



HAL
open science

Deterministic limit of a PDMP model of epithelial tissue interacting with diffusing chemicals and application to the intestinal crypt

Léo Darrigade, Simon Labarthe, Béatrice Laroche

► To cite this version:

Léo Darrigade, Simon Labarthe, Béatrice Laroche. Deterministic limit of a PDMP model of epithelial tissue interacting with diffusing chemicals and application to the intestinal crypt. 2023. hal-04336174v2

HAL Id: hal-04336174

<https://hal.science/hal-04336174v2>

Preprint submitted on 13 Dec 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.



Distributed under a Creative Commons Attribution 4.0 International License

Deterministic limit of a PDMP model of epithelial tissue interacting with diffusing chemicals and application to the intestinal crypt

Léo Darrigade ^{*1}, Simon Labarthe², and Béatrice Laroche ^{†3}

¹Université Paris-Saclay, INRAE, MaIAGE, 78350, Jouy-en-Josas, France, leo.darrigade@inria.fr

²Université Paris-Saclay, INRAE, MaIAGE, 78350, Jouy-en-Josas, France, Univ. Bordeaux, INRAE, BIOGECO, 33610, Cestas, France, Inria, INRAE, Pléiade, 33400, Talence, France, simon.labarthe@inrae.fr

³Université Paris-Saclay, INRAE, MaIAGE, 78350, Jouy-en-Josas, France, Université Paris-Saclay, Inria, 91120, Palaiseau, France, beatrice.laroche@inrae.fr

December 13, 2023

Abstract

Mathematical models of biological tissues are a promising tool for multi-scale data integration, computational experiments and system biology approaches. While some data and insights are rooted at the cellular level, macroscopic mechanisms emerge and are observed at the tissue scale, rendering tissue modeling an inherently multi-scale process. As a consequence, tissue models can be broadly categorized as either individual-based or continuous population-based.

In this paper, we first introduce a generic individual-based model of epithelial tissue including the main regulation processes such as cell division, differentiation, migration and death, together with cell-cell mechanical interactions. The coupling with diffusing molecules is also considered. This model is a measure-valued piecewise-deterministic Markov processes, coupled with reaction-diffusion PDEs. The well-posedness of the model is assessed and large population deterministic limit is rigorously derived. Finally, numerical experiments are conducted: this model is applied to the context of epithelial tissues in the intestinal crypt and the convergence towards the deterministic model are numerically illustrated.

Keywords: biological tissue model; individual-based model (IBM); large population asymptotics.

AMS Subject Classification: 60J85, 92D25, 92C15.

Contents

1	Introduction	2
2	Preliminary notations and definitions	3
3	Definition of the process	5
4	Rescaled process	9
5	Deterministic limit of the rescaled process by a large population approximation.	12
6	Numerical convergence of the PDMP towards a deterministic limit	21
7	Conclusive discussion	26
A	Metrics on $\mathcal{M}_{F^+}(\mathcal{X})$, properties of $\mathcal{D}([0, T], (\mathcal{M}_{F^+}(\mathcal{X}), d_{PR}))$	29
B	Reaction-diffusion equations	32

*Current address : Université Paris-Saclay, Inria, 91120, Palaiseau, France

†Corresponding author

1 Introduction

Biological tissues are meso-scale biological structures between micro-scale individual cells and macro-scale organs. Tissues are characterized by interacting cells with similar physiology embedded into an extracellular matrix. Thus, tissues have their own regulation mechanisms including cell metabolism, cell-cell communications, cell movement and mechanical interactions [MFK09]. In the framework of cell-population studies at the tissue scale, mathematical models of biological tissues are particularly useful for integrating data at different scales or for testing hypothesis with high-throughput in silico experiments [FO22]. A large set of tissue models have then been developed in different contexts as diverse as surgical robotics [FS08], embryogenesis [BNL19; Ros+20] or mathematical oncology [Roc+19].

Usually, tissue models are either individual-based models (IBM) or continuous population models. IBMs treat cells as distinct entities, aggregating a large number of individual cell models. These models are interconnected through cell-to-cell interactions, facilitating the emergence of patterns at the tissue scale [Osb+17]. IBMs offer several key advantages. Primarily, they align well with biological observations frequently made at the cellular level, such as through microscopy or single-cell experiments, making them suitable for integrating data and knowledge. Additionally, they are particularly well adapted to object-oriented programming, allowing to leverage high-performance and parallel computing. Furthermore, IBMs efficiently incorporate stochastic events at the individual level, allowing for the exploration of inter-individual variability within populations. However, the computational and memory requirements for IBM scale at least linearly with the number of cell models, posing challenges for large cell numbers that approach real tissue scales. In such cases, the synthetic data generated by IBMs can be as complex to analyze as actual experimental data [Osb+17].

On the other hand, continuous population models consider the fate of the cell distribution at the population level. They focus on the dynamics of averaged quantities across the population, such as the mean of a quantitative trait or higher order momentum, using ordinary or partial differential equations (ODE or PDE). These models, however, lack important properties of the IBMs. They are most often fully deterministic (although stochastic events can technically be incorporated) whereas stochasticity is an important aspect of cellular behaviour [LD08]. Additionally, they do not facilitate the modelling of individual trajectories hindering the comparison with individual-based data. The main advantage of continuous population models is their ability to analyze population-wide traits independent of the actual cell count, allowing scalability for large populations. Moreover, their relative simplicity compared to large scale IBMs make them more amenable to mathematical analysis.

Tissue models can integrate both formalism. For instance, IBMs often include smaller-scale elements such as nutrients through a continuous model coupled to the IBM [Smi+12; Osb+17]. Furthermore, continuous population models can be directly derived from IBMs using large population approximations [EK86]. This approach is particularly attractive for several reasons. It establishes a formal link between micro-scale IBM parameters and macro-scale population model parameters, facilitating the connection of measurements across scales. Additionally, the deterministic asymptotic model derived from the IBM can be used to accelerate computations in computationally intensive tasks such as parameter inference or sensitivity analysis. If qualitative mathematical analysis is impractical on IBMs due to technical constraints, it might be feasible on the deterministic approximation. This qualitative behavior may then be extrapolated to the IBM for large enough populations.

The goal of this paper is twofold: we aim to propose a versatile stochastic individual-based model of epithelial tissues, that depicts a population of cells interacting with diffusing chemicals, alongside a rigorous deterministic approximation. We want the IBM general enough to be applicable to a broad range of tissues, encapsulating essential aspects of biological tissues turnover. The model is structured by cell types and spatial dimensions, enabling the representation of population distributions along these axes. Additionally, it accounts for cell fate regulation by diffusing extracellular molecules and the associated spatial variations in chemical concentrations. We therefore model : 1) spatial movement of cells driven by mechanical interactions with neighbouring cells, 2) stochastic cell division, extrusion and differentiation regulated by chemical concentrations, cell location and local cell density, 3) spatial diffusion of chemicals, 4) consumption or production of chemicals by cell metabolism.

