*.

**Theorem 11.** *Under hypotheses (5.5) through (5.18), for all  $k \geq 1$  and any non-negative functions  $v^0 \in W^{k, \infty}(\mathbb{R}^d)$  with compact support, there exists a unique non-negative weak solution  $v \in \mathcal{C}([0, T], \mathcal{C}_c^{k-1}(\mathbb{R}^d))$  to problem (5.19). Furthermore, such solution satisfies*

$$\sup_{t \in [0, T]} \|v(t, \cdot)\|_{L^1(\mathbb{R}^d)} \leq \max\{\|v^0\|_{L^1(\mathbb{R}^d)}, \frac{I^*}{\underline{\psi}_g}\}, \quad (5.28)$$

$$\sup_{t \in [0, T]} \|v(t, \cdot)\|_{W^{k-1, 1}(\mathbb{R}^d)} \leq C_T \|v^0\|_{W^{k-1, 1}(\mathbb{R}^d)}, \quad (5.29)$$

and, for  $k \geq 2$ ,  $v \in \mathcal{C}^1([0, T], \mathcal{C}_c^{k-1}(\mathbb{R}^d))$ .

*Proof.* Directly from Morrey's inequality, we get the relation  $v^0 \in W^{k, \infty}(\mathbb{R}^d) \subset \mathcal{C}^{k-1, 1}(\mathbb{R}^d)$ .

If  $k \geq 2$ , we are able to apply Theorem 10 in order to get the desired result.

Consider now  $k = 1$ . The compact support of  $v^0$  and the  $W^{1, \infty}(\mathbb{R}^d)$  regularity imply that  $v^0 \in W^{k, 1}(\mathbb{R}^d)$ . This means that, there exists a sequence of compactly supported functions  $v_\varepsilon^0 \in \mathcal{C}_c^1(\mathbb{R}^d)$  such that

$$\begin{aligned} \text{supp}(v_\varepsilon^0) &\subset \text{supp}(v^0), \\ \|v_\varepsilon^0\|_{W^{1, \infty}(\mathbb{R}^d)} &\leq \|v^0\|_{W^{1, \infty}(\mathbb{R}^d)}, \\ \lim_{\varepsilon \rightarrow 0} \|v^0 - v_\varepsilon^0\|_{W^{1, 1}(\mathbb{R}^d)} &= 0. \end{aligned}$$

We denote as  $v_\varepsilon$  the solution of problem (5.19) associated to  $v_\varepsilon^0$ , and we claim that  $v_\varepsilon$  is a Cauchy sequence in  $\mathcal{C}([0, T], L^1(\mathbb{R}^d))$ .

We recall that for all  $\varepsilon$ ,

$$\sup_{t \in [0, T]} \|v_\varepsilon\|_{L^1(\mathbb{R}^d)} \leq \bar{\rho}_\varepsilon := \max\{\|v_\varepsilon^0\|_{L^1(\mathbb{R}^d)}, \frac{I^*}{\underline{\psi}_g}\}.$$

Furthermore, the equality  $v_\varepsilon = \Phi v_\varepsilon$  holds true, where  $\Phi$  was defined on the proof of Theorem 10. Consequently, for  $\varepsilon_1, \varepsilon_2 > 0$  we have the relation

$$\begin{aligned} \Delta V_{\varepsilon_1 \varepsilon_2} &:= v_{\varepsilon_1} - v_{\varepsilon_2} \\ &= \Phi v_{\varepsilon_1} - \Phi v_{\varepsilon_2} \\ &= \left( v_{\varepsilon_1}^0(X_{v_{\varepsilon_1}}^{-1}(t, x)) - v_{\varepsilon_2}^0(X_{v_{\varepsilon_2}}^{-1}(t, x)) \right) \exp\left( \int_0^t \mathcal{G}_1(\tau, t, x) d\tau \right) \\ &\quad - v_{\varepsilon_2}^0(X_{v_{\varepsilon_2}}^{-1}(t, x)) \left( \Delta E(0, t, x) \right) \\ &\quad + \int_0^t \int_{\mathbb{R}^d} \left( \mathcal{M}_1(s, t, x, z) - \mathcal{M}_2(s, t, x, z) \right) v_{\varepsilon_1}(s, z) dz \exp\left( \int_s^t \mathcal{G}_1(\tau, t, x) d\tau \right) ds \\ &\quad + \int_0^t \int_{\mathbb{R}^d} \mathcal{M}_2(s, t, x, z) \left( \Delta V_{\varepsilon_1 \varepsilon_2}(s, z) \right) dz \exp\left( \int_s^t \mathcal{G}_1(\tau, t, x) d\tau \right) ds \\ &\quad + \int_0^t \int_{\mathbb{R}^d} \mathcal{M}_2(s, t, x, z) v_{\varepsilon_2}(s, z) dz \left( \Delta E(s, t, x) \right) ds, \end{aligned}$$

where

$$\begin{aligned}\mathcal{G}_i(\tau, t, x) &= \mathcal{G}_{v_{\varepsilon_i}}(\tau, X_{v_{\varepsilon_i}}^{-1}(t, x)), \\ \Delta E(s, t, x) &:= \exp\left(\int_s^t \mathcal{G}_1(\tau, t, x) d\tau\right) - \exp\left(\int_s^t \mathcal{G}_2(\tau, t, x) d\tau\right), \\ \mathcal{M}_i(s, t, x, z) &:= m(s, X_{v_{\varepsilon_i}}(s, X_{v_{\varepsilon_i}}^{-1}(t, x)), z, (Id_{v_{\varepsilon_i}})(s, X_{v_{\varepsilon_i}}(s, X_{v_{\varepsilon_i}}^{-1}(t, x)))).\end{aligned}$$

We write

$$\begin{aligned}v_{\varepsilon_1}^0(X_{v_{\varepsilon_1}}^{-1}(t, x)) - v_{\varepsilon_2}^0(X_{v_{\varepsilon_2}}^{-1}(t, x)) &= v_{\varepsilon_1}^0(X_{v_{\varepsilon_1}}^{-1}(t, x)) - v_{\varepsilon_2}^0(X_{v_{\varepsilon_1}}^{-1}(t, x)) \\ &\quad + v_{\varepsilon_2}^0(X_{v_{\varepsilon_1}}^{-1}(t, x)) - v_{\varepsilon_2}^0(X_{v_{\varepsilon_2}}^{-1}(t, x)).\end{aligned}\tag{5.30}$$

Thanks to the change of variables  $y = X_{v_{\varepsilon_1}}^{-1}(t, x)$  and the relations

$$\begin{aligned}|J_{X_{v_{\varepsilon_1}}}(t, y)| &= e^{\int_0^t \operatorname{div} \mathcal{A}_{v_{\varepsilon_1}}(s, X_{v_{\varepsilon_1}}(t, y)) ds}, \\ \mathcal{G}(s, X_{v_{\varepsilon_1}}^{-1}(t, x)) &\leq \gamma := \|R\|_{L_{t,x,I}^\infty} + \|a\|_{W_x^{1,\infty} L_{t,I}^\infty} + \|a\|_{W_I^{1,\infty} L_{t,x}^\infty} \|\psi_a\|_{W_x^{1,\infty} L_y^\infty} \bar{\rho}, \\ \operatorname{div} \mathcal{A}_{v_{\varepsilon_1}}(s, X_{v_{\varepsilon_1}}(t, y)) &\leq \tilde{a} := \|a\|_{W_x^{1,\infty} L_{t,I}^\infty} + \|a\|_{W_I^{1,\infty} L_{t,x}^\infty} \|\psi_a\|_{W_x^{1,\infty} L_y^\infty} \bar{\rho},\end{aligned}$$

we conclude that

$$\begin{aligned}&\int_{\mathbb{R}^d} \left( v_{\varepsilon_1}^0(X_{v_{\varepsilon_1}}^{-1}(t, x)) - v_{\varepsilon_2}^0(X_{v_{\varepsilon_1}}^{-1}(t, x)) \right) \exp\left(\int_0^t \mathcal{G}_1(\tau, t, x) d\tau\right) dx \\ &= \int_{\mathbb{R}^d} \left( v_{\varepsilon_1}^0(y) - v_{\varepsilon_2}^0(y) \right) \exp\left(\int_0^t \mathcal{G}_{v_{\varepsilon_1}}(\tau, y) d\tau\right) |J_{X_{v_{\varepsilon_1}}^{-1}}(t, y)| dy \\ &\leq e^{(\gamma + \tilde{a})T} \|v_{\varepsilon_1}^0 - v_{\varepsilon_2}^0\|_{L^1(\mathbb{R})}.\end{aligned}\tag{5.31}$$

On the other hand, the compactness of the support of  $v_{\varepsilon_2}^0$ , together with the relation

$$|X_{v_{\varepsilon_2}}(t, y) - y| \leq \|a\|_{L^\infty} t,$$

implies that

$$v_{\varepsilon_2}^0(X_{v_{\varepsilon_1}}^{-1}(t, x)) - v_{\varepsilon_2}^0(X_{v_{\varepsilon_2}}^{-1}(t, x)) = 0, \text{ for all } x \notin O_t := \operatorname{supp} v^0 + B_{r_t},$$

where  $B_{r_t}$  is the ball of radius  $\|a\|_{L^\infty} t$ . Hence, we obtain

$$\begin{aligned}&\int_{\mathbb{R}^d} \left( v_{\varepsilon_2}^0(X_{v_{\varepsilon_1}}^{-1}(t, x)) - v_{\varepsilon_2}^0(X_{v_{\varepsilon_2}}^{-1}(t, x)) \right) \exp\left(\int_0^t \mathcal{G}_1(\tau, t, x) d\tau\right) dx \\ &\leq |O_t| \|v_{\varepsilon_2}^0\|_{W^{1,\infty}(\mathbb{R}^d)} e^{\gamma T} |X_{v_{\varepsilon_1}}^{-1}(t, x) - X_{v_{\varepsilon_2}}^{-1}(t, x)| \\ &\leq \tilde{C}_T \|v_{\varepsilon_1} - v_{\varepsilon_2}\|_1,\end{aligned}\tag{5.32}$$

where we have used (5.24) on the second line<sup>1</sup>.

From the definition of  $\mathcal{G}_u(t, x)$ , we observe that

$$\begin{aligned}|\Delta E(s, t, x)| &\leq e^{\gamma T} \int_s^t |\mathcal{G}_1(\tau, t, x) - \mathcal{G}_2(\tau, t, x)| d\tau \\ &\leq \tilde{C}_T \|v_{\varepsilon_1} - v_{\varepsilon_2}\|_1,\end{aligned}\tag{5.33}$$

<sup>1</sup>This term is responsible for the possible loss of regularity for  $t > 0$ : In order to prove that  $v_\varepsilon$  is a Cauchy sequence in  $\mathcal{C}([0, T], W^{1,1}(\mathbb{R}^d))$ , we would need a  $W^{2,\infty}(\mathbb{R}^d)$  estimate over  $v_\varepsilon^0$ , which we do not have.

where  $\tilde{C}_T$  depends on  $\bar{\rho}$ ,  $T$ , the derivatives of  $R$ ,  $a$  and  $\psi_a$ , and on the constant appearing in (5.24).

Using the change of variables  $y = X_{v_{\varepsilon_2}}(s, X_{v_{\varepsilon_2}}^{-1}(t, x))$  and recalling that

$$|J_{X_{v_{\varepsilon_2}}(t, X_{v_{\varepsilon_2}}^{-1}(s, x))}| = e^{\int_0^t \operatorname{div} \mathcal{A}_u(\tau, X_{v_{\varepsilon_2}}(\tau, X_{v_{\varepsilon_2}}^{-1}(s, x))) d\tau - \int_0^s \operatorname{div} \mathcal{A}_{v_{\varepsilon_2}}(\tau, x) d\tau},$$

we see that

$$\begin{aligned} \int_{\mathbb{R}^d} \mathcal{M}_2(s, t, x, z) dx &= \int_{\mathbb{R}^d} m(s, y, z, (I_d v_{\varepsilon_2})(s, y)) |J_{X_{v_{\varepsilon_2}}(t, X_{v_{\varepsilon_2}}^{-1}(s, y))}|^{-1} dy \\ &\leq e^{2\bar{a}T} \int_{\mathbb{R}^d} \sup_{s, z, I} m_{\varepsilon_i}(s, y, z, I) dy \leq e^{2\bar{a}T} |\mathcal{K}| \bar{M}. \end{aligned}$$

Therefore, we have the bounds

$$\int_{\mathbb{R}^d} v^0(X_{v_{\varepsilon_2}}^{-1}(t, x)) |\Delta E(0, t, x)| dx \leq e^{\bar{a}T} \|v^0\|_{L^1(\mathbb{R}^d)} \tilde{C}_T \|v_{\varepsilon_1} - v_{\varepsilon_2}\|_1, \quad (5.34)$$

$$\begin{aligned} \int_0^t \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathcal{M}_2(s, t, x, z) \left( \Delta V_{\varepsilon_1 \varepsilon_2}(s, z) \right) dz \exp \left( \int_s^t \mathcal{G}_1(\tau, t, x) d\tau \right) dx ds &\leq e^{(\gamma+2\bar{a})T} \\ &\times |\mathcal{K}| \bar{M} \|v_{\varepsilon_1} - v_{\varepsilon_2}\|_1, \end{aligned} \quad (5.35)$$

$$\int_0^t \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathcal{M}_2(t, x, z) v_{\varepsilon_j}(s, z) dz \left( |\Delta E(s, t, x)| \right) dx ds \leq e^{2\bar{a}T} |\mathcal{K}| \bar{M} \bar{\rho} T \tilde{C}_T \|v_{\varepsilon_1} - v_{\varepsilon_2}\|_1. \quad (5.36)$$

The function  $m(t, x, z, I)$  having a compact support on the  $x$  variable, leads to

$$\mathcal{M}_1(s, t, x, z) - \mathcal{M}_2(s, t, x, z) = 0, \text{ for all } x \notin \mathcal{K} + B_{2r_t}.$$

On the other hand, the function  $m$  being differentiable and Lipschitz, implies that, for all  $x \in \mathcal{K} + B_{2r_t}$

$$\begin{aligned} |\mathcal{M}_1(s, t, x, z) - \mathcal{M}_2(s, t, x, z)| &\leq \|m\|_{W_x^{1, \infty}} |X_{v_{\varepsilon_1}}(s, X_{v_{\varepsilon_1}}^{-1}(t, x)) - X_{v_{\varepsilon_2}}(s, X_{v_{\varepsilon_2}}^{-1}(t, x))| \\ &\quad + \mu |(I_d v_{\varepsilon_1})(s, X_{v_{\varepsilon_1}}(s, X_{v_{\varepsilon_1}}^{-1}(t, x))) - (I_d v_{\varepsilon_2})(s, X_{v_{\varepsilon_2}}(s, X_{v_{\varepsilon_2}}^{-1}(t, x)))| \\ &\leq (\|m\|_{W_x^{1, \infty}} + \|\psi_d\|_{W_x^{1, \infty}} \bar{\rho}) |X_{v_{\varepsilon_1}}(s, X_{v_{\varepsilon_1}}^{-1}(t, x)) - X_{v_{\varepsilon_2}}(s, X_{v_{\varepsilon_2}}^{-1}(t, x))| \\ &\quad + \|\psi_d\|_{L^\infty} \|v_{\varepsilon_1} - v_{\varepsilon_2}\|_{L^1(\mathbb{R}^d)}. \end{aligned}$$

Using first Lemma 11 and then Lemma 14, we conclude that there exists a constant  $\tilde{C}_T$  such that

$$\begin{aligned} |\mathcal{M}_1(s, t, x, z) - \mathcal{M}_2(s, t, x, z)| &\leq \tilde{C}_T \left( |X_{v_{\varepsilon_1}}^{-1}(t, x) - X_{v_{\varepsilon_2}}^{-1}(t, x)| + \|v_{\varepsilon_1} - v_{\varepsilon_2}\|_1 + \|v_{\varepsilon_1} - v_{\varepsilon_2}\|_{L^1(\mathbb{R}^d)} \right) \\ &\leq \tilde{C}_T \left( \|v_{\varepsilon_1} - v_{\varepsilon_2}\|_1 + \|v_{\varepsilon_1} - v_{\varepsilon_2}\|_{L^1(\mathbb{R}^d)} \right). \end{aligned}$$

Therefore, we have the bound

$$\begin{aligned} & \int_0^t \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \left( \mathcal{M}_1(s, t, x, z) - \mathcal{M}_2(s, t, x, z) \right) v_{\varepsilon_1}(s, z) dz \exp \left( \int_s^t \mathcal{G}_1(\tau, t, x) d\tau \right) dx ds \\ & \leq |\mathcal{K} + B_{2r_T}| \tilde{\rho} \tilde{C}_T e^{\gamma T} \|v_{\varepsilon_1} - v_{\varepsilon_2}\|_1. \end{aligned} \quad (5.37)$$

Putting together the bounds (5.31) through (5.37), we get

$$\|v_{\varepsilon_1} - v_{\varepsilon_2}\|_{L^1(\mathbb{R}^d)} \leq \tilde{C}_T \left( \|v_{\varepsilon_1}^0 - v_{\varepsilon_2}^0\|_{L^1(\mathbb{R}^d)} + \int_0^T \|v_{\varepsilon_1} - v_{\varepsilon_2}\|_{L^1(\mathbb{R}^d)} ds \right).$$

Thanks to Grönwall's lemma, we have then the relation

$$\sup_{t \in [0, T]} \|v_{\varepsilon_1} - v_{\varepsilon_2}\|_{L^1(\mathbb{R}^d)} \leq \tilde{C}_T \|v_{\varepsilon_1}^0 - v_{\varepsilon_2}^0\|_{L^1(\mathbb{R}^d)}$$

for some  $\tilde{C}_T$  independent of  $\varepsilon_1$  and  $\varepsilon_2$ , which proves that, up to the extraction of a sub-sequence,  $v_\varepsilon$  is a Cauchy sequence in  $\mathcal{C}([0, T], L^1(\mathbb{R}^d))$ . Therefore, there exists  $v \in \mathcal{C}([0, T], L^1(\mathbb{R}^d))$  such that

$$\lim_{\varepsilon \rightarrow 0} \sup_{t \in [0, T]} \|v_\varepsilon - v\|_{L^1(\mathbb{R}^d)} = 0.$$

Furthermore, such function satisfies the bounds (5.25) and (5.27).

We claim now that the sequence  $L_{v_\varepsilon}^* \varphi$  converges to  $L_v^* \varphi$  in  $L^\infty([0, T] \times \mathbb{R}^d)$  for all  $\varphi \in \mathcal{C}_c^1([0, T] \times \mathbb{R}^d)$ . This is a direct consequence of the relation

$$|L_{v_\varepsilon}^* \varphi - L_v^* \varphi| \leq (L_r \|\psi_g\|_{L^\infty} + \mu \|\psi_d\|_{L^\infty}) \|\varphi\|_{L^\infty(\mathbb{R}^d)} \|v^\varepsilon - v\|_{L^1(\mathbb{R}^d)}.$$

In order to conclude, we recall that all classical solutions are weak solutions, and therefore, for all  $\varepsilon > 0$

$$\int_0^T \int_{\mathbb{R}^d} v_\varepsilon L_{v_\varepsilon}^* \varphi dx dt = \int_{\mathbb{R}^d} v_\varepsilon^0 \varphi dx,$$

and taking the limit when  $\varepsilon$  goes to 0 we see that  $v$  is a weak solution of problem (5.19).  $\square$

Considering initial data with compact support might be enough in order to model most of the biological scenarios found in nature. However, the hypothesis  $v^0 \in W^{k, \infty}(\mathbb{R}^d)$  might be too restrictive for some real life scenarios. Furthermore, the study of the problem when more general initial conditions are present, is of theoretical interest. We show below that, when  $\partial_I a = 0$ , a solution exists for any initial data  $v^0 \in W^{k, 1}(\mathbb{R}^d)$ ,  $k \geq 1$ .

**Theorem 12.** *Under hypothesis (5.15) through (5.18), if  $\partial_I a = 0$ , for all non-negative functions  $v^0 \in W^{k, 1}(\mathbb{R}^d)$ , there exists a unique non-negative weak solution  $v \in \mathcal{C}([0, T], W^{k, 1}(\mathbb{R}^d))$  of problem (5.19). Furthermore, such a solution satisfies*

$$\sup_{t \in [0, T]} \|v\|_{L^1(\mathbb{R}^d)} \leq \max \left\{ \|v^0\|_{L^1(\mathbb{R}^d)}, \frac{I^*}{\underline{\psi}_g} \right\}, \quad (5.38)$$

$$\sup_{t \in [0, T]} \|v\|_{W^{k, \infty}(\mathbb{R}^d)} \leq C_T \|v^0\|_{W^{k, 1}(\mathbb{R}^d)}. \quad (5.39)$$

*Proof.* As in the proof for  $k = 1$  when  $\partial_I a \neq 0$ , we can approximate any function  $v^0 \in W^{k, 1}$  by a smooth, compactly supported sequence of functions  $v_\varepsilon^0$ . The same arguments as in the previous proof will show that  $v_\varepsilon$ , the sequence of solutions associated to  $v_\varepsilon^0$ , is a Cauchy sequence in  $\mathcal{C}([0, T], L^1(\mathbb{R}^d))$ . Furthermore, given that the second term in (5.30), which is responsible for the possible loss of regularity in the previous case, is equal 0 when  $\partial_I a = 0$ , we show that  $v_\varepsilon$  is a Cauchy sequence in  $\mathcal{C}([0, T], W^{k, 1}(\mathbb{R}^d))$  as well. We prove as in Theorem 11 that the limit of  $v_\varepsilon$  is the required weak solution.  $\square$



Given that the regularity of the solution varies depending on whether  $\partial_I a = 0$  or  $\partial_I a \neq 0$ , and that such regularity will be of importance in the upcoming sections, we define the parameter

$$\kappa := \begin{cases} k - 1, & \text{if } \partial_I a \neq 0, \\ k, & \text{if } \partial_I a = 0, \end{cases}$$

which encompasses the information over said regularity.

We remark that if we had  $\partial_I m = 0$ , we might obtain existence for a larger class of mutation functions  $m$ . For conciseness, we will however not consider such cases in the present work.

### 5.3 Particle Method

The particle method basically consists in searching for an approximate solution of problem (5.19) which is a sum of weighted Dirac masses.

Throughout the following section we suppose

$$\psi_a \in \mathcal{C} \left( [0, T] \times \mathbb{R}_x^d, W^{1,\infty}(\mathbb{R}^d) \right) \cap \mathcal{C} \left( [0, T] \times \mathbb{R}_y, \mathcal{C}^2(\mathbb{R}_x^d) \cap W^{2,\infty}(\mathbb{R}_x^d) \right), \quad (5.40)$$

$$0 < \underline{\psi}_g \leq \psi_g \in \mathcal{C} \left( [0, T] \times \mathbb{R}_x^d, W^{1,\infty}(\mathbb{R}_y^d) \right) \cap \mathcal{C} \left( [0, T] \times \mathbb{R}_y^d, \mathcal{C}^1(\mathbb{R}_x^d) \cap W^{1,\infty}(\mathbb{R}_x^d) \right) \quad (5.41)$$

$$\psi_d \in \mathcal{C} \left( [0, T] \times \mathbb{R}_x^d, W^{1,\infty}(\mathbb{R}_y^d) \right) \cap \mathcal{C} \left( [0, T] \times \mathbb{R}_y^d, \mathcal{C}^1(\mathbb{R}_x^d) \cap W^{1,\infty}(\mathbb{R}_x^d) \right). \quad (5.42)$$

Notice that, unlike the set of hypotheses (5.15)-(5.17), we have imposed  $W^{1,\infty}(\mathbb{R}^d)$  regularity for the  $y$  variable, which is needed in order to approximate the integral terms by sums over a countable set.

Consider as well  $m \geq 0$ ,

$$m \in \mathcal{C} \left( [0, T] \times \mathbb{R}_x^d \times \mathbb{R}_y^d, W^{1,\infty}(\mathbb{R}_I) \right) \cap \mathcal{C} \left( [0, T] \times \mathbb{R}_x^d \times \mathbb{R}_I, W^{1,\infty}(\mathbb{R}_y^d) \right) \cap \mathcal{C} \left( [0, T] \times \mathbb{R}_y^d \times \mathbb{R}_I, \mathcal{C}_c^1(\mathbb{R}_x^d) \right) \quad (5.43)$$

satisfying hypotheses (5.7) through (5.18).

For  $h > 0$ , consider a countable set of indices  $\mathcal{J}_h \in \mathbb{Z}^d$ , points  $x_i^0 \in \mathbb{R}^d$  and weights  $w_i^0$  for  $i \in \mathcal{J}_h$ . The weights  $w_i^0$  can be regarded as the respective masses of a collection of subsets  $\Omega_i^0 \subset \mathbb{R}^d$  satisfying

$$\Omega_i^0 \cap \Omega_j^0 = \emptyset, \text{ if } i \neq j, \text{ and } \bigcup_{i \in \mathcal{J}_h} \Omega_i^0 = \mathbb{R}^d. \quad (5.44)$$

For example, we may choose the  $\Omega_i^0$  as the set of all non intersecting cubes of side length equal  $h$  having the points  $hi$  as centers,  $i \in \mathbb{Z}^d$ . This way,  $w_i^0 = h^d$ , with each of the  $x_i^0$  being a point in  $\Omega_j^0$ . In general we assume that there exist positive constants  $c$  and  $C$  such that

$$ch \leq |x_i^0 - x_j^0| \leq Ch, \quad \forall i \neq j, \quad (5.45)$$

$$ch^d \leq w_i^0 \leq Ch^d, \quad \forall i \in \mathcal{J}_h. \quad (5.46)$$

Following [46], the particle method then consists in looking for a measure  $\nu_h$  of the form

$$\nu_h(t) = \sum_{i \in \mathcal{J}_h} \nu_i(t) w_i(t) \delta_{x_i(t)},$$

where  $(\nu := \{\nu_i(t)\}_{i \in \mathcal{J}_h}, w := \{w_i(t)\}_{i \in \mathcal{J}_h}, \bar{x} := \{x_i(t)\}_{i \in \mathcal{J}_h})$ , is the solution of the following system

$$\begin{cases} \dot{x}_i(t) &= A_{\nu,w}(t, x_i), \\ \dot{w}_i(t) &= \operatorname{div} A_{\nu,w}(t, x_i(t))w_i(t), \\ \dot{\nu}_i(t) &= (-\operatorname{div} A_{\nu,w}(t, x_i(t)) + R(t, x_i(t), I_g(t, x_i(t), \nu, w)))\nu_i(t) \\ &\quad + \sum_{j \in \mathcal{J}_h} w_j(t)\nu_j(t)m(t, x_i(t), x_j(t), I_d(t, x_i(t), \nu, w)), \\ x_i(0) &= x_i^0, w_i(0) = w_i^0, \nu_i(0) = \nu_i^0(x_i^0), \end{cases} \quad (5.47)$$

where

$$A_{\nu,w}(t, x) = a(t, x, I_a(t, x, \nu, w)),$$

and

$$I_l(t, x, \nu, w) := \sum_{j \in \mathcal{J}_h} \nu_j(t)w_j(t)\psi_l(t, x, x_j(t)),$$

with  $l \in \{a, g, d\}$ .

In what follows we assume that  $h$  and  $x_k^0$  are chosen in such a way that

$$\|v^0\|_{1,h} := \sum_{i \in \mathcal{J}_h} v^0(x_i^0)w_i^0 < \infty. \quad (5.48)$$

We define the subset of indices

$$\mathcal{J}_h^m := \{i \in \mathcal{J}_h : x_i^0 \in \operatorname{supp} m + B_{\|a\|_{L^\infty T}}\},$$

where  $B_{\|a\|_{L^\infty T}}$  is the ball of radius  $\|a\|_{L^\infty T}$ . The compact support of  $m$  implies that  $|\mathcal{J}_h^m| < \infty$ .

For a positive value of  $h$  and a set of indexes  $\mathcal{J}_h$  we define the functional spaces

$$\begin{aligned} l^1(\mathcal{J}_h) &:= \{u = \{u_i\}_{i \in \mathcal{J}_h} : \sum_{i \in \mathcal{J}_h} |u_i| < +\infty\}, \\ l^\infty(\mathcal{J}_h) &:= \{w = \{w_i\}_{i \in \mathcal{J}_h} : \sup_{i \in \mathcal{J}_h} |w_i| < +\infty\}. \end{aligned}$$

We equip these spaces with the norms

$$\|u\|_{l^1} := \sum_{i \in \mathcal{J}_h} |u_i| \quad \text{and} \quad \|w\|_{l^\infty} := \sup_{i \in \mathcal{J}_h} |w_i|$$

respectively. It is clear that for all  $u \in l^1(\mathcal{J}_h)$  and  $v \in l^\infty(\mathcal{J}_h)$ , then  $uv \in l^1(\mathcal{J}_h)$ . For  $T > 0$  we define as well the spaces

$$X_h^T := \mathcal{C}([0, T], l^1(\mathcal{J}_h)) \quad \text{and} \quad Y_h^T := \mathcal{C}([0, T], l^\infty(\mathcal{J}_h)),$$

equipped with the norms

$$\|u\|_{1,h} := \sup_{t \in [0, T]} \|u(t)\|_{l^1} \quad \text{and} \quad \|w\|_{\infty,h} := \sup_{t \in [0, T]} \|w(t)\|_{l^\infty}.$$

Problem (5.47) is a strongly coupled system of ODEs, with an infinite number of unknowns and equations. In some cases the system becomes uncoupled (for example if  $\partial_I a = 0$ ) or with a finite number of equations and unknowns (for example if  $v^0$ ,  $a$  and  $m$  have compact support), however, for the sake of generality, we present below the proof of existence of solution in the general case, and later discuss briefly these particular scenarios.

We start by giving two results that will be of great use for the proof of existence for problem (5.47). First, we deal with the existence of solution for a simpler system of infinite equations with infinitely many unknowns:

**Lemma 17.** *Consider  $a \in \mathcal{C}([0, T], (W^{1, \infty}(\mathbb{R}^{d+1})))^d$ ,  $u \in X_h^T$ ,  $w \in Y_h^T$  and  $\psi_a$  satisfying hypothesis (5.40). Then there exists a unique family of functions  $x := \{x_i\}_{i \in \mathcal{J}_h}$ ,  $x_i \in \mathcal{C}^1([0, T])$  for all  $i \in \mathcal{J}_h$  which is solution of the system of equations*

$$\dot{x}_i(t) = A_{u,w}(t, x_i), \quad t \in [0, T], \quad x_i(0) = x_i^0. \quad (5.49)$$

When  $\partial_I a = 0$ , system (5.49) becomes uncoupled, each individual equation has a solution, thanks to the classic Cauchy-Lipschitz theory. The proof of the general case is given in Appendix C.2.

The second auxiliary result comes from approximation theory, and it will also be of great use in Section 5.4:

**Lemma 18.**

$$\forall \varphi \in W^{k,1}(\mathbb{R}^d), \quad \left| \int_{\mathbb{R}^d} \varphi(x) dx - \sum_{i \in \mathcal{J}_h} w_i(t) \varphi(x_i(t)) \right| \leq Ch^k \|\varphi\|_{k,1},$$

where  $C$  is a constant which depends on  $a$ ,  $\psi_a$ ,  $\|vw\|_{1,h}$  and  $T$ .

This result is a direct corollary of Lemma 8 in [115]. More details regarding its proof are given in Appendix C.3.

From now on, we suppose  $h$  to be small enough so that for any  $t \in [0, T]$ ,

$$\sum_{i \in \mathcal{J}_h} w_i(t) m(t, x_i(t), y, I_d(t, x_i(t), \nu, w)) < K + \frac{r^*}{2}, \quad (5.50)$$

where the values of  $K$  and  $r^*$  are given in (5.11) and (5.18) respectively. Such a choice is always possible thanks to Lemma 18.

**Theorem 13.** *Under hypothesis (5.5) through (5.18) and (5.40) through (5.46), for all  $T > 0$  and all non-negative initial data  $v^0 \in l^1(\mathcal{J}_h, \Omega^0)$  there exists a unique solution  $x_i \in \mathcal{C}^1([0, T])$ , for all  $i \in \mathcal{J}_h$ ,  $w := \{w_i(\cdot)\}_{i \in \mathcal{J}_h} \in \mathcal{C}([0, T], l^\infty(\mathcal{J}_h))$  and  $0 \leq \nu := \{\nu_i(\cdot)\}_{i \in \mathcal{J}_h} \in \mathcal{C}([0, T], l^1(\mathcal{J}_h))$  of problem (5.47). Furthermore, there exist positive constants  $c_T$  and  $C_T$  such that the solution satisfies, for all  $t \in [0, T]$*

$$c_T h \leq |x_i(t) - x_j(t)| \leq C_T h, \quad \forall i, j \in \mathcal{J}_h, \quad i \neq j, \quad (5.51)$$

$$c_T h^d \leq w_i(t) \leq C_T h^d, \quad \forall i \in \mathcal{J}_h, \quad (5.52)$$

$$\|vw\|_{1,h} \leq \max\{\|v^0 h^d\|_{l^1}, \frac{I^*}{\underline{\psi}_g}\}. \quad (5.53)$$

*Proof.* Consider  $v^0 \in l^1(\mathcal{J}_h, \Omega^0)$ , satisfying  $v^0 \geq 0$ . Consider as well  $\alpha > 1$ , and define

$$\bar{\rho}_\alpha := \max\{\alpha h^d \|v^0\|_{l^1}, \frac{I^*}{\underline{\psi}_g}\},$$

$$\tilde{a} := \|a\|_{W_x^{1,\infty} L_{I,t}^\infty} + \|a\|_{W_I^{1,\infty} L_{t,x}^\infty} \|\psi_a\|_{W_x^{1,\infty}} C_\alpha.$$

For  $T > 0$  we define the set

$$D_\alpha^T := \{(u, w) \in X_h^T \times Y_h^T : \|uw\|_{1,h} \leq \bar{\rho}_\alpha, \forall t \in [0, T], u(t) \geq 0, w(t) \geq 0, h^d e^{-\bar{a}t} \leq w_k(t) \leq h^d e^{\bar{a}t}\}.$$

For any  $(u, w) \in D_\alpha^T$  we introduce the problem, for  $t \in [0, T]$ ,

$$\begin{cases} \dot{x}_i(t) &= A_{u,w}(t, x_i), \\ \dot{\omega}_i(t) &= \operatorname{div} A_{u,w}(t, x_i(t))\omega_i(t), \\ \dot{\nu}_i(t) &= (-\operatorname{div} A_{u,w}(t, x_i(t)) + R(t, x_i(t), I_g(t, x_i(t), u, w)))\nu_i(t) \\ &\quad + \sum_{j \in \mathcal{J}_h} \omega_j(t)u_j(t)m(t, x_i(t), x_j(t), I_d(t, x_i(t), u, w)), \\ x_i(0) &= x_i^0, \omega_i(0) = w_i^0, \nu_i(0) = v^0(x_i^0). \end{cases} \quad (5.54)$$

We denote  $(\nu, \omega) = \Phi(u, w)$ .