Although these cellular functions are broadly applicable, our primary focus is on the epithelium of the intestinal colon crypt, a tissue with unique characteristics that make it an ideal test case for our

model. First, the gut epithelium is one of the tissue with the highest renewal rate in the human body: the cell turn over is in the range of a few days and includes strong dynamical processes such as differentiation and cell migration [Bar14]. Next, a lot of different cell-cell interactions occur, including mechanical repression of division (contact inhibition), contact signaling [Bat+02] and longer scale interactions with diffusing molecules [Wan+18]. Furthermore, complex interactions with the cell environment take place, in particular the uptake of gut-microbes-derived metabolites [Kai+16]. Models of the gut crypt already exist, both in the IBMs [DNO13; Bus+12; Tha+18] and continuous population model [Joh+07; Mur+11] frameworks. However, these IBMs lack a rigorous mathematical foundation, particularly in their stochastic components, limiting the scope for mathematical asymptotic analysis. We note a previous derivation of a population model from IBMs [Mur+11] but the derivation is purely formal. Our paper addresses these gaps by defining an IBM within the PDMP (Piecewise Deterministic Markov Process) framework and proving convergence to the population model using functional analysis tools. This mathematical framework was first introduced in theoretical ecology and evolution [FM04; CFM07; Tra06]. It was then adapted to other biological context with less emphasis on evolutionary dynamics such as the chemostat [CF14] or the ovarian follicle [Mic17]. Our work extends previous studies by accounting for a greater diversity of biological phenomena. It provides a strong theoretical grounding for the models used in the community for mathematical oncology, morphogenesis and gut microbiota studies [Dar+22].

Our individual based PDMP model (PDMP-IBM) represents the population of cells as a point measure and concentrations of molecules as smooth densities. In this class of IBM continuous processes modeled by deterministic equations (like ODEs) are punctuated by stochastic events: the random process determines the time occurrence and the magnitude of stochastic jumps in the state space, with individual trajectories driven by the continuous deterministic model between jumps [Tra08; CF14; Ben+15]. Our model's deterministic part includes cell-to-cell interactions and the reaction-diffusion of chemicals, while the stochastic part models cell division, extrusion and differentiation. Custom notations used throughout the paper are defined in section 2. The PDMP is introduced in section 3, with its deterministic part presented in section 3.1 and its stochastic part in 3.2. The well-posedness and stability of the process are shown in 4. We then prove in section 5 weak convergence of this stochastic cell-based model to a deterministic model of self-aggregation type, under the hypothesis of large population. This part is inspired by previous works [FM04; Tra06; CF14] with adaptations to take into account mechanical interactions between cells and the strong coupling of the cell population with reaction-diffusion PDEs. Finally, section 6 assesses the convergence numerically, using a model that simulates epithelial turn-over in an intestinal crypt.

2 Preliminary notations and definitions

We consider a finite set \mathcal{T} of cell types, and a spatial domain \mathcal{Z} which is a compact subset of \mathbb{R}^d , with $d \leq 3$. Cells are characterised by their position and their type $x = (z, l) \in \mathcal{Z} \times \mathcal{T}$. We model a population of cells ν as a point measure on $\mathcal{X} = \mathcal{Z} \times \mathcal{T}$:

$$\nu = \sum_{i=1}^n \delta_{x_i}(dx) \quad (1)$$

where $\delta_x(dx)$ is the Dirac measure on x and $n = \langle \nu, 1 \rangle$ is the total number of cells in the population. We also model the spatial and temporal evolution of c , the vector of concentrations of $N_m \in \mathbb{N}^*$ chemicals on \mathcal{Z} .

2.1 Topology and metric on \mathcal{Z} and \mathcal{X}

We endow $\mathcal{Z} \subset \mathbb{R}^d$ with the topology associated to the euclidean norm and denote $\mathcal{B}(\mathcal{Z})$ the corresponding Borel σ -algebra, and endow \mathcal{T} with the topology generated by $\mathcal{P}(\mathcal{T})$, the set of its parts. The product space $\mathcal{X} = \mathcal{Z} \times \mathcal{T}$ is equipped with the product topology for which an associated metric $d_{\mathcal{X}}$ is defined as

$$d_{\mathcal{X}} : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}^+ \\ ((z_1, l_1), (z_2, l_2)) \mapsto \begin{cases} |z_1 - z_2| & \text{if } l_1 = l_2, \\ \beta & \text{otherwise,} \end{cases} \quad (2)$$

with $\beta > \text{diam}(\mathcal{Z})$ and $|\cdot|$ the euclidean norm on \mathcal{Z} . Then the product σ -algebra $\mathcal{B}(\mathcal{X}) = \mathcal{B}(\mathcal{Z}) \otimes \mathcal{P}(\mathcal{T})$ is the Borel σ -algebra on $(\mathcal{X}, d_{\mathcal{X}})$.

Moreover, we sometimes use $\hat{\mathcal{Z}} := \{z - y : (z, y) \in \mathcal{Z}^2\}$.

2.2 Functions defined on \mathcal{X}

Let f be a function defined on \mathcal{X} . For all l in \mathcal{T} , we denote $f_l : z \mapsto f(z, l)$ the restriction of f to $\mathcal{Z} \times \{l\}$. Then f is:

- measurable (resp. continuous) iff, for all $l \in \mathcal{T}$, f_l is measurable (resp. continuous) with respect to the Borel σ -algebra (resp. the euclidian norm) on \mathcal{Z} .
- $f \in \mathcal{C}^{k, \bullet}(\mathcal{X}, \mathbb{R})$ iff, for all $l \in \mathcal{T}$, f_l is continuous of class k in \mathcal{Z} ($f_l \in \mathcal{C}^k$), meaning that it is k times differentiable on \mathcal{Z} , with k -th derivative continuous on \mathcal{Z} . When $k \geq 1$, we denote $\nabla_z f(x) \in \mathbb{R}^d$ the gradient operator with respect to the space variable z .
- $f \in \mathcal{C}^{k_1, k_2, \bullet}(\mathbb{R}^+ \times \mathcal{X}, \mathbb{R})$ iff, for all $l \in \mathcal{T}$, $f_l(t, z)$ is continuous of class k_1 with respect to $t \in \mathbb{R}^+$ and continuous of class k_2 with respect to $z \in \mathcal{Z}$.

For $f : \mathcal{X} \rightarrow \mathbb{R}$, we write

$$\|f\|_\infty = \sup_{x \in \mathcal{X}} |f(x)| \quad \text{and} \quad \|f\|_{Lip} = \sup_{(x, y) \in \mathcal{X}, x \neq y} \left\{ \frac{|f(x) - f(y)|}{d_{\mathcal{X}}(x, y)} \right\}.$$

We have the following elementary proposition.

Proposition 1. $f : \mathcal{X} \rightarrow \mathbb{R}$ is Lipschitz on \mathcal{X} iff, $\forall l \in \mathcal{T}$, f_l is Lipschitz on \mathcal{Z} . Moreover,

$$\|f\|_{Lip} \leq \max \left(\max_{l \in \mathcal{T}} \|f_l\|_{Lip(\mathcal{Z})}, \frac{2\|f\|_\infty}{\beta} \right)$$

with $\|f_l\|_{Lip(\mathcal{Z})}$ is the Lipschitz constant of f_l on \mathcal{Z} and β is defined in (2).

For f Lipschitz on \mathcal{X} , we define

$$\|f\|_\vee = \max(\|f\|_\infty, \|f\|_{Lip}) \quad (3)$$

and introduce the space of functions bounded by 1 with Lipschitz constant lower than 1

$$\mathcal{Lip}_{1,1}(\mathcal{X}, \mathbb{R}) = \{f : \mathcal{X} \rightarrow \mathbb{R} : \|f\|_\vee \leq 1\}.$$

2.3 Measures on \mathcal{X}

We denote $\mathcal{M}_F(\mathcal{X})$ ($\mathcal{M}_{F^+}(\mathcal{X})$) the space of finite (resp. finite positive) measures on $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$, and $\mathcal{M}_P(\mathcal{X})$ the subset of finite positive point measures:

$$\mathcal{M}_P(\mathcal{X}) = \left\{ \sum_{i=1}^n \delta_{x^i}; n \in \mathbb{N}, x^i = (z^i, l^i) \in \mathcal{Z} \times \mathcal{T} \right\},$$

with $\sum_{i=1}^0 \delta_{x^i}(x)$ the null measure by convention.

Let ν be a measure of $\mathcal{M}_F(\mathcal{X})$, then, for all $l \in \mathcal{T}$, we define $\nu^l \in \mathcal{M}_F \mathcal{Z}$ as

$$\nu^l(Z) = \nu(Z \times \{l\})$$

with Z any Borel set of \mathcal{Z} . Then:

$$\nu(dz, dl) = \sum_{k \in \mathcal{T}} \delta_k(dl) \nu^k(dz) \quad (4)$$

as for any set X of $\mathcal{B}(\mathcal{X})$, $X = \cup_{l \in \mathcal{T}} Z_l \times \{l\}$ with Z_l a Borel set of \mathcal{Z} .

For a measurable function $f : \mathcal{X} \rightarrow \mathbb{R}$ and a measure $\nu \in \mathcal{M}_F(\mathcal{X})$, we define

$$\langle \nu, f \rangle := \int_{\mathcal{X}} f(x) \nu(dx) = \sum_{l \in \mathcal{T}} \int_{\mathcal{Z}} f(z, l) \nu(dz, l). \quad (5)$$

If $\nu \in \mathcal{M}_P(\mathcal{X})$, that is to say $\nu = \sum_{i=1}^n \delta_{x^i}$ for some $n \in \mathbb{N}$,

$$\langle \nu, f \rangle = \sum_{i=1}^n f(x^i) = \sum_{i=1}^n f_l(z^i).$$

In particular,

$$\langle \nu, 1 \rangle = n.$$

2.4 Convolution

For a measurable function $D : \mathbb{R}^d \times \mathcal{T} \rightarrow \mathbb{R}$ and $\nu \in \mathcal{M}_F(\mathcal{X})$, we set for $x = (z, l) \in \mathcal{X}$

$$\nu * D(x) := \sum_{l \in \mathcal{T}} \int_{\mathcal{Z}} D(z - y, l) \nu(dy, l). \quad (6)$$

If $\nu \in \mathcal{M}_P(\mathcal{X})$,

$$\nu * D(x) = \sum_{i=1}^{\langle \nu, 1 \rangle} D(z - z^i, l^i).$$

3 Definition of the process

We now describe our generic individual based cell model, expressed as a piecewise-deterministic (or hybrid) Markov process.

3.1 Deterministic movement of cells and reaction-diffusion of chemicals

3.1.1 Movement of cells

For each cell i with mass $m_i \ll 1$, the position of its center $z^i(t)$ evolves continuously in the spatial domain \mathcal{Z} , in a non-inertial regime:

$$\underbrace{m_i \frac{d^2 z^i}{dt^2}(t)}_{\text{inertia}} + \underbrace{\frac{dz^i}{dt}(t)}_{\text{friction}} = \underbrace{\mathcal{F}(z^i(t), \nu_t * F(z^i(t)))}_{\text{other forces}} \quad (7)$$

with

$$F : \mathbb{R}^d \rightarrow \mathbb{R}^d \text{ and } \mathcal{F} : \mathcal{Z} \times \mathbb{R}^d \rightarrow \mathbb{R}^d.$$

The first variable $z^i(t)$ in \mathcal{F} is introduced to take into account local modulations (if any) in the total force (apart from friction). The second variable is the total force exerted by other cells. The cell j at position z^j exerts on cell i at position z^i a force $F(z^i - z^j)$. This force only depends on spatial distance and direction between cells, and not on cell types. The total force exerted on cell i is

$$\sum_{j=1, j \neq i}^n F(z^i - z^j) = \sum_{j=1}^n F(z^i - z^j) = \nu * F(z^i). \quad (8)$$

We assume that

Hypothesis 2. 1. $\mathcal{F} \in \mathcal{C}^1(\mathcal{Z} \times \mathbb{R}^d, \mathbb{R}^d)$, $F \in \mathcal{C}^1(\mathbb{R}^d, \mathbb{R}^d)$,

2. $F(0) = 0$,

3. $\mathcal{F}(z, \cdot) = 0 \quad \forall z \in \partial \mathcal{Z}$.

We have the following standard result.

Proposition 3 (Cauchy-Lipschitz). *Thanks to hypothesis (3) and regularity assumptions on F and \mathcal{F} , the system of coupled equations (7) has a unique solution on \mathbb{R}^+ , taking values in \mathcal{Z}^n , for any initial condition in \mathcal{Z}^n .*

Remark 4. Proposition 3 holds under the weaker hypothesis that \mathcal{F} and F are locally Lipschitz, the \mathcal{C}^1 regularity will be used later in section 5.

Remark 5. Assumption (2) deserves an explanation. It means that two cells can occupy the same location without exerting a force on each other. Such situation occurs when the space dimension is reduced due to symmetry assumptions, in which case cells are allowed to form vertical or spherical layers corresponding to the same altitude or radial distance, see for instance [Dar+22]. More generally, we assume that when they get very close, cells exert a repulsive force on each other. However, when the distance between cells is too small, this correspond to a high compression and cells are damaged, so that they cannot exert a force on each other any more. This usually comes with a high probability of extrusion (expulsion of damaged cells), as reminded in the description of cell fate events below.

Remark 6. We could take into account an effect of cell types in mechanical interaction by using for the force $F(z^i - z^j, l^i, l^j)$ with $F : \mathbb{R}^d \times \mathcal{T} \times \mathcal{T} \rightarrow \mathbb{R}^d$, and l^i, l^j are respectively the cell types of cell i and cell j . By extending hypotheses done on F to $F(\cdot, l^i, l^j)$, all of the following results are still valid.

3.1.2 Reaction-diffusion of chemicals

Let a be the typical diameter of a cell and $B_{\mathbb{R}^d}(z, r)$ be the open ball of \mathbb{R}^d centered on z of radius r . We define

$$\mathcal{Z}_a := \bigcup_{z \in \mathcal{Z}} B_{\mathbb{R}^d} \left(z, \frac{1}{2}a \right)$$

the spatial domain we use for modelling concentrations of chemicals. It is an enlarged version of \mathcal{Z} which contains not only the position of the center of cells, as \mathcal{Z} , but also the whole cells area (or volume), which is appropriate to model surfacic absorption of chemicals. We assume the following:

Hypothesis 7. \mathcal{Z}_a is an open set of \mathbb{R}^d of regularity C^∞ , with boundary Γ .

Let $c(z)$ be the vector of concentrations of the N_m chemicals located at z . Each cell type $l \in \mathcal{T}$ can degrade or produce chemical m , $1 \leq m \leq N_m$, at speed $\gamma_{ml}(c(z))$. A cell i located at z^i can catalyze a reaction only if it is physically close to z . Therefore, we weight the reaction speed $\gamma_{ml}(c(z))$ with a term $\psi_a(z - z^i)$ where

Hypothesis 8. ψ_a is a positive function with radial symmetry, Lipschitz on \mathbb{R}^d with support contained in $B_{\mathbb{R}^d}(0, \frac{1}{2}a)$ and its integral is 1.