For each pair  $(u, w)$ , the existence and uniqueness of  $x_i$  is immediate from Lemma 17. Furthermore, for all values of  $i$ , we have the following explicit expression for  $\omega_i$

$$\omega_i(t) = w_i^0 e^{\int_0^t \operatorname{div} A_{u,w}(s, x_i(s)) ds},$$

which satisfies, for any  $t \in [0, T]$

$$h^d e^{-\bar{a}t} \leq \omega_i(t) \leq h^d e^{\bar{a}t}.$$

On the other hand, for all  $(u, w) \in D_\alpha^T$  and all values of  $i$ , the right-hand side of the differential equation in (5.54) is well defined, as we have for all  $t \in [0, T]$

$$\sum_{j \in \mathcal{J}_h} w_j(t)u_j(t)m(t, x_i(t), x_j(t), I_d(t, x_i(t), u, w)) \leq \bar{M} \sum_{j \in \mathcal{J}_h} w_j(t)u_j(t) \leq \bar{M}\bar{\rho}_\alpha.$$

Therefore, the expression for  $\nu_i$  is given by

$$\nu_i(t) = v^0(x_i^0) e^{\int_0^t \mathcal{G}_i(s) ds} + \int_0^t \sum_{j \in \mathcal{J}_h} w_j(s)u_j(s)m(s, x_i(s), x_j(s), I_d(s, x_i(s), u, w)) e^{\int_s^t \mathcal{G}_i(\tau) d\tau} ds, \quad (5.55)$$

where

$$\mathcal{G}_i(t) := -\operatorname{div} A_{u,w}(t, x_i(t)) + R(t, x_i(t), I_g(t, x_i(t), u(t), w(t))),$$

satisfies

$$\sup_{t \in [0, T]} |\mathcal{G}_i(t)| \leq \gamma := \|R\|_{L_{t,x,I}^\infty} + \|a\|_{W_x^{1,\infty} L_{t,I}^\infty} + \|a\|_{W_I^{1,\infty} L_{t,x}^\infty} \|\psi_a\|_{W_x^{1,\infty} L_y^\infty} \bar{\rho}_\alpha.$$

The positiveness of  $\nu$  is immediate from the positiveness of  $v^0$  and  $m$ .

Furthermore, given that  $|x_i(t) - x_i^0| \leq \|a\|_{L^\infty} T$ , for all  $k \in \mathcal{J}_h$  and  $t \in [0, T]$ , we have

$$m(t, x_i(s), y, I) = 0,$$

for all  $i \notin \mathcal{J}_h^m$ ,  $t \in [0, T]$ ,  $y \in \mathbb{R}^d$  and  $I \in \mathbb{R}$ . As a result, multiplying (5.55) by  $\omega_i(t)$  for each  $i$  and adding for all values of  $i$ , we obtain

$$\begin{aligned}
\sum_{i \in \mathcal{J}_h} \nu_i(t) \omega_i(t) &\leq e^{(\gamma + \tilde{\alpha})T} \left( \sum_{i \in \mathcal{J}_h} v^0(x_i^0) h^d \right. \\
&\quad \left. + h^d \int_0^t \sum_{i \in \mathcal{J}_h} \sum_{j \in \mathcal{J}_h} w_j(s) u_j(s) m(s, x_i(s), x_j(s), I_d(s, x_i(s), u, w)) ds \right) \\
&= e^{(\gamma + \tilde{\alpha})T} \left( \sum_{i \in \mathcal{J}_h} v^0(x_i^0) h^d \right. \\
&\quad \left. + h^d \int_0^t \sum_{i \in \mathcal{J}_h^m} \sum_{j \in \mathcal{J}_h} w_j(s) u_j(s) m(s, x_i(s), x_j(s), I_d(s, x_i(s), u, w)) ds \right) \\
&\leq e^{(\gamma + \tilde{\alpha})T} \left( \sum_{i \in \mathcal{J}_h} v^0(x_i^0) h^d + h^d |\mathcal{J}_h^m| \overline{M} \int_0^t \sum_{j \in \mathcal{J}_h} w_j(s) u_j(s) ds \right) \\
&\leq e^{(\gamma + \tilde{\alpha})T} \left( \frac{1}{\alpha} + TK_h \right) \bar{\rho}_\alpha,
\end{aligned}$$

where  $K_h := h^d |\mathcal{J}_h^m| \overline{M}^2$ . Thanks to the condition  $\alpha > 1$  there exists  $T_\alpha$  (only depending on  $\alpha$  and on the coefficients of the problem) such that  $\Phi : D_\alpha^T \rightarrow D_\alpha^T$ , for all  $T \leq T_\alpha$ . We now prove that there exists  $T \in (0, T_\alpha)$  such that  $\Phi$  is a contraction over  $D_\alpha^T$ .

*Step 1: Bounds over  $x = \{x_i\}_{i \in \mathcal{J}_h}$*

Let  $(u^1, w^1)$  and  $(u^2, w^2)$  be two pairs in  $D_\alpha^T$ , and let  $x^1, x^2$  be the respective solutions of

$$\dot{x}_i^j = A_{uj, w^j}(t, x_i^j).$$

By following the same ideas as in the proof of Lemma 11 (see Appendix C.1), we obtain that for all  $t \in [0, T]$ ,

$$\|x^1(t) - x^2(t)\|_{\infty, h} \leq C(T, h) (\|w^1 - w^2\|_{\infty, h} + \|u^1 - u^2\|_{1, h}),$$

where the constant  $C(T, h)$  satisfies  $\lim_{T \rightarrow 0} C(T, h) = 0$ .

*Step 2: Bounds over  $\omega = \{\omega_i\}_{i \in \mathcal{J}_h}$*

From the expression for  $\omega$ , we get, for all  $t \in [0, T]$

$$\begin{aligned}
|\omega_i^1(t) - \omega_i^2(t)| &\leq h^d T e^{\tilde{\alpha}T} \left( \|a\|_{W_{x, I}^{2, \infty}} (1 + |\partial_x I_a(t, x_i^2, u^2, w^2)|) \|x^1 - x^2\|_{\infty, h} \right. \\
&\quad \left. + \|a\|_{W_{x, I}^{2, \infty}} (1 + |\partial_x I_a(t, x_i^2, u^2, w^2)|) |I_a(t, x_i^1, u^1, w^1) - I_a(t, x_i^2, u^2, w^2)| \right. \\
&\quad \left. + \|a\|_{W_I^{1, \infty}} |\partial_x I_a(t, x_i^1, u^1, w^1) - \partial_x I_a(t, x_i^2, u^2, w^2)| \right).
\end{aligned}$$

On the other hand we have

$$\begin{aligned}
|\partial_x I_a(t, x_i^2, u^2, w^2)| &\leq \|\psi_a\|_{W_{x, y}^{1, \infty}} \bar{\rho}_\alpha, \\
|I_a(t, x_i^1, u^1, w^1) - I_a(t, x_i^2, u^2, w^2)| &\leq \bar{\rho}_\alpha \|\psi_a\|_{W_{x, y}^{1, \infty}} \|x^1 - x^2\|_{\infty, h}, \\
&\quad + \|\psi_a\|_{L_{x, y}^\infty} e^{\tilde{\alpha}T} (h^d \|u^1 - u^2\|_{1, \infty} + \frac{\bar{\rho}_\alpha}{h^d} \|w^1 - w^2\|_{\infty, h}),
\end{aligned}$$

<sup>2</sup>Notice that  $h^d |\mathcal{J}_h^m| \approx |\text{supp } m + B|_{\|a\|_{L^\infty T}}$

and

$$|\partial_x I_\alpha(t, x_i^1, u^1, w^1) - \partial_x I_\alpha(t, x_i^2, u^2, w^2)| \leq \bar{\rho}_\alpha \|\psi_\alpha\|_{W_{x_i}^{2,\infty}} \|x^1 - x^2\|_{\infty,h} \\ + \|\psi_\alpha\|_{W_x^{1,\infty}} e^{\bar{\alpha}T} (h^d \|u^1 - u^2\|_{1,\infty} + \frac{\bar{\rho}_\alpha}{h^d} \|w^1 - w^2\|_{\infty,h}).$$

In conclusion, there exists a constant  $C(T, h)$ , satisfying  $\lim_{T \rightarrow 0} C(T, h) = 0$  such that

$$\|\omega^1 - \omega^2\|_{\infty,h} \leq C(T, h) (\|w^1 - w^2\|_{\infty,h} + \|u^1 - u^2\|_{1,h}).$$

*Step 3: Bounds over  $\nu = \{\nu_i\}_{i \in \mathcal{J}_h}$*

Using the expression for  $\nu$ , the regularity of  $m$  and bounds similar to those used for  $\omega$ , we see that there exists a constant  $C(T, h)$  satisfying  $\lim_{T \rightarrow 0} C(T, h) = 0$  such that

$$\|\nu^1 - \nu^2\|_{1,h} \leq C(T, h) (\|w^1 - w^2\|_{\infty,h} + \|u^1 - u^2\|_{1,h}).$$

Consequently, there exists a constant  $C(T, h)$  satisfying  $\lim_{T \rightarrow 0} C(T, h) = 0$ , such that

$$\|\omega^1 - \omega^2\|_{\infty,h} + \|\nu^1 - \nu^2\|_{1,h} \leq C(T, h) (\|w^1 - w^2\|_{\infty,h} + \|u^1 - u^2\|_{1,h}),$$

which implies that, for  $0 < T_1 \leq T_\alpha$  small enough,  $\Phi$  is a contraction over  $D_\alpha^{T_1}$ , and therefore it has a unique fixed point. Such fixed point is a solution of problem (5.47) over  $[0, T_1]$ .

We now claim that the solution exists for  $T$  arbitrary, and furthermore, it satisfies the relation (5.53). Let  $T_f$  be the maximal time of existence of solution. Suppose that there exists  $t_0 \in (0, T_f]$  such that  $\|\nu w\|_{1,h} > \bar{\rho}_\alpha$ . This implies that there exist  $\delta \geq 0$  and  $t^* > 0$  such that for a certain finite subset of  $\mathcal{J}_h$ , that we denote as  $\mathcal{K}_h$ , the following statements are true:

$$\sum_{i \in \mathcal{K}_h} \nu_i(t) w_i(t) \leq \bar{\rho}_\alpha, \quad \forall t \in [t^* - \delta, t^*], \\ \sum_{i \in \mathcal{K}_h} \nu_i(t) w_i(t) > \bar{\rho}_\alpha, \quad \forall t \in (t^*, t^* + \delta].$$

This implies the existence of  $t_1 \in [t^*, t^* + \delta]$  such that the following properties are satisfied simultaneously

$$\sum_{i \in \mathcal{K}_h} \nu_i(t_1) w_i(t_1) \geq \bar{\rho}_\alpha \quad \text{and} \quad \left( \sum_{i \in \mathcal{K}_h} \nu_i w_i \right)' (t_1) \geq 0. \quad (5.56)$$

Multiplying the equation satisfied by  $\nu_i(t)$  by  $w_i(t)$  we get the relation

$$\dot{\nu}_i(t) w_i(t) = (-\operatorname{div} A_{\nu,w}(t, x_i(t)) + R(t, x_i(t), I_g(t, x_i(t), \nu, w))) \nu_i(t) w_i(t) \\ + w_i(t) \sum_{j \in \mathcal{J}_h} w_j(t) \nu_j(t) m(t, x_i(t), x_j(t), I_g(t, x_i(t), \nu, w)),$$

while, directly from the equation for  $w_i(t)$  we deduce

$$\nu_i(t) \dot{w}_i(t) = \operatorname{div} A_{\nu,w}(t, x_i(t)) \nu_i(t) w_i(t).$$

Therefore, adding both relations for  $i \in \mathcal{K}_h$  and using (5.50), we get

$$\begin{aligned} \left( \sum_{i \in \mathcal{K}_h} \nu_i w_i \right)' (t) &= \left( \sum_{i \in \mathcal{K}_h} R(t, x_i(t), I_g(t, x_i(t), \nu, w)) \nu_i(t) w_i(t) \right) \\ &\quad + \sum_{j \in \mathcal{J}} w_j(t) \nu_j(t) \sum_{i \in \mathcal{K}_h \cap \mathcal{J}_h^m} m(t, x_i(t), x_j(t), I_d(t, x_i(t), \nu, w)) w_i(t) \\ &\leq \left( \sum_{i \in \mathcal{K}_h} R(t, x_i(t), I_g(t, x_i(t))) \nu_i(t) w_i(t) \right) + (K + \frac{r^*}{2}) \sum_{j \in \mathcal{K}_h} w_j(t) \nu_j(t). \end{aligned}$$

Given that  $\|\nu(t_1)w(t_1)\|_{l^1} \geq \bar{\rho}_\alpha \geq \frac{I^*}{\underline{\psi}_g}$ , then  $I_g(t_1, x_i(t_1), \nu(t_1), w(t_1)) \geq I^*$ , and consequently

$$R(t_1, x_i(t_1), I_g(t_1, x_i(t_1), \nu(t_1), w(t_1))) < -r^* - K,$$

for all values of  $i$ , which in turn implies that

$$\left( \sum_{i \in \mathcal{K}_h} \nu_i w_i \right)' (t_1) \leq -\frac{r^2}{2} \sum_{i \in \mathcal{K}_h} \nu_i(t_1) w_i(t_1) < 0,$$

which contradicts (5.56). Therefore,  $\|\nu w\|_{1,h} \leq \bar{\rho}_\alpha$  for all values of  $\alpha > 1$  and for all  $t \in (0, T_f)$ . We can then iterate the arguments used to prove existence of a solution, and conclude that the solution can be extended to any interval  $[0, T]$ . As  $\|\nu w\|_{1,h} \leq \bar{\rho}_\alpha$  for all  $t$ , independently of  $\alpha$ , taking the limit when  $\alpha$  goes to 1, we obtain (5.53).  $\square$

## 5.4 Convergence of the numerical solution towards a weak solution

We study now the conditions under which a solution of problem (5.54) converges towards a solution of problem (5.19), in a certain sense that will be defined later. We split our analysis in two cases: first, the study of convergence on a finite interval of time  $[0, T]$ . We will see that for any  $T > 0$ , the solution obtained through the particle method converges towards the solution of the PDE (5.19). However, the speed of convergence might be affected by the value of  $T$ . The second case we study is the asymptotic proximity of both solutions when  $t$  goes to  $\infty$ . This is a far more complex and interesting issue, and we show different examples exposing some of the behaviours that can be observed.

Directly from the study of existence of solutions for each problem, we notice that the sets of hypotheses we have used, do not coincide. We give a set of hypotheses which simultaneously guarantees the existence of solution for both problems, while taking into account the distinction of cases involved in the definition of  $\kappa$ .

For a certain  $T > 0$  and  $k > 0$ , we consider the functions

$$\psi_a \in \mathcal{C}([0, T] \times \mathbb{R}_x^d, W^{1,\infty}(\mathbb{R}_y^d)) \cap \mathcal{C}([0, T] \times \mathbb{R}_y^d, \mathcal{C}^{k+1}(\mathbb{R}_x^d) \cap W^{k+1,\infty}(\mathbb{R}_x^d)), \quad (5.57)$$

$$0 < \underline{\psi}_g \leq \psi_g \in \mathcal{C}([0, T] \times \mathbb{R}_x, W^{1,\infty}(\mathbb{R}_y^d)) \cap \mathcal{C}([0, T] \times \mathbb{R}_y, \mathcal{C}^\kappa(\mathbb{R}_x^d) \cap W^{\kappa,\infty}(\mathbb{R}_x^d)), \quad (5.58)$$

$$\psi_d \in \mathcal{C}([0, T] \times \mathbb{R}_x^d, W^{1,\infty}(\mathbb{R}_y^d)) \cap \mathcal{C}([0, T] \times \mathbb{R}_y^d, \mathcal{C}^\kappa(\mathbb{R}_x^d) \cap W^{\kappa,\infty}(\mathbb{R}_x^d)). \quad (5.59)$$

As in Section 5.2 we introduce

$$a \in \mathcal{C}([0, T], W^{k+1, \infty}(\mathbb{R}^{d+1})), \quad (5.60)$$

$$R \in \mathcal{C}\left([0, T] \times \mathbb{R}_x, W_{loc}^{\kappa, \infty}(\mathbb{R}_y^d)\right) \cap \mathcal{C}\left([0, T] \times \mathbb{R}_y, \mathcal{C}^\kappa(\mathbb{R}_x^d) \cap W^{\kappa, \infty}(\mathbb{R}_x^d)\right) \quad (5.61)$$

We consider as well  $m \geq 0$ ,

$$m \in \mathcal{C}\left([0, T] \times \mathbb{R}_x^d \times \mathbb{R}_y^d, W^{\kappa, \infty}(\mathbb{R}_I)\right) \cap \mathcal{C}\left([0, T] \times \mathbb{R}_x^d \times \mathbb{R}_I^d, W^{\kappa, \infty}(\mathbb{R}_y^d)\right) \cap \mathcal{C}\left([0, T] \times \mathbb{R}_y^d \times \mathbb{R}_I^d, \mathcal{C}_c^\kappa(\mathbb{R}_x^d)\right). \quad (5.62)$$

satisfying hypothesis (5.7) through (5.18).

Finally, we consider  $v^0 \in W^{k, 1}(\mathbb{R}^d)$  if  $\partial_I a = 0$  and  $v^0 \in W^{k, 1}(\mathbb{R}^d) \cap W^{k, \infty}(\mathbb{R}^d)$  with compact support otherwise.

### 5.4.1 Convergence on a finite time interval

We recall that the function  $v$  represents the solution of problem (5.19) while  $x_i$ ,  $w_i$  and  $\nu_i$ ,  $i \in \mathcal{J}_h$  represents that of problem (5.54). We recall as well that

$$\max\{\|v\|, \|\nu w\|_{1, h}\} \leq \max\{\|v^0\|_{L^1(\mathbb{R}^d)}, \|v^0 h^d\|_{L^1}, \frac{I^*}{\psi_g}\} =: \bar{\rho}.$$

Let  $\varepsilon > 0$ ,  $r \in \mathbb{R}$  and  $\varphi \in \mathcal{C}_c(\mathbb{R}^d)$  satisfy the following conditions

$$\int_{\mathbb{R}^d} \varphi(x) dx = 1, \quad (5.63)$$

$$\int_{\mathbb{R}^d} x^\alpha \varphi(x) dx = 0, \quad \forall \alpha \in \mathbb{N}^n, |\alpha| \leq r - 1. \quad (5.64)$$

. We define, for all  $t \in [0, T]$ ,  $x \in \mathbb{R}^d$

i)

$$v^h(t, x) = \sum_{i \in \mathcal{J}_h} \nu_i(t) w_i(t) \delta(x - x_i(t)), \quad (5.65)$$

a time dependent measure obtained as a sum of weighted Dirac deltas at  $x_i(t)$ ,

ii)

$$v_\varepsilon^h(t, x) = (v^h(t) * \varphi_\varepsilon)(x) = \sum_{i \in \mathcal{J}_h} \nu_i(t) w_i(t) \varphi_\varepsilon(x - x_i(t)), \quad (5.66)$$

a regular function obtained as the space convolution of  $v^h(t, x)$  and  $\varphi_\varepsilon(x)$ , where

$$\varphi_\varepsilon := \frac{1}{\varepsilon^d} \varphi\left(\frac{\cdot}{\varepsilon}\right).$$

We also introduce the following operator, for any function  $v \in L^\infty(0, T; L^1(\mathbb{R}^d))$ :

$$\left(\Pi_\varepsilon^h(t)v\right)(x) = \sum_{i \in \mathcal{J}_h} w_i(t) v(t, x_i(t)) \varphi_\varepsilon(x - x_i(t)).$$

We recall a direct corollary of the **Theorem 3** in [115]:

**Proposition 29.** *Let  $k, r$  be two integers, with  $k > d$ , and let us assume that  $a \in L^\infty(0, T; W^{k+1, \infty}(\mathbb{R}^d)^d)$ , and that  $\varphi \in \mathcal{C}_c^1(\mathbb{R}^d) \cap W^{k+1, 1}(\mathbb{R}^d)$  satisfies conditions (5.63) and (5.64). Then, for any  $p \in [1, +\infty]$ , there exists  $C = C(T) > 0$  such that, for any  $u \in W^{\mu, 1}(\mathbb{R}^d)$  ( $\mu = \max(r, k)$ ),*

$$\|u - \Pi_\varepsilon^h(t)u\|_{L^p(\mathbb{R}^d)} \leq C(\varepsilon^r \|u\|_{W^{r, p}(\mathbb{R}^d)} + \left(\frac{\varepsilon}{h}\right)^k \|u\|_{W^{k, p}(\mathbb{R}^d)}).$$



We seek to prove the following approximation result between  $v_\varepsilon^h$  and  $v$ , the solution of problem (5.19).

**Theorem 14.** *Assume that hypotheses (5.57) through (5.62) are satisfied, and that  $\varphi \in \mathcal{C}_c^1(\mathbb{R}^d) \cap W^{k+1,1}(\mathbb{R}^d)$  satisfies (5.63) and (5.64). Then, there exists  $C = C(T, a, R, m, \bar{\rho}) > 0$ , a positive constant which depends on  $T$ ,  $a$ ,  $R$ ,  $m$  and  $\bar{\rho}$  such that*

$$\|v - v_\varepsilon^h\|_{L^1(\mathbb{R}^d)} \leq C(\varepsilon^r + (\frac{h}{\varepsilon})^\kappa + h^\kappa) \|v^0\|_{W^{\mu,1}(\mathbb{R}^d)}, \quad \forall 0 \leq t \leq T,$$

where  $\mu = \max(r, \kappa)$ .

The proof of Theorem 14 strongly relies on Proposition 29 and the following result:

**Proposition 30.** *Under hypotheses (5.57) through (5.62), there exists a constant  $C_T > 0$ , depending only on  $T$  and on the parameters of problems (5.19) and (5.54), such that their respective solutions satisfy, for all  $t \in [0, T]$ ,*

$$\sum_{i \in \mathcal{J}_h} |v(t, x_i(t)) - \nu_i(t)| w_i(t) \leq C_T h^{k-1} \|v^0\|_{W^{\mu,1}(\mathbb{R}^d)}. \quad (5.67)$$

*Proof.* Consider  $\beta_\varepsilon$  as in (C.1). We define  $e = \{e_i(\cdot)\}_{i \in \mathcal{J}}$  where for all  $i \in \mathcal{J}$  and  $t \in [0, T]$ ,

$$\begin{aligned} e_i(t) &:= v(t, x_i(t)) - \nu_i(t), \\ e_{\varepsilon,i}(t) &:= \beta_\varepsilon(e_i(t)) w_i(t), \end{aligned}$$

and compute

$$\dot{e}_{\varepsilon,i}(t) = \beta_\varepsilon'(e_i(t)) \dot{e}_i(t) w_i(t) + \beta_\varepsilon(e_i(t)) \dot{w}_i(t). \quad (5.68)$$

We recall that

$$\begin{aligned} \dot{e}_i(t) &= \left( G_{\nu,w}(t, x_i(t)) - \mathcal{G}_v(t, x_i(t)) \right) \nu_i(t) + \left( A_{\nu,w}(t, x_i(t)) - \mathcal{A}_v(t, x_i(t)) \right) \nabla v(t, x_i(t)) \\ &\quad - \mathcal{G}_v(t, x_i(t)) e_i(t) + \Delta \mathcal{M}(t, x_i(t)), \end{aligned}$$

where

$$\begin{aligned} G_{\nu,w}(t, x_i(t)) &:= \operatorname{div} A_{\nu,w}(t, x_i(t)) - R(t, x_i(t), I_g(t, x_i(t), \nu, w)), \\ \mathcal{G}_v(t, x_i(t)) &:= \operatorname{div} \mathcal{A}_v(t, x_i(t)) + R(t, x_i(t), (I_g v)(t, x_i(t))), \\ \Delta \mathcal{M}(t, x_i(t)) &:= \int_{\mathbb{R}^d} m(t, x_i(t), y, (I_d v)(t, x_i)) v(t, y) dy \\ &\quad - \sum_{j \in \mathcal{J}_h} w_j(t) \nu_j(t) m(t, x_i(t), x_j(t), I_d(t, x_i(t), \nu, w)). \end{aligned}$$

The functions  $a$  and  $R$  being Lipschitz, there exists a constant  $C$  depending on the parameters of the problem and the value  $\bar{\rho}$ , such that

$$\begin{aligned} \left| G_{\nu,w}(t, x_i(t)) - \mathcal{G}_v(t, x_i(t)) \right| &\leq C \left( |(I_a v)(t, x_i(t)) - I_a(t, x_i(t), \nu, w)| \right. \\ &\quad \left. + |\partial_x (I_a v)(t, x_i(t)) - \partial_x I_a(t, x_i(t), \nu, w)| \right. \\ &\quad \left. + |(I_g v)(t, x_i(t)) - I_g(t, x_i(t), \nu, w)| \right). \end{aligned}$$

Notice that

$$\begin{aligned}
|(I_a v)(t, x_i(t)) - I_a(t, x_i(t), \nu, w)| &= \left| \int_{\mathbb{R}^d} \psi_a(t, x_i(t), y) v(y) dy - \sum_{j \in \mathcal{J}_h} \psi_a(t, x_i(t), x_j(t)) \nu_j(t) w_j(t) \right| \\
&\leq \left| \int_{\mathbb{R}^d} \psi_a(t, x_i(t), y) v(y) dy - \sum_{j \in \mathcal{J}_h} \psi_a(t, x_i(t), x_j(t)) v(t, x_j(t)) w_j(t) \right| \\
&\quad + \left| \sum_{j \in \mathcal{J}_h} \psi_a(t, x_i(t), x_j(t)) \left( v(t, x_j(t)) - \nu_j(t) \right) w_j(t) \right| \\
&\leq C \left( h^\kappa \|v\|_{W^{\kappa,1}(\mathbb{R}^d)} + \sum_{j \in \mathcal{J}_h} |e_j(t)| w_j(t) \right),
\end{aligned}$$

where in the last line we have used Lemma 18 and the  $W^{\kappa,1}(\mathbb{R}^d)$  regularity of  $v$ . Similar results are true for  $|\partial_x(I_a v)(t, x_i(t)) - \partial_x I_a(t, x_i(t), \nu, w)|$  and  $|(I_g v)(t, x_i(t)) - I_g(t, x_i(t), \nu, w)|$ . In conclusion, thanks to (5.27), there exists a constant  $C_T$ , only depending on  $T$ , the parameters of the problem and the value  $\bar{\rho}$ , such that

$$\left| G_{\nu, w}(t, x_i(t)) - \mathcal{G}_v(t, x_i(t)) \right| \leq C_T \left( h^\kappa \|v^0\|_{W^{\kappa,1}(\mathbb{R}^d)} + \sum_{j \in \mathcal{J}_h} |e_j(t)| w_j(t) \right). \quad (5.69)$$

Again, using the Lipschitz regularity of  $a$ , we see that

$$\begin{aligned}
\left| A_{\nu, w}(t, x_i(t)) - \mathcal{A}_v(t, x_i(t)) \right| &\leq C |(I_a v)(t, x_i(t)) - I_a(t, x_i(t), \nu, w)| \\
&\leq C_T \left( h^\kappa \|v^0\|_{W^{\kappa,1}(\mathbb{R}^d)} + \sum_{j \in \mathcal{J}_h} |e_j(t)| w_j(t) \right). \quad (5.70)
\end{aligned}$$

The boundedness of  $a$  and  $R$  implies the existence of a constant  $\bar{G}$  such that for all  $i \in \mathcal{J}$  and  $t \in [0, T]$ ,

$$|G_v(t, x_i(t))| \leq \bar{G}. \quad (5.71)$$

We recall from the previous section that

$$\Delta \mathcal{M}(t, x_i(t)) = 0, \quad \forall i \notin \mathcal{J}_h^m,$$

where the set of indexes  $\mathcal{J}_h^m$  has a finite number of elements, which depends on  $T$ . On the other hand, for those  $i \in \mathcal{J}_h^m$ , we have

$$\begin{aligned}
|\Delta \mathcal{M}(t, x_i(t))| &\leq \left| \int_{\mathbb{R}^d} m(t, x_i(t), y, (I_d v)(t, x_i(t))) v(t, y) dy \right. \\
&\quad - \sum_{j \in \mathcal{J}_h} w_j(t) v(t, x_j(t)) m(t, x_i(t), x_j(t), (I_d v)(t, x_i(t))) \left. \right| \\
&\quad + \sum_{j \in \mathcal{J}_h} w_j(t) v(t, x_j(t)) \left| m(t, x_i(t), x_j(t), (I_d v)(t, x_i(t))) \right. \\
&\quad \left. - m(t, x_i(t), x_j(t), I_d(t, x_i(t), \nu, w)) \right| \\
&\quad + \sum_{j \in \mathcal{J}_h} w_j(t) |e_j(t)| m(t, x_i(t), x_j(t), (I_d v)(t, x_i(t))) \\
&\leq C \left( h^\kappa \|v\|_{W^{\kappa,1}(\mathbb{R}^d)} + \mu |(I_d v)(t, x_i(t)) - I_d(t, x_i(t), \nu, w)| \sum_{j \in \mathcal{J}_h} w_j(t) v(t, x_j(t)) \right. \\
&\quad \left. + \bar{M} \sum_{j \in \mathcal{J}_h} |e_j(t)| w_j(t) \right),
\end{aligned}$$

where we have used again Lemma 18, the  $W^{\kappa,1}(\mathbb{R}^d)$  regularity of  $m(t, x, y, I)v(t, y)$  with respect to the  $y$  variable, and the Lipschitz regularity of  $m$ . Furthermore, Lemma 18 gives us the bound

$$\sum_{j \in \mathcal{J}_h} w_j(t)v(t, x_j(t)) \leq \|v\|_{L^1(\mathbb{R}^d)} + Ch^\kappa \|v\|_{W^{\kappa,1}(\mathbb{R}^d)},$$

which together with manipulations similar to those made for  $|(I_a v)(t, x_i(t)) - I_a(t, x_i(t), \nu, w)|$ , and the bound (5.27), gives

$$|\Delta \mathcal{M}(t, x_i(t))| \leq C_T \left( h^\kappa \|v^0\|_{W^{\kappa,1}(\mathbb{R}^d)} + \sum_{j \in \mathcal{J}_h} |e_j(t)|w_j(t) \right). \quad (5.72)$$

We denote as  $\mathcal{K}$  an arbitrary finite subset of  $\mathcal{J}_h$ . If we add (5.68) for all values of  $i \in \mathcal{K}$ , and use bounds (5.69) through (5.72), together with the equation for  $w_i$ , we get

$$\left( \sum_{i \in \mathcal{K}} e_{\varepsilon,i}(t) \right)' \leq C_T B(t) \left( h^\kappa \|v^0\|_{W^{\kappa,1}(\mathbb{R}^d)} + \sum_{j \in \mathcal{J}_h} |e_j(t)|w_j(t) \right),$$

where

$$B(t) := \sum_{i \in \mathcal{K}} (\nu_i(t) + |\nabla v(t, x_i(t))|)w_i(t) + \bar{G} + |\mathcal{K} \cap \mathcal{J}_h^m| h^d e^{\bar{a}T} + \tilde{a}.$$

Given that

$$\sum_{i \in \mathcal{K}} \nu_i(t)w_i(t) \leq \bar{\rho},$$

and

$$\sum_{i \in \mathcal{K}} w_j(t)|\nabla v(t, x_j(t))| \leq \|\nabla v\|_{L^1(\mathbb{R}^d)} + Ch^{\kappa-1} \|\nabla v\|_{W^{\kappa-1,1}(\mathbb{R}^d)},$$

thanks to (5.27), we conclude that there exists a constant  $B_T$ , independent of the choice of  $\mathcal{K}$  and  $h$ , such that  $B(t) \leq B_T$ . Consequently, for all values of  $t \in [0, T]$ ,  $h$  small enough and any finite subset of  $\mathcal{J}_h$ , we have the relation

$$\left( \sum_{i \in \mathcal{K}} e_{\varepsilon,i}(t) \right)' \leq C_T \left( h^\kappa + \sum_{j \in \mathcal{J}_h} |e_j(t)|w_j(t) \right).$$

Integrating between 0 and  $t$ , taking the limit when  $\varepsilon$  goes to zero, and using Grönwall's lemma, we obtain that there exists a constant  $C_T$ , independent of  $\mathcal{K}$  such that

$$\sum_{i \in \mathcal{K}} |e_i(t)|w_i(t) \leq C_T h^\kappa \|v^0\|_{W^{\kappa,1}(\mathbb{R}^d)}.$$

Being  $C_T$  independent of  $\mathcal{K}$  and  $h$ , (5.67) is immediate.  $\square$

In other words, we proved in Proposition 30 that the piece-wise constant functions that take values  $v(t, x_k(t))$  and  $v_k(t)$  respectively over the intervals  $\Omega_k(t)$  are close in  $L^1(\mathbb{R}^d)$ .

*Proof of Theorem 14.* According to the triangle inequality,

$$\|v - v_\varepsilon^h\|_{L^1(\mathbb{R}^d)} \leq \|v - \Pi_\varepsilon^h(t)v\|_{L^1(\mathbb{R}^d)} + \|\Pi_\varepsilon^h(t)v - v_\varepsilon^h\|_{L^1(\mathbb{R}^d)}, \quad (5.73)$$

it only remains to bound both terms on the right hand side.

i) According to Proposition 29 with  $p = 1$ , and bound (5.27)

$$\|v - \Pi_\varepsilon^h(t)v\|_{L^1(\mathbb{R}^d)} \leq C(\varepsilon^r + \left(\frac{\varepsilon}{h}\right)^\kappa) \|v\|_{\mu,p} \leq C_T(\varepsilon^r + \left(\frac{\varepsilon}{h}\right)^\kappa) \|v^0\|_{\mu,p}.$$

ii) On the other hand, one computes

$$\begin{aligned} \|\Pi_\varepsilon^h(t)v - v_\varepsilon^h\|_{L^1(\mathbb{R}^d)} &= \int_{\mathbb{R}^d} \left| \sum_{i \in \mathcal{J}_h} w_i(t) \varphi_\varepsilon(x - x_i(t)) (v(t, x_i(t)) - \nu_i(t)) \right| dx \\ &\leq \sum_{i \in \mathcal{J}_h} \left( w_i(t) |v(t, x_i(t)) - \nu_i(t)| \int_{\mathbb{R}^d} |\varphi_\varepsilon(x - x_i(t))| dx \right). \end{aligned}$$

According to the definition of  $\varphi_\varepsilon$ , with the change of variable  $x' = \frac{x - x_k(t)}{\varepsilon}$ , we note that

$$\int_{\mathbb{R}^d} |\varphi_\varepsilon(x - x_k(t))| dx = \int_{\mathbb{R}^d} |\varphi(x)| dx < +\infty,$$

by hypothesis on  $\varphi$ . We have then, according to Proposition 30, that

$$\|\Pi_\varepsilon^h(t)v - v_\varepsilon^h\|_{L^1(\mathbb{R}^d)} \leq C_T h^\kappa \|v^0\|_{\kappa,1} \leq C h^\kappa \|v^0\|_{\mu,1},$$

which concludes the proof of Theorem 14. □

### 5.4.2 Asymptotic preserving properties

The study of the asymptotic behaviour of the solution for adaptive dynamics models, such as (5.19), is one of the main interests often treated in the literature (see [10, 13, 21, 39, 42, 48, 59, 86, 106]). For this reason, the design of numerical methods which preserve the asymptotic behaviour, or at least, the identification of the problems for which the asymptotics are preserved under a certain numerical scheme, is a priority. In other words, given that  $v(t, \cdot)$  converges to a measure  $\mu$  when  $t$  goes to infinity, we expect to identify the conditions under which  $\lim_{h \rightarrow 0} \lim_{t \rightarrow +\infty} v_{\varepsilon(h)}^h(t, \cdot) = \mu$ , that is, ensuring the commutativity of diagram (5.74):

$$\begin{array}{ccc} v(t, \cdot) & \xrightarrow{t \rightarrow \infty} & \mu \\ \uparrow h \rightarrow 0 & & \uparrow h \rightarrow 0 \\ v_{\varepsilon(h)}^h(t) & \xrightarrow{t \rightarrow \infty} & \mu_h \end{array} \quad (5.74)$$

In what follows, we formally define the concept of an asymptotic preserving approximation, and give examples and counter-examples of this concept.

We recall that, according to the Riesz representation theorem, the space of finite Radon measures can be identified with the topological dual space of  $\mathcal{C}_c(\mathbb{R}^d)$ . Hence, we say that a sequence of finite Radon measures  $\{\mu_n\}_{n \in \mathbb{N}}$  converges weakly to a finite Radon measure  $\mu$  (denoted  $\mu_n \xrightarrow{n \rightarrow +\infty} \mu$ ) if for all  $\phi \in \mathcal{C}_c(\mathbb{R}^d)$ ,

$$\int_{\mathbb{R}^d} \phi(x) d\mu_n(x) \xrightarrow{n \rightarrow +\infty} \int_{\mathbb{R}^d} \phi(x) d\mu(x).$$

This leads us to introduce the following definition:

**Definition 6.** We say that the particle solution  $v_\varepsilon^h$  defined in (5.66) is an *asymptotic preserving approximation* of  $v$ , the solution to (5.19), if for all  $\varepsilon : (0, 1] \rightarrow \mathbb{R}_+^*$  which converges to 0 when  $h$  goes to 0, and all  $\phi \in \mathcal{C}_c(\mathbb{R}^d)$ ,

$$\limsup_{t \rightarrow +\infty} \left| \int_{\mathbb{R}^d} \phi(x) v_{\varepsilon(h)}^h(t, x) dx - \int_{\mathbb{R}^d} \phi(x) v(t, x) dx \right| \xrightarrow{h \rightarrow 0} 0.$$

The following lemma ensures that, in the previous definition,  $v_{\varepsilon(h)}^h$ , introduced in (5.66), can be replaced by  $v^h$ , introduced in (5.65).

**Lemma 19.** *The function  $v_{\varepsilon}^h$  is an asymptotic preserving approximation of  $v$  if and only if*

$$\limsup_{t \rightarrow +\infty} \left| \int_{\mathbb{R}^d} \phi(x) dv^h(t, x) - \int_{\mathbb{R}^d} \phi(x) v(t, x) dx \right| \xrightarrow{h \rightarrow 0} 0.$$

*Proof.* Let us prove that for all  $\varepsilon : (0, 1] \rightarrow \mathbb{R}_+^*$  which converges to 0 as  $h$  goes to 0, and all  $\phi \in \mathcal{C}_c(\mathbb{R}^d)$ ,

$$\limsup_{t \rightarrow +\infty} \left| \int_{\mathbb{R}^d} \phi(x) v_{\varepsilon(h)}^h(t, x) dx - \int_{\mathbb{R}^d} \phi(x) dv^h(t, x) \right| \xrightarrow{h \rightarrow 0} 0.$$

Let  $h > 0$ . According to the definitions of  $v^h$  and  $v_{\varepsilon}^h$ , and since  $\int_{\mathbb{R}^d} \varphi_{\varepsilon(h)}(x - x_i(t)) dx = 1$  for all  $i \in \mathcal{J}_h$  we get, for all  $t \geq 0$ ,

$$\begin{aligned} \left| \int_{\mathbb{R}^d} \phi(x) v_{\varepsilon(h)}^h(t, x) dx - \int_{\mathbb{R}^d} \phi(x) dv^h(t, x) \right| &= \left| \int_{\mathbb{R}^d} \sum_{i \in \mathcal{J}_h} \nu_i(t) w_i(t) (\phi(x) - \phi(x_i(t))) \varphi_{\varepsilon(h)}(x - x_i(t)) dx \right| \\ &\leq \sum_{i \in \mathcal{J}_h} |\nu_i(t) w_i(t)| \int_{\mathbb{R}^d} |(\phi(x) - \phi(x_i(t))) \varphi_{\varepsilon(h)}(x - x_i(t))| dx. \end{aligned}$$

With the change of variable ' $y = \frac{x - x_i(t)}{\varepsilon(h)}$ ', we get, for all  $i \in \mathcal{J}_h$ ,

$$\int_{\mathbb{R}^d} |(\phi(x) - \phi(x_i(t))) \varphi_{\varepsilon(h)}(x - x_i(t))| dx = \int_K |(\phi(\varepsilon(h)y + x_i(t)) - \phi(x_i(t))) \varphi(y)| dy,$$

where  $K$  is the support of  $\varphi$ . Let  $\eta > 0$ . Since  $\phi$  is continuous with a compact support, and thus uniformly continuous, then  $|\phi(\varepsilon(h)x + x_i(t)) - \phi(x_i(t))| \leq \eta$  for all  $i \in \mathcal{J}_h$ ,  $x \in K$ ,  $t \geq 0$  and any  $h$  small enough. Therefore, for any  $h$  small enough,

$$\left| \int_{\mathbb{R}^d} \phi(x) v_{\varepsilon(h)}^h(t, x) dx - \int_{\mathbb{R}^d} \phi(x) dv^h(t, x) \right| \leq \eta \sum_{i \in \mathcal{J}_h} \nu_i(t) w_i(t),$$

which concludes the proof, since there exists  $\bar{\rho} > 0$  such that  $0 \leq \sum_{i \in \mathcal{J}_h} \nu_i(t) w_i(t) \leq \bar{\rho}$  for all  $h > 0$ ,  $t \geq 0$ , as proved in Theorem 13.  $\square$

The problem of determining if  $v_{\varepsilon}^h$  is an asymptotic preserving approximation of  $v$  is generally a difficult question. In what follows, we deal with cases where we are able to determine the asymptotic behaviour of both  $v$  and  $v^h$ , and we check if the necessary and sufficient condition from Lemma 19 holds. From now on, we assume that  $a$  is local and not time dependent, *i.e.*  $a(t, x, I) = a(x)$  and that the functions  $m$ ,  $R$ ,  $\psi_g$  and  $\psi_d$  are not time-dependent. We assume as well that the function  $(x, y, I) \mapsto m(x, y, I)$  is not only uniformly compactly supported as a function of the  $x$  variable, but relative to the  $y$  variable as well. That is, there exist two compact sets  $K_x$  and  $K_y$  such that  $\sup_{y, I} m(x, y, I) = 0$  for all  $y$  outside of  $K_x$  and  $\sup_{x, I} m(x, y, I) = 0$  for all  $x$  outside of  $K_y$ . We denote  $K_{xy} := K_x \cup K_y$ . Finally, we assume  $\psi_g$  and  $\psi_d$  to be compactly supported as functions of the  $y$  variable.

### Necessary conditions of convergence towards a Radon measure

With the help of necessary conditions, we would be able to rule out those cases where  $v$  does not converge towards certain types of Radon measures, which are the object of our interest. We start by giving a general result, involving the necessary conditions of convergence towards any Radon measure.

**Lemma 20.** *Let us assume that  $v(t, \cdot) \xrightarrow[t \rightarrow +\infty]{} \mu$  in the weak sense in the space of finite Radon measures. Then, for  $\gamma \in \{g, d\}$ ,*

$$I_\alpha(t, x) \xrightarrow[t \rightarrow +\infty]{} \int_{\mathbb{R}^d} \psi_\alpha(x, y) d\mu(y) =: \bar{I}_\alpha(x) \quad \forall x \in \mathbb{R}^d,$$

and for all  $\phi \in \mathcal{C}_0^1(\mathbb{R}^d)$ ,

$$\int_{\mathbb{R}^d} \left( a(x) \cdot \nabla \phi(x) + R(x, \bar{I}_g(x)) \phi(x) \right) d\mu(x) + \int_{\mathbb{R}^d} \left( \int_{\mathbb{R}^d} m(x, y, \bar{I}_d(x)) \phi(x) dx \right) d\mu(y) = 0.$$

*Proof.* Let us assume that  $v(t, \cdot) \xrightarrow[t \rightarrow +\infty]{} \mu$ . By definition of the weak convergence in the space of finite Radon measure, for any  $\phi \in \mathcal{C}_c^1(\mathbb{R}^d)$ ,

$$\int_{\mathbb{R}^d} \phi(x) v(t, x) dx \xrightarrow[t \rightarrow +\infty]{} \int_{\mathbb{R}^d} \phi(x) d\mu(x).$$

The first identity is thus a direct consequence of the definition of the weak convergence, applied with  $\phi = \psi_\gamma(x, \cdot)$ ,  $\gamma \in \{g, d\}$ .

Let  $\phi \in \mathcal{C}_c(\mathbb{R}^d)$ . One computes

$$\begin{aligned} \frac{d}{dt} \int_{\mathbb{R}^d} \phi(x) v(t, x) dx &= - \int_{\mathbb{R}^d} \phi(x) \nabla \cdot (a(x) v(t, x)) + \int_{\mathbb{R}^d} R(x, I_g(t, x)) \phi(x) v(t, x) dx \\ &\quad + \int_{\mathbb{R}^d} \left( \int_{\mathbb{R}^d} m(x, y, I_d(t, x)) v(t, y) \right) \phi(x) dx \\ &= \int_{\mathbb{R}^d} \left( a(x) \cdot \nabla \phi(x) + R(x, I_g(t, x)) \phi(x) \right) v(t, x) dx \\ &\quad + \int_{\mathbb{R}^d} \left( \int_{\mathbb{R}^d} m(x, y, I_d(t, x)) \phi(x) dx \right) v(t, y) dy \\ &\xrightarrow[t \rightarrow +\infty]{} \int_{\mathbb{R}^d} \left( a(x) \cdot \nabla \phi(x) + R(x, \bar{I}_g(x)) \phi(x) \right) d\mu(x) \\ &\quad + \int_{\mathbb{R}^d} \left( \int_{\mathbb{R}^d} m(x, y, \bar{I}_d(x)) \phi(x) dx \right) d\mu(y). \end{aligned}$$

Thus,  $t \mapsto \int_{\mathbb{R}^d} \phi(x) v(t, x) dx$  is a convergent function with a convergent derivative, which ensures that the limit of its derivative is zero, which concludes the proof.  $\square$

The following proposition provides a necessary condition for the convergence to a sum of Dirac masses.

**Proposition 31.** *Let us assume that  $v(t, \cdot) \xrightarrow[t \rightarrow +\infty]{} \sum_{i=1}^N C_i \delta_{x_i}$ , with  $x_1, \dots, x_N \in \mathbb{R}^d$ ,  $C_1, \dots, C_N >$*

*0. Then, for  $\alpha = g, d$ ,  $I_\alpha(t, x) \xrightarrow[t \rightarrow +\infty]{} \sum_{i=1}^N C_i \psi_\alpha(x, x_i) =: \bar{I}_\alpha(x)$ . Moreover, for all  $i \in \{1, \dots, N\}$ ,  $a(x_i) = 0$ ,  $R(x_i, \bar{I}_g(x_i)) = 0$ , and for all  $x \in \mathbb{R}^d$ ,  $m(x, x_i, \bar{I}_d(x)) = 0$ .*

*Proof.* According to the previous lemma, for all  $\phi \in \mathcal{C}_c^1(\mathbb{R}^d)$ ,

$$\sum_{i=1}^N C_i a(x_i) \cdot \nabla \phi(x_i) + \sum_{i=1}^N C_i R(x_i, \bar{I}_g(x_i)) \phi(x_i) + \sum_{i=1}^N C_i \int_{\mathbb{R}^d} m(x, x_i, \bar{I}_d(x)) \phi(x) dx = 0.$$

Let  $\varepsilon > 0$ . For any non-negative function  $\phi_\varepsilon \in \mathcal{C}_c^1(\mathbb{R}^d)$  with a support on  $\mathbb{R}^d \setminus \bigcup_{i=1}^N B(x_i, \varepsilon)$ , we have that

$$\int_{\mathbb{R}^d} \sum_{i=1}^N C_i m(x, x_i, \bar{I}_d(x)) \phi_\varepsilon(x) dx = 0,$$

which proves that  $x \mapsto \sum_{i=1}^N C_i m(x, x_i, \bar{I}_d(x))$  is 0 on  $\mathbb{R}^d \setminus \bigcup_{i=1}^N B(x_i, \varepsilon)$ . Since  $m$  is a non-negative function, for all  $i \in \{1, \dots, N\}$ ,  $x \mapsto m(x, x_i, \bar{I}_d(x))$  is 0 as well on  $\mathbb{R}^d \setminus \bigcup_{i=1}^N B(x_i, \varepsilon)$ , and therefore on  $\mathbb{R}^d$ , since the result holds for any  $\varepsilon > 0$ . Hence,

$$\sum_{i=1}^N C_i a(x_i) \cdot \nabla \phi(x_i) + \sum_{i=1}^N C_i R(x_i, \bar{I}_g(x_i)) \phi(x_i) = 0,$$

for any  $\phi \in \mathcal{C}_c^1(\mathbb{R}^d)$ . By choosing  $\phi_j^1$  such that

$$\phi_j^1(x_i) = \delta_{ij} \text{ and } \nabla \phi_j^1(x_i) = 0, \quad i, j \in \{1, \dots, N\},$$

where  $\delta_{ij}$  represents the Kronecker delta, one proves that  $R(x_i, \bar{I}_g(x_i)) = 0$  for all  $i \in \{1, \dots, N\}$ .

Finally, by choosing  $\phi_{ij}^2$  such that

$$\nabla \phi_{jl}^2(x_i) = \delta_{ij} e_l, \quad i, j \in \{1, \dots, N\}, \quad l \in \{1, \dots, d\},$$

where  $\{e_l\}_{l=1}^d$  represents the euclidean basis in  $\mathbb{R}^d$ , one proves that  $a(x_i) = 0$ , for any  $i \in \{1, \dots, N\}$ .  $\square$

### Limit identification and asymptotic preserving approximations

In some cases, it is possible to guarantee the existence of a limit for  $v^h$  and identify it. Assume that there exists  $\hat{x} \in \mathbb{R}^d$  an asymptotically stable equilibrium for the ODE ' $\dot{x} = a(x)$ ' and that there exists  $C, \delta > 0$  such that

$$\forall y \in \text{supp}(n^0), t \geq 0, \quad \|X(t, y) - \hat{x}\| \leq C e^{-\delta t}. \quad (5.75)$$

Moreover, let us assume that there exist positive values  $D, I^m$  and  $I^M$  such that

$$R(x, I^m) \geq 0, \quad R(x, I^M) \leq 0 \text{ and } \partial_I R(x, I) \leq -D, \quad \forall x \in \text{supp}(v^0). \quad (5.76)$$

Then, we can compute the limit of  $v^h$  when  $t$  goes to  $+\infty$ , whatever the value of  $m$ , as stated in the following proposition.

**Proposition 32.** *Let us assume that  $\text{supp}(v^0)$  is a compact set such that hypotheses (5.75) and (5.76) hold. Then,  $v^h$  converges to  $\hat{\rho}_h \delta_{\hat{x}}$  in the weak sense in the space of Radon measures, where  $\hat{\rho}_h$  is the unique solution of*

$$R(\hat{x}, \psi_g(\hat{x}, \hat{x}) \hat{\rho}_h) = 0.$$

The following lemma, proved in Appendix C.4, is required in the proof of this result.

**Lemma 21.** *Let  $u \in \mathcal{C}^2(\mathbb{R}_+, \mathbb{R})$  be a bounded function, and let us assume that there exist  $p_0 > 0$ ,  $p : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  a function which satisfies  $p \geq p_0$  and  $B \in L^1(\mathbb{R}_+)$  an integrable function such that*

$$\ddot{u}(t) \geq -p(t)\dot{u}(t) + B(t).$$

*Then, there exists  $u_\infty \in \mathbb{R}$  such that  $\lim_{t \rightarrow +\infty} u(t) = u_\infty$ .*

*Proof of Proposition 32.* Given that  $\text{supp}(v^0) \cup K_{xy}$  is a compact set which is strictly contained in the basin of attraction of  $\hat{x}$  we will have the existence of  $\mathcal{J}_h^0 \subset \mathcal{J}_h$ , with  $|\mathcal{J}_h^0| < +\infty$  such that  $\nu^h(t) \neq 0$  only for  $i \in \mathcal{J}_h^0$ .

Let us denote, for all  $i \in \mathcal{J}_h^0$ ,  $\alpha_i(t) := \nu_i(t)w_i(t)$ , and

$$\rho_h(t) := \sum_{i \in \mathcal{J}_h} \alpha_i(t) = \sum_{i \in \mathcal{J}_h^0} \alpha_i(t).$$

Let us note that, according to the hypotheses on  $a$ , for all  $i \in \mathcal{J}_h^0$ ,  $x_i(t)$  converges to  $\hat{x}$ . Thus,

$$v^h(t) - \rho_h(t)\delta_{\hat{x}} \xrightarrow{t \rightarrow +\infty} 0.$$

Hence, it only remains to prove that  $\rho_h$  converges to the expected limit. According to the definition of  $\rho_h$ ,

$$\dot{\rho}_h(t) = \sum_{i \in \mathcal{J}_h^0} R(x_i(t), I_g(t, x_i(t)))\alpha_i(t) + \underbrace{\sum_{i,j \in \mathcal{J}_h^0} w_i(t)m(x_i(t), x_j(t), I_d(t, x_i(t)))\alpha_j(t)}_{:=\varepsilon(t)}. \quad (5.77)$$

According to hypothesis 5.75, there exist  $C, \delta > 0$  such that

$$\|x_i(t) - \hat{x}\| \leq Ce^{-\delta t}, \quad \forall i \in \mathcal{J}_h^0, \quad \forall t \geq 0.$$

Thus, for all  $i \in \mathcal{J}_h^0$ ,

$$w_i(t) = e^{\int_0^t \text{div} a(x_i(s))ds} w_i(0) \leq \underbrace{e^{\int_0^t \text{div} a(x_i(s)) - \text{div} a(\hat{x})ds}}_{\leq \tilde{C}} e^{\text{div} a(\hat{x})t} w_i(0),$$

which proves, since  $\text{div} a(\hat{x}) < 0$ , that there exist  $C', \delta'$  such that for all  $t \geq 0$ ,

$$0 \leq \max_{i \in \mathcal{J}_h^0} w_i(t) \leq C'e^{-\delta' t}. \quad (5.78)$$

In particular, it proves that  $t \mapsto \varepsilon(t)$  defined in (5.77), converges to zero with an exponential speed, since

$$|\varepsilon(t)| \leq \max_{i \in \mathcal{J}_h} w_i(t) |\mathcal{J}_h^0| \|m\|_{L^\infty} \rho_h(t),$$

and  $\rho_h$  is bounded, according to Theorem 13.

Moreover, according to the hypothesis on  $\psi_g$ ,  $\underline{\psi}_g \rho(t) \leq I_g(t, x_i(t)) \leq \|\psi_g\|_\infty \rho(t)$ . Thus, according to the hypotheses (5.76) on  $R$ , the relation

$$\dot{\rho}_h(t) \geq \left( \min_{i \in \mathcal{J}_h^0} (R(x_i(t), I_g(t, x_i(t)))) \right) \rho_h(t)$$



implies that, as soon as  $\rho_h$  becomes small, and so does  $I_g$ ,  $\rho_h$  becomes increasing, which proves that  $\rho_h$  is lower bounded by a positive constant. Moreover, since  $\rho_h$  and  $\varepsilon$  are bounded, and

$$\begin{aligned} \|\dot{\alpha}\|_{l_1} &= \sum_{i \in \mathcal{J}_h^0} |\dot{\alpha}_i(t)| \leq \sum_{i \in \mathcal{J}_h^0} (2|\operatorname{div} a(x_i(t))| + |R(x_i(t), I_g(t, x_i(t)))|) \alpha_i(t) + \varepsilon(t) \\ &\leq \left(2\|a\|_{W^{1,\infty}(\mathbb{R}^d)} + \bar{R}\right) \rho_h(t) + \varepsilon(t), \end{aligned}$$

where

$$\bar{R} := \max_{t \in \mathbb{R}, i \in \mathcal{J}_h^0} |R(x_i(t), I_g(t, x_i(t)))|,$$

then  $\|\dot{\alpha}\|_{l_1}$  is also bounded.

Now, let us prove that  $\rho_h$  satisfies the equality of Lemma 21. First, for  $\gamma \in \{g, d\}$ , we compute

$$\begin{aligned} \left| \frac{d}{dt} I_\gamma(t, x_i(t)) - \psi_\gamma(\hat{x}, \hat{x}) \dot{\rho}_h(t) \right| &\leq \left| \sum_{j \in \mathcal{J}_h^0} \left( a(x_i(t)) \partial_x \psi_\gamma(x_i(t), x_j(t)) + a(x_j(t)) \partial_y \psi_\gamma(x_i(t), x_j(t)) \right) \alpha_j(t) \right| \\ &\quad + \left| \sum_{j \in \mathcal{J}_h^0} (\psi_\gamma(x_i(t), x_j(t)) - \psi_\gamma(\hat{x}, \hat{x})) \dot{\alpha}_j(t) \right| \\ &\leq 2 \max_{i \in \mathcal{J}_h^0} |a(x_i(t))| \|\psi_\gamma\|_{W^{1,\infty}(\mathbb{R}^d)} \rho_h(t) \\ &\quad + \max_{i,j \in \mathcal{J}_h^0} |\psi_\gamma(x_i(t), x_j(t)) - \psi_\gamma(\hat{x}, \hat{x})| \|\dot{\alpha}\|_{l_1}. \end{aligned}$$

Since hypothesis (5.75), is satisfied, the functions  $t \mapsto \max_{i \in \mathcal{J}_h^0} |a(x_i(t))|$  and  $t \mapsto \max_{i,j \in \mathcal{J}_h^0} |\psi_\gamma(x_i(t), x_j(t)) - \psi_\gamma(\hat{x}, \hat{x})|$  converge to zero with an exponential speed. Since  $|\mathcal{J}_h^0| < +\infty$ , this proves that, for  $\gamma \in \{g, d\}$ ,

$$\sum_{i \in \mathcal{J}_h^0} \left| \frac{d}{dt} I_\gamma(t, x_i(t)) - \psi_\gamma(\hat{x}, \hat{x}) \dot{\rho}_h(t) \right| = O(e^{-\delta t}), \quad (5.79)$$

for a certain  $\delta > 0$ .

Thus, by differentiating (5.77), we get

$$\begin{aligned} \dot{\rho}_h(t) &= \sum_{i \in \mathcal{J}_h^0} a(x_i(t)) \partial_x R(x_i(t), I_g(t, x_i(t))) \alpha_i(t) + \sum_{i \in \mathcal{J}_h^0} \left( \frac{d}{dt} I_g(t, x_i(t)) \right) \partial_I R(x_i(t), I_g(t, x_i(t))) \alpha_i(t) \\ &\quad + \sum_{i \in \mathcal{J}_h^0} \dot{\alpha}_i(t) R(x_i(t), I_g(t, x_i(t))) + \sum_{i,j \in \mathcal{J}_h^0} w_i(t) m(x_i(t), x_j(t), I_d(t, x_i(t))) \dot{\alpha}_j(t) \\ &\quad + \sum_{i,j \in \mathcal{J}_h^0} \dot{w}_i(t) m(x_i(t), x_j(t), I_d(t, x_i(t))) \alpha_j(t) + \sum_{i,j \in \mathcal{J}_h^0} w_i(t) \frac{d}{dt} m(x_i(t), x_j(t), I_d(t, x_i(t))) \alpha_j(t), \end{aligned}$$

where

i)  $\sum_{i \in \mathcal{J}_h^0} a(x_i(t)) \partial_x R(x_i(t), I_g(t, x_i(t))) \alpha_i(t) = O(e^{-\delta t})$ , since  $\max_{i \in \mathcal{J}_h} |a(x_i(t))|$  converges to zero with exponential speed, and  $\partial_x R$  and  $\rho_h$  are bounded,

ii) according to (5.79),  $\sum_{i \in \mathcal{J}_h^0} \left( \frac{d}{dt} I_g(t, x_i(t)) \right) \partial_I R(x_i(t), I_g(t, x_i(t))) \alpha_i(t) = -p(t) \dot{\rho}_h(t) + O(e^{-\delta t})$ , where

$$p(t) = -\psi_g(\hat{x}, \hat{x}) \sum_{i \in \mathcal{J}_h^0} \partial_I R(x_i(t), I_g(t, x_i(t))) \alpha_i(t) \geq D \underline{\psi}_g \min_{t \geq 0} \rho_h(t) > 0,$$

iii)

$$\begin{aligned} \sum_{i \in \mathcal{J}_h^0} \dot{\alpha}_i(t) R(x_i(t), I_g(t, x_i(t))) &= \underbrace{\sum_{i \in \mathcal{J}_h^0} R(x_i(t), I_g(t, x_i(t)))^2 \alpha_i(t)}_{:=P(t) \geq 0} \\ + \underbrace{\sum_{i, j \in \mathcal{J}_h^0} w_i(t) R(x_i(t), I_g(t, x_i(t))) \alpha_j(t) m(t, x_i(t), x_j(t), I_d(t, x_i(t)))}_{=O(e^{-\delta t})}. \end{aligned}$$

where the relation for the second term was proved thanks to the bound

$$\left| \sum_{i, j \in \mathcal{J}_h^0} w_i(t) R(x_i(t), I_g(t, x_i(t))) \alpha_j(t) m(t, x_i(t), x_j(t), I_d(t, x_i(t))) \right| \leq \bar{M} \bar{R} \rho_h(t) |\mathcal{J}_h^0| \max_{i \in \mathcal{J}_h} w_i(t),$$

the boundedness of  $\rho_h$  and inequality (5.78),

iv)

$$\begin{aligned} \sum_{i, j \in \mathcal{J}_h^0} w_i(t) m(x_i(t), x_j(t), I_d(t, x_i(t))) \dot{\alpha}_j(t) &+ \sum_{i, j \in \mathcal{J}_h^0} \dot{w}_i(t) m(x_i(t), x_j(t), I_d(t, x_i(t))) \alpha_j(t) \\ &+ \sum_{i, j \in \mathcal{J}_h^0} w_i(t) \frac{d}{dt} m(x_i(t), x_j(t), I_d(t, x_i(t))) \alpha_j(t) \\ &= O(e^{-\delta t}), \end{aligned}$$

since for all  $t \geq 0$ ,

$$\left| \sum_{i, j \in \mathcal{J}_h^0} w_i(t) m(x_i(t), x_j(t), I_d(t, x_i(t))) \dot{\alpha}_j(t) \right| \leq \bar{M} |\mathcal{J}_h^0| \|\dot{\alpha}\|_{l^1} \max_{i \in \mathcal{J}_h^0} w_i(t),$$

$$\left| \sum_{i, j \in \mathcal{J}_h} \dot{w}_i(t) m(x_i(t), x_j(t), I_d(t, x_i(t))) \alpha_j(t) \right| \leq \bar{M} \rho_h(t) |\mathcal{J}_h^0| \|a\|_{W^{1, \infty}(\mathbb{R}^d)} \max_{i \in \mathcal{J}_h^0} w_i(t),$$

and

$$\begin{aligned} &\left| \sum_{i, j \in \mathcal{J}_h} w_i(t) \frac{d}{dt} m(x_i(t), x_j(t), I_d(t, x_i(t))) \alpha_j(t) \right| \\ &\leq \bar{M} \left( \|a\|_{L^\infty(\mathbb{R})} + \left| \frac{d}{dt} I_g(t, x_i(t)) \right| \right) \rho_h(t) |\mathcal{J}_h^0| \max_{i \in \mathcal{J}_h^0} w_i(t), \end{aligned}$$

where  $\frac{d}{dt} I_d(t, x_i(t))$  is bounded thanks to (5.79). For these three inequalities, we conclude with (5.78).

Hence,  $\dot{\rho}_h(t) \geq -p(t)\dot{\rho}_h(t) + O(e^{-\delta t})$ . According to Lemma 21,  $\rho_h$  has a limit when  $t$  goes to  $\infty$ , which we denote  $\hat{\rho}_h$ . Since  $\dot{\rho}_h(t) = \sum_{i \in \mathcal{J}_h^0} R(x_i(t), I_g(t, x_i(t))) \alpha_i(t) + \varepsilon(t)$ , which converges to  $R(\hat{x}, \psi_g(\hat{x}, \hat{x}) \hat{\rho}_h) \hat{\rho}_h$ , we deduce that  $R(\hat{x}, \psi_g(\hat{x}, \hat{x}) \hat{\rho}_h) = 0$ , since  $\hat{\rho}_h > 0$ .  $\square$

When  $m \equiv 0$  and under the same assumptions for the remaining coefficients of the problem as in Proposition 32, we are able to identify the limit of  $v$  and prove that it coincides with the limit of  $v^h$ . According to Lemma 19, this ensures that  $v_\varepsilon^h$  is an asymptotic preserving approximation of  $v$ .

**Theorem 15.** *Let us assume that there exists  $\hat{x} \in \mathbb{R}^d$  which is an asymptotically stable equilibrium for the ODE  $\dot{x} = a(x)$  such that hypothesis (5.75) holds. We assume as well that  $m \equiv 0$  and that hypotheses (5.76) hold. Then,  $v$  converges to  $\hat{\rho} \delta_{\hat{x}}$  in the weak sense in the space of Radon measures, where  $\hat{\rho}$  is the unique solution of*

$$R(\hat{x}, \psi_g(\hat{x}, \hat{x})\hat{\rho}) = 0.$$

Consequently,  $v_\varepsilon^h$  is an asymptotic preserving approximation of  $v$ .