On the boundary of \mathcal{Z}_a , concentrations are influenced by the external media through quantities f_m which depend of time and space. Then, concentrations c on \mathcal{Z}_a evolve according to the following system of coupled reaction-diffusion equations

$$\begin{aligned} \forall 1 \leq m \leq N_m, \quad \partial_t c_m - \sigma \Delta c_m &= \sum_{l \in \mathcal{T}} \gamma_{ml}(c) \nu_t^l * \psi_a(z), \\ \partial_n c_m|_{[0, T] \times \Gamma} + b_m c_m|_{[0, T] \times \Gamma} &= f_m, \\ c_m(\cdot, 0) &= c_{0m}, \end{aligned} \tag{9}$$

where ∂_n is the normal derivative on Γ .

For an open set $U \subset \mathbb{R}^d$ and $\alpha \geq 0$, let

$$W_{L^2(U)}^{2\alpha, [\alpha]} := \left\{ f \in L^2([0, T], H^{2\alpha}(U)); \forall j \in \{1, \dots, [\alpha]\} \frac{\partial^j f}{\partial t^j} \in L^2([0, T], H^{2\alpha-2j}(U)) \right\}$$

with $[\cdot]$ the ceiling function.

We assume that

Hypothesis 9. For all $1 \leq m \leq N_m$,

1. b_m is in $C^\infty(\overline{\mathcal{Z}_a})$,
2. $\forall l \in \mathcal{T}$, γ_{ml} is bounded and uniformly Lipschitz from $\mathbb{R}_+^{N_m}$ to \mathbb{R} ,
3. the initial condition c_{0m} is in $H^1(\mathcal{Z}_a)$,
4. f_m is in $W_{L^2(\Gamma)}^{\frac{1}{2}, 1}$.

With the notations introduced in section B, we denote S_m the analytic semi-group S on $L^2(\mathcal{Z}_a)$ generated by $\sigma \Delta$ with boundary conditions of chemical m , D_m the related lifting operator (D in the appendix) and λ_m the associated scalar (λ in the appendix). Moreover, for $\nu \in \mathcal{M}_F(\mathcal{X})$, let $\|\nu\|_{TV}$ be the total variation norm defined in equation (63) of the appendix. Then the following proposition is verified.

Proposition 10. Under hypothesis 9, for all ν measurable on $[0, T]$ with values in $\mathcal{M}_{F^+}(\mathcal{X})$ endowed with the weak convergence topology and the associated σ -algebra, and such that $\sup_{t \in [0, T]} \|\nu_t\|_{TV} < +\infty$, the non-linear evolution problem (9) has a unique solution c in $\left(W_{L^2(\mathcal{Z}_a)}^{2, 1} \right)^{N_m} \cap \mathcal{C}([0, T], H^1(\mathcal{Z}_a)^{N_m})$, verifying for $1 \leq m \leq N_m$:

$$\begin{aligned} c_m(\cdot, t) &= S_m(t)c_{0m} + (\lambda_m I - \Delta) \left(\int_0^t S_m(t-s) D_m(f_m)(\cdot, s) ds \right) \\ &\quad + \int_0^t S_m(t-s) \sum_{l \in \mathcal{T}} \gamma_{ml}(c(s)) \nu_s^l * \psi_a(\cdot) ds \end{aligned} \tag{10}$$

Proof. Let us define the function G on $[0, T] \times L^2(\mathcal{Z}_a)^{N_m}$ as

$$G(t, c) := \left(\sum_{l \in \mathcal{T}} \gamma_{1l}(c) \nu_t^l * \psi_a(\cdot), \dots, \sum_{l \in \mathcal{T}} \gamma_{N_m l}(c) \nu_t^l * \psi_a(\cdot) \right) \in L^\infty([0, T], L^2(\mathcal{Z}_a)^{N_m}).$$

G is Lipschitz with respect to c on $L^2(\mathcal{Z}_a)^{N_m}$, uniformly for $t \in [0, T]$. Indeed, as γ_{ml} are Lipschitz under hypothesis 2,

$$\|\gamma\|_{Lip} = \sup_{m,l} \|\gamma_{ml}\|_{Lip} < \infty. \quad (11)$$

Let c_1, c_2 in $L^2(\mathcal{Z}_a)^{N_m}$, then

$$\begin{aligned} \sup_{t \in [0, T]} \|G(t, c_1) - G(t, c_2)\|_{L^2(\mathcal{Z}_a)^{N_m}} &\leq \sum_{m=1}^{N_m} \sum_{l \in \mathcal{T}} \sup_{t \in [0, T]} \|\nu_t^l\|_{TV} \|\gamma_{ml}(c_1) - \gamma_{ml}(c_2)\|_{L^2(\mathcal{Z}_a)} \\ &\leq \sup_{t \in [0, T]} \|\nu_t\|_{TV} \times N_m |\mathcal{T}| K \|c_1 - c_2\|_{L^2(\mathcal{Z}_a)^{N_m}}. \end{aligned}$$

With a standard fixed point argument (see [CZ12; Ama93]) on the system (10) and the semi-group of $L^2(\mathcal{Z}_a)^{N_m}$ defined by $S = \text{diag}(S_1, \dots, S_{N_m})$, we prove the existence of a weak solution \tilde{c} in $\mathcal{C}([0, T], L^2(\mathcal{Z}_a)^{N_m})$.

Then

$$u(z, t) = \sum_{l \in \mathcal{T}} \gamma_{ml}(\tilde{c}(z, t)) \nu_t^l * \psi_a(z) \in L^\infty([0, T], L^2(\mathcal{Z}_a)) \subset W_{L^2(\mathcal{Z}_a)}^{0,0}.$$

Indeed, γ_{ml} are bounded lipschitz on $\mathbb{R}_+^{N_m}$ under hypothesis 2, so $\gamma_{ml}(\tilde{c}(z, t)) \in \mathcal{C}([0, T], L^2(\mathcal{Z}_a))$ as $\tilde{c} \in \mathcal{C}([0, T], L^2(\mathcal{Z}_a)^{N_m})$ according to the previous step. Moreover, $\nu_t^l * \psi_a(z) \in L^\infty([0, T], L^2(\mathcal{Z}_a))$ because ψ_a is bounded and $\sup_{t \in [0, T]} \|\nu_t\|_{TV} < \infty$.

We obtain the adequate regularity result by observing that c_m is a solution of a problem of type (73) with $c_{0m} \in H^1(\mathcal{Z}_a)$, f_m verifying the regularity conditions stated in proposition 35 and $u \in W_{L^2(\mathcal{Z}_a)}^{0,0}$. Therefore $c_m \in W_{L^2(\mathcal{Z}_a)}^{2,1} \cap \mathcal{C}([0, T], H^1(\mathcal{Z}_a))$. \square

The following lemma will be used later. It bounds the distance between two populations of cells ν^1 and ν^2 , by a quantity depending on the distance between the two corresponding solutions of (9).