*Proof.* Let us recall that

$$\rho(t) = \int_{\mathbb{R}^d} v(t, x) dx.$$

We recall as well that, the function  $a$  being only dependent of  $x$ , the characteristic lines  $X_t(x) := X(t, x)$  satisfy  $X_t^{-1}(x) := X^{-1}(t, x) = X_{-t}(x)$  and  $X_t(X_s(x)) = X_{t+s}(x)$ . By using the fact that, for any  $x \in \mathbb{R}^d$ ,

$$v(t, x) = v^0(X_{-t}(x))e^{\mathcal{G}(0, t, x)},$$

where

$$\mathcal{G}(s, t, x) := \int_s^t R(X_{\tau-t}(x), (I_g v)(\tau, X_{\tau-t}(x))) - \operatorname{div} a(X_{\tau-t}(x)) d\tau,$$

and that, by hypothesis,  $K := \operatorname{supp}(v^0)$  is a compact set included in the basin of attraction of  $\hat{x}$ , one proves that  $\operatorname{supp}(v(t, \cdot))$  is the image of  $\operatorname{supp}(v^0)$  by  $X_t(\cdot)$ . Since  $X_t(y)$  converges to  $\hat{x}$  for all  $y \in \operatorname{supp}(v^0)$ , we prove that  $K_t := \operatorname{supp}(v(t, \cdot)) = X_t(\operatorname{supp}(v^0)) = X_t(K)$  is a compact set included in the basin of attraction of  $\hat{x}$ , for all  $t \geq 0$ . By (5.75), there exist  $C > 0$  and  $\delta > 0$  such that

$$\|X_t(y) - \hat{x}\| \leq C e^{-\delta t}, \quad \forall y \in K, \quad \forall t \geq 0.$$

Let  $\phi \in \mathcal{C}_c(\mathbb{R}^d)$ . By definition of  $\rho$  and by using the change of variable ' $x = X_t(y)$ ' we get

$$\begin{aligned} \left| \int_{\mathbb{R}^d} \phi(x) v(t, x) - \rho(t) \phi(\hat{x}) \right| &\leq \int_{K_t} |\phi(x) - \phi(\hat{x})| v(t, x) dx \\ &= \int_K |\phi(X_t(y)) - \phi(\hat{x})| v(t, X_t(y)) e^{\int_0^t \operatorname{div} a(X_t(y)) dy} dy \\ &\leq \max_{y \in K} |\phi(X_t(y)) - \phi(\hat{x})| \int_K v(t, X_t(y)) e^{\int_0^t \operatorname{div} a(X_t(y)) dy} dy \\ &= \max_{y \in K} |\phi(X_t(y)) - \phi(\hat{x})| \rho(t). \end{aligned}$$

Since  $\rho$  is bounded, this proves that

$$v(t, \cdot) - \rho(t) \delta_{\hat{x}} \xrightarrow[t \rightarrow +\infty]{} 0. \quad (5.80)$$

Hence, it only remains to prove that  $\rho$  converges to the expected limit. We see that

$$\dot{\rho}(t) = \int_{\mathbb{R}^d} R(x, (I_g v)(t, x)) v(t, x) dx. \quad (5.81)$$

First, let us note that  $\rho$  has a positive lower bound. Indeed, according to the hypothesis on  $\psi_g$ , for all  $x \in \mathbb{R}^d$ ,  $\underline{\psi}_g \rho(t) \leq (I_g v)(t, x) \leq \|\psi_g\|_{L^\infty} \rho(t)$ . Thus,

$$\dot{\rho}(t) \geq \left( \min_{x \in K_t} (R(x, (I_g v)(t, x))) \right) \rho(t).$$

Hence, as soon as  $\rho(t)$  becomes small, and so does  $I_g(t, x)$ ,  $\rho$  becomes increasing, which proves that  $\rho$  is lower bounded by a positive constant. We get an upper bound for  $|\dot{\rho}(t)|$  from the relation

$$|\dot{\rho}(t)| \leq \left( \max_{x \in K_t} (R(x, (I_g v)(t, x))) \right) \rho(t),$$

and the boundedness of  $\rho(t)$ .

We introduce now the function

$$\tilde{v}(t, y) = v(t, X_t(y)) e^{\int_0^t \operatorname{div} a(X_s(y)) ds},$$

which satisfies

$$\int_K \tilde{v}(t, y) dy = \int_{K_t} v(t, x) dx = \rho(t).$$

Moreover,

$$\partial_t \tilde{v}(t, y) = R(X_t(y), (I_g v)(t, X_t(y))) \tilde{v}(t, y). \quad (5.82)$$

Before proving that  $\rho$  satisfies the equality of Lemma 21, we observe that the following relation holds for all  $y \in K$ :

$$\begin{aligned} \left| \frac{d}{dt} (I_g v)(t, X_t(y)) - \psi_g(\hat{x}, \hat{x}) \dot{\rho}(t) \right| &\leq \left| \int_{K_t} a(X_t(y)) \cdot \nabla_x \psi_g(X_t(y), z) v(t, z) dz \right| \\ &+ \left| \int_{K_t} (\psi_g(X_t(y), z) - \psi_g(\hat{x}, \hat{x})) \partial_t v(t, z) dz \right|. \end{aligned} \quad (5.83)$$

From the equation satisfied by  $v$ , we see that

$$\begin{aligned} \int_{K_t} (\psi_g(X_t(y), z) - \psi_g(\hat{x}, \hat{x})) \partial_t v(t, z) dz &= \int_{K_t} \nabla_z \psi_g(X_t(y), z) a(z) v(t, z) dz \\ &+ \int_{K_t} (\psi_g(X_t(y), z) - \psi_g(\hat{x}, \hat{x})) R(z, (I_g v)(t, z)) v(t, z) dz \\ &= \int_K \nabla_z \psi_g(X_t(y), X_t(\bar{z})) a(X_t(\bar{z})) \tilde{v}(t, \bar{z}) d\bar{z} \\ &+ \int_K (\psi_g(X_t(y), X_t(\bar{z})) - \psi_g(\hat{x}, \hat{x})) R(X_t(\bar{z}), (I_g v)(t, X_t(\bar{z}))) \tilde{v}(t, \bar{z}) d\bar{z} \end{aligned}$$

which allows us to conclude that

$$\begin{aligned} \left| \int_{K_t} (\psi_g(X_t(y), z) - \psi_g(\hat{x}, \hat{x})) \partial_t v(t, z) dz \right| &\leq \max_{\bar{z} \in K} \|a(X_t(\bar{z}))\| \|\psi_g\|_{W^{1, \infty}(\mathbb{R}^d)} \rho(t) \\ &+ \max_{y, \bar{z} \in K} |\psi_g(X_t(y), X_t(\bar{z})) - \psi_g(\hat{x}, \hat{x})| \bar{R} \rho(t). \end{aligned}$$

Using this relation in (5.83) gives

$$\begin{aligned} \left| \frac{d}{dt} (I_\gamma v)(t, X_t(y)) - \psi_\gamma(\hat{x}, \hat{x}) \dot{\rho}(t) \right| &\leq 2 \max_{\bar{z} \in K} \|a(X_t(\bar{z}))\| \|\psi_\gamma\|_{W^{1, \infty}(\mathbb{R}^d)} \rho(t) \\ &+ \max_{y, \bar{z} \in K} |\psi_\gamma(X_t(y), X_t(\bar{z})) - \psi_\gamma(\hat{x}, \hat{x})| \bar{R} \rho(t), \end{aligned}$$

which proves, according to hypothesis (5.75), that for all  $y \in K$

$$\frac{d}{dt}I_g(t, X(t, y)) = \psi_g(\hat{x}, \hat{x})\dot{\rho}(t) + O(e^{-\delta t}), \quad (5.84)$$

for a certain  $\delta > 0$ .

Differentiating (5.81) we get

$$\begin{aligned} \ddot{\rho}(t) &= \frac{d}{dt} \left( \int_{K_t} R(x, (I_g v)(t, x)) v(t, x) dx \right) \\ &= \frac{d}{dt} \left( \int_K R(X_t(y), (I_g v)(t, X_t(y))) \tilde{v}(t, y) dy \right) \\ &= \int_K a(X_t(y)) \cdot \nabla_x R(X_t(y), (I_g v)(t, X_t(y))) \tilde{v}(t, y) dy \\ &\quad + \int_K \frac{d}{dt} (I_g v)(t, X_t(y)) \partial_I R(X_t(y), (I_g v)(t, X_t(y))) \tilde{v}(t, y) dy \\ &\quad + \int_K R(X_t(y), (I_g v)(t, X_t(y))) \partial_t \tilde{v}(t, y) dy. \end{aligned}$$

Let us note that

i) Since  $a(X_t(y))$  converges uniformly to zero with an exponential speed, then

$$\int_{\mathbb{R}^d} a(X_t(y)) \cdot \nabla_x R(X_t(y), (I_g v)(t, X_t(y))) \tilde{v}(t, y) dy = O(e^{-\delta t}),$$

thanks to the boundedness of  $\nabla_x R$  and  $\rho(t)$ .

ii) According to (5.84),

$$\int_K \frac{d}{dt} (I_g v)(t, X_t(y)) \partial_I R(X_t(y), (I_g v)(t, X_t(y))) \tilde{v}(t, y) dy = -p(t)\dot{\rho}(t) + O(e^{-\delta t}),$$

with

$$p(t) := -\psi_g(\hat{x}, \hat{x}) \int_K \partial_I R(X_t(y), (I_g v)(t, X_t(y))) \tilde{v}(t, y) dy \geq \psi_g(\hat{x}, \hat{x}) D \min_{t \geq 0} \rho(t) > 0.$$

iii) Directly from (5.82),

$$\begin{aligned} \int_K R(X_t(y), (I_g v)(t, X_t(y))) \partial_t \tilde{v}(t, y) dy &= \int_K \left( R(X_t(y), (I_g v)(t, X_t(y))) \right)^2 \tilde{v}(t, y) dy \\ &=: P(t) \geq 0. \end{aligned}$$

Thus,  $\ddot{\rho}(t) \geq -p(t)\dot{\rho}(t) + O(e^{-\delta t})$ , hence,  $\rho$  converges, thanks to Lemma 21.

Recalling (5.80),  $v(t, \cdot)$  thus converges to  $\hat{\rho}\delta_{\hat{x}}$ , where  $\hat{\rho}$  is the limit of  $\rho$ . We conclude, according to Proposition 31, that  $\hat{\rho}$  satisfies the expected equality.

Having proved that  $v$  and  $v^h$  share the same limit is enough then to conclude, thanks to Lemma 19, that  $v_\varepsilon^h$  is an asymptotic preserving approximation.  $\square$

If  $m$  is not 0, under very specific hypotheses over its support, we can extend the result of Theorem 15. The explanation behind this is simple: as long as the population is composed of traits that are not prone to mutations, it will evolve as in the case where mutations are not possible at all.

**Theorem 16.** *Let us assume that there exists  $\hat{x} \in \mathbb{R}^d$  which is an asymptotically stable equilibrium for the ODE  $\dot{x} = a(x)$  such that (5.75) holds. Moreover, let us assume that  $\text{supp}(v^0) \cup K_x$  is a compact set such that there exist  $C', \delta' > 0$  such that*

$$\forall y \in \text{supp}(v^0) \cup K_x, \quad t \geq 0, \quad \|X(t, y) - \hat{x}\| \leq C' e^{-\delta' t},$$

that  $\bigcup_{s \geq 0} (X_s(\text{supp}(v^0) \cup K_x)) \cap K_y = \emptyset$  and that hypothesis (5.76) holds. Then,  $v$  converges to  $\hat{\rho} \delta_{\hat{x}}$  in the weak sense in the space of Radon measures, where  $\hat{\rho}$  is the unique solution of

$$R(\hat{x}, \psi_g(\hat{x}, \hat{x})\hat{\rho}) = 0.$$

Consequently,  $v_\varepsilon^h$  is an asymptotic preserving approximation of  $v$ .

*Proof.* By using the fact that, for any  $x \in \mathbb{R}^d$ ,

$$v(t, x) = v^0(X_{-t}(x))e^{\mathcal{G}(0,t,x)} + \int_0^t \int_{\mathbb{R}^d} m(X_{s-t}(x), z, (I_d v)(s, X_{s-t}(x)))v(s, z)dz e^{\mathcal{G}(s,t,x)} ds,$$

where

$$\mathcal{G}(s, t, x) := \int_s^t R(X_{\tau-t}(x), (I_g v)(\tau, X_{\tau-t}(x))) - \text{div } a(X_{\tau-t}(x))d\tau,$$

we observe that  $\text{supp}(v(t, x)) \subset X_t(\text{supp}(v^0)) \cup \bigcup_{0 \leq s \leq t} X_s(K_x) \subset \bigcup_{s \geq 0} (X_s(\text{supp}(v^0) \cup K_x))$ ,

therefore

$$\begin{aligned} v(t, x) &= v^0(X_{-t}(x))e^{\mathcal{G}(0,t,x)} + \int_0^t \int_{\mathbb{R}^d} m(X_{s-t}(x), z, (I_d v)(s, X_{s-t}(x)))v(s, z)dz e^{\mathcal{G}(s,t,x)} ds \\ &= v^0(X_{-t}(x))e^{\mathcal{G}(0,t,x)} \\ &\quad + \int_0^t \int_{\text{supp}(v(t,x)) \cap K_y} m(X_{s-t}(x), z, (I_d v)(s, X_{s-t}(x)))v(s, z)dz e^{\mathcal{G}(s,t,x)} ds \\ &= v^0(X_{-t}(x))e^{\mathcal{G}(0,t,x)}. \end{aligned}$$

Therefore, we can replicate the proof for the case  $m \equiv 0$ . □

The result of Theorem 15 does not generalize when  $\text{supp}(v^0)$  is not strictly contained in the basin of attraction of  $x_s$ , as shown in the following result:

**Proposition 33.** *Let us consider the one-dimensional PDE*

$$\begin{cases} \partial_t v(t, x) + \nabla_x \cdot (a(x)v(t, x)) = (r(x) - \rho(t))v(t, x), \\ \rho(t) = \int_{\mathbb{R}} v(t, x)dx, \\ n(0, \cdot) = n^0(\cdot), \end{cases} \quad (5.85)$$

which is a particular case of (5.19), and let us assume that there exist  $x_u < x_s$  such that  $a(x_u) = a(x_s) = 0$ ,  $a'(x_u) > 0$ ,  $a'(x_s) < 0$ ,  $\text{supp}(n^0) \subset [x_u, x_s]$ ,  $n^0(x_u) = 0$ , and that there exists  $\alpha > 0$  such that  $n^0(x) = O_{x \rightarrow x_u^+}((x - x_u)^\alpha)$  and that  $r(x_u) - (1 + \alpha)f'(x_u) > r(x_s)$ .

Then,  $v_\varepsilon^h$  is not an asymptotic preserving approximation of  $v$ .

*Proof.* The long-time behaviour of the solution of (5.85) has been studied in detail in Chapter 4, and it has been proved, under the hypotheses of Proposition 33, that  $v$  converges to a function in  $L^1$ . Let us now compute the limit of  $v^h$ . Since  $n^0(x_u) = 0$ , we can assume, without loss of generality, that for all  $i \in \mathcal{J}_h$ ,  $x_i^0 \in (0, 1]$ . Thus, since  $a > 0$  on  $(x_u, x_s)$ , for all  $t \geq 0$ ,  $x_i(t)$  converges to  $x_s$ . As seen in the proof of Proposition 32,  $v^h$  therefore converges to  $r(x_s)\delta_{x_s}$ , and  $v_\varepsilon^h$  is therefore not a asymptotic preserving approximation of  $v$ . □

## 5.5 Simulations

In this section, we present some simulations obtained with the particle method developed throughout the chapter. In Figure 5.1, we deal with the non-local advection equation presented in [59], which writes

$$\partial_t v(t, x) + \nabla_x (a(t, I_1 v(t, x), I_2 v(t, x))) = 0, \quad x \in \mathbb{R}^d, t \geq 0, \quad (5.86)$$

with  $I_j v(t, x) = \int_{\mathbb{R}^2} x_j v(t, x) dx$ , for  $j \in \{1, 2\}$ . Note that this equation is not exactly a particular case of (5.1), since there are two non-local terms involved for advection, but the particle method can straightforwardly be adapted to this case. As in this paper, we show that, depending on the parameters, the solution of this PDE can converge to a single Dirac mass, to a sum of two Dirac masses, or to the sum of four Dirac masses. The parameters used for the simulations are the same as the one detailed in Figures 9, 10 and 11 of [59].

Figure 5.2 illustrates different scenarios for the equation

$$\begin{cases} \partial_t v(t, x) + \nabla_x (a(x)v(t, x)) = (r(x) - \rho(t))v(t, x), \\ \rho(t) = \int_{\mathbb{R}} v(t, x) dx, \\ n(0, x) = n^0(x), \end{cases} \quad (5.87)$$

with  $a(x) = x(1 - x)$ , and an initial solution supported in  $[0, 1]$ . This equation has been studied in Chapter 4 where it has been proved that its solution can either converge to a function in  $L^1$ , (which depends on the initial condition), or to a Dirac mass on 1, depending on the functions  $r$  and  $n^0$ .

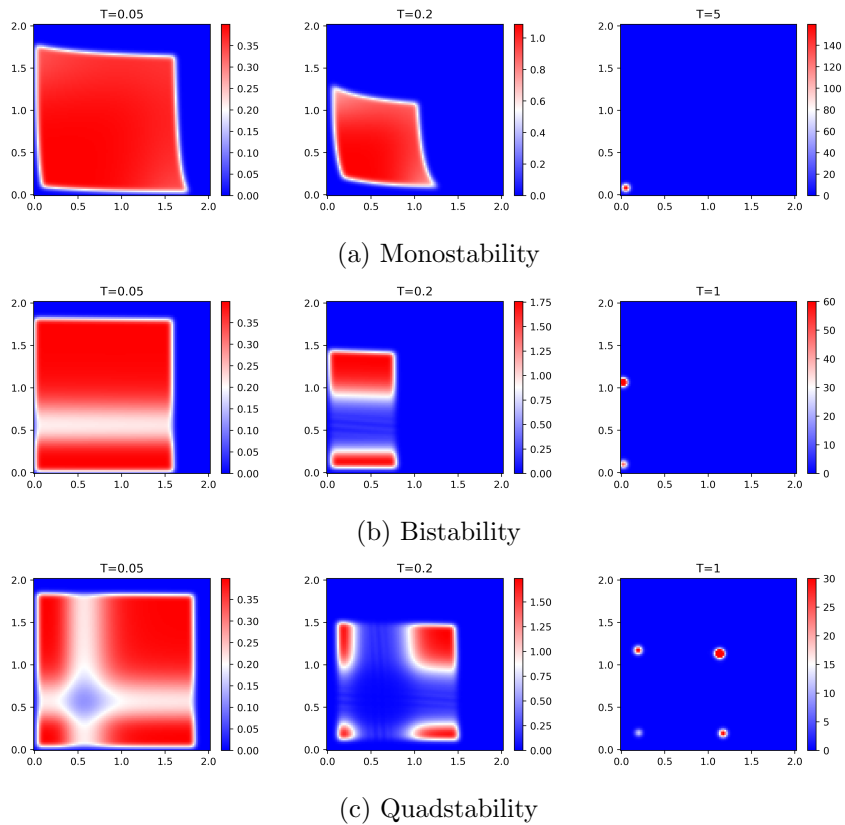


Figure 5.1: The three possible regimes of convergence for equation (5.86), obtained with the particle method. The lines (a), (b) and (c) respectively show the convergence to a single Dirac mass, two Dirac masses and four Dirac masses, and have been obtained by choosing the parameters of Figures 9, 10 and 11 of [59]. In the three cases, we have chosen  $N = 100$ ,  $h = 2./100$ ,  $\varepsilon = h^{0.8}$ , and the cut off function  $\varphi$  is a Gaussian.



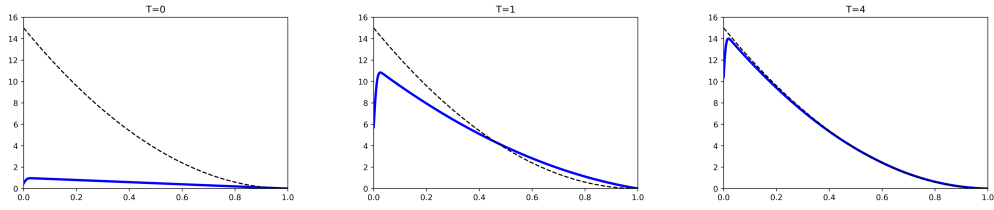
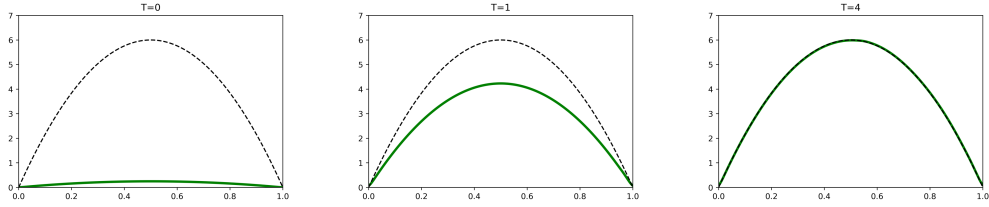
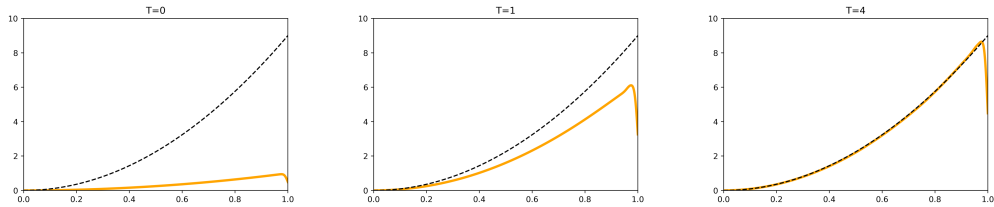
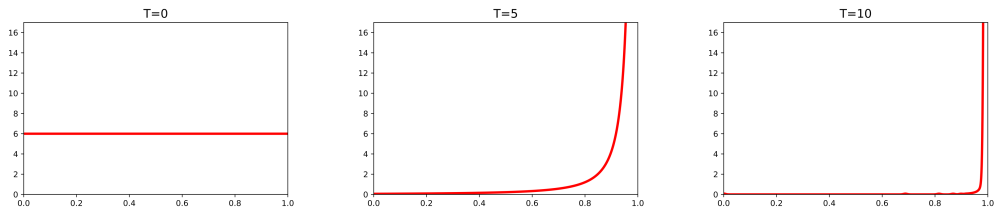
(a) Solution of (5.87) at different time steps, with  $n^0(x) = 1 - x$ ,  $r(x) = 6 - 4x$ .(b) Solution of (5.87) at different time steps, with  $n^0(x) = x(1 - x)$ ,  $r(x) = 6 - 4x$ .(c) Solution of (5.87) at different time steps, with  $n^0(x) = x^2$ ,  $r(x) = 6 - 4x$ .(d) Solution of (5.87) at different time steps, with  $n^0(x) = 6$ ,  $r(x) = 6 - 0.5x$ .

Figure 5.2: Different possible regimes of convergence for the solution of (5.87). The first three lines (green, blue and orange curves), show the convergence to a function in  $L^1$ , which can be explicitly computed (see Chapter 4), and is represented by a black dashed line. Note that the limit function is different when the initial condition changes. The last line (red curves) shows the convergence to a Dirac mass in 1. In the four cases, we have chosen  $a(x) = x(1 - x)$ ,  $N = 5000$ ,  $h = \frac{1}{N}$ ,  $\varepsilon = \sqrt{h}$  and the cut-off function  $\varphi$  is a Gaussian.





## Chapter 6

# An advection-selection-mutation model for epithelial-mesenchymal plasticity

Epithelial-mesenchymal plasticity is a cell differentiation phenomenon which plays a crucial role in cancer development, by allowing them to form metastases (when in the epithelial state), and to spread in the organism (when in the mesenchymal state). Starting from an ODE model which accounts for the cell phenotype of a single cancer cell, we develop an advection-selection-mutation PDE model, in line with the previous chapters, that we solve numerically with a particle method. We first verify that the proposed model reproduces the known behaviours of the associated ODE model, such as hysteresis. We then explore the parameter space for the full model, uncovering the intertwined roles played by different phenomena in a heterogeneous population of cells.

This work, carried out with Paras Jain, Mohit Kumar Jolly, Camille Pouchol and Nastassia Pouradier Duteil, is the subject of an article.

## 6.1 Introduction

Intra-tumour heterogeneity – the co-existence of multiple distinct cellular phenotypes in a tumour – is being increasingly reported to associate with poor patient outcomes. It contributes to both metastasis and therapy resistance – two major unsolved clinical challenges [88, 114]. Such heterogeneity can arise at a genetic level over the course of tumour evolution and manifests as different clonal populations. However, over the last two decades, non-genetic or phenotypic heterogeneity among cancer cells has been identified as a key driver of disease aggressiveness [11, 133]. Such heterogeneity is often characterised by single-cell measurements showing diversity among cells in a population at different levels. A canonical example of non-genetic heterogeneity is along the Epithelial-Mesenchymal (E-M) phenotypic spectrum. Given the implications of E-M heterogeneity in cancer metastasis and patient outcomes, various *in vitro*, *in vivo* and *in silico* attempts have been focused on understanding its underlying mechanisms [17, 89, 95, 141].

An iterative cross-talk among *in silico* and *in vitro* studies has contributed enormously to understanding how non-genetic heterogeneity emerges in a population. Mathematical modelling of regulatory networks underlying Epithelial - Mesenchymal Plasticity (EMP) have reported the co-existence of multiple cellular phenotypes – Epithelial (E), Mesenchymal (M) and one or more hybrid (E/M) cell states [56, 71, 82, 135, 148]. Their co-existence has been experimentally reported in varying ratios in multiple cell lines and primary tumours [61, 95]. The relative stability of cells in different phenotypes, and consequently the phenotypic distribution at a given time, is governed by the underlying topology of the regulatory network involving transcriptional and translational control [72]. Thus, a diversity of regulatory interactions within and among cells can contribute to shaping the E-M heterogeneity patterns in a cell population.

Another milieu of factors such as asymmetric cell division [80, 156], stochastic biochemical noise [123], differences in cellular microenvironment [128], and variable cell-cycle dynamics [108] can amplify dynamic heterogeneity in a cellular population. These factors alter cell-to-cell variability in protein levels in a population, thus contributing to their functional heterogeneity when these cells are exposed to external factors. Despite extensive efforts in investigating above-mentioned regulatory and stochastic processes in E-M plasticity and heterogeneity, only a few computational models have incorporated these processes within a growing and dividing heterogeneous cellular population. Broadly speaking, two modelling approaches are employed:

1. Agent-based models, where each cell is modelled individually,
2. PDE models, which describe densities of cells.

The former approach is more precise since one can follow the fate of each cell, but simulating agent-based models leads to prohibitive computational costs for large number of cells.

We here follow the latter approach to study the population dynamics of E-M heterogeneity. We note that PDE models have widely been adopted because their output (cell density) can be related directly to flow cytometry experiments conducted at multiple timepoints for a population.

Using PDE models, we demonstrate how heterogeneity along the Epithelial - Mesenchymal axis emerges at a population level. First, we show that, as expected, the advection equation reproduces previously reported dynamical features of hysteresis for cells undergoing an Epithelial-Mesenchymal Transition (EMT) followed by a Mesenchymal-Epithelial Transition (MET). Second, we report how cellular heterogeneity depends on the characteristic of epimutations leading to stochastic changes in cell phenotype, whether in the

case of the external EMT-inducing signal (SNAIL), or the E-M state variable (miR200 or ZEB levels). Third, we highlight how differences in the relative growth rates among E, E/M and M states impact how the cell density evolves over time.

All simulations are performed following the particle method detailed in Chapter 5.

## 6.2 Advection equation and hysteresis

### 6.2.1 ODE model and corresponding advection equation

In vitro and in silico studies on EMP have demonstrated a hysteretic behaviour as cancer cells undergo one cycle of EMT and MET, i.e., an asymmetry in EMT and MET trajectories. For our further analyses, we choose a minimal EMT regulatory network with canonical epithelial (microRNA-200 (miR200)) and mesenchymal (ZEB) players that mutually inhibit each other. An EMT-inducing transcription factor SNAIL that activates ZEB and inhibits miR-200 represents the cumulative effect of the extracellular environment [110]. The model writes

$$\begin{cases} \dot{\mu}_{200} &= g_{\mu_{200}} H_{Z, \mu_{200}}(Z) H_{S, \mu_{200}}(S) - g_{m_Z} H_{Z, m_Z}(Z) H_{S, m_Z}(S) Q(\mu_{200}) - k_{\mu_{200}} \mu_{200} \\ \dot{Z} &= g_Z g_{m_Z} H_{Z, m_Z}(Z) H_{S, m_Z}(S) P(\mu_{200}) - k_Z Z \end{cases} \quad (6.1)$$

We denote  $F$  the right-hand side of this ODE model, which models interactions between a transcription factor ZEB (denoted  $Z$ ), and a micro-RNA miR-200 (denoted  $\mu_{200}$ ). The variable  $S$  represents a third molecule, SNAIL, which is seen in our case as an external signal characterising the extracellular environment. All the parameters of this model are given in Appendix A.

Within this framework, the three cell states (E, E/M, and M) are characterised as follows:

- The epithelial state corresponds to a high level of miR-200, and a low level of ZEB;
- The mesenchymal state corresponds to a low level of miR-200 and high level of ZEB;
- The hybrid state corresponds to a medium level of miR-200 and ZEB.

Depending on the value of  $S$ , system (6.1) exhibits monostability, bistability or tristability, as illustrated by the bifurcation diagram (Figure 6.1 A).

The bifurcation diagram depicts the different possible stable states, each characterised by specific ranges of miR200 levels (solid lines) for increasing levels of SNAIL, resulting from the network dynamics (Figure 6.1 A). As a cell undergoes EMT (i.e. SNAIL levels increase), it switches from high to intermediate to low levels of miR200 which corresponds to the E, E/M and M states respectively. However, during MET, the cell switches directly from low (M) to high (E) miR200 levels without passing through the hybrid E/M state, thus displaying hysteresis.

**Homogeneous population.** To reproduce the hysteretic behaviour of EMP, *i.e.* asymmetry in EMT and MET trajectories, we first consider the advection equation

$$\partial_t u(t, y) + \nabla \cdot (f(t, y) u(t, y)) = 0,$$

where  $y = (\mu_{200}, Z)$ , and  $f(t, \mu_{200}, Z) = F(\mu_{200}, Z, S(t))$ , and  $t \mapsto S(t)$  is the piecewise-affine function which connects the points  $(0, 160K)$ ,  $(5000, 240K)$  and  $(10000, 160K)$ , as represented in Figure 6.1 B.

**Heterogeneous population.** In order to account for heterogeneity within the population, and more specifically for the fact that the signal  $S$  can be interpreted in a different way by each cell, we incorporate  $S$  within the structure variable, which becomes  $y = (\mu_{200}, Z, S)$ . SNAIL level variation then impacts the advection term, which becomes  $f(t, y) = (F(y), f_S(S))$ , where  $f_S$  is the step function corresponding to the derivative of the function  $S$  introduced in the previous paragraph, *i.e.*  $f_S(S) = 40$  if  $S \in [0, 5000)$ , and  $f_S(S) = -40$  if  $S \in (5000, 10000]$ . As initial condition, we take  $u^0(\mu_{200}, Z, S) = \frac{1}{\sigma}G(\frac{S-160K}{\sigma})$ , with  $\sigma = 20K$ , and  $G$  denotes the Gaussian function.

## 6.2.2 Results

When sticking to the exact PDE analogue of the ODE (6.1), *i.e.*, the advection equation, we recover the hysteretic phenomenon at the cell population level. In this case, we insist that cell growth and epimutations are not taken into account. The SNAIL dynamics used to obtain hysteresis are shown by the blue curve (Figure 6.1 B).

We simulate the dynamics for homogeneous and heterogeneous cell populations with respect to their distribution of SNAIL levels. For a homogeneous cell population (Figure 6.1 C, i), we see that cells reside in three distinct miR200 states (high, intermediate, and low) for time 0 to 5000 hours, but make a quick transition from low to high miR200 levels for time 5000 to 10000 units without spending much time in the intermediate state (decreasing SNAIL levels). The unexpected intermediate miR200 levels seen during MET can be explained by considering it as a sample timepoint where miR200 level is responding to changes in SNAIL levels before settling to its equilibrium (high) state.

Similar observation of hysteresis is made while considering a heterogeneous cell population (Figure 6.1 C, ii). Particularly, the cell distribution along ZEB and miR200 axis during MET shows that the transient intermediate miR200 peak in homogeneous population turns into little dispersed transient peaks which are clearly distinct from the intermediate peak arising during EMT (Figure A.1). This observation recapitulates the experimental data on different partial states seen during EMT vs. during MET [95]. Also, in the heterogeneous population case, some cells fail to complete a full EMT but rather undergo a partial transition and then return to an epithelial state upon reduced SNAIL levels (Figure 6.1 C, ii).

## 6.3 Effect of epimutations and growth

In a population exhibiting growth, the cell-state (characterised by levels of a set of specific biomolecules) is dynamically evolving due to stochasticity in biochemical reactions, thus causing heterogeneity. To evaluate how these stochastic processes influence the population distribution of E, M and E/M states, we next observe cell state distribution as a result of epimutations in a regulatory network for a population where cell division rate is independent of cell-state (*i.e.* EMT does not impact cell cycle). Because the levels of EMT-inducing signal SNAIL are also evolving due to noise, its levels are distributed in the population around the mean environmental characteristics  $S_0$  (Figure 6.2 A, i).

### 6.3.1 Model

**Reducing the dimensions of the structure variable.** When incorporating growth and mutations into the model, computation times becomes significant. To ease the burden, we use dimension reduction by further simplifying the advection function: we consider the phenotype  $y = (x, S)$  rather than  $(\mu_{200}, Z, S)$ , where  $x$  is roughly equivalent to  $\mu_{200}$ .

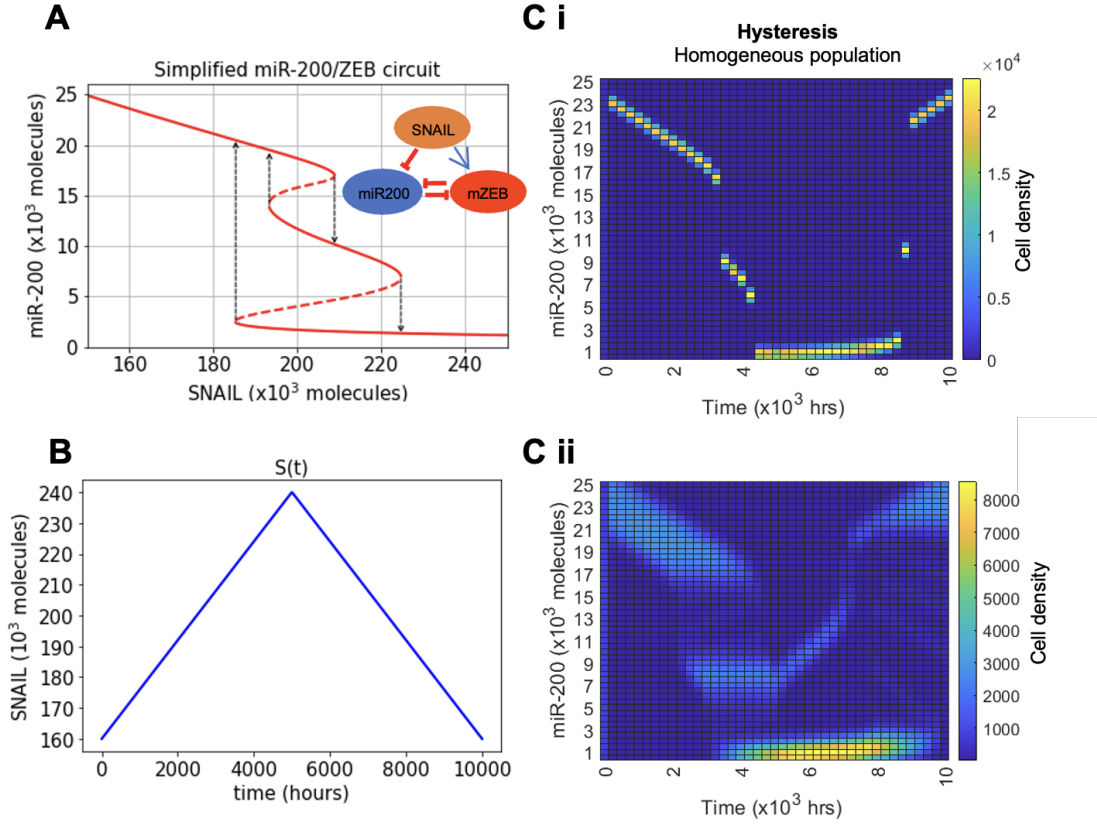


Figure 6.1: The advection equation reproduces hysteresis at the population level. A) EMT gene regulatory network (inset) and the bifurcation of cell states resulting from the network dynamics with increasing input signal (SNAIL) levels, B) Variation in the external input (SNAIL) levels with time to capture hysteresis regulation of cell states in C. C) Hysteresis (non-symmetric trajectories) in cell state transition during one cycle of EMT and MET by varying SNAIL levels as shown in Figure 1B while considering – i) Homogeneous population and ii) Heterogeneous population.

This requires to choose a function  $f_r$  such that the dynamics of  $x$  for any  $S$  is given by  $\dot{x} = f_r(x, S)$ . Our main requirement in choosing this new function is to preserve the bifurcation diagram given in Figure 6.1 A, which means that, for a given value of  $S \in [150K, 250K]$ ,  $f_r(\cdot, S)$  has the same number of zeros as  $F(\cdot, S)$ , and are such that  $f_r(x, S) = 0$  if and only if there exists  $Z > 0$  such that  $F(x, Z, S) = 0$ .

Since infinitely many functions satisfy this property, one must make further assumptions in order to select an adapted one. For simplicity and to avoid overfitting, we assume that, for any  $S$ ,  $f_r(\cdot, S)$  is piecewise affine (with one segment by root) and that the rate of change is constant on each piece. Under these constraints, the function is defined up to a multiplicative constant, which has been chosen in order to minimise the value

$$\int_0^T \int_{150K}^{250K} \int_0^{25K} \int_0^{800K} |x(t, S, x_0) - \mu(t, S, x_0, Z_0)| dt dS dx_0 dZ_0, \quad (6.2)$$

where for all  $x_0 \in [0, 25K]$ ,  $S \in [150K, 250K]$ ,  $x(\cdot, S, x_0)$  solves

$$\begin{cases} \dot{x}(t, S, x_0) = f_r(x(t, S, x_0)) \\ x(0, S, x_0) = x_0 \end{cases},$$



and for all  $Z_0 \in [0, 800K]$ ,  $(\mu(\cdot, S, x_0, Z_0), Z(\cdot, S, x_0, Z_0))$  satisfies (6.1).

The value of the multiplicative constant which minimises (6.2) depends on  $T$  but we establish that, for  $T \in \{10, 100, 1000\}$  its value is rather insensitive to that of  $T$ : it is close to 0.02, which is the value that we select. The obtained function is shown in Figure A.2 for various  $S$  values.

We first assess the accuracy of our dimension reduction approach, establishing similarity between the dynamics of system (6.1) with two variables (full EMT) and that of the reduced system. We compare the distribution of miR200 levels for a given SNAIL distribution with the distribution of the cell population along the state  $x$  (Figure 6.2 A and Figure A.2 B). For example, with a SNAIL distribution with a mean value ( $S_0$ ) of 200K molecules, the bifurcation diagram depicts the possibility for the cell population to be distributed in all three states (Figure 6.2 A, i). We observe that the asymptotic distribution of the state variable  $x$  from the reduced system dynamics exhibits tri-modality, showing the co-existence of all three states, irrespective of the initial condition (Figure 6.2 A, ii – initial condition: all cells as epithelial, Figure A.2 B – initial condition: all cells as hybrid or mesenchymal). Similarly for a SNAIL distribution with mean  $S_0 \in \{190K, 225K, 150K, 250K\}$  molecules, we observe respective combinations of phenotypes as in bifurcation diagram of the full EMT network – co-existence of E and E/M (bimodal), co-existence of hybrid E/M and M (bimodal), epithelial (unimodal) and mesenchymal (unimodal) (Figure 6.2 A), thus providing further evidence that EMT dynamics are preserved with the sole variable  $x$ .

**Considering growth and epimutations.** The full model incorporating growth and epimutations writes

$$\begin{cases} \partial_t u(t, y) + \nabla \cdot (f(y)u(t, y)) = (r(y) - d(y)\rho(t)) u(t, y) \\ \quad + \int_{\mathbb{R}^2} M(y, z)u(t, z)dz - \int_{\mathbb{R}^2} M(z, y)dz u(t, y) \\ \rho(t) = \int_{\mathbb{R}^2} u(t, y)dy \end{cases}, \quad (6.3)$$

where

- The structure variable  $y \in \mathbb{R}^2$  writes  $y = (x, S)$ .
- The advection function  $f$  writes  $f(y) = f(x, S) = (f_r(x, S), f_S(S))$ , where  $f_S(S) = \delta(1 - \frac{S}{S_0})$ , with  $S_0 \in [150K, 250K]$  corresponding to the mean of the SNAIL distribution,  $\delta := \frac{S_0 \ln(2)}{\alpha}$ , and  $\alpha > 0$  representing the characteristic time of convergence of SNAIL to the mean  $S_0$ .
- $r(y)$  and  $d(y)$  respectively represent the birth and death rate of a cell of phenotype  $y$ . In all our simulations, the death rate is constant ( $d \equiv 1.82 \times 10^{-7}$  cell/hr).
- The mutation function  $M$  writes  $M(y, z) = r(z)P(y - z)$ . Here,  $P(x) = P(y, S) := \frac{1}{\eta_x \eta_S} G(\frac{x}{\eta_x}) G(\frac{S}{\eta_S})$ , where  $G$  is the Gaussian function. Variables  $\eta_x$  and  $\eta_S$  are the standard deviations for  $x$  and  $S$  respectively.

### 6.3.2 Focus on epimutations

We now turn to simulations for the full model (6.3) where we investigate the effect of epimutations; growth functions  $r$  and  $d$  are kept constant.

For multi-modal distributions of state variable  $x$ , the exact phenotypic composition depends on relative stability of the multiple co-existing phenotypes. For example, we see a reduced share of hybrid E/M (intermediate  $x$  levels) cells for distribution of SNAIL levels that overlap significantly with those that have mean SNAIL levels corresponding

to monostable E or monostable M regions ( $S_0 = 190\text{K}$ ,  $225\text{K}$  molecules respectively), especially for a reduced standard deviation  $\eta_x$  of epimutations in state  $x$  (Figure 6.2 Bi, ii). Similarly, for distribution of SNAIL levels with mean  $S_0 = 175\text{K}$  molecules, although both E and M states co-exist (Figure 6.2 A, i), the relative stability of the E state is much greater than that of the M state, thus disallowing cells to make transition to the M state even at higher levels of  $\eta_x$  (Figure 6.2 B, iii). As mentioned previously, the distribution of SNAIL levels and correspondingly that of cell state  $x$  in a population can be attributed to stochasticity in biochemical reactions.

Another type of perturbation in cellular variables (here,  $x$  and SNAIL) can arise when certain subpopulations are isolated and re-cultured independently. For instance, when the E, M and hybrid E/M prostate cancer subpopulations are segregated, they exhibit very different distributions after two weeks [139]. Similarly, the segregated EpCAM-high and EpCAM-low subpopulations in breast cancer have varied recovery dynamics [12]. Thus, in case of either internal (epimutations) or external (microenvironmental factors) perturbations, the rate at which the cellular variables recover towards the characteristic distribution can differ even though they eventually converge to the same equilibrium distribution.

Thus, we next modulate the rate of recovery of SNAIL levels to mimic the scenario of extrinsic perturbation to the cell population by isolating distinct subpopulations and simulating (re-culturing) them independently. The rate of recovery to perturbations in SNAIL levels is inversely proportional the parameter  $\alpha$  in our model, which captures the characteristic time of convergence of SNAIL levels to its equilibrium  $S_0$  [90, 145]. Starting from a mostly mesenchymal population (high values of  $x$ ) the convergence to a mostly epithelial population slows down with increasing values of  $\alpha$  (Figure 6.2 C).

Furthermore, the slowed down dynamics increases cellular heterogeneity by making the population to be distributed in all three states for a considerable interval of time, as quantified by Renyi entropy (Figure 6.2 D), defined for a positive function  $u \in L^1(\mathbb{R}^d)$  as  $-\int_{\mathbb{R}^d} \tilde{u}(x) \ln(\tilde{u}(x)) dx$ , with  $\tilde{u} := u/\|u\|_{L^1}$ . In the example shown, the population heterogeneity first increases as the population shifts from majority epithelial to more uniformly distributed among the three phenotypes in the intermediate time points and then decreases as the population turns to majority M cells. Similar observations can be made for any combination of initial condition and mean  $S_0$  values of SNAIL distribution (Figure A.3, Figure A.4).

Overall, the interplay between deterministic and stochastic dynamics of cellular biomolecules shapes the population heterogeneity. This is done both by distributing cells in all plausible states permitted by the underlying ODE model, and by slowing down the kinetics of cells towards equilibrium when perturbed by the external signal  $S$ .

### 6.3.3 Focus on growth

So far, we considered all three phenotypic states (E, M and hybrid E/M) to divide at equal rates. To observe the additional influence of growth rate differences on E-M heterogeneity we considered three possible scenarios – a) case ‘r1’: all the three phenotypes divide at same rate, b) case ‘r2’: both E and hybrid E/M cells divide at equal rates, while M cells divide at half the rate of E cells; and c) case ‘r3’: both E/M and M divide at equal rates, which is half the rate of division of E cells. In practice, this means considering three different possible piecewise-constant functions  $r$ , as illustrated in Figure 6.3 A ii.

Across all these cases, the E state divides at either an equal or a faster rate than hybrid E/M and/or M cells. This constraint recapitulates the current experimental understanding that EMT may suppress cell cycle to varying extents, thus reducing the division rate of hybrid E/M and/or M cells.

First, we investigate how the phenotypic composition of the population changes with

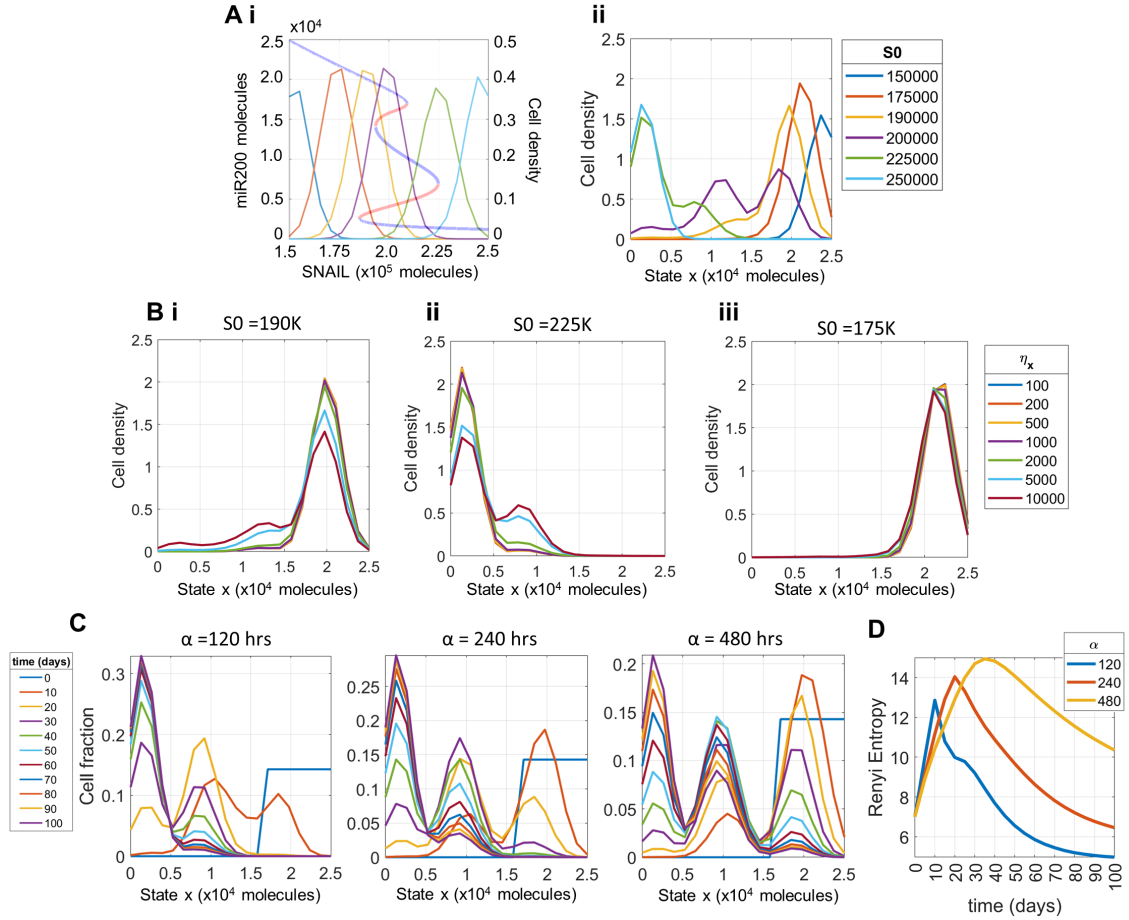


Figure 6.2: Simulations of the advection-selection-mutation (PDE (6.3)) with reduced function  $f_r$ . Empirical reduced single cell state variable  $x$ . Ai) Bifurcation diagram showing stable (blue) and unstable (red) levels of miR200 ( $y$  axis on left) for increasing levels of SNAIL levels; and distribution of input SNAIL levels among the cells ( $y$  axis on right) for different mean characteristics levels  $S_0$ . Aii) Distribution of cells along the state  $x$  of reduced EMT system for different distributions of SNAIL shown in Ai. B) Changes in population's phenotypic distribution with increasing levels of epigenetic noise  $\eta_x$  for SNAIL distributions with mean  $S_0 = 190K$  (Bi) and  $225K$  (Bii). For  $S_0 = 175K$  (Bii), the population remains invariant to increasing noise. C) Temporal changes in cell state distribution of the population for decreasing value of input signal SNAIL's perturbation recovery rate (increasing values of the characteristic time ' $\alpha$ '). D) Changes in population's heterogeneity (measured by Renyi entropy) with time for different  $\alpha$  parameter values. Parameters used to generate plots, unless stated otherwise, are  $\alpha = 120hrs$ ,  $\eta_x = 1000$ , ini pop Epi, time point 20,  $S_0 = 225K$  molecules, and per-capita growth rate  $r$  is constant across phenotypes, given by  $r = 0.0182/hr$ .

different growth rate scenarios (Figure 6.3 , Figure A.5). For input condition as SNAIL distribution with mean  $S_0 = 190K$  or  $200K$  molecules that predominantly enables an E state with/without hybrid E/M state, the reduced growth of M cells has very slight effects on phenotypic composition over the time course, as expected (Figure 6.3 B, i). The initial peak in hybrid cell fraction for the 'r2' growth scenario is the combined effect of M to hybrid E/M state transition and a relatively higher growth rate of hybrid E/M cells. However, when hybrid E/M cells also have reduced proliferation (growth scenario

‘r3’), we see a lasting change in the phenotypic composition as E cells become dominant because of higher division frequency (Figure 6.3 B, i -  $S_0 = 200K$ ). For the input SNAIL distribution with a mean value  $S_0$  of 225K molecules that majorly supports hybrid E/M and M phenotypes, in the case ‘r2’, the growth advantage provided to hybrid E/M cells enables their dominance in the population on the long run, when compared to the growth scenarios of ‘r1’ or ‘r3’ where both E/M and M cells proliferate at equal rates (Figure 6.3 Bi,  $S_0 = 200K$ ). The initial peaks in hybrid E/M fractions are combined effects of E to hybrid E/M transitions with either growth similarity or advantage of hybrid E/M cells over M cells.

The overall change in phenotypic composition can be calculated using Renyi entropy as a heterogeneity score (Figure 6.3 B, ii). Although, growth scenarios ‘r1’ and ‘r2’ have the same phenotypic composition and an equal heterogeneity score eventually, the growth scenario ‘r1’ shows a much smoother change in heterogeneity values from the initial levels because of all three phenotypes being equally proliferative.

Next, we look at the effects of increasing level of epigenetic noise level  $\eta_x$  in cell state  $x$  on phenotypic composition laid down by growth rate differences (Figure A.6 A-D). For  $S_0 = 190K$  and  $200K$ , where hybrid E/M state is less dominant than the E state (Figure 6.3 B), increasing the noise levels (from  $\eta_x = 1000$  to  $\eta_x = 5000$ ) in state  $x$  causes more cell-state transitions, raising the frequency of hybrid E/M phenotype in population for all growth scenarios (Figure A.6 B). However, as for  $S_0 = 225K$  and growth scenario ‘r2’ where hybrid phenotype is the dominant state, increasing the noise  $\eta_x$  level (from  $\eta_x = 1000$  to  $\eta_x = 5000$ ) raises M fractions in the population even though M cells are dividing slowly (Figure 6.3 B,i; Figure A.6 B-C). Overall, we observe an increase in population heterogeneity with higher noise levels in state variable ‘ $x$ ’, irrespective of the growth scenario (Figure 6.3 Bii, Figure A.6 D).

After observing changes in phenotypic composition for different growth scenarios, we move on to see how total cell population grows for combinations of initial conditions and growth scenarios. We consider six different conditions – isolated E, hybrid E/M, and M population, uniform mixture of either E and M or E, E/M and M cells, and uniform distribution of cells in all possible cell states  $x$  and input SNAIL levels.

When all the phenotypic states are dividing at equal rates, the total number of cells does not vary across different initial conditions. However, with M dividing slower than E and hybrid E/M cells (growth scenario ‘r2’), we see that an initially mesenchymal population has slower population growth compared to other initial conditions. Similarly, with both hybrid E/M and M cells dividing slower than E (growth scenario ‘r3’), initial conditions having component of either E/M or M cells divide slower than isolated (pure) E cells. Also, as transition from hybrid E/M to E is much probable than M to E transitions, presence of hybrid cells in the populations increases the overall growth rate (Figure 6.3 C compare ‘hyb vs mes’, and ‘epi mes’ vs ‘epi hyb mes’ initial conditions for  $S_0 = 200K$ , ‘r3’ growth scenario). Further, increasing the level of epigenetic noise (from  $\eta_x = 1000$  to  $\eta_x = 5000$ ) causes more state transitions rendering lesser variability in the growth dynamics by quickly equalising the effect of differences in the initial conditions (Figure A.6 E).

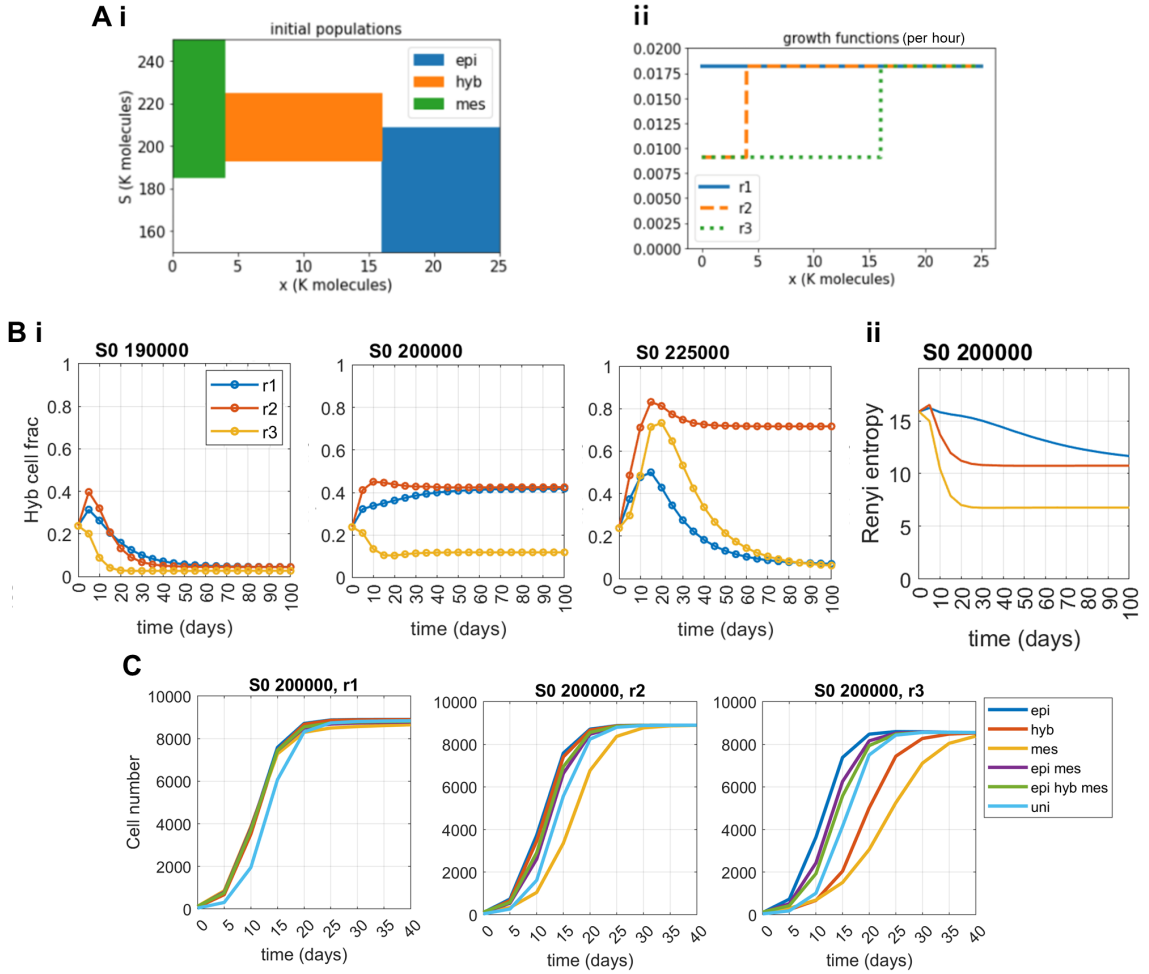


Figure 6.3: Effect of growth rate differences among E, hybrid E/M and M phenotypes on population's heterogeneity and overall growth dynamics. Ai) Support of the different initial conditions: In each case, the initial condition is uniformly distributed on its support, and such that the total population is equal to 100 cells. Aii) Profile of different growth functions: 'r1': All three phenotypes divide at same rate; 'r2': E and E/M divide at equal rates, while M divide at half the rate of E cells; and 'r3': Both E/M and M divide at equal but half the rate of E cells. Bi) Temporal changes in hybrid cell fraction in the population for different growth scenarios among phenotypes. Bii) Changes in population's heterogeneity (measured by Renyi entropy) with time. C) Population growth dynamics for different combinations of growth scenarios and initial condition; 'epi hyb mes' corresponds to an initial condition uniformly distributed on the three colored domains of A1), and uni to a uniform population on the whole rectangle  $[0, 25K] \times [150K, 250K]$ . For panel B, the initial condition is uniformly distributed in all states. The input SNAIL mean ( $S_0$ ) levels used are mentioned for all the individual plots. Other parameters used to generate plots are  $\alpha = 120hrs$ ,  $\eta_x = 1000$ , and the per-capita growth rate ( $r$ ) of the epithelial phenotype is  $0.0182/hr$ .





# Appendix A

## Appendix of Chapter 6

### A.1 Supplementary figures

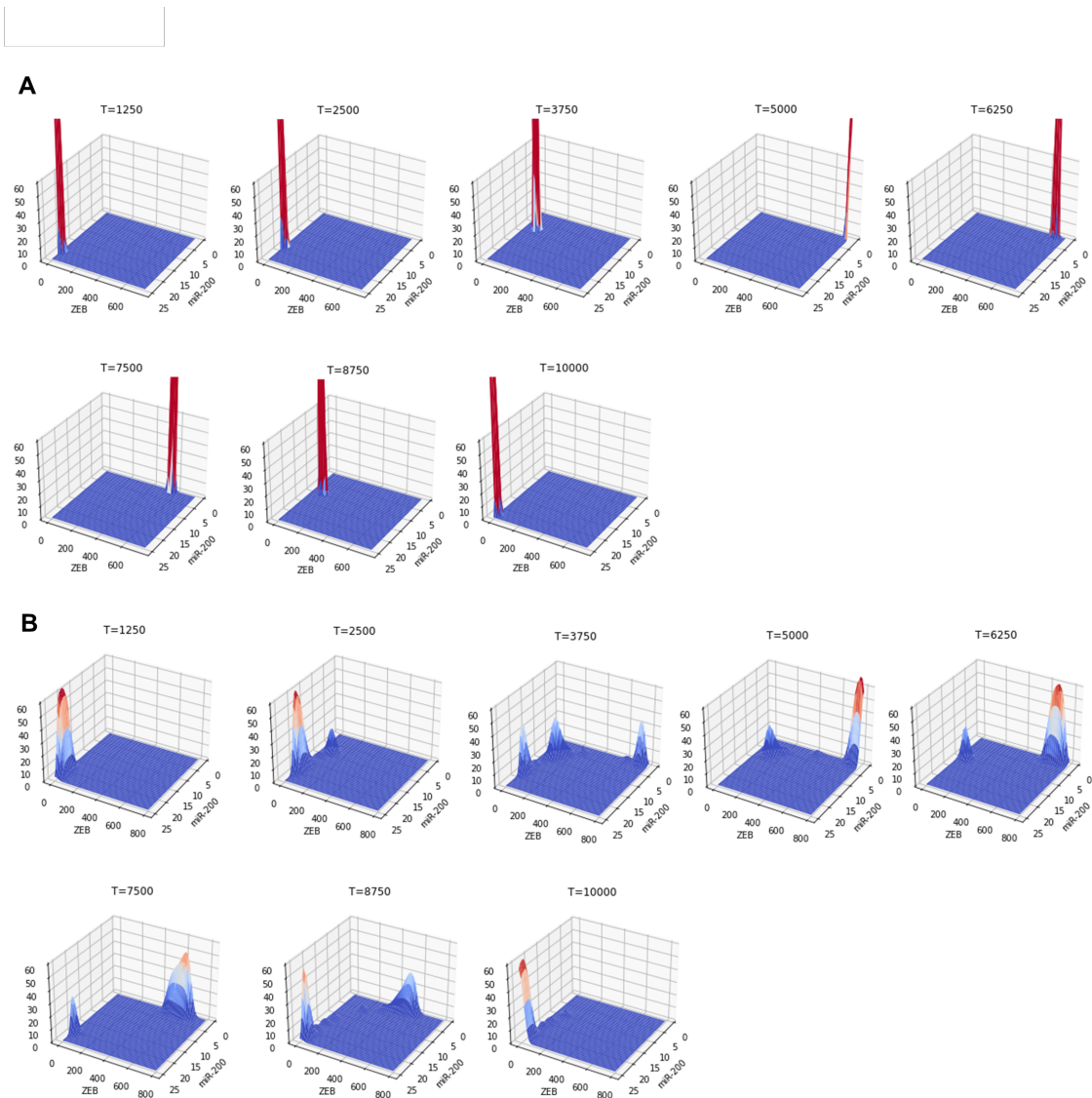


Figure A.1: Hysteresis (non-symmetric transition) in cell density (z-axis) along the two cell state variables miR200 and ZEB during one cycle of EMT and MET caused by increasing and decreasing levels of input SNAIL levels (Figure 1B blue curve) for A) homogeneous and B) heterogeneous cell population.



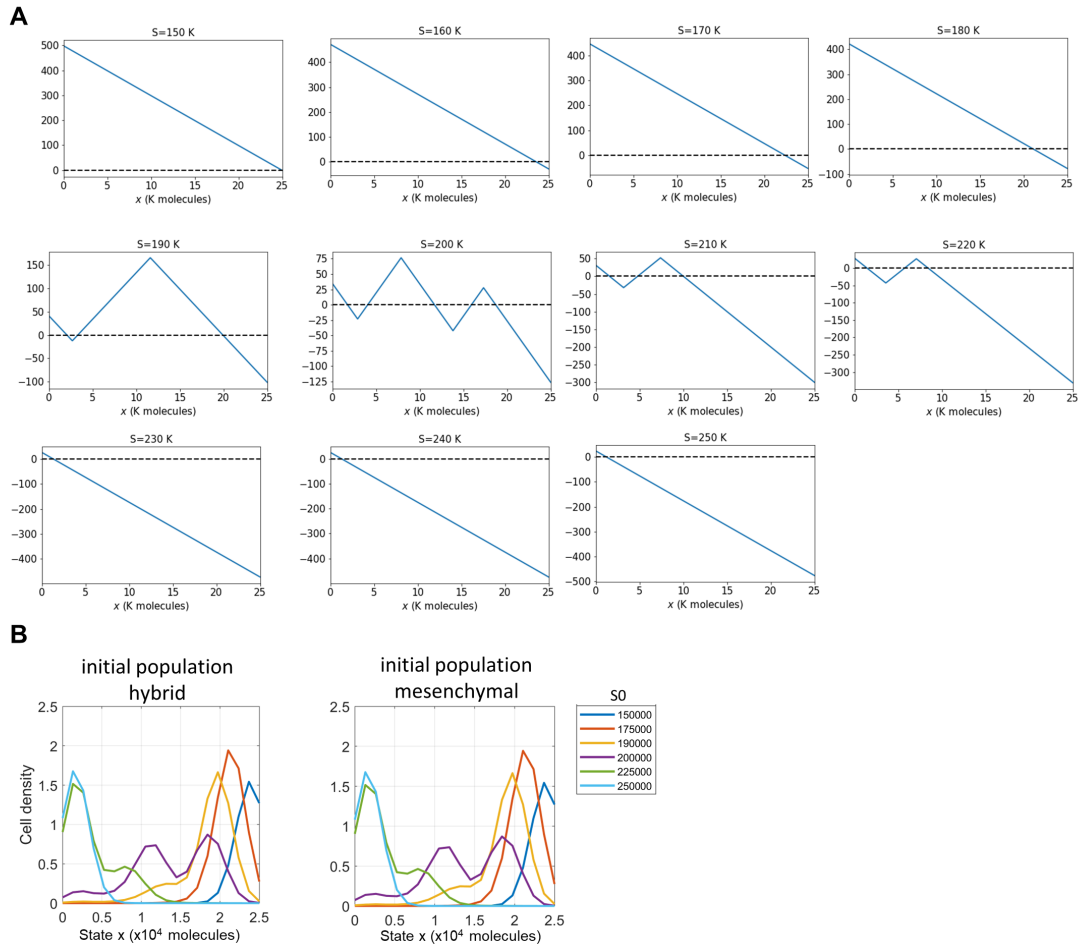


Figure A.2: Reduced function  $f_r$  underlying the ODE with single variable cell state  $x$  and single input SNAIL; and distribution of the cell population for different distributions of SNAIL shown in Figure 2Ai. A) Function  $x \mapsto f_r(x, S)$ , showing the existence of mono-, bi- and tri-stable states for varying levels of SNAIL inputs  $S$ . B) Cell population distribution at the end point of simulations started with hybrid and mesenchymal populations for increasing levels of input signal SNAIL's mean ( $S_0$ ) levels. Parameters used to generate the above plots, unless stated otherwise, are  $\alpha = 120$  hrs,  $\eta_x = 1000$ , ini pop Epi,  $S_0 = 200K$  molecules, and the per-capita growth rate ( $r$ ) is constant =  $0.0182 \backslash hr$ .