Lemma 11. *Let ν^1, ν^2 be two measurable functions from $[0, T]$ to $\mathcal{M}_{F^+}(\mathcal{X})$ (endowed with the topology of weak convergence and the associated Borel σ -algebra) such that*

$$\sup_{t \in [0, T]} \|\nu_t^1\|_{TV} \leq B < +\infty \text{ and } \sup_{t \in [0, T]} \|\nu_t^2\|_{TV} \leq B < +\infty.$$

Let c^1 et c^2 be the associated solutions of (9), with c_0 and f verifying the regularity hypothesis 9. Then, $\forall t \in [0, T]$,

$$\|c^1(\cdot, t) - c^2(\cdot, t)\|_{L^2(\mathcal{Z}_a)^{N_m}} \leq N_m e^{Ct} \gamma^\infty \int_0^t \sum_{l \in \mathcal{T}} \|\nu_s^{1l} * \psi_a(\cdot) - \nu_s^{2l} * \psi_a(\cdot)\|_{L^2(\mathcal{Z}_a)} ds, \quad (12)$$

where $C = N_m \|\psi_a\|_\infty \|\gamma\|_{Lip} B$ and γ^∞ are positive constants. We can deduce that

$$\|c^1(\cdot, t) - c^2(\cdot, t)\|_{L^2(\mathcal{Z}_a)^{N_m}} \leq N_m \gamma^\infty \|\psi_a\|_\infty |\mathcal{Z}_a|^{1/2} |\mathcal{T}| e^{Ct} \int_0^t d_F(\nu_s^1, \nu_s^2) ds, \quad (13)$$

where d_F is the flat metric (see A) defined on $\mathcal{M}_{F^+}(\mathcal{X})$ by

$$d_F(\nu, \xi) = \sup_{f \in \mathcal{L}ip_{1,1}(\mathcal{X}, \mathbb{R})} \langle \nu - \xi, f \rangle.$$

Proof. By formula (10),

$$\begin{aligned}
\|c_m^1(\cdot, t) - c_m^2(\cdot, t)\|_{L^2(\mathcal{Z}_a)} &= \left\| \int_0^t S_m(t-s) \sum_{l \in \mathcal{T}} \left[\gamma_{ml}(c^1(\cdot, s)) \nu_s^{1l} * \psi_a(\cdot) \right. \right. \\
&\quad \left. \left. - \gamma_{ml}(c^2(\cdot, s)) \nu_s^{2l} * \psi_a(\cdot) \right] ds \right\|_{L^2(\mathcal{Z}_a)} \\
&= \left\| \int_0^t S_m(t-s) \sum_{l \in \mathcal{T}} \left[\gamma_{ml}(c^1(\cdot, s)) (\nu_s^{1l} * \psi_a(\cdot) - \nu_s^{2l} * \psi_a(\cdot)) \right. \right. \\
&\quad \left. \left. - (\gamma_{ml}(c^2(\cdot, s)) - \gamma_{ml}(c^1(\cdot, s))) \nu_s^{2l} * \psi_a(\cdot) \right] ds \right\|_{L^2(\mathcal{Z}_a)} \\
&\leq \int_0^t \sum_{l \in \mathcal{T}} \left[\left\| \gamma_{ml}(c^1(\cdot, s)) (\nu_s^{1l} * \psi_a(\cdot) - \nu_s^{2l} * \psi_a(\cdot)) \right\|_{L^2(\mathcal{Z}_a)} \right. \\
&\quad \left. + \left\| (\gamma_{ml}(c^2(\cdot, s)) - \gamma_{ml}(c^1(\cdot, s))) \nu_s^{2l} * \psi_a(\cdot) \right\|_{L^2(\mathcal{Z}_a)} \right] ds
\end{aligned}$$

because S_m is a contraction semi-group (see B). Moreover, there exists a constant γ^∞ such that $\forall m, l$, $|\gamma_{ml}(\cdot)| \leq \gamma^\infty < \infty$. Therefore,

$$\begin{aligned}
\|c_m^1(\cdot, t) - c_m^2(\cdot, t)\|_{L^2(\mathcal{Z}_a)} &\leq \gamma^\infty \int_0^t \sum_{l \in \mathcal{T}} \left\| \nu_s^{1l} * \psi_a(\cdot) - \nu_s^{2l} * \psi_a(\cdot) \right\|_{L^2(\mathcal{Z}_a)} ds \\
&\quad + \|\psi_a\|_\infty \|\gamma\|_{Lip} \sup_{[0, T]} \|\nu_s^2\|_{TV} \int_0^t \|c^1(\cdot, s) - c^2(\cdot, s)\|_{L^2(\mathcal{Z}_a)^{N_m}} ds.
\end{aligned}$$

We set $C = N_m \|\psi_a\|_\infty \|\gamma\|_{Lip} B$ and deduce from the previous inequality that

$$\begin{aligned}
\|c^1(\cdot, t) - c^2(\cdot, t)\|_{L^2(\mathcal{Z}_a)^{N_m}} &\leq N_m \gamma^\infty \int_0^t \sum_{l \in \mathcal{T}} \left\| \nu_s^{1l} * \psi_a(\cdot) - \nu_s^{2l} * \psi_a(\cdot) \right\|_{L^2(\mathcal{Z}_a)} ds \\
&\quad + C \int_0^t \|c^1(\cdot, s) - c^2(\cdot, s)\|_{L^2(\mathcal{Z}_a)^{N_m}} ds
\end{aligned}$$

By Gronwall lemma, $\forall t \in [0, T]$,

$$\|c^1(\cdot, t) - c^2(\cdot, t)\|_{L^2(\mathcal{Z}_a)^{N_m}} \leq N_m e^{Ct} \gamma^\infty \int_0^t \sum_{l \in \mathcal{T}} \left\| \nu_s^{1l} * \psi_a(\cdot) - \nu_s^{2l} * \psi_a(\cdot) \right\|_{L^2(\mathcal{Z}_a)} ds, \quad (14)$$

which is the first result of the lemma. We obtain the second as for all $0 \leq s \leq t$,

$$\left\| \nu_s^{1l} * \psi_a(\cdot) - \nu_s^{2l} * \psi_a(\cdot) \right\|_{L^2(\mathcal{Z}_a)} \leq \|\psi_a\|_\infty |\mathcal{Z}_a|^{1/2} d_F(\nu_s^1, \nu_s^2)$$

where $|\mathcal{Z}_a|$ is the Lebesgue measure of \mathcal{Z}_a . □

3.2 Stochastic model of cell fate

Cells can divide symmetrically, differentiate or die. These phenomena are modeled as independent random events, indexed by \mathcal{E} the finite set of event types. When at t the cell i undergoes a jump of type k , the population is modified according to

$$\nu_t = \nu_{t^-} + \mu_k(x^i(t^-))$$

where

$$\mu_k : \mathcal{X} \rightarrow \mathcal{M}_{F^+}(\mathcal{X}), \quad (15)$$

and

- if k is an *extrusion*: $\mu_k(x) = -\delta_x$,
- if k is a *differentiation* of a cell of type l into a type l^* : $\mu_k(z, l) = -\delta_{(z, l)} + \delta_{(z, l^*)}$,

- if k is a *division*: we introduce a new daughter cell at a distance $\lambda(z)$ from mother cell that is to say $\mu_k(z, l) = \delta_{(z+\lambda(z), l)}$. We assume the following

Hypothesis 12. – $\lambda : \mathcal{Z} \rightarrow \mathbb{R}^d$ is continuous and such that, $\forall z \in \mathcal{Z}, z + \lambda(z) \in \mathcal{Z}$,
– λ is locally Lipschitz from \mathcal{Z} to \mathbb{R}^d .

Remark 13. We conjecture that our results could be extended to a stochastic division shift λ , as done in [FM04; Tra06; CF14].