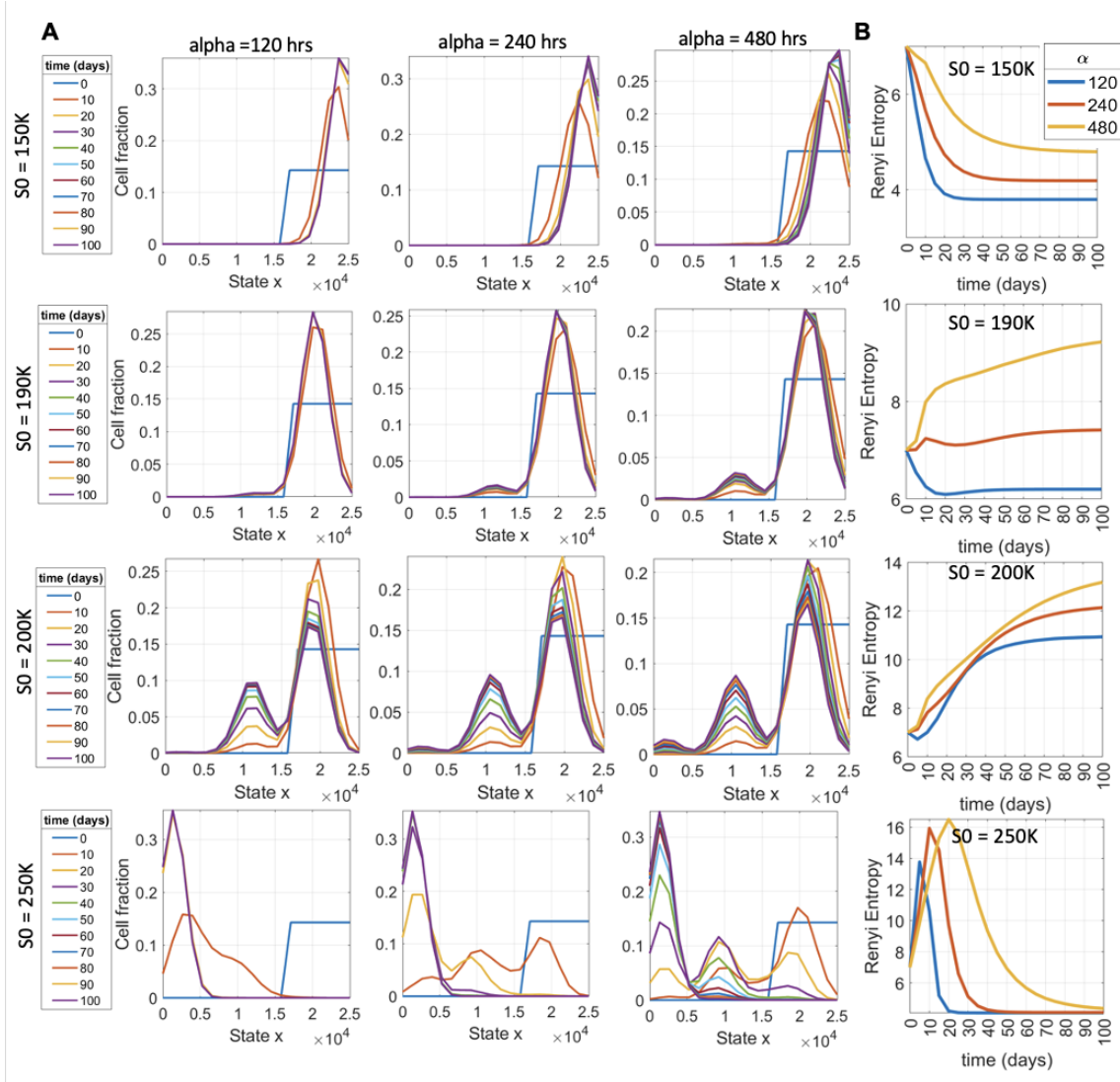


Figure A.3: Population dynamics and changes in population heterogeneity (measured using Renyi Entropy) with time for several combinations of  $S_0$  and  $\alpha$  parameters. The intermediate  $S_0$  values of 200K molecules where a cell can either attain a stable epithelial, hybrid, and mesenchymal, has highest heterogeneity score. Further, as we go towards smaller or larger values of  $S_0$  where only epithelial, and mesenchymal states are possible, respectively, the population has least heterogeneity. And, by increasing the residence time of cells in a state  $x$ , cells get enough time to populate the cell state they reside in, and thereby, contribute to a significant fraction in the overall population. For the above plots, the initial condition population is epithelial, and  $\eta_x = 1000$ , and per-capita growth rate  $r$  is constant =  $0.0182/hr$ .

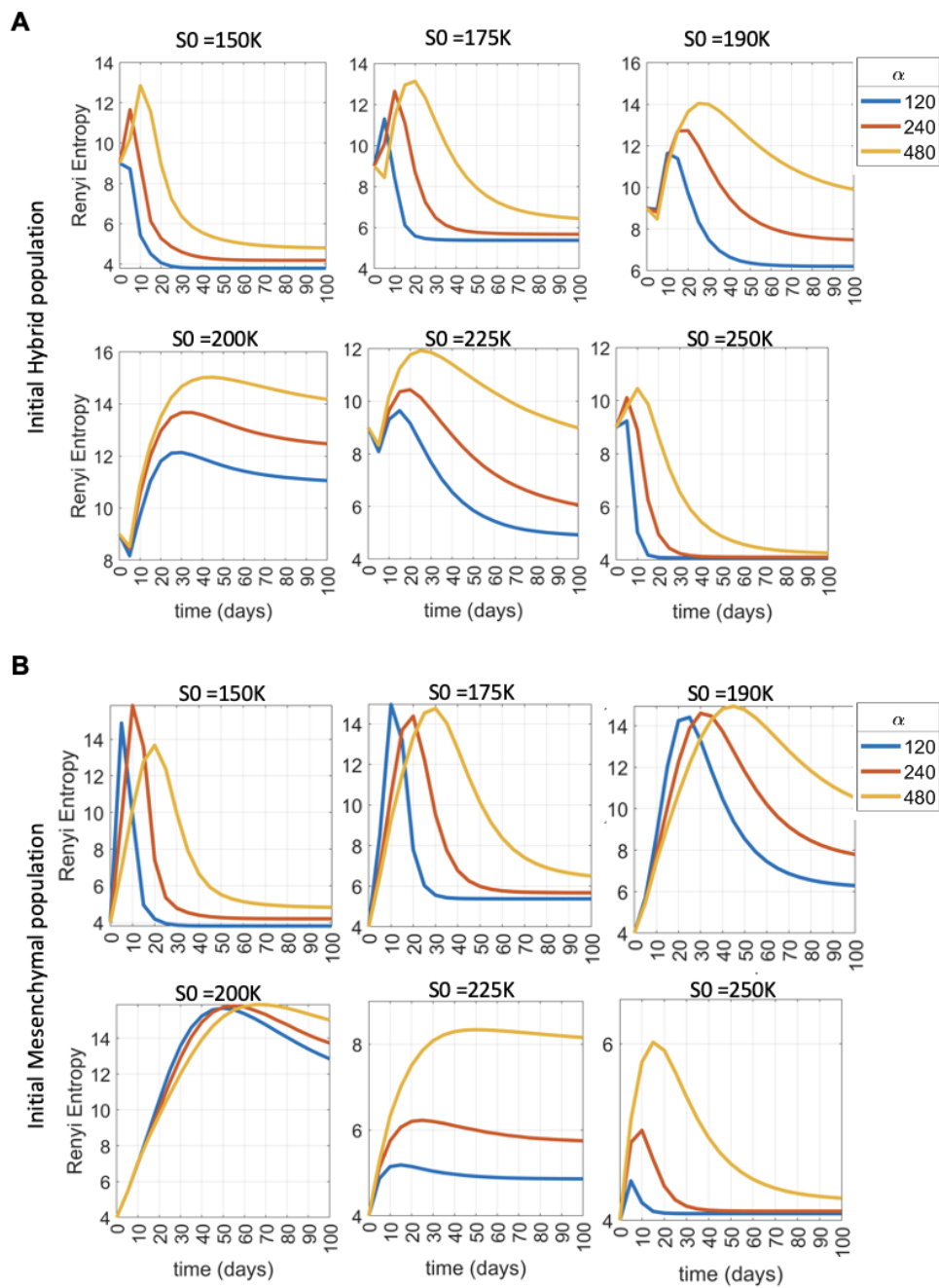


Figure A.4: Changes in population heterogeneity (measured using Renyi Entropy) with time for several combinations of  $S_0$  and  $\alpha$  values while starting with a population of A) Hybrid cells, and B) Mesenchymal cells. For the above plots,  $\eta_x = 1000$ , and per-capita growth rate  $r$  of all subpopulation = 0.0182/hr.

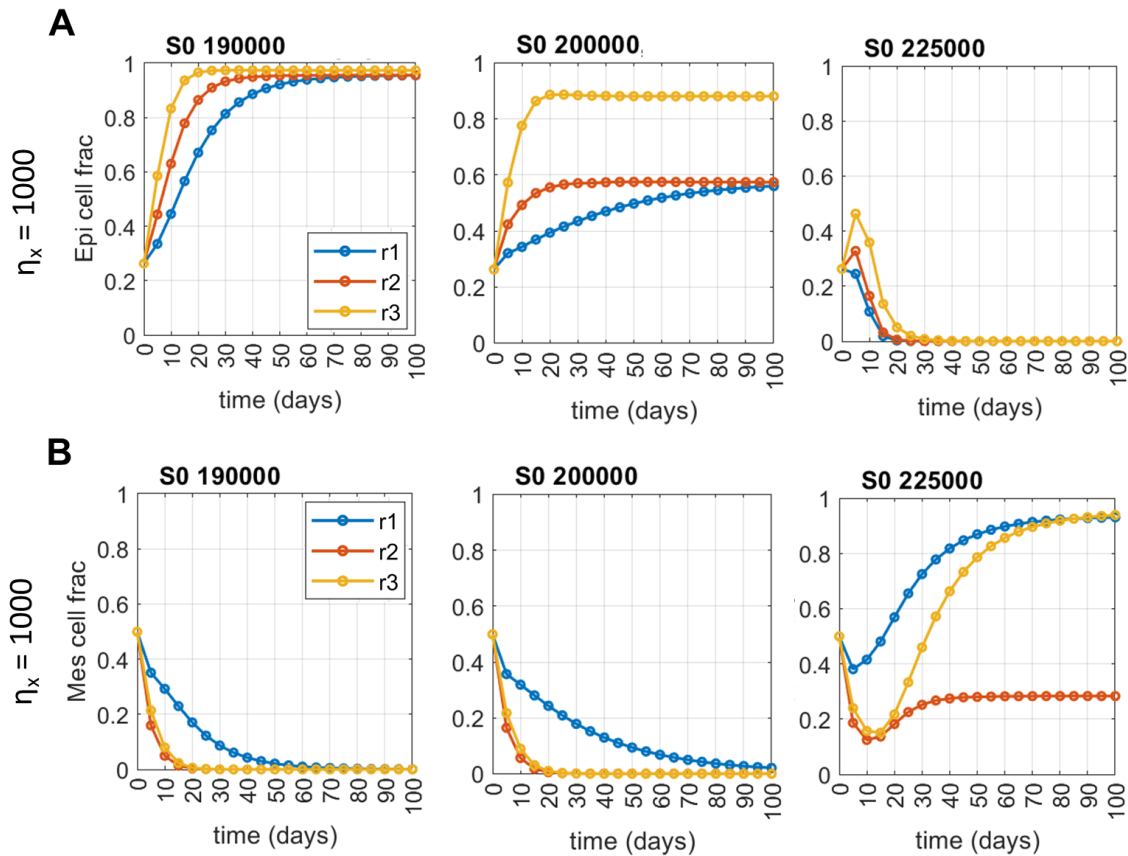


Figure A.5: Effect of growth rate differences among E, hybrid E/M and M phenotypes on population's heterogeneity. Temporal changes in E cell fraction (panel A) and M cell fraction (panel B) in the population for different growth scenarios among phenotypes – ‘r1’: All three phenotypes divide at same rate; ‘r2’: E and E/M divide at equal rates, while M divide at half the rate of E cells; and ‘r3’: Both E/M and M divide at equal but half the rate of E cells. For the above results, the initial condition is uniformly distributed in E, hybrid E/M and M state. The input SNAIL mean  $S_0$  levels used are mentioned for all the individual plots. Other parameters used to generate plots are  $\alpha = 120$  hrs,  $\eta_x = 1000$ , and per-capita growth rate  $r$  of epithelial phenotype is 0.0182/hr.

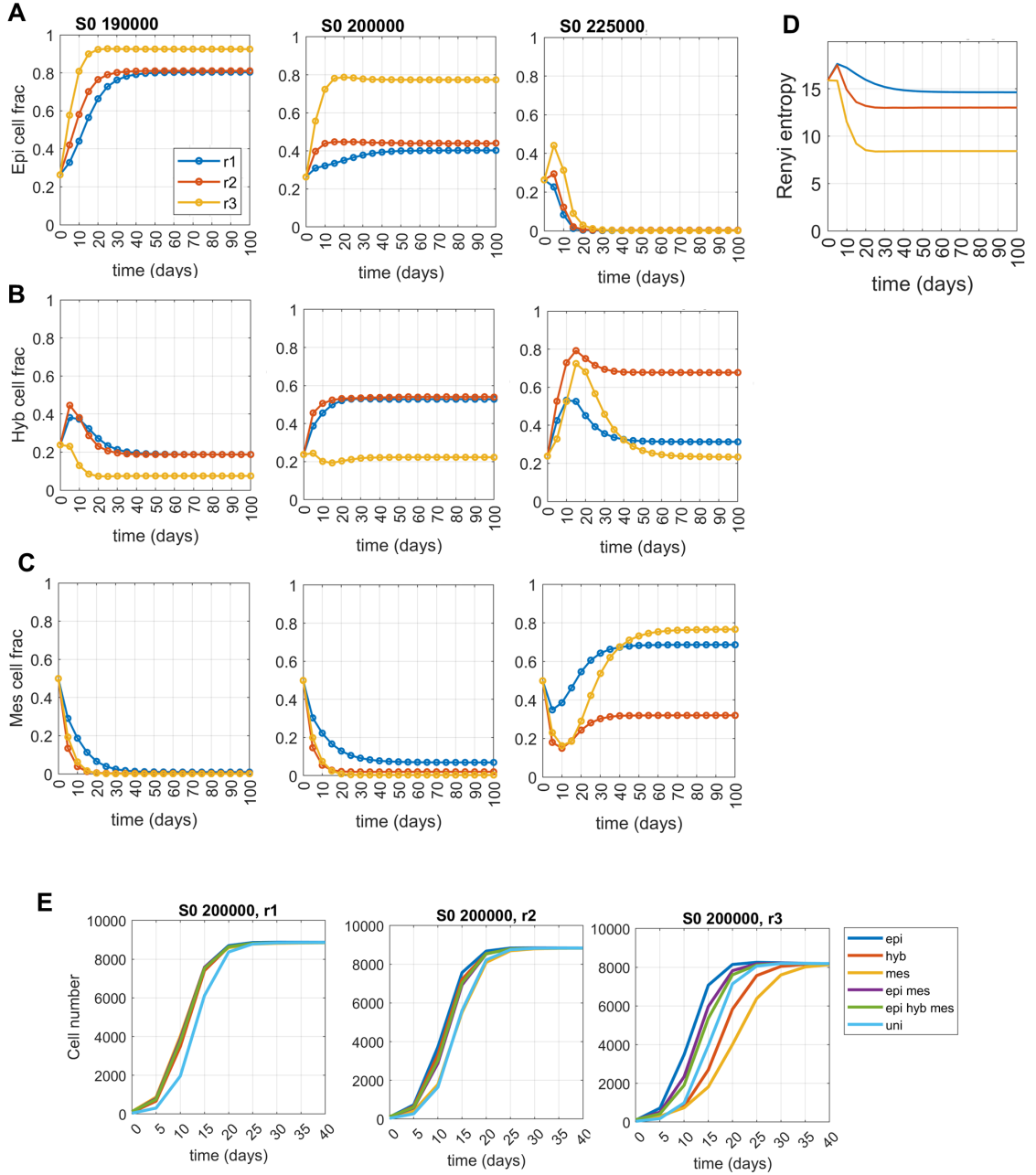


Figure A.6: Effect of growth rate differences among E, hybrid E/M and M phenotypes on population's heterogeneity and overall growth dynamics. A-C) Temporal changes in E, hybrid E/M, and M cell fraction in the population for different growth scenarios among phenotypes – ‘r1’: All three phenotypes divide at same rate; ‘r2’: E and E/M divide at equal rates, while M divide at half the rate of E cells; and ‘r3’: Both E/M and M divide at equal but half the rate of E cells. D) Changes in population's heterogeneity (measured by Renyi entropy) with time. E) Population growth dynamics for different combinations of growth scenarios and initial conditions. For panel A-D, the initial condition is uniformly distributed in E, hybrid E/M and M state. The input SNAIL mean  $S_0$  level used are mentioned for all the individual plots. Other parameters used to generate plots are  $\alpha = 120$  hrs,  $\eta_x = 5000$ , and per-capita growth rate  $r$  of epithelial phenotype is 0.0182/hr.

## A.2 Parameters for ODE (6.1)

We detail the parameters underlying ODE (6.1):

$$\begin{cases} \dot{\mu}_{200} &= g_{\mu_{200}} H_{Z,\mu_{200}}(Z) H_{S,\mu_{200}}(S) - g_{m_Z} H_{Z,m_Z}(Z) H_{S,m_Z}(S) Q(\mu_{200}) - k_{\mu_{200}} \mu_{200} \\ \dot{Z} &= g_Z g_{m_Z} H_{Z,m_Z}(Z) H_{S,m_Z}(S) P(\mu_{200}) - k_Z Z \end{cases}$$

The functions  $H_{Z,\mu_{200}}$ ,  $H_{S,\mu_{200}}$ ,  $H_{Z,m_Z}$  and  $H_{S,m_Z}$  are shifted Hill functions which writes under the form

$$H(X) = \frac{1 + \lambda \left(\frac{X}{X_0}\right)^n}{1 + \left(\frac{X}{X_0}\right)^n}.$$

The associated parameters are given in Table A.1.

				Molecules		Molecules.Hour <sup>-1</sup>		Hour <sup>-1</sup>	
$n$	6			$\mu_0$	10K				
$n_{Z,\mu_{200}}$	3	$\lambda_{Z,\mu_{200}}$	0.1	$Z_{\mu_{200}}^0$	220K	$g_{\mu_{200}}$	2.1K	$k_{\mu_{200}}$	0.05
$n_{S,\mu_{200}}$	2	$\lambda_{S,\mu_{200}}$	0.1	$S_{\mu_{200}}^0$	180K	$g_Z$	0.1K	$k_Z$	0.1
$n_{Z,m_Z}$	2	$\lambda_{Z,m_Z}$	7.5	$Z_{m_Z}^0$	25K	$g_{m_Z}$	11	$k_{m_Z}$	0.5
$n_{S,m_Z}$	2	$\lambda_{S,m_Z}$	10	$S_{m_Z}^0$	180K				

Table A.1: Parameters of the Hill functions.

The functions  $Y_\mu$ ,  $Y_m$  and  $L$  are defined by

$$\begin{aligned} Y_\mu(\mu) &:= \sum_{i=1}^n i \gamma_{\mu_i} \binom{n}{i} M_n^i(\mu) \\ Y_m(\mu) &:= \sum_{i=1}^n \gamma_{m_i} \binom{n}{i} M_n^i(\mu), \\ L(\mu) &:= \sum_{i=0}^n l_i \binom{n}{i} M_n^i(\mu), \end{aligned}$$

where  $M_n^i := \frac{(\mu/\mu_0)^i}{(1+\mu/\mu_0)^n}$ .

$$\begin{aligned} P(\mu) &:= \frac{L(\mu)}{Y_m(\mu) + k_{m_Z}} \\ Q(\mu) &:= \frac{Y_\mu(\mu)}{Y_m(\mu) + k_{m_Z}} \end{aligned}$$

The parameters of these three functions are given in Table A.2.

i	0	1	2	3	4	5	6
$l_i$	1.0	0.6	0.3	0.1	0.05	0.05	0.05
$\gamma_{mi}$		0.04	0.2	1.0	1.0	1.0	1.0
$\gamma_{\mu i}$		0.005	0.05	0.5	0.5	0.5	0.5

Table A.2: Parameters for the functions  $L$ ,  $Y_\mu$  and  $Y_m$

## Appendix B

# Appendix of Chapter 2

### B.1 Complement of Proposition 2.5

Here is the complement and the proof of Proposition 2.5.

**Proposition 34.** *We define:  $\phi : \lambda \mapsto \frac{1-\sqrt{\lambda}}{1+\sqrt{\lambda}}$  and  $\rho := ab\phi(\lambda_1)\phi(\lambda_2)$ . If  $\rho < 1$ , then  $E_1^> = \emptyset$ .  
If  $\rho \geq 1$ , we define:*

- If  $\lambda_1 > 0$  and  $\lambda_2 > 0$  :

$$\begin{aligned}
 x^\pm &:= \left[ \frac{1}{\sqrt{\lambda_2}} + \frac{1}{2\lambda_2} (1 + \sqrt{\lambda_2})^2 \sqrt{\rho-1} \left( \sqrt{\rho-1} \pm \sqrt{\rho - \phi(\lambda_2)^2} \right) \right]^{1/b} \\
 y^\pm &:= \left[ \frac{1}{\sqrt{\lambda_1}} + \frac{1}{2\lambda_1} (1 + \sqrt{\lambda_1})^2 \sqrt{\rho-1} \left( \sqrt{\rho-1} \pm \sqrt{\rho - \phi(\lambda_1)^2} \right) \right]^{1/a} \\
 \alpha(x) &:= (1 + \lambda_1) - (1 + \sqrt{\lambda_1})^2 (1 + \sqrt{\lambda_2})^2 \rho \frac{x^b}{(1+x^b)(1+\lambda_2 x^b)} \\
 \beta(y) &:= (1 + \lambda_2) - (1 + \sqrt{\lambda_1})^2 (1 + \sqrt{\lambda_2})^2 \rho \frac{y^a}{(1+y^a)(1+\lambda_1 y^a)} \\
 r^\pm(x) &:= \left[ \frac{-\alpha(x) \pm \sqrt{\alpha(x)^2 - 4\lambda_1}}{2\lambda_1} \right]^{1/a} \\
 s^\pm(y) &:= \left[ \frac{-\beta(y) \pm \sqrt{\beta(y)^2 - 4\lambda_2}}{2\lambda_2} \right]^{1/b}.
 \end{aligned}$$

Then,

$$\begin{aligned}
 E_1^> &= \{(x, y) \in (x^-, x^+) \times (y^-, y^+) : r^-(x) < y < r^+(x)\} \\
 &= \{(x, y) \in (x^-, x^+) \times (y^-, y^+) : s^-(y) < x < s^+(y)\}.
 \end{aligned}$$

- If  $\lambda_1 = 0$  and  $\lambda_2 > 0$  :

$$\begin{aligned}
 x^\pm &:= \left[ \frac{1}{\sqrt{\lambda_2}} + \frac{1}{2\lambda_2} (1 + \sqrt{\lambda_2})^2 \sqrt{\rho - 1} \left( \sqrt{\rho - 1} \pm \sqrt{\rho - \phi(\lambda_2)^2} \right) \right]^{1/b} \\
 y^- &:= \left( \frac{1}{\rho - 1} \right)^{1/a} \\
 r^-(x) &:= \left[ \frac{(1 + x^b)(1 + \lambda_2 x^b)}{-\lambda_2 x^{2b} + (ab(1 - \lambda_2) - (1 + \lambda_2))x^b - 1} \right]^{1/a} \\
 \beta(y) &:= (1 + \lambda_2) - (1 + \sqrt{\lambda_1})^2 (1 + \sqrt{\lambda_2})^2 \rho \frac{y^a}{(1 + y^a)(1 + \lambda_1 y^a)} \\
 s^\pm(y) &:= \left[ \frac{-\beta(y) \pm \sqrt{\beta(y)^2 - 4\lambda_2}}{2\lambda_2} \right]^{1/b}
 \end{aligned}$$

$$\begin{aligned}
 E_1^> &= \{(x, y) \in (x^-, x^+) \times (y^-, +\infty) : y > r^-(x)\} \\
 &= \{(x, y) \in (x^-, x^+) \times (y^-, +\infty) : s^-(y) < x < s^+(y)\}.
 \end{aligned}$$

- If  $\lambda_1 = \lambda_2 = 0$  :

$$\begin{aligned}
 x^- &:= \left( \frac{1}{\rho - 1} \right)^{1/b}, \quad y^- := \left( \frac{1}{\rho - 1} \right)^{1/a} \\
 r^-(x) &:= \left( \frac{1 + x^{n_2}}{(\rho - 1)x^b - 1} \right)^{1/a}, \quad s^-(y) := \left( \frac{1 + y^a}{(\rho - 1)y^a - 1} \right)^{1/b}
 \end{aligned}$$

$$\begin{aligned}
 E_1^> &= \{(x, y) \in (x^-, +\infty) \times (y^-, +\infty) : y > r^-(x)\} \\
 &= \{(x, y) \in (x^-, +\infty) \times (y^-, +\infty) : x > s^-(y)\}.
 \end{aligned}$$

In order to prove this proposition, we will need this very simple lemma.

**Lemma 22.** *Let  $P(X) := aX^2 + bX + c$  be a polynomial of degree two. We assume that  $a > 0$  and  $c > 0$ . If  $P$  has two real roots, then they have the same sign. Moreover,  $P$  has two positive roots if and only if  $b + 2\sqrt{ac} < 0$ .*

*Proof of Proposition 2.5.* Let  $(x, y) \in \mathbb{R}_+^2$ .

$$\frac{xg'(x)}{g(x)} \frac{yf'(y)}{f(y)} = ab(1 - \lambda_1)(1 - \lambda_2) \frac{x^b y^a}{(1 + x^b)(1 + \lambda_2 x^b)(1 + y^a)(1 + \lambda_1 y^a)}.$$

We denote

$$\begin{aligned}
 \omega &:= x^b \\
 \delta &:= y^a \\
 C &:= ab(1 - \lambda_1)(1 - \lambda_2) = (1 + \sqrt{\lambda_1})^2 (1 + \sqrt{\lambda_2})^2 \rho \\
 C_1(\delta) &:= (1 + \delta)(1 + \lambda_1 \delta) \\
 C_2(\omega) &:= (1 + \omega)(1 + \lambda_2 \omega) \\
 P(\omega, \delta) &:= \lambda_1 C_2(\omega) \delta^2 + (C_2(\omega)(1 + \lambda_1) - C\omega)\delta + C_2(\omega) \\
 &= \lambda_2 C_1(\delta) \omega^2 + (C_1(\delta)(1 + \lambda_2) - C\delta)\omega + C_1(\delta).
 \end{aligned}$$



Then,

$$(x, y) \in E_1^> \iff \frac{xg'(x)}{g(x)} \frac{yf'(y)}{f(y)} \iff P(\omega, \delta) < 0.$$

We now assume that  $\lambda_1 > 0$  and  $\lambda_2 > 0$ : the proof for  $\lambda_1 = 0$  or  $\lambda_2 = 0$  is similar and simpler, as  $P$  becomes a polynomial of degree 1 for the variable  $\omega$  or  $\gamma$ .

According to lemma 22, there exist  $\omega, \delta > 0$  such that  $P(\omega, \delta) < 0$  if and only if

$$\begin{aligned} Q_2(\omega) &:= C_2(\omega)(1 + \lambda_1) - C\omega + 2\sqrt{\lambda_1}C_2(\omega) < 0 \\ &\iff C_2(\omega)(1 + \sqrt{\lambda_1})^2 - C\omega < 0 \\ &\iff \lambda_2\omega^2 + \left(1 + \lambda_2 - \frac{C}{(1 + \sqrt{\lambda_1})^2}\right)\omega + 1 < 0 \\ &\iff \lambda_2\omega^2 + (1 + \lambda_2 - (1 + \sqrt{\lambda_2})^2\rho)\omega + 1 < 0. \end{aligned}$$

According to lemma 22, there exists  $\omega > 0$  such that  $Q_2(\omega) < 0$  if and only if

$$(1 + \sqrt{\lambda_2})^2 - (1 + \sqrt{\lambda_2})^2\rho < 0 \iff \rho > 1.$$

Therefore,  $Q_2(\omega) < 0$  if and only if  $\rho > 1$  and  $\omega \in (\omega^-, \omega^+)$ , where  $\omega^-, \omega^+$  (which depend on  $n_1, n_2, \lambda_1$  and  $\lambda_2$ ), denote the two positive roots of  $Q_2$ .

Lastly,  $P(\omega, \delta) < 0$  if and only if  $\rho > 1$ ,  $\omega \in (\omega^-, \omega^+)$ , and  $\delta \in (\tilde{r}^-(\omega), \tilde{r}^+(\omega))$ , where, for all  $\omega \in (\omega^-, \omega^+)$ ,  $\tilde{r}^-(\omega), \tilde{r}^+(\omega)$  denote the two positive roots of  $\delta \mapsto P(\omega, \delta)$ .

With the same reasoning, by denoting  $\delta_-, \delta_+$  the two positive roots of  $Q_1(\delta) := C_1(\delta)(1 + \lambda_2) - C\delta + \sqrt{\lambda_2}C_1(\delta) < 0$ , and, for all  $\delta \in (\delta^-, \delta^+)$ ,  $\tilde{s}^-(\delta)$  and  $\tilde{s}^+(\delta)$ , the two positive roots of  $\omega \mapsto P(\omega, \delta)$ ,  $P(\omega, \delta) < 0$  if and only if  $\rho > 1$ ,  $\delta \in (\delta^-, \delta^+)$ , and  $\omega \in (\tilde{s}^-(\delta), \tilde{s}^+(\delta))$ .

We denote

$$\begin{aligned} x^- &:= (\omega^-)^{1/b}, x^+ := (\omega^+)^{1/b} \\ y^- &:= (\delta^-)^{1/a}, y^+ := (\delta^+)^{1/a} \\ r^-(x) &:= (\tilde{r}^-(\omega))^{1/a}, r^+(\omega) := (\tilde{r}^+(\omega))^{1/a} \\ s^-(y) &:= (\tilde{s}^-(\delta))^{1/b}, s^+(\delta) := (\tilde{s}^+(\delta))^{1/b}. \end{aligned}$$

The values of  $x^-, x^+, y^-, y^+, r^+, r^-, s^-, s^+$  are directly computed from the formula of  $Q_1, Q_2$  and  $P$ .  $\square$

## B.2 Complement of Proposition 12

**Proposition 35.** *Let us denote  $\lambda_0^- := \left(\frac{a-1}{a+1}\right)^2$  and  $\lambda_0^+ = \left(\frac{a+1}{a-1}\right)^2$ .*

- If  $\lambda \in (\lambda_0^-, \lambda_0^+)$ , then  $E_s = \emptyset$ . In other words, system (2.4) is monostable, for any value of  $g$ .
- If  $\lambda = 0$ , we define:  $\alpha_0^- := z_0 \left(\frac{1}{a-1}\right)^{1/a} \frac{a}{a-1}$ . Then,  $E_s^> = (\alpha_0^-, +\infty)$ .

- If  $\lambda \in (0, \lambda_0^-]$ , we define:

$$\begin{aligned}\omega_\lambda^- &:= \frac{a-1-(a+1)(\lambda + \sqrt{(1-\lambda)(\lambda_0^- - \lambda)})}{2\lambda} \\ \omega_\lambda^+ &:= \frac{a-1-(a+1)(\lambda - \sqrt{(1-\lambda)(\lambda_0^- - \lambda)})}{2\lambda} \\ \alpha_\lambda^- &:= z_0(\omega_\lambda^-)^{1/a} \frac{1+\omega_\lambda^-}{1+\lambda\omega_\lambda^-}, \quad \alpha_\lambda^+ := z_0(\omega_\lambda^+)^{1/a} \frac{1+\omega_\lambda^+}{1+\lambda\omega_\lambda^+}.\end{aligned}$$

Then,  $G_s(E_s^>) = (\alpha_\lambda^-, \alpha_\lambda^+)$ .

- If  $\lambda \geq \lambda_0^+$ , we define:

$$\begin{aligned}\tau_\lambda^- &:= \frac{-(a+1) + (a-1)(\lambda - \sqrt{(\lambda-1)(\lambda-\lambda_0^+)})}{2\lambda} \\ \tau_\lambda^+ &:= \frac{-(a+1) + (a-1)(\lambda + \sqrt{(\lambda-1)(\lambda-\lambda_0^+)})}{2\lambda} \\ \beta_\lambda^- &:= z_0(\tau_\lambda^+)^{1/a} \frac{1+\tau_\lambda^+}{1+\lambda\tau_\lambda^+}, \quad \beta_\lambda^+ := z_0(\tau_\lambda^-)^{1/a} \frac{1+\tau_\lambda^-}{1+\lambda\tau_\lambda^-}.\end{aligned}$$

Then,  $G_s(E_s^>) = (\beta_\lambda^-, \beta_\lambda^+)$ .

*Proof.* With the same reasoning as for the asymmetric system, we can assume that  $z_0 = 1$ . Let  $x \in \mathbb{R}_+$ . We denote:  $\omega := x^a$  and

$$P(\omega) := (1 + \lambda\omega)(1 + \omega) - n|\lambda - 1|\omega = \lambda\omega^2 + (\lambda + 1 - a|\lambda - 1|)\omega + 1.$$

Since

$$\left| \frac{xf'(x)}{f(x)} \right| = \frac{a|\lambda - 1|x^a}{(1 + \lambda x^a)(1 + x^a)},$$

we clearly have:  $x \in E_s^> \iff \left| \frac{xf'(x)}{f(x)} \right| > 1 \iff P(\omega) < 0$ .

If  $\lambda = 0$ , then:  $P(\omega) = 1 - (a-1)\omega$ . Thus,  $P(\omega) < 0 \iff \omega > \frac{1}{a-1} \iff x > (\frac{1}{a-1})^{1/a}$ .

Since  $G_s((\frac{1}{a-1})^{1/a}) = (\frac{1}{a-1})^{1/a} \frac{a}{a-1} := a_0^-$ , the result follows.

We now assume that  $\lambda > 0$ . Denoting  $\Delta$  the discriminant of P, we get

$$\Delta = (\lambda + 1)^2 + a^2(\lambda - 1)^2 - 2a(\lambda + 1)|\lambda - 1| - 4\lambda = |\lambda - 1|(|\lambda - 1|(1 + a^2) - 2a(\lambda + 1)).$$

We distinguish two cases:

- If  $\lambda < 1$ , recalling that  $\lambda_0^- = (\frac{a-1}{a+1})^2$ , we have

$$\Delta = (1 - \lambda)((1 - a)^2 - \lambda(1 + a)^2) = -(1 - \lambda)(1 + a)^2(\lambda - \lambda_0^-).$$

Hence, if  $\lambda > \lambda_0^-$ , then  $P(\omega) > 0$ , and thus  $x \notin E_s^>$ .

If  $\lambda \in (0, \lambda_0^-]$ , then  $P$  has two roots, and we have

$$P(\omega) < 0 \iff \omega \in (\omega_\lambda^-, \omega_\lambda^+) \iff x \in ((\omega_\lambda^-)^{1/a}, (\omega_\lambda^+)^{1/a}).$$

Hence,  $E_s^> = ((\omega_\lambda^-)^{1/a}, (\omega_\lambda^+)^{1/a})$ . Since  $G_s$  is increasing, we get

$$G_s(E_s) = (G_s((\omega_\lambda^-)^{1/a}), G_s((\omega_\lambda^+)^{1/a})) = (\alpha_\lambda^-, \alpha_\lambda^+).$$

- If  $\lambda > 0$ , the reasoning is exactly the same. In this case

$$\Delta = (\lambda - 1)(\lambda(a - 1)^2 - (a + 1)^2) = (\lambda - 1)(a - 1)^2(\lambda - \lambda_0^+),$$

where  $\lambda_0^+ = \left(\frac{a+1}{a-1}\right)^2$ . We get  $E_s^> = ((\tau_\lambda^-)^{1/a}, (\tau_\lambda^+)^{1/a})$ , and, lastly, since  $G_s$  is decreasing in this case,

$$G_s(E_s^>) = (G(\tau_\lambda^+), G(\tau_\lambda^-)) = (\beta_\lambda^-, \beta_\lambda^+),$$

which ends the proof. □

### B.3 Other sigmoid functions which are $\gamma - 1/2$ -convex

The following sigmoid functions are strictly  $\gamma^{1/2}$ -convex:

- (i) The logistic function, and more generally, the general logistic functions, *i.e.*, all the functions of the shape

$$x \mapsto \left( \frac{1}{1 + e^{-x}} \right)^\alpha, \quad \alpha > 0.$$

- (ii) The hyperbolic tangent:

$$x \mapsto \frac{e^x - e^{-x}}{e^x + e^{-x}}.$$

- (iii) The arctan function.

- (iv) The Gudermannian function:

$$x \mapsto \int_0^x \frac{1}{\cosh(t)} dt.$$

- (v) The error function

$$x \mapsto \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt.$$

*Proof.* (i) We note that the result is immediate if we assume  $\alpha \geq 1$ , because the function happens to be the composite of  $\gamma^{1/2}$ -convex functions in this case. Surprisingly, the results holds true if  $\alpha < 1$ . This result is equivalent to showing that  $f : x \mapsto \frac{-1}{\alpha}(1 + e^x)^{-\alpha}$  is strictly  $\gamma^{1/2}$ -convex, for any  $\alpha > 0$  (the result is in fact true for any  $\alpha \neq 1$ ). We indeed compute

$$f'(x) = e^x(1+e^x)^{-(\alpha+1)}, \quad f''(x) = \frac{1 - \alpha e^x}{1 + e^x} f'(x), \quad f^{(3)}(x) = \frac{\alpha e^{2x} - (3\alpha + 1)e^x + 1}{(1 + e^x)^2} f'(x)^2,$$

and find

$$f^{(3)}(x)f'(x) - \frac{3}{2}f''(x)^2 = -\frac{f'(x)^2}{2(1 + e^x)}(\alpha^2 e^{2x} + 2e^x + 1) < 0.$$

(ii) We rewrite

$$\frac{e^x - e^{-x}}{e^x + e^{-x}} = \frac{(e^x)^2 - 1}{(e^x)^2 + 1}.$$

The hyperbolic tangent is thus the composite function of the exponential function, the square function and a homographic function, which proves that it is strictly  $\gamma^{1/2}$ -convex as the composition of  $\gamma^{1/2}$ -convex functions, one of which is strictly  $\gamma^{1/2}$ -convex.

(iii) It is equivalent to proving that the tangent function is strictly  $\gamma^{1/2}$ -concave on  $(-\pi/2, \pi/2)$ . Since  $\tan' = \frac{1}{\cos^2}$ , the result follows according to the positivity and the concavity of the cosine function on this interval.

(iv) Since this function may be rewritten as

$$x \mapsto \int_0^x \frac{1}{\cosh(t)} dt = 2 \arctan(\tanh(\frac{x}{2})),$$

this result is true according to the  $\gamma^{1/2}$ -convexity of  $\arctan$  and  $\tanh$ .

(v) The result is obvious.

□

# Appendix C

## Appendix of Chapter 5

### C.1 Proof of the results over the characteristics

In order to prove results which involve the use of absolute values, we introduce a smooth re-normalizing sequence of functions. Consider a sequence of smooth positive functions  $\beta_\varepsilon$  satisfying  $\beta_\varepsilon(0) = 0$ ,  $\beta_\varepsilon(s) > 0$  for all  $s \neq 0$ ,  $\beta_\varepsilon(s) \leq |s|$ ,  $\beta_\varepsilon(s) \rightarrow |s|$  almost everywhere,  $|\dot{\beta}_\varepsilon(s)| \leq 1$  and  $s\dot{\beta}_\varepsilon(s) \rightarrow |s|$  almost everywhere. For example we may choose

$$\beta_\varepsilon(s) = \begin{cases} -s - \varepsilon(1 - \frac{2}{\pi}) & \text{if } s \leq -\varepsilon, \\ \frac{2\varepsilon}{\pi} (1 - \cos(\frac{\pi}{2\varepsilon}s)) & \text{if } -\varepsilon < s < \varepsilon, \\ s - \varepsilon(1 - \frac{2}{\pi}) & \text{if } s \geq \varepsilon. \end{cases} \quad (\text{C.1})$$

*Proof of Lemma 11.* We introduce the notation

$$\Delta X_j(t) := X_{u_1}^j(t, y_1) - X_{u_2}^j(t, y_2).$$

For all  $t \in [0, T]$ , the function

$$U_\varepsilon(t) := \sum_{j=1}^d \beta_\varepsilon(\Delta X_j(t))$$

satisfies then the relation

$$\begin{aligned} \dot{U}_\varepsilon(t) &= \sum_{j=1}^d \dot{\beta}_\varepsilon(\Delta X_j(t)) (a_j(t, X_{u_1}(t, y_1), (I_a u_1)(t, X_{u_1}(t, y_1))) - a_j(t, X_{u_2}(t, y_2), (I_a u_2)(t, X_{u_2}(t, y_2)))) \\ &\leq \sum_{j=1}^d \left( \|a_j\|_{W_x^{1,\infty}} \sum_{i=1}^d |\Delta X_i(t)| + \|a_j\|_{W_T^{1,\infty}} |(I_a u_1)(t, X_{u_1}(t, y_1)) - (I_a u_2)(t, X_{u_2}(t, y_2))| \right) \\ &\leq d \|a\|_{W_{x,I}^{1,\infty}} \left( \sum_{i=1}^d |\Delta X_i(t)| + |(I_a u_1)(t, X_{u_1}(t, y_1)) - (I_a u_2)(t, X_{u_2}(t, y_2))| \right) \\ &\leq d \|a\|_{W_{x,I}^{1,\infty}} \left( (1 + \|\psi_a\|_{W_{x,y}^{1,\infty} L_y^\infty} \|u_1\|) \sum_{i=1}^d |\Delta X_i(t)| + \|\psi_a\|_{L^\infty} \|u_1 - u_2\|_{L^1(\mathbb{R}^d)} \right). \end{aligned}$$

Integrating between 0 and  $t$  we get

$$\begin{aligned} U_\varepsilon(t, y) - U_\varepsilon(0, y) &\leq d \|a\|_{L_t^\infty W_{x,I}^{1,\infty}} \left( (1 + \|\psi_a\|_{L_{t,y}^\infty W_x^{1,\infty} L_y^\infty} \|u_1\|) \int_0^t \sum_{i=1}^d |\Delta X_i(s)| ds + \|\psi_a\|_{L^\infty} \int_0^t \|u_1 - u_2\|_{L^1(\mathbb{R}^d)} ds \right). \end{aligned}$$

Taking the limit when  $\varepsilon$  goes to 0 and applying Grönwall's lemma we get the desired result.  $\square$

*Proof of Lemma 12.* We explicitly give the proof for  $k = 1$ . The proof for higher values of  $k$  follows the same ideas.

Thanks to the hypothesis over  $a$ , the function  $X_u$  is one time differentiable with respect to  $y$ , and directly from (5.20) we get the system of equations

$$\begin{cases} \partial_{y_i} \dot{X}_u(t, y) = J_a(t, X_u(t, y)) \partial_{y_i} X_u(t, y), & t \in [0, T], \\ \partial_{y_i} X_u(0, y) = e_i, \end{cases} \quad (\text{C.2})$$

for all values of  $i \in \{1, \dots, d\}$ , where

$$[J_a(t, x)]_{ij} := \partial_{x_i} a_j(t, x, (I_a u)(t, x)) + \partial_I a_j(t, x, (I_a u)(t, x)) \int_{\mathbb{R}^d} \partial_{x_i} \psi_a(t, x, y) u(t, y) dy, \quad (\text{C.3})$$

is the Jacobian matrix of the function  $a(t, x, (I_a u)(t, x))$  and the  $e_i$  represent the canonical basis of  $\mathbb{R}^d$ .

The function

$$V_\varepsilon(t, y) := \sum_{i,j=1}^n \beta_\varepsilon(\partial_{y_i} X_u^j(t, y))$$

satisfies then

$$\dot{V}_\varepsilon(t, y) = \sum_{i,j=1}^n \dot{\beta}_\varepsilon(\partial_{y_i} X_u^j(t, y)) \sum_{k=1}^d [J_a(t, X_u(t, y))]_{kj} \partial_{y_i} X_u^k(t, y)$$

and consequently

$$\dot{V}_\varepsilon(t, y) \leq d \|a\|_{W_{x,I}^{1,\infty}} (1 + \|\psi_a\|_{W_x^{1,\infty} L_y^\infty} \|u\|) \sum_{i,j=1}^n |\partial_{y_i} X_u^j(t, y)|.$$

Integrating between 0 and  $t$ , we obtain the relation

$$V_\varepsilon(t, y) - V_\varepsilon(0, y) \leq d \tilde{\alpha}_1 \int_0^t \sum_{i,j=1}^n |\partial_{y_i} X_u^j(s, y)| ds$$

with  $\tilde{\alpha}_1 := \sup_{t \in [0, T]} \|a\|_{W_{x,I}^{1,\infty}} (1 + \|\psi_a\|_{W_x^{1,\infty} L_y^\infty} \|u\|)$ , which after taking the limit when  $\varepsilon$  goes to 0 leads to

$$\sum_{i,j=1}^n |\partial_{y_i} X_u^j(t, y)| \leq d + d \tilde{\alpha}_1 \int_0^t \sum_{i,j=1}^n |\partial_{y_i} X_u^j(s, y)| ds.$$

We obtain (5.21) thanks to Grönwall's lemma.

In order to prove (5.22) we adopt the notation

$$\Delta \partial_{y_k} X^j(t) := \partial_{y_k} X_{u_1}^j(t, y_1) - \partial_{y_k} X_{u_2}^j(t, y_2)$$

and define

$$[J_{X_u}(t, y)]_{jk} := \partial_{y_k} X_u^j(t, y),$$

which satisfies the relation

$$\dot{J}_{X_u}(t, y) = J_a(t, X_u(t, y)) J_{X_u}(t, y), \quad J_{X_u}(0, y) = I_d,$$

with  $J_a(t, x)$  as defined on (C.3). Consequently, we have that

$$D_\varepsilon(t) := \sum_{j=1}^d \sum_{k=1}^d \beta_\varepsilon (\Delta \partial_{y_k} X^j(t))$$

satisfies for all  $t \in [0, T]$

$$\begin{aligned} \dot{D}_\varepsilon(t) &= \sum_{j=1}^d \sum_{k=1}^d \dot{\beta}_\varepsilon (\Delta \partial_{y_k} X^j(t)) \sum_{i=1}^d ([J_a(t, X_{u_1}(t, y_1))]_{ij} \partial_{y_k} X_{u_1}^i(t, y_1) - [J_a(t, X_{u_2}(t, y_2))]_{ij} \partial_{y_k} X_{u_2}^i(t, y_2)) \\ &\leq \sum_{j=1}^d \sum_{k=1}^d \sum_{i=1}^d |[J_a(t, X_{u_1}(t, y_1))]_{ij} \partial_{y_k} X_{u_1}^i(t, y_1) - [J_a(t, X_{u_2}(t, y_2))]_{ij} \partial_{y_k} X_{u_2}^i(t, y_2)| \\ &\leq \sum_{j=1}^d \sum_{k=1}^d \sum_{i=1}^d |[J_a(t, X_{u_1}(t, y_1))]_{ij}| |\Delta \partial_{y_k} X^i(t)| \\ &\quad + \sum_{j=1}^d \sum_{k=1}^d \sum_{i=1}^d |[J_a(t, X_{u_1}(t, y_1))]_{ij} - [J_a(t, X_{u_2}(t, y_2))]_{ij}| |\partial_{y_k} X_{u_2}^i(t, y_2)|. \end{aligned}$$

From the hypothesis over  $a$  and  $\psi_a$  we see that

$$|[J_a(t, X_{u_1}(t, y_1))]_{ij}| \leq \|a\|_{W_{x,I}^{1,\infty}} (1 + \|\psi_a\|_{W_x^{1,\infty} L_y^\infty} \|u_1\|).$$

Furthermore, from the definition of  $J_a(t, x)$  we conclude that there exists a constant  $C$ , depending only on  $\|a\|_{W_{x,I}^{2,\infty}}$ ,  $\|\psi_a\|_{W_x^{2,\infty} L_y^\infty}$  and  $\|u_i\|$  such that

$$\begin{aligned} |[J_a(t, X_{u_1}(t, y_1))]_{ij} - [J_a(t, X_{u_2}(t, y_2))]_{ij}| &\leq C \left( \sum_{j=1}^d |X_{u_1}^j(t, y_1) - X_{u_2}^j(t, y_2)| + \|u_1 - u_2\|_{L^1(\mathbb{R}^d)} \right) \\ &\leq C(|y_1 - y_2| + \|u_1 - u_2\|_1 + \|u_1 - u_2\|_{L^1(\mathbb{R}^d)}), \end{aligned}$$

where we have used the results from Lemma 11 on the second line.

Putting all estimates together, we conclude that there exist constants  $C_1$  and  $C_2$  only depending on  $\|a\|_{W_{x,I}^{2,\infty}}$ ,  $\|\psi_a\|_{W_x^{2,\infty} L_y^\infty}$  and  $\|u_i\|$ , such that

$$\dot{D}_\varepsilon(t) \leq C_1 \sum_{j=1}^d \sum_{k=1}^d |\partial_{y_k} X_{u_1}^j(t, y_1) - \partial_{y_k} X_{u_2}^j(t, y_2)| + C_2 (|y_1 - y_2| + \|u_1 - u_2\|_1 + \|u_1 - u_2\|_{L^1(\mathbb{R}^d)}).$$

Integrating in time, using Grönwall's lemma and taking the limit when  $\varepsilon$  goes to zero, we obtain (5.22).  $\square$

*Proof of Lemma 13.* We explicitly give the proof for  $k = 1$ . The proof for higher values of  $k$  follows the same ideas.

Differentiating once each component of the equality  $X_u(t, X_u^{-1}(t, x)) = x$  with respect to each of the variables  $x_k$ , we obtain the family of relations

$$\sum_{i=1}^d \partial_{y_i} X_u^j(t, X_u^{-1}(t, x)) \partial_{x_k} (X_u^{-1})^i(t, x) = \delta_{jk}, \quad j, k = 1, \dots, d,$$

where  $\delta_{jk}$  represents the Kronecker's delta. Written in matrix form, this equality reads

$$J_{X_u}(t, X_u^{-1}(t, x)) J_{X_u^{-1}}(t, x) = I_d.$$

It is known that the matrix  $J_{X_u}(t, y)$  is invertible for all values of  $x$ , furthermore, its determinant is given by the expression

$$\det(J_{X_u}(t, y)) = e^{\int_0^t \nabla_x \cdot a(s, y, (I_a u)(s, y)) + \partial_I a(s, y, (I_a u)(s, y)) \cdot \int_{\mathbb{R}^d} \nabla_x \psi_a(s, y, z) u(s, z) dz ds} \geq c_T > 0$$

for all values of  $t \in [0, T]$  and  $y \in \mathbb{R}^d$ .

We conclude by writing

$$J_{X_u^{-1}}(t, x) = J_{X_u}^{-1}(t, X_u^{-1}(t, x)),$$

and noticing that all of the components of  $J_{X_u}^{-1}(t, X_u^{-1}(t, x))$  are a combination of sums and multiplications of the components of  $J_{X_u}(t, X_u^{-1}(t, x))$ , divided by  $\det(J_{X_u}(t, y))$ . The bound (5.21) from Lemma 12, together with the lower bound for the determinant of  $J_{X_u}(t, y)$  gives the bound (5.23) over the components of  $J_{X_u^{-1}}(t, x)$ .  $\square$

*Proof of Lemma 14.* We explicitly give the proof for  $k = 1$ . The proof for higher values of  $k$  follows the same ideas.

Differentiating with respect to  $t$  the relation  $X_{u_i}(t, X_{u_i}^{-1}(t, x)) = x$ , for  $i = 1, 2$ , we see that

$$a(t, x, (I_a u_i)(t, x)) + J_{X_{u_i}}(t, X_{u_i}^{-1}(t, x)) \dot{X}_{u_i}^{-1}(t, x) = 0,$$

which gives

$$\dot{X}_{u_i}^{-1}(t, x) = -J_{X_{u_i}}^{-1}(t, X_{u_i}^{-1}(t, x)) a(t, x, (I_a u_i)(t, x)). \quad (\text{C.4})$$

From now on we adopt the notations

$$\begin{aligned} A^i(t, x) &:= a(t, x, (I_a u_i)(t, x)), \\ K^i(t, x) &:= -J_{X_{u_i}}^{-1}(t, X_{u_i}^{-1}(t, x)), \\ \Delta X_j^{-1}(t, x) &:= (X_{u_1}^{-1})^j(t, x) - (X_{u_2}^{-1})^j(t, x). \end{aligned}$$

The function

$$W_\varepsilon(t, x) := \sum_{j=1}^d \beta_\varepsilon(\Delta X_j^{-1}(t, x))$$

satisfies the relation

$$\begin{aligned} \dot{W}_\varepsilon(t, x) &= \sum_{j=1}^d \sum_{k=1}^d \dot{\beta}_\varepsilon(\Delta X_j^{-1}(t, x)) (K_{jk}^1(t, x) A_k^1(t, x) - K_{jk}^2(t, x) A_k^2(t, x)) \\ &\leq \sum_{j=1}^d \sum_{k=1}^d |K_{jk}^1(t, x) A_k^1(t, x) - K_{jk}^2(t, x) A_k^2(t, x)|. \end{aligned}$$

From Lemma 13 we know that all components of  $K^i$  are uniformly bounded by a constant only depending on  $T$  and  $\|u_i\|$ . We deduce from the hypothesis over  $a$  that the components of  $A^i$  are uniformly bounded by  $\tilde{a} := \|a\|_{L^\infty}$ . Therefore

$$\dot{W}_\varepsilon(t, x) \leq d\tilde{C}(T, \|u_1\|) \sum_{k=1}^d |A_k^1(t, x) - A_k^2(t, x)| + \tilde{a} \sum_{j=1}^d \sum_{k=1}^d |K_{jk}^1(t, x) - K_{jk}^2(t, x)|.$$

The function  $a$  being  $L$ -Lipschitz with respect to the  $I$  variable, we have that, for all values of  $k$

$$|A_k^1(t, x) - A_k^2(t, x)| \leq L \|\psi_a\|_{L^\infty} \|u_1 - u_2\|_{L^1(\mathbb{R}^d)}.$$



On the other hand, from the definition of  $K^i$  and Lemma 12 we conclude that there exists  $C$ , depending on  $T$ ,  $\|a\|_{W_{x,I}^{2,\infty}}$ ,  $\|\psi_a\|_{W_x^{2,\infty}L_y^\infty}$  and  $\|u_i\|$  such that

$$|K_{jk}^1(t, x) - K_{jk}^2(t, x)| \leq \begin{cases} 0, & \text{if } \partial_I a = 0, \\ C \left( \sum_{j=1}^d |(X_{u_1}^{-1})^j(t, x) - (X_{u_2}^{-1})^j(t, x)| + \|u_1 - u_2\|_1 \right), & \text{if } \partial_I a \neq 0. \end{cases}$$

Putting everything together, integrating between 0 and  $t$ , taking the limit when  $\varepsilon$  goes to 0 and applying Grönwall's lemma we get (5.24).  $\square$

## C.2 Existence of solution for a system of ODEs with infinitely many unknowns and equations

*Proof of Lemma 17.* For all  $u \in X_h^T$ , there exists a sequence of elements  $u^\delta \in X_h^T$  such that:

- 1)  $\mathcal{K}_h^\delta := \{k \in \mathcal{J}_h : u_k^\delta \neq 0\}$  has a finite number of elements.
- 2)  $\lim_{\delta \rightarrow 0} \|u - u^\delta\|_{1,h} = 0$ .

We denote  $K^\delta := |\mathcal{K}_h^\delta|$  and notice that the system

$$\dot{x}_k^\delta(t) = A_{u^\delta, w}(t, x_k^\delta), \quad x_k^\delta(0) = x_k^0, \quad (\text{C.5})$$

is composed of a coupled system of  $K^\delta$  equations and unknowns (corresponding to those  $k \in \mathcal{K}_h^\delta$ ), and an uncoupled infinite number of equations, corresponding to those  $k \notin \mathcal{K}_h^\delta$ . Therefore, thanks to the classic Cauchy-Lipschitz theory, the system (C.5) has a unique solution  $x_k^\delta \in \mathcal{C}^1([0, T])$ ,  $k \in \mathcal{J}_h$ .

We claim that for all values of  $k$ , the sequence  $x_k^{\delta_1} - x_k^{\delta_2}$  is a Cauchy sequence in  $\mathcal{C}^1([0, T])$ , therefore it has a limit that we will call  $x_k(t)$ , which is solution to (5.49).

We first remark that  $x^{\delta_1} - x^{\delta_2} \in Y_h^T$  due to the fact that  $|x_k^\delta(t) - x_k^0| \leq \|a\|_{L^\infty T}$  for all values of  $k$  and  $\delta$ .

Consider now  $\beta_\varepsilon$  as defined in (C.1), then

$$\begin{aligned} \dot{\beta}_\varepsilon(x_k^{\delta_1} - x_k^{\delta_2}) &\leq |A_{u^{\delta_1}, w}(t, x_k^{\delta_1}) - A_{u^{\delta_2}, w}(t, x_k^{\delta_2})| \\ &\quad \|a\|_{W_x^{1,\infty}} |x_k^{\delta_1} - x_k^{\delta_2}| + \|a\|_{W_I^{1,\infty}} |I_a(t, x_k^{\delta_1}, u^{\delta_1}, w) - I_a(t, x_k^{\delta_1}, u^{\delta_2}, w)|. \end{aligned}$$

Noticing that

$$\begin{aligned} |I_a(t, x_k^{\delta_1}, u^{\delta_1}, w) - I_a(t, x_k^{\delta_1}, u^{\delta_2}, w)| &= \left| \sum_{j \in \mathcal{J}_h} \left( u_j^{\delta_1}(t) \psi_a(t, x_k^{\delta_1}, x_j^{\delta_1}(t)) - u_j^{\delta_2}(t) \psi_a(t, x_k^{\delta_1}, x_j^{\delta_2}(t)) \right) w_j(t) \right| \\ &\leq (\|\psi_a\|_{L^\infty} \|u^{\delta_1} - u^{\delta_2}\|_{1,h} \\ &\quad + \|u^{\delta_2}\|_{1,h} \|\psi_a\|_{W_{x,y}^{1,\infty}} \|x^{\delta_1} - x^{\delta_2}\|_{\infty, h}) \|w\|_{\infty, h}, \end{aligned}$$

we deduce the existence of two constants<sup>1</sup>,  $C_1$  and  $C_2$ , only depending on  $a$ ,  $\psi_a$ ,  $u$  and  $w$ , such that

$$\dot{\beta}_\varepsilon(x_k^{\delta_1} - x_k^{\delta_2}) \leq C_1 \|x^{\delta_1} - x^{\delta_2}\|_{\infty, h} + C_2 \|u^{\delta_1} - u^{\delta_2}\|_{1,h}.$$

<sup>1</sup>Notice that in order to obtain the estimate over  $I_a$ , we used the hypothesis of differentiability over both variables on  $\psi_a$

Integrating between 0 and  $t$ , taking the maximum over  $k$  and  $t$  and using Grönwall's lemma, we conclude that there exists a constant  $C_T$ , only depending on  $T$  and the coefficients of the problem, such that

$$\|x^{\delta_1} - x^{\delta_2}\|_{\infty, h} \leq C_T \|u^{\delta_1} - u^{\delta_2}\|_{1, h}.$$

Proceeding in a similar way with the absolute value of  $\dot{x}^{\delta_1} - \dot{x}^{\delta_2}$  we obtain that

$$\|\dot{x}^{\delta_1} - \dot{x}^{\delta_2}\|_{\infty, h} \leq C_T \|u^{\delta_1} - u^{\delta_2}\|_{1, h}.$$

Recalling that  $u^\delta$  is a Cauchy sequence on  $X_h^T$ , then so it is  $x_k^\delta$  on  $\mathcal{C}^1([0, T])$ , for each  $k$ . Let  $x := \{x_k\}_{k \in \mathcal{J}_h}$  be the limit of  $x^\delta$  when  $\delta$  goes to 0. With a simple continuity argument we conclude that  $x$  is a solution of (5.49) over  $[0, T]$ . The uniqueness can be obtained by assuming the existence of two solutions, deriving the equation satisfied by the difference and using Grönwall's lemma to conclude that they have to be equal.  $\square$

### C.3 A result from approximation theory

As mentioned before, Lemma 18 is a direct corollary of Lemma 8 in [115], that we recall here

**Lemma 23.** *Let  $k > d$  an integer. Assume that*

$$a \in (L^\infty(0, T; W^{k+1, \infty}(\mathbb{R}^d)))^d.$$

*Then, there exists a constant  $C > 0$  such that for all functions  $\varphi \in W^{k, p}(\mathbb{R}^d)$ ,  $1 \leq p \leq +\infty$ , and  $t \in [0, T]$ ,*

$$\|\varphi - \sum_{i \in \mathcal{J}_h} w_i(t) \varphi(x_i(t)) \delta(\cdot - x_i(t))\|_{W^{-k, p}(\mathbb{R}^d)} \leq Ch^k \|\varphi\|_{W^{k, p}(\mathbb{R}^d)}.$$

Given that for fixed functions  $\nu \in X_h^T$  and  $w \in Y_h^T$  we have the inclusion

$$A_{\nu, w} : (t, x) \mapsto a(t, x, I_a(t, x, \nu, w)) \in (L^\infty(0, T; W^{k+1, \infty}(\mathbb{R}^d)))^d,$$

then Lemma 23 holds true as well for the values of  $x_i$  obtained in Section 5.3.

*Proof of Lemma 18.* We recall that  $W^{-k, 1}(\mathbb{R}^d)$  is the dual space of  $W^{k, \infty}(\mathbb{R}^d)$ . Thus, for any  $\psi \in W^{-k, 1}(\mathbb{R}^d)$  we have

$$\|\psi\|_{-k, 1} = \sup_{f \in W^{k, \infty}(\mathbb{R}^d)} \frac{|\langle \psi, f \rangle|}{\|f\|_{k, \infty}}.$$

Since the function  $f \equiv 1$  belongs to  $W^{k, \infty}(\mathbb{R}^d)$ , and has norm equal to 1 in this space, we get for all  $\varphi \in W^{k, p}(\mathbb{R}^d)$

$$\begin{aligned} \left| \int_{\mathbb{R}^d} \varphi(x) dx - \sum_{i \in \mathcal{J}_h} w_i(t) \varphi(x_i(t)) \right| &= \left| \langle \varphi - \sum_{i \in \mathcal{J}_h} w_i(t) \varphi(x_i(t)) \delta(\cdot - x_i(t)), 1 \rangle \right| \\ &\leq \|\varphi - \sum_{i \in \mathcal{J}_h} w_i(t) \varphi(x_i(t)) \delta(\cdot - x_i(t))\|_{-k, 1}. \end{aligned}$$

We conclude by applying Lemma 23 with  $p = 1$ .  $\square$

## C.4 Proofs of convergence results from ODE theory

This appendix is dedicated to the proofs of lemma 21, used in subsection 5.4.2. In order to prove this lemma, we use the following result:

**Lemma 24.** *Let  $\alpha > 0$  and  $B \in L^1(\mathbb{R}_+)$ . Then, all the solutions of the ODE*

$$\dot{u}(t) = -\alpha u(t) + B(t)$$

are in  $L^1(\mathbb{R}^+)$ .

*Proof of lemma 24.* The solution of this ODE is explicitly given by

$$u(t) = u(0)e^{-\alpha t} + \int_0^t e^{-\alpha(t-s)} B(s) ds.$$

Hence,

$$\int_0^{+\infty} |u(t)| dt \leq |u(0)| \int_0^{+\infty} e^{-\alpha t} dt + \iint_{\mathbb{R}_+^2} e^{-\alpha(t-s)} |B(s)| \mathbb{1}_{\{s \leq t\}} ds dt.$$

With the change of variables  $y = s, z = t - s$ , we get

$$\iint_{\mathbb{R}_+^2} e^{-\alpha(t-s)} |B(s)| \mathbb{1}_{\{s \leq t\}} ds dt \leq \int_0^{+\infty} |B(y)| dy \int_0^{+\infty} e^{-\alpha z} dz,$$

which concludes the proof. □

*Proof of Lemma 21.* First, let us note that if  $\dot{u}$  is a BV function, *i.e.* if  $\int_0^{+\infty} |\dot{u}(t)| dt < +\infty$ , then  $u$  is a Cauchy function, and thus converges. Let us denote  $v := \dot{u}$ . Since  $u$  is assumed to be bounded, and  $|v| = v + 2v^-$ , where  $v^-$  denotes the negative part of  $v$ , it is enough to prove that  $v^- \in L^1(\mathbb{R}^d)$ . By hypothesis,

$$\dot{v}(t) = -p(t)v(t) + P(t) + B(t),$$

which implies that

$$\dot{v}^-(t) \leq -p_0 v^-(t) + B(t).$$

We conclude, according to lemma 24, that  $v^- \in L^1(\mathbb{R}^d)$ , which implies that  $u$  converges. □





# Bibliography

- [1] Matthieu Alfaro, Henri Berestycki, and Gaël Raoul. The effect of climate shift on a species submitted to dispersion, evolution, growth, and nonlocal competition. *SIAM Journal on Mathematical Analysis*, 49(1):562–596, 2017.
- [2] Luís Almeida, Patrizia Bagnerini, Giulia Fabrini, Barry D Hughes, and Tommaso Lorenzi. Evolution of cancer cell populations under cytotoxic therapy and treatment optimisation: insight from a phenotype-structured model. *ESAIM: Mathematical Modelling and Numerical Analysis*, 53(4):1157–1190, 2019.
- [3] Uri Alon. *An introduction to systems biology: design principles of biological circuits*. CRC press, 2019.
- [4] FE Alvarez, JA Carrillo, and J Clairambault. Evolution of a structured cell population endowed with plasticity of traits under constraints on and between the traits. *J. Math. Biol.*, 85(64), 2022.
- [5] David Angeli, James E Ferrell Jr, and Eduardo D Sontag. Detection of multistability, bifurcations, and hysteresis in a large class of biological positive-feedback systems. *Proceedings of the National Academy of Sciences*, 101(7):1822–1827, 2004.
- [6] Gábor Balázsi, Alexander Van Oudenaarden, and James J Collins. Cellular decision making and biological noise: from microbes to mammals. *Cell*, 144(6):910–925, 2011.
- [7] HT Banks and JM Mahaffy. Stability of cyclic gene models for systems involving repression. *Journal of Theoretical Biology*, 74(2):323–334, 1978.
- [8] Albert-Laszlo Barabasi and Zoltan N Oltvai. Network biology: understanding the cell’s functional organization. *Nature reviews genetics*, 5(2):101–113, 2004.
- [9] Loic Barbaroux, Philippe Michel, Mostafa Adimy, and Fabien Crauste. Multi-scale modeling of the cd8 immune response. In *AIP Conference Proceedings*, volume 1738, page 320002. AIP Publishing LLC, 2016.
- [10] Guy Barles, Sepideh Mirrahimi, and Benoît Perthame. Concentration in lotka-volterra parabolic or integral equations: a general convergence result. *Methods and Applications of Analysis*, 16(3):321–340, 2009.
- [11] C.C. Bell and O. Gilan. Principles and mechanisms of non-genetic resistance in cancer. *British Journal of Cancer*, 122:465–472, 2020.
- [12] S. Bhatia, J. Monkman, T. Blick, C. Pinto, A. Waltham, S.H. Nagaraj, and E.W. Thompson. Interrogation of phenotypic plasticity between epithelial and mesenchymal states in breast cancer. *J Clin Med*, 8(6):893, 2019.