For a cell at $x \in \mathcal{X}$, in a population $\nu \in \mathcal{M}_{F^+}(\mathcal{X})$, the individual jump rate associated to jump $k \in \mathcal{E}$ takes the form

$$q_k^\infty \times q_k(x, \nu * D(z), c * \psi_a(z)) \quad (16)$$

with

$$D : \mathbb{R}^d \rightarrow \mathbb{R}^+, \quad q_k : \mathcal{X} \times \mathbb{R} \times \mathbb{R}^+ \rightarrow \mathbb{R}^+,$$

ψ_a the function representing the cell volume defined in hypothesis 8, and $q_k^\infty < +\infty$ are constants. Moreover,

$$c * \psi_a(z) = \left(\int_{\mathcal{Z}_a} c_1(y) \psi_a(z - y) dy, \dots, \int_{\mathcal{Z}_a} c_{N_m}(y) \psi_a(z - y) dy \right)$$

is the vector of each coordinate of c convoluted with ψ_a . Therefore, we model a cell population where jump rates depend of the cell state x , the local density of the population through the term $\nu * D(z)$, and concentrations of chemicals near the cell through the term $c * \psi_a(z)$. For instance, as mentioned earlier, a high local density could promote extrusion events as observed experimentally [Eis+12].

We assume the following

Hypothesis 14. • For all $k \in \mathcal{E}, q_k \leq 1$,

- For all $k \in \mathcal{E}$, for all $l \in \mathcal{T}, (z, y, c) \mapsto q_k((z, l), y, c)$ is Lipschitz on $\mathcal{Z} \times \mathbb{R} \times \mathbb{R}$. We denote $K_{k,l}$ be a Lipschitz constant and set $C(q_k) = \max_{l \in \mathcal{T}} K_{k,l}$.
- $D \in \mathcal{C}^1(\mathbb{R}^d, \mathbb{R}^+)$

4 Rescaled process

We introduce a sequence of processes $(c_t^N, \nu_t^N)_{N \in \mathbb{N}^*}$ similar to the model described in sections 3.1-3.2 with parameters $\mathcal{F}^N, D^N, \gamma^N$ such that

$$F^N = \frac{F}{N}, \quad D^N = \frac{D}{N}, \quad \gamma^N = \frac{\gamma}{N}, \quad (17)$$

leaving the other model parameters unchanged. Note that $N = 1$ corresponds to the process described in sections 3.1-3.2. The rescaled process is defined as

$$\bar{\nu}_t^N := \frac{1}{N} \nu_t^N.$$

Remark 15. Obviously, the global existence and uniqueness properties of cell spatial trajectories and concentrations of chemicals as stated in propositions 3 and 10 readily extended to $(c_t^N, \nu_t^N)_{N \in \mathbb{N}^*}$, for all $N \in \mathbb{N}^*$.

We consider independent Poisson random measure on $[0, 1] \times \mathbb{N}^* \times \mathbb{R}^+$, one for each type of jump $k \in \mathcal{E}$, denoted $N_k(d\theta, di, dt)$, with intensity

$$n_k(d\theta, di, ds) = q_k^\infty d\theta dt \sum_{k \in \mathbb{N}^*} \delta_k(di).$$

From hypothesis 17, for all $z \in \mathcal{Z}$,

$$\nu_t^N * F^N(z) = \nu_t^N * \frac{F(z)}{N} = \bar{\nu}_t^N * F(z),$$

$$\nu_t^N * D^N(z) = \bar{\nu}_t^N * D(z),$$

and, for all $(m, l) \in \{1, \dots, N_m\} \times \mathcal{T}$,

$$\gamma_{ml}(c^N) \nu_t^{lN} * \psi_a(z) = \gamma_{ml}(c^N) \bar{\nu}_t^{lN} * \psi_a(z).$$

Provided trajectories exist on \mathbb{R}^+ (which will be proved in proposition 17 below), the process $(c_t^N, \bar{\nu}_t^N)_t \in \mathcal{C}([0, T], H^1(\mathcal{Z}_a)^{N_m}) \times \mathcal{D}([0, T], \mathcal{M}_{F^+}(\mathcal{X}))$ is solution of, $\forall f \in \mathcal{C}^{1,1,\bullet}(\mathbb{R}^+ \times \mathcal{X}, \mathbb{R})$,

$$\begin{aligned} \langle \bar{\nu}_t^N, f_t \rangle &= \langle \bar{\nu}_0^N, f_0 \rangle \\ &+ \int_0^t \langle \bar{\nu}_s^N, \mathcal{F}(\cdot, \bar{\nu}_s^N * F) \nabla_z f_s + \partial_s f_s \rangle ds \\ &+ \sum_{k \in \mathcal{E}} \int_0^t \int_{\mathbb{N}^*} \int_0^1 \left[\langle \bar{\nu}_{s^-}^N + \frac{\mu_k}{N}(x_{s^-}^{Ni}), f_s \rangle - \langle \bar{\nu}_{s^-}^N, f_s \rangle \right] \\ &\quad \times \mathbb{1}_{\{i \leq \langle \nu_{s^-}, 1 \rangle\}}(i) \mathbb{1}_{\{\theta \leq q_k(x_{s^-}^{Ni}, \bar{\nu}_{s^-}^N * D(z_{s^-}^{Ni}), c_s^N * \psi_a(z_{s^-}^{Ni}))\}}(\theta) N_k(d\theta, di, ds), \end{aligned} \quad (18)$$

and, for chemical $m, 1 \leq m \leq N_m$,

$$\begin{aligned} \partial_t c_m^N - \sigma \Delta c_m^N &= \sum_{l \in \mathcal{T}} \gamma_{ml}(c^N) \bar{\nu}_t^{lN} * \psi_a(z), \\ \partial_n c_{|[0, T] \times \Gamma}^N + b_m c_{|[0, T] \times \Gamma}^N &= f_m, \\ c_m^N(\cdot, 0) &= c_{0m}. \end{aligned} \quad (19)$$

Remark 16. The motivation for employing test functions f_t depending on both space and time comes from our proof of the convergence theorem 21, more specifically in lemma 26, second step.

Without stochastic jumps (or between jumps), the deterministic flow of the process is defined on \mathbb{R}^+ for all $N \in \mathbb{N}^*$ (see remark 15). For $(c_t^N, \bar{\nu}_t^N)$ to have trajectories on \mathbb{R}^+ when stochastic events occur, the process needs to be stable meaning that an infinite number of jumps cannot accumulate in finite time. This is shown in the next proposition.

Proposition 17. 1. (Control of the population size) *Let $t \geq 0$ and $p \in \mathbb{N}^*$. If $\sup_{N \in \mathbb{N}^*} \mathbb{E} [\langle \bar{\nu}_0^N, 1 \rangle^p] < +\infty$, then*

$$\sup_{N \in \mathbb{N}^*} \mathbb{E} \left[\sup_{s \leq t} \langle \bar{\nu}_s^N, 1 \rangle^p \right] \leq C(p, t) < +\infty \quad (20)$$

with $C(p, t)$ a constant that depends of p and t .

2. (Stability) *For all $N \in \mathbb{N}^*$, the process $(c_t^N, \bar{\nu}_t^N)$ solution of (18-19) is stable and trajectories exist on \mathbb{R}^+ .*

Proof. (1) : Let $p, N \in \mathbb{N}^*$. For all $n \in \mathbb{N}$, we define the stopping-time $\tau_n^N = \inf\{t \geq 0, \langle \bar{\nu}_t^N, 1 \rangle = n\}$. The process $(c_{t \wedge \tau_n^N}^N, \bar{\nu}_{t \wedge \tau_n^N}^N)_t$ is well-defined as the cell population is bounded by n and, by property 15, c^N exists.

Note that the determinist evolution does not change population size, and jumps modify it at most by one. Let $t \geq 0$, applying Itô formula (theorem 5.1 chapter II in [IW89]) to $\langle \bar{\nu}_{t \wedge \tau_n^N}^N, 1 \rangle^p$, we get

$$\begin{aligned} \langle \bar{\nu}_{t \wedge \tau_n^N}^N, 1 \rangle^p &\leq \langle \bar{\nu}_0^N, 1 \rangle^p \\ &+ \sum_{k \in \mathcal{E}} \int_0^{t \wedge \tau_n^N} \int_{\mathbb{N}^*} \int_0^1 \left[\left(\langle \bar{\nu}_{s^-}^N, 1 \rangle + \frac{1}{N} \right)^p - \langle \bar{\nu}_{s^-}^N, 1 \rangle^p \right] \mathbb{1}_{\{i \leq \langle \nu_{s^-}, 1 \rangle\}}(i) \\ &\quad \mathbb{1}_{\{\theta \leq q_k(\cdot)\}}(\theta) N_k(d\theta, di, ds). \end{aligned}$$

The integral is increasing with time so we can switch to $\sup_{s \leq t \wedge \tau_n^N}$ on the left hand-side. For positive x ,

$$\left(x + \frac{1}{N}\right)^p - x^p \leq \frac{C(p)}{N^p} ((Nx)^{p-1} + 1) \leq \frac{C(p)}{N} (x^{p-1} + 1)$$

with $C(p) = \sum_{k=1}^{p-1} \binom{p}{k}$. Therefore,

$$\sup_{s \leq t \wedge \tau_n} \langle \bar{\nu}_s^N, 1 \rangle^p \leq \langle \bar{\nu}_0^N, 1 \rangle^p + \sum_{k \in \mathcal{E}} \int_0^{t \wedge \tau_n} \int_{\mathbb{N}^*} \int_0^1 \frac{C(p)}{N} (\langle \bar{\nu}_{s^-}^N, 1 \rangle^{p-1} + 1) \mathbf{1}_{\{i \leq \langle \nu_{s^-}^N, 1 \rangle\}}(i) \mathbf{1}_{\{\theta \leq q_k(\cdot)\}}(\theta) N_k(d\theta, di, ds).$$

Taking expectation:

$$\begin{aligned} \mathbb{E} \left[\sup_{s \leq t \wedge \tau_n} \langle \bar{\nu}_s^N, 1 \rangle^p \right] &\leq \mathbb{E} [\langle \bar{\nu}_0^N, 1 \rangle^p] \\ &\quad + \sum_{k \in \mathcal{E}} \mathbb{E} \left[\int_0^{t \wedge \tau_n} \sum_{i=1}^{\langle \nu_{s^-}^N, 1 \rangle} \frac{C(p)}{N} (\langle \bar{\nu}_{s^-}^N, 1 \rangle^{p-1} + 1) q_k^\infty q_k(\cdot) ds \right] \\ &\leq \mathbb{E} [\langle \bar{\nu}_0^N, 1 \rangle^p] + \sum_{k \in \mathcal{E}} \mathbb{E} \left[\int_0^t q_k^\infty C(p) (\langle \bar{\nu}_{s \wedge \tau_n}^N, 1 \rangle^p + \langle \bar{\nu}_{s \wedge \tau_n}^N, 1 \rangle) ds \right] \\ &\leq \mathbb{E} [\langle \bar{\nu}_0^N, 1 \rangle^p] + 2C(p) \left(\sum_{k \in \mathcal{E}} q_k^\infty \right) \int_0^t \mathbb{E} \left[\sup_{u \leq s \wedge \tau_n} \langle \bar{\nu}_u^N, 1 \rangle^p \right] ds. \end{aligned}$$

as for $p \geq 1$, $n + n^p \leq 2n^p$. Using Gronwall lemma and our hypotheses, we can find a bound which does not depend of n nor N :

$$\mathbb{E} \left[\sup_{s \leq t \wedge \tau_n} \langle \bar{\nu}_s^N, 1 \rangle^p \right] \leq \mathbb{E} [\langle \bar{\nu}_0^N, 1 \rangle^p] \exp \left(2C(p) \left(\sum_{k \in \mathcal{E}} q_k^\infty \right) t \right) = C(p, t). \quad (21)$$

Therefore $\tau_n \rightarrow \infty$ a.s. Then, by Fatou lemma:

$$\begin{aligned} \mathbb{E} \left[\sup_{s \leq t} \langle \bar{\nu}_s^N, 1 \rangle^p \right] &= \mathbb{E} \left[\liminf_{n \rightarrow \infty} \sup_{s \leq t \wedge \tau_n} \langle \bar{\nu}_s^N, 1 \rangle^p \right] \\ &\leq \liminf_{n \rightarrow \infty} \mathbb{E} \left[\sup_{s \leq t \wedge \tau_n} \langle \bar{\nu}_s^N, 1 \rangle^p \right] \leq C(p, t). \end{aligned}$$

(2) : A direct consequence of (1) is that the population does not explode in finite time a.s. Individual jump rates being bounded, as long as the population is finite, the total jump rate is finite and only a finite number of jumps occurs on finite time interval. \square

We now establish a martingale property which will be used in our convergence study. To do so, we introduce for each jump type $k \in \mathcal{E}$ the compensated random Poisson measure:

$$\tilde{N}_k(d\theta, di, ds) = N_k(d\theta, di, ds) - n_k(d\theta, di, ds).$$

Proposition 18. *We suppose that $\mathbb{E} [\langle \bar{\nu}_0^N, 1 \rangle^2] < \infty$. For all $f \in \mathcal{C}^{1,1,\bullet}(\mathbb{R}^+ \times \mathcal{X}, \mathbb{R})$ and $t \geq 0$,*

$$\begin{aligned} \langle \bar{\nu}_t^N, f_t \rangle &= \langle \bar{\nu}_0^N, f_0 \rangle + \int_0^t \langle \bar{\nu}_s^N, \mathcal{F}(\cdot, \bar{\nu}_s^N * F) \nabla_z f_s + \partial_s f_s \rangle ds \\ &\quad + \sum_{k \in \mathcal{E}} q_k^\infty \int_0^t \int_{\mathcal{X}} \langle \mu_k(x), f_s \rangle q_k(x, \bar{\nu}_s^N * D(z), c_s^N * \psi_\alpha(z)) \bar{\nu}_s^N(dx) ds \\ &\quad + Z_t^{f,N} \end{aligned} \quad (22)$$

where $Z_t^{f,N}$ is the square integrable martingale

$$Z_t^{f,N} = \sum_{k \in \mathcal{E}} \int_0^t \int_{\mathbb{N}^*} \int_0^1 \langle \mu_k(x_s^{Ni}), \frac{1}{N} f_s \rangle \mathbf{1}_{\{i \leq \langle \nu_{s^-}^N, 1 \rangle\}}(i) \mathbf{1}_{\{\theta \leq q_k(\cdot)\}}(\theta) \tilde{N}_k(d\theta, di, ds), \quad (23)$$

with quadratic variation

$$\langle Z_t^{f,N} \rangle = \frac{1}{N} \sum_{k \in \mathcal{E}} q_k^\infty \int_0^t \int_{\mathcal{X}} \langle \mu_k(x), f_s \rangle^2 q_k(x, \bar{\nu}_s^N * D(z), c_s^N * \psi_\alpha(z)) \bar{\nu}_s^N(dx) ds. \quad (24)$$