- 
- [13] Olivier Bonnefon, Jérôme Coville, and Guillaume Legendre. Concentration phenomenon in some non-local equation. *arXiv preprint arXiv:1510.01971*, 2015.
- [14] Emeric Bouin, Vincent Calvez, Nicolas Meunier, Sepideh Mirrahimi, Benoît Perthame, Gaël Raoul, and Raphaël Voituriez. Invasion fronts with variable motility: phenotype selection, spatial sorting and wave acceleration. *Comptes Rendus Mathématique*, 350(15-16):761–766, 2012.
- [15] Nicolas Bourbaki. *Elements of mathematics*. Springer Berlin Heidelberg, 1966.
- [16] Cameron P Bracken, Philip A Gregory, Natasha Kolesnikoff, Andrew G Bert, Jun Wang, M Frances Shannon, and Gregory J Goodall. A double-negative feedback loop between zeb1-sip1 and the microrna-200 family regulates epithelial-mesenchymal transition. *Cancer research*, 68(19):7846–7854, 2008.
- [17] M.S. Brown, B. Abdollahi, O.M. Wilkins, H. Lu, P. Chakraborty, N.B. Ognjenovic, K.E. Muller, M.K. Jolly, B.C. Christensen, S. Hassanpour, and D.R. Pattabiraman. Phenotypic heterogeneity driven by plasticity of the intermediate emt state governs disease progression and metastasis in breast cancer. *Science Advances*, 8(31), 2022.
- [18] Reinhard Bürger. *The mathematical theory of selection, recombination, and mutation*. John Wiley & Sons, 2000.
- [19] Olguta Buse, Rodrigo Pérez, and Alexey Kuznetsov. Dynamical properties of the repressilator model. *Physical Review E*, 81(6):066206, 2010.
- [20] Angel Calsina and Sílvia Cuadrado. Stationary solutions of a selection mutation model: The pure mutation case. *Mathematical Models and Methods in Applied Sciences*, 15(07):1091–1117, 2005.
- [21] Àngel Calsina, Sílvia Cuadrado, Laurent Desvillettes, and Gaël Raoul. Asymptotics of steady states of a selection–mutation equation for small mutation rate. *Proceedings of the Royal Society of Edinburgh Section A: Mathematics*, 143(6):1123–1146, 2013.
- [22] José Antonio Carrillo, Katy Craig, and Francesco S Patacchini. A blob method for diffusion. *Calculus of Variations and Partial Differential Equations*, 58:1–53, 2019.
- [23] Nicolas Champagnat, Régis Ferrière, and Sylvie Méléard. Unifying evolutionary dynamics: from individual stochastic processes to macroscopic models. *Theoretical population biology*, 69(3):297–321, 2006.
- [24] Nicolas Champagnat, Régis Ferrière, and Sylvie Méléard. From individual stochastic processes to macroscopic models in adaptive evolution. *Stochastic Models*, 24(sup1):2–44, 2008.
- [25] Britton Chance, Ronald W Estabrook, and A Ghosh. Damped sinusoidal oscillations of cytoplasmic reduced pyridine nucleotide in yeast cells. *Proceedings of the National Academy of Sciences*, 51(6):1244–1251, 1964.
- [26] Joshua L Cherry and Frederick R Adler. How to make a biological switch. *Journal of theoretical biology*, 203(2):117–133, 2000.
- [27] Alina Chertock. A practical guide to deterministic particle methods. In *Handbook of numerical analysis*, volume 18, pages 177–202. Elsevier, 2017.

- [28] Vijay Chickarmane, Carl Troein, Ulrike A Nuber, Herbert M Sauro, and Carsten Peterson. Transcriptional dynamics of the embryonic stem cell switch. *PLoS Comput Biol*, 2(9):e123, 2006.
- [29] Vi Ley Chin and Chooi Ling Lim. Epithelial-mesenchymal plasticity—engaging stemness in an interplay of phenotypes. *Stem Cell Investigation*, 6, 2019.
- [30] Rebecca H Chisholm, Tommaso Lorenzi, and Alexander Lorz. Effects of an advection term in nonlocal lotka–volterra equations. *Communications in mathematical sciences*, 14(4):1181–1188, 2016.
- [31] Rebecca H Chisholm, Tommaso Lorenzi, Alexander Lorz, Annette K Larsen, Luís Neves de Almeida, Alexandre Escargueil, and Jean Clairambault. Emergence of drug tolerance in cancer cell populations: an evolutionary outcome of selection, nongenetic instability, and stress-induced adaptation. *Cancer research*, 75(6):930–939, 2015.
- [32] Rebecca H. Chisholm, Tommaso Lorenzi, Alexander Lorz, Annette K. Larsen, Luís Neves de Almeida, Alexandre Escargueil, and Jean Clairambault. Emergence of drug tolerance in cancer cell populations: An evolutionary outcome of selection, nongenetic instability, and stress-induced adaptation. *Cancer Research*, 75(6):930–939, 2015.
- [33] JP Choquin and S Huberson. Particles simulation of viscous flow. *Computers & fluids*, 17(2):397–410, 1989.
- [34] Jason J Christiansen and Ayyappan K Rajasekaran. Reassessing epithelial to mesenchymal transition as a prerequisite for carcinoma invasion and metastasis. *Cancer research*, 66(17):8319–8326, 2006.
- [35] Dominique Chu, Nicolae Radu Zabet, and Boris Mitavskiy. Models of transcription factor binding: sensitivity of activation functions to model assumptions. *Journal of Theoretical Biology*, 257(3):419–429, 2009.
- [36] Hans Clevers, Susanne Rafelski, Michael Elowitz, Allon Klein, Jay Shendure, Cole Trapnell, Ed Lein, Emma Lundberg, Matthias Uhlen, Alfonso Martinez-Arias, et al. What is your conceptual definition of “cell type” in the context of a mature organism? *Cell Systems*, 4(3):255–259, 2017.
- [37] John L Cloudsley-Thompson. Rhythmic activity in animal physiology and behaviour. 1961.
- [38] Albert Cohen and Benoit Perthame. Optimal approximations of transport equations by particle and pseudoparticle methods. *SIAM Journal on Mathematical Analysis*, 32(3):616–636, 2000.
- [39] Daniel B Cooney and Yoichiro Mori. Long-time behavior of a pde replicator equation for multilevel selection in group-structured populations. *Journal of Mathematical Biology*, 85(2):1–67, 2022.
- [40] Georges-Henri Cottet and Pierre-Arnaud Raviart. On particle-in-cell methods for the vlasov-poisson equations. *Transport Theory and Statistical Physics*, 15(1-2):1–31, 1986.
- [41] GH Cottet and S Mas-Gallic. A particle method to solve the navier-stokes system. *Numerische Mathematik*, 57(1):805–827, 1990.



- 
- [42] Jerome Coville. Convergence to equilibrium for positive solutions of some mutation-selection model. *arXiv preprint arXiv:1308.6471*, 2013.
- [43] Francis Crick. Central dogma of molecular biology. *Nature*, 227(5258):561–563, 1970.
- [44] Francis H Crick. On protein synthesis. In *Symp Soc Exp Biol*, volume 12, page 8, 1958.
- [45] Eric H Davidson. *The regulatory genome: gene regulatory networks in development and evolution*. Elsevier, 2010.
- [46] Pierre Degond and S Mas-Gallic. The weighted particle method for convection-diffusion equations. i. the case of an isotropic viscosity. *Mathematics of computation*, 53(188):485–507, 1989.
- [47] Pierre Degond and Francisco-José Mustieles. A deterministic approximation of diffusion equations using particles. *SIAM Journal on Scientific and Statistical Computing*, 11(2):293–310, 1990.
- [48] Laurent Desvillettes, Pierre Emmanuel Jabin, Stéphane Mischler, and Gaël Raoul. On selection dynamics for continuous structured populations. *Communications in Mathematical Sciences*, 6(3):729–747, 2008.
- [49] Ulf Dieckmann and Richard Law. The dynamical theory of coevolution: a derivation from stochastic ecological processes. *Journal of mathematical biology*, 34(5):579–612, 1996.
- [50] Odo Diekmann, Pierre-Emanuel Jabin, Stéphane Mischler, and Benoit Perthame. The dynamics of adaptation: an illuminating example and a hamilton–jacobi approach. *Theoretical population biology*, 67(4):257–271, 2005.
- [51] Ronald J DiPerna and Pierre-Louis Lions. Ordinary differential equations, transport theory and sobolev spaces. *Inventiones mathematicae*, 98(3):511–547, 1989.
- [52] LNM Duysens and J Amesz. Fluorescence spectrophotometry of reduced phosphopyridine nucleotide in intact cells in the near-ultraviolet and visible region. *Biochimica et biophysica acta*, 24:19–26, 1957.
- [53] Michael B Elowitz and Stanislas Leibler. A synthetic oscillatory network of transcriptional regulators. *Nature*, 403(6767):335–338, 2000.
- [54] James E Ferrell and Eric M Machleder. The biochemical basis of an all-or-none cell fate switch in xenopus oocytes. *Science*, 280(5365):895–898, 1998.
- [55] Gerald B Folland. *Real analysis: modern techniques and their applications*, volume 40. John Wiley & Sons, 1999.
- [56] F. Font-Clos, S. Zapperi, and C.A.M. Porta. Topography of epithelial–mesenchymal plasticity. *Proceedings of the National Academy of Sciences*, 115(23):5902–5907, 2018.
- [57] Alex Fraser and Jawahar Tiwari. Genetical feedback-repression: Ii. cyclic genetic systems. *Journal of theoretical biology*, 47(2):397–412, 1974.
- [58] Peter Friedl and Darren Gilmour. Collective cell migration in morphogenesis, regeneration and cancer. *Nature reviews Molecular cell biology*, 10(7):445–457, 2009.

- [59] Avner Friedman, Chiu-Yen Kao, and Chih-Wen Shih. Asymptotic phases in a cell differentiation model. *Journal of Differential Equations*, 247(3):736–769, 2009.
- [60] Timothy S Gardner, Charles R Cantor, and James J Collins. Construction of a genetic toggle switch in *escherichia coli*. *Nature*, 403(6767):339–342, 2000.
- [61] J.T. George, M.K. Jolly, S. Xu, J.A. Somarelli, and H. Levine. Survival outcomes in cancer patients predicted by a partial emt gene expression scoring metric. *Cancer Research*, 77(22):6415–6428, 2017.
- [62] Stefan AH Geritz, Eva Kisdi, Johan AJ Metz, et al. Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree. *Evolutionary ecology*, 12(1):35–57, 1998.
- [63] Rudolf Gesztelyi, Judit Zsuga, Adam Kemeny-Beke, Balazs Varga, Bela Juhasz, and Arpad Tosaki. The hill equation and the origin of quantitative pharmacology. *Archive for history of exact sciences*, 66:427–438, 2012.
- [64] Brian C Goodwin. Oscillatory behavior in enzymatic control processes. *Advances in enzyme regulation*, 3:425–437, 1965.
- [65] Brian C Goodwin et al. Temporal organization in cells. a dynamic theory of cellular control processes. *Temporal organization in cells. A dynamic theory of cellular control processes.*, 1963.
- [66] Celso Grebogi, Edward Ott, and James A Yorke. Chaos, strange attractors, and fractal basin boundaries in nonlinear dynamics. *Science*, 238(4827):632–638, 1987.
- [67] Raúl Guantes and Juan F Poyatos. Multistable decision switches for flexible control of epigenetic differentiation. *PLoS Comput Biol*, 4(11):e1000235, 2008.
- [68] Jules Guilberteau, Camille Pouchol, and Nastassia Pouradier Duteil. Monostability and bistability of biological switches. *Journal of Mathematical Biology*, 83(6-7):65, 2021.
- [69] Mats Gyllenberg and Géza Meszéna. On the impossibility of coexistence of infinitely many strategies. *Journal of mathematical biology*, 50(2):133–160, 2005.
- [70] TG Hallam and CE Clark. Non-autonomous logistic equations as models of populations in a deteriorating environment. *Journal of Theoretical Biology*, 93(2):303–311, 1981.
- [71] K. Hari, B. Sabuwala, B.V. Subramani, C.A.M. Porta, S. Zapperi, F. Font-Clos, and M.K. Jolly. Identifying inhibitors of epithelial–mesenchymal plasticity using a network topology-based approach. *Npj Systems Biology and Applications 2020 6:1*, 6(1):1–12, 2020.
- [72] K. Hari, V. Ullanat, A. Balasubramanian, A. Gopalan, and M.K. Jolly. Landscape of epithelial mesenchymal plasticity as an emergent property of coordinated teams in regulatory networks. *ELife*, 11, 2022.
- [73] Philip Hartman. *Ordinary differential equations*. SIAM, 2002.
- [74] F Hermeline. A deterministic particle method for transport diffusion equations: Application to the fokker-planck equation. *Journal of Computational Physics*, 82(1):122–146, 1989.

- 
- [75] Archibald Vivian Hill. The possible effects of the aggregation of the molecules of hemoglobin on its dissociation curves. *j. physiol.*, 40:iv–vii, 1910.
- [76] WGS Hines. Evolutionary stable strategies: a review of basic theory. *Theoretical Population Biology*, 31(2):195–272, 1987.
- [77] Morris W Hirsch. Systems of differential equations which are competitive or cooperative: I. limit sets. *SIAM Journal on Mathematical Analysis*, 13(2):167–179, 1982.
- [78] Morris W Hirsch. Differential equations and convergence almost everywhere in strongly monotone semiflows. *Contemp. Math*, 17:267–285, 1983.
- [79] Morris W Hirsch. The dynamical systems approach to differential equations. *Bulletin of the American mathematical society*, 11(1):1–64, 1984.
- [80] M. Hitomi, A.P. Chumakova, D.J. Silver, A.M. Knudsen, W.D. Pontius, S. Murphy, N. Anand, B.W. Kristensen, and J.D. Lathia. Asymmetric cell division promotes therapeutic resistance in glioblastoma stem cells. *JCI Insight*, 6(3), 2021.
- [81] Roger W Hockney and James W Eastwood. *Computer simulation using particles*. crc Press, 2021.
- [82] T. Hong, K. Watanabe, C.H. Ta, A. Villarreal-Ponce, Q. Nie, and X. Dai. An *ovol2-zeb1* mutual inhibitory circuit governs bidirectional and multi-step transition between epithelial and mesenchymal states. *PLOS Computational Biology*, 11(11):1004569, 2015.
- [83] Thomas Y Hou. Convergence of a variable blob vortex method for the euler and navier–stokes equations. *SIAM Journal on Numerical Analysis*, 27(6):1387–1404, 1990.
- [84] Sui Huang, Yan-Ping Guo, Gillian May, and Tariq Enver. Bifurcation dynamics in lineage-commitment in bipotent progenitor cells. *Developmental biology*, 305(2):695–713, 2007.
- [85] Witold Hurewicz and Henry Wallman. *Dimension Theory (PMS-4), Volume 4*, volume 63. Princeton university press, 2015.
- [86] Pierre-Emmanuel Jabin and Gaël Raoul. On selection dynamics for competitive interactions. *Journal of mathematical biology*, 63(3):493–517, 2011.
- [87] François Jacob and Jacques Monod. Genetic regulatory mechanisms in the synthesis of proteins. *Journal of molecular biology*, 3(3):318–356, 1961.
- [88] V. Jacquemin, M. Antoine, J.E. Dumont, G. Dom, V. Detours, and C. Maenhaut. Dynamic cancer cell heterogeneity: Diagnostic and therapeutic implications. *Cancers*, 14(2):280, 2022.
- [89] P. Jain, S. Bhatia, E.W. Thompson, and M.K. Jolly. Population dynamics of epithelial-mesenchymal heterogeneity in cancer cells. *Biomolecules*, 12(3):348, 2022.
- [90] Paras Jain, Sophia Corbo, Kulsoom Mohammad, Sarthak Sahoo, Santhalakshmi Ranganathan, Jason T George, Herbert Levine, Joseph Taube, Michael Toneff, and Mohit Kumar Jolly. Epigenetic memory acquired during long-term emt induction governs the recovery to the epithelial state. *Journal of the Royal Society Interface*, 20(198):20220627, 2023.

- [91] Dongya Jia, Mohit Kumar Jolly, William Harrison, Marcelo Boareto, Eshel Ben-Jacob, and Herbert Levine. Operating principles of tristable circuits regulating cellular differentiation. *Physical biology*, 14(3):035007, 2017.
- [92] Robert J Johnston, Sarah Chang, John F Etchberger, Christopher O Ortiz, and Oliver Hobert. Micrnas acting in a double-negative feedback loop to control a neuronal cell fate decision. *Proceedings of the National Academy of Sciences*, 102(35):12449–12454, 2005.
- [93] Mohit Kumar Jolly, Marcelo Boareto, Bin Huang, Dongya Jia, Mingyang Lu, Eshel Ben-Jacob, José N Onuchic, and Herbert Levine. Implications of the hybrid epithelial/mesenchymal phenotype in metastasis. *Frontiers in oncology*, 5:155, 2015.
- [94] Raghu Kalluri, Robert A Weinberg, et al. The basics of epithelial-mesenchymal transition. *The Journal of clinical investigation*, 119(6):1420–1428, 2009.
- [95] L.G. Karacosta, B. Anchang, N. Ignatiadis, S.C. Kimmey, J.A. Benson, J.B. Shrager, R. Tibshirani, S.C. Bendall, and S.K. Plevritis. Mapping lung cancer epithelial-mesenchymal transition states and trajectories with single-cell resolution. *Nature Communications*, 10:5587, 2019.
- [96] M Karin. Too many transcription factors: positive and negative interactions. *The new biologist*, 2(2):126–131, 1990.
- [97] Peter Laslo, Chauncey J Spooner, Aryeh Warmflash, David W Lancki, Hyun-Jun Lee, Roger Sciammas, Benjamin N Gantner, Aaron R Dinner, and Harinder Singh. Multilineage transcriptional priming and determination of alternate hematopoietic cell fates. *Cell*, 126(4):755–766, 2006.
- [98] David S Latchman. Inhibitory transcription factors. *The international journal of biochemistry & cell biology*, 28(9):965–974, 1996.
- [99] David S Latchman. Transcription factors: an overview. *The international journal of biochemistry & cell biology*, 29(12):1305–1312, 1997.
- [100] Anthony Leonard. Vortex methods for flow simulation. *Journal of Computational Physics*, 37(3):289–335, 1980.
- [101] Anthony Leonard. Computing three-dimensional incompressible flows with vortex elements. *Annual Review of Fluid Mechanics*, 17(1):523–559, 1985.
- [102] Jie Li and Weinian Zhang. Transition between monostability and bistability of a genetic toggle switch in escherichia coli. *Discrete & Continuous Dynamical Systems-Series B*, 25(5), 2020.
- [103] Ronghui Li, Jialiang Liang, Su Ni, Ting Zhou, Xiaobing Qing, Huapeng Li, Wenzhi He, Jiekai Chen, Feng Li, Qiang Zhuang, et al. A mesenchymal-to-epithelial transition initiates and is required for the nuclear reprogramming of mouse fibroblasts. *Cell stem cell*, 7(1):51–63, 2010.
- [104] Tommaso Lorenzi, Rebecca H Chisholm, Laurent Desvillettes, and Barry D Hughes. Dissecting the dynamics of epigenetic changes in phenotype-structured populations exposed to fluctuating environments. *Journal of theoretical biology*, 386:166–176, 2015.

- 
- [105] Tommaso Lorenzi, Benoît Perthame, and Xinran Ruan. Invasion fronts and adaptive dynamics in a model for the growth of cell populations with heterogeneous mobility. *European Journal of Applied Mathematics*, 33(4):766–783, 2022.
- [106] Tommaso Lorenzi and Camille Pouchol. Asymptotic analysis of selection-mutation models in the presence of multiple fitness peaks. *Nonlinearity*, 33(11):5791, 2020.
- [107] Alexander Lorz, Tommaso Lorenzi, Michael E Hochberg, Jean Clairambault, and Benoît Perthame. Populational adaptive evolution, chemotherapeutic resistance and multiple anti-cancer therapies. *ESAIM: Mathematical Modelling and Numerical Analysis*, 47(2):377–399, 2013.
- [108] S. Lovisa, V.S. LeBleu, B. Tampe, H. Sugimoto, K. Vadnagara, J.L. Carstens, C.C. Wu, Y. Hagos, B.C. Burkhardt, T. Pentcheva-Hoang, H. Nischal, J.P. Allison, M. Zeisberg, and R. Kalluri. Epithelial-to-mesenchymal transition induces cell cycle arrest and parenchymal damage in renal fibrosis. *Nature Medicine*, 21(9):998–1009, 2015.
- [109] Mingyang Lu, Mohit Kumar Jolly, Ryan Gomoto, Bin Huang, Jose Onuchic, and Eshel Ben-Jacob. Tristability in cancer-associated microRNA-tf chimera toggle switch. *The journal of physical chemistry B*, 117(42):13164–13174, 2013.
- [110] Mingyang Lu, Mohit Kumar Jolly, Herbert Levine, José N Onuchic, and Eshel Ben-Jacob. MicroRNA-based regulation of epithelial–hybrid–mesenchymal fate determination. *Proceedings of the National Academy of Sciences*, 110(45):18144–18149, 2013.
- [111] B Lucquin-Desreux and S Mas-Gallic. A conservative particle approximation for a boundary advection-diffusion problem. *ESAIM: Mathematical Modelling and Numerical Analysis*, 26(6):757–791, 1992.
- [112] Javier Macía, Stefanie Widder, and Ricard Solé. Why are cellular switches boolean? general conditions for multistable genetic circuits. *Journal of theoretical biology*, 261(1):126–135, 2009.
- [113] John Mallet-Paret and Hal Smith. The poincaré-bendixson theorem for monotone cyclic feedback systems. *Journal of Dynamics and Differential Equations*, 2(4):367–421, 1990.
- [114] A. Marusyk, M. Janiszewska, and K. Polyak. Intratumor heterogeneity: The rosetta stone of therapy resistance. *Cancer Cell*, 37(4):471–484, 2020.
- [115] S Mas-Gallic and PA Raviart. A particle method for first-order symmetric systems. *Numerische Mathematik*, 51(3):323–352, 1987.
- [116] AG McKendrick and M Kesava Pai. Xlv.—the rate of multiplication of microorganisms: a mathematical study. *Proceedings of the Royal Society of Edinburgh*, 31:649–655, 1912.
- [117] Johan AJ Metz, Stefan AH Geritz, Géza Meszéna, Frans JA Jacobs, and Joost S Van Heerwaarden. Adaptive dynamics: a geometrical study of the consequences of nearly faithful reproduction. 1995.
- [118] Philippe Michel, Stéphane Mischler, and Benoît Perthame. General relative entropy inequality: an illustration on growth models. *Journal de mathématiques pures et appliquées*, 84(9):1235–1260, 2005.

- [119] John Milnor. On the concept of attractor. *Communications in Mathematical Physics*, 99:177–195, 1985.
- [120] Jacques Monod, Jean-Pierre Changeux, and François Jacob. Allosteric proteins and cellular control systems. *Journal of molecular biology*, 6(4):306–329, 1963.
- [121] Samantha A Morris. The evolving concept of cell identity in the single cell era. *Development*, 146(12):dev169748, 2019.
- [122] Stefan Müller, Josef Hofbauer, Lukas Endler, Christoph Flamm, Stefanie Widder, and Peter Schuster. A generalized model of the repressilator. *Journal of mathematical biology*, 53:905–937, 2006.
- [123] B. Munsky, Z. Fox, and G. Neuert. Integrating single-molecule experiments and discrete stochastic models to understand heterogeneous gene transcription dynamics. *Methods*, 85:12–21, 2015.
- [124] M Angela Nieto. Epithelial plasticity: a common theme in embryonic and cancer cells. *Science*, 342(6159):1234850, 2013.
- [125] Bela Novak and John J Tyson. Numerical analysis of a comprehensive model of m-phase control in xenopus oocyte extracts and intact embryos. *Journal of cell science*, 106(4):1153–1168, 1993.
- [126] Hans G Othmer. The qualitative dynamics of a class of biochemical control circuits. *Journal of mathematical biology*, 3(1):53–78, 1976.
- [127] Ertugrul M Ozbudak, Mukund Thattai, Han N Lim, Boris I Shraiman, and Alexander Van Oudenaarden. Multistability in the lactose utilization network of escherichia coli. *Nature*, 427(6976):737–740, 2004.
- [128] D. Pally, S. Goutham, and R. Bhat. Extracellular matrix as a driver for intratumoral heterogeneity. *Physical Biology*, 19(4):043001, 2022.
- [129] Dmitri Papatsenko and Michael Levine. The drosophila gap gene network is composed of two parallel toggle switches. *PLoS One*, 6(7):e21145, 2011.
- [130] Lawrence Perko. *Differential equations and dynamical systems*, volume 7. Springer Science & Business Media, 2013.
- [131] Benoît Perthame. *Transport equations in biology*. Springer Science & Business Media, 2006.
- [132] Benoît Perthame and Guy Barles. Dirac concentrations in lotka-volterra parabolic pdes. *Indiana University Mathematics Journal*, pages 3275–3301, 2008.
- [133] M. Pillai, E. Hojel, M.K. Jolly, and Y. Goyal. Unraveling non-genetic heterogeneity in cancer with dynamical models and computational tools. *Nature Computational Science 2023 3:4*, 3(4):301–313, 2023.
- [134] Camille Pouchol and Emmanuel Trélat. Global stability with selection in integro-differential lotka-volterra systems modelling trait-structured populations. *Journal of Biological Dynamics*, 12(1):872–893, 2018.
- [135] M. Rashid, K. Hari, J. Thampi, N.K. Santhosh, and M.K. Jolly. Network topology metrics explaining enrichment of hybrid epithelial/mesenchymal phenotypes in metastasis. *PLoS Computational Biology*, 18(11):1010687, 2022.

- 
- [136] Pierre-Arnaud Raviart. An analysis of particle methods. In *Numerical methods in fluid dynamics*, pages 243–324. Springer, 1985.
- [137] Jean Richelle. Boucles de rétroaction négatives: un parallèle entre diverses analyses formelles. *Bulletins de l'Académie Royale de Belgique*, 63(1):534–546, 1977.
- [138] Ingo Roeder and Ingmar Glauche. Towards an understanding of lineage specification in hematopoietic stem cells: a mathematical model for the interaction of transcription factors gata-1 and pu. 1. *Journal of theoretical biology*, 241(4):852–865, 2006.
- [139] M. Ruscetti, E.L. Dadashian, W. Guo, B. Quach, D.J. Mulholland, J.W. Park, L.M. Tran, N. Kobayashi, D. Bianchi-Frias, Y. Xing, P.S. Nelson, and H. Wu. Hdac inhibition impedes epithelial-mesenchymal plasticity and suppresses metastatic, castration-resistant prostate cancer. *Oncogene*, 35(29):3781–3795, 2016.
- [140] Giovanni Russo. Deterministic diffusion of particles. *Communications on Pure and Applied Mathematics*, 43(6):697–733, 1990.
- [141] S. Sahoo, A. Mishra, H. Kaur, K. Hari, S. Muralidharan, S. Mandal, and M.K. Jolly. A mechanistic model captures the emergence and implications of non-genetic heterogeneity and reversible drug resistance in er+ breast cancer cells. *NAR Cancer*, 3(3), 2021.
- [142] Payman Samavarchi-Tehrani, Azadeh Golipour, Laurent David, Hoon-ki Sung, Tobias A Beyer, Alessandro Datti, Knut Woltjen, Andras Nagy, and Jeffrey L Wrana. Functional genomics reveals a bmp-driven mesenchymal-to-epithelial transition in the initiation of somatic cell reprogramming. *Cell stem cell*, 7(1):64–77, 2010.
- [143] James F Selgrade. Asymptotic behavior of solutions to single loop positive feedback systems. *Journal of Differential Equations*, 38(1):80–103, 1980.
- [144] Helge Siemens, Rene Jackstadt, Sabine Hünten, Markus Kaller, Antje Menssen, Ursula Götz, and Heiko Hermeking. mir-34 and snail form a double-negative feedback loop to regulate epithelial-mesenchymal transitions. *Cell cycle*, 10(24):4256–4271, 2011.
- [145] A. Sigal, R. Milo, A. Cohen, N. Geva-Zatorsky, Y. Klein, Y. Liron, N. Rosenfeld, T. Danon, N. Perzov, and U. Alon. Variability and memory of protein levels in human cells. *Nature*, 444(7119):643–646, 2006.
- [146] Hal Smith. Oscillations and multiple steady states in a cyclic gene model with repression. *Journal of mathematical biology*, 25(2):169–190, 1987.
- [147] Hal L Smith. Systems of ordinary differential equations which generate an order preserving flow. a survey of results. *SIAM review*, 30(1):87–113, 1988.
- [148] S.N. Steinway, J.G.T. Zañudo, P.J. Michel, D.J. Feith, T.P. Loughran, and R. Albert. Combinatorial interventions inhibit  $\text{tgf}\beta$ -driven epithelial-to-mesenchymal transition and support hybrid cellular phenotypes. *Npj Systems Biology and Applications*, 1:15014, 2015.
- [149] René Thomas. Boolean formalization of genetic control circuits. *Journal of theoretical biology*, 42(3):563–585, 1973.
- [150] René Thomas. Laws for the dynamics of regulatory networks. *International Journal of Developmental Biology*, 42(3):479–485, 2002.

- [151] René Thomas. *Kinetic logic: a Boolean approach to the analysis of complex regulatory systems: proceedings of the EMBO course “formal analysis of genetic regulation”, held in Brussels, September 6–16, 1977*, volume 29. Springer Science & Business Media, 2013.
- [152] René Thomas and Richard d’Ari. *Biological feedback*. CRC press, 1990.
- [153] René Thomas and Jean Richelle. Positive feedback loops and multistationarity. *Discrete Applied Mathematics*, 19(1-3):381–396, 1988.
- [154] Jawahar Tiwari, Alex Fraser, and Richard Beckman. Genetical feedback repression i. single locus models. *Journal of Theoretical Biology*, 45(2):311–326, 1974.
- [155] Cole Trapnell. Defining cell types and states with single-cell genomics. *Genome research*, 25(10):1491–1498, 2015.
- [156] S. Tripathi, P. Chakraborty, H. Levine, and M.K. Jolly. A mechanism for epithelial-mesenchymal heterogeneity in a population of cancer cells. *PLoS Computational Biology*, 16(2):1–27, 2020.
- [157] John J Tyson and Bela Novak. A dynamical paradigm for molecular cell biology. *Trends in Cell Biology*, 30(7):504–515, 2020.
- [158] H Edwin Umbarger. Evidence for a negative-feedback mechanism in the biosynthesis of isoleucine. *Science*, 123(3202):848–848, 1956.
- [159] Cruz Vargas-De-León. Lyapunov functions for two-species cooperative systems. *Applied Mathematics and Computation*, 219(5):2493–2497, 2012.
- [160] PF Verhulst. La loi d’accroissement de la population. *Nouveaux Memories de l’Académie Royale des Sciences et Belles-Lettres de Bruxelles*, 18:14–54, 1845.
- [161] Stefanie Widder, Josef Schicho, and Peter Schuster. Dynamic patterns of gene regulation i: simple two-gene systems. *Journal of theoretical biology*, 246(3):395–419, 2007.
- [162] Roderick Wong. *Asymptotic approximations of integrals*. SIAM, 2001.
- [163] Bo Xia and Itai Yanai. A periodic table of cell types. *Development*, 146(12):dev169854, 2019.
- [164] Andrew Yates, Robin Callard, and Jaroslav Stark. Combining cytokine signalling with t-bet and gata-3 regulation in th1 and th2 differentiation: a model for cellular decision-making. *Journal of theoretical biology*, 231(2):181–196, 2004.
- [165] Jingyu Zhang, Xiao-Jun Tian, Hang Zhang, Yue Teng, Ruoyan Li, Fan Bai, Subbiah Elankumaran, and Jianhua Xing. Tgf- $\beta$ -induced epithelial-to-mesenchymal transition proceeds through stepwise activation of multiple feedback loops. *Science signaling*, 7(345):ra91–ra91, 2014.