Proof. The integrands in (23) are left-continuous and therefore predictable. From proposition 17, for all $t \geq 0$,

$$\begin{aligned} & \mathbb{E} \left[\sum_{k \in \mathcal{E}} \int_0^t \int_{\mathbb{N}^*} \int_0^1 \left| \langle \mu_k(x_s^{Ni}), \frac{1}{N} f_s \rangle \right| \mathbb{1}_{\{i \leq \langle \nu_s^N, 1 \rangle\}}(i) \mathbb{1}_{\{\theta \leq q_k(\cdot)\}}(\theta) n_k(d\theta, di, ds) \right] \\ &= \sum_{k \in \mathcal{E}} \int_0^t \int_{\mathcal{X}} \mathbb{E} \left[q_k^\infty \left| \langle \mu_k(x), \frac{1}{N} f_s \rangle \right| q_k(x, \bar{\nu}_s^N * D(z), c_s^N * \psi_a(z)) \bar{\nu}_s^N(dx) ds \right] \\ &< +\infty \end{aligned}$$

and

$$\begin{aligned} & \mathbb{E} \left[\sum_{k \in \mathcal{E}} \int_0^t \int_{\mathbb{N}^*} \int_0^1 \left| \langle \mu_k(x_s^{Ni}), \frac{1}{N} f_s \rangle \right|^2 \mathbb{1}_{\{i \leq \langle \nu_s^N, 1 \rangle\}}(i) \mathbb{1}_{\{\theta \leq q_k(\cdot)\}}(\theta) n_k(d\theta, di, ds) \right] \\ &< +\infty. \end{aligned}$$

We can then apply the classical result given p.62 of [IW89]. \square

Remark 19. Note that results presented in this section are valid when the rate functions q_k are measurable. The stronger hypothesis 14 will prove necessary in section 5

5 Deterministic limit of the rescaled process by a large population approximation.

Let T be a positive time. Let $\mathcal{D}([0, T], \mathcal{M}_{F^+}(\mathcal{X}))$ be the space of càdlàg functions of $[0, T]$ to $\mathcal{M}_{F^+}(\mathcal{X})$. We now prove the weak convergence of the rescaled process sequence $(c_t^N, \bar{\nu}_t^N)_{N \in \mathbb{N}^*}$ defined in section 4 towards a deterministic process in the space

$$\mathcal{C}([0, T], H^1(\mathcal{Z}_a)^{N_m}) \times \mathcal{D}([0, T], \mathcal{M}_{F^+}(\mathcal{X})).$$

We endow $\mathcal{D}([0, T], \mathcal{M}_{F^+}(\mathcal{X}))$ with the Skorokhod metric, using the Prokhorov distance d_{PR} on $\mathcal{M}_{F^+}(\mathcal{X})$. The topology associated to the Prokhorov metric is the topology of the weak convergence on $\mathcal{M}_{F^+}(\mathcal{X})$. For more details on these points, see A.

We suppose that hypotheses 2, 7, 8, 9, 12, 14 are verified. Additionally, we assume that :

Hypothesis 20. • *the sequence $\bar{\nu}_0^N = \frac{1}{N} \nu_0^N$, seen as a sequence of random variables on $\mathcal{M}_{F^+}(\mathcal{X})$, converges weakly to $\xi_0 \in \mathcal{M}_{F^+}(\mathcal{X})$ where ξ_0 can be random.*

- *There exists an integer $p \geq 2$ such that $\sup_{N \in \mathbb{N}^*} \mathbb{E} [\langle \bar{\nu}_0^N, 1 \rangle^p] < +\infty$.*

We can now state the main result of the paper.

Theorem 21. *Under all the hypotheses mentioned above, the process $(c_t^N, \bar{\nu}_t^N)_{t \geq 0}$ converges weakly on $\mathcal{C}([0, T], H^1(\mathcal{Z}_a)^{N_m}) \times \mathcal{D}([0, T], \mathcal{M}_{F^+}(\mathcal{X}))$ to the solution of*

$$\begin{aligned} \langle \xi_t, f_t \rangle &= \langle \xi_0, f_0 \rangle + \int_0^t \langle \xi_s, \mathcal{F}(\cdot, \xi_s * F) \nabla_z f_s + \partial_t f_s \rangle ds \\ &\quad + \sum_{k \in \mathcal{E}} \int_0^t \sum_{l \in \mathcal{T}} \int_{\mathcal{Z}} \langle \mu_k(z, l), f_s \rangle q_k^\infty q_k((z, l), \xi_s * D(z), c_s * \psi_a(z)) \xi_s(dz, l) ds \end{aligned} \quad (25)$$

for all $f \in \mathcal{C}^{1,1,\bullet}(\mathbb{R}^+ \times \mathcal{X}, \mathbb{R})$, and $\forall 1 \leq m \leq N_m$,

$$\begin{aligned} \partial_t c_m - \sigma \Delta c_m &= \sum_{l \in \mathcal{T}} \gamma_{ml}(c) \xi_t^l * \psi_a(z), \\ \partial_n c_{m,|[0,T] \times \Gamma} + b_m c_{m,|[0,T] \times \Gamma} &= f_m, \\ c_m(\cdot, 0) &= c_{0m}, \end{aligned} \quad (26)$$

The proof of theorem 21 is established through three successive steps:

- the sequence of laws of $(\bar{\nu}^N)_{N \in \mathbb{N}^*}$, considered as probability measures on $\mathcal{D}([0, T], \mathcal{M}_{F^+}(\mathcal{X}))$, is tight (lemma 22).
- All accumulation points of $(\bar{\nu}^N)_{N \in \mathbb{N}^*}$ are almost surely solutions of (25), and the corresponding accumulation points in the concentration space are solutions of (26) (lemma 25).
- Uniqueness of the solution of (25-26) in lemma 26.

Lemma 22. *The sequence of laws of $(\bar{\nu}^N)_{N \in \mathbb{N}^*}$, considered as probability measures on $\mathcal{D}([0, T], \mathcal{M}_{F^+}(\mathcal{X}))$, is tight.*

Proof. Based on a theorem from Roelly (Thm. 2.1 in [Cop86]), the tightness of $(\bar{\nu}^N)_{N \in \mathbb{N}^*}$ is equivalent to the tightness of laws of $(\langle \bar{\nu}^N, f \rangle)_{N \in \mathbb{N}^*}$ on $\mathcal{D}([0, T], \mathbb{R})$, for all f in a dense set of $\mathcal{C}(\mathcal{X}, \mathbb{R})$ endowed with the topology of uniform convergence. Here, we choose $f \in \mathcal{C}^{1, \bullet}(\mathcal{X}, \mathbb{R})$.

A sufficient condition for this property is given by the Aldous-Rebolledo criteria (Cor. 2.3.3 in [JM86]) according to which we have to establish that

- 1) for $t \geq 0$, the sequence $(\langle \bar{\nu}_t^N, f \rangle)_{N \in \mathbb{N}^*}$ is tight on \mathbb{R} ,
- 2) considering the semi-martingale decomposition defined at Prop. 18

$$\langle \bar{\nu}_t^N, f \rangle = \langle \bar{\nu}_0^N, f \rangle + A_t^{f, N} + Z_t^{f, N}$$

with $A_t^{f, N}$ a finite variation process and $Z_t^{f, N}$ a martingale, then for $t \geq 0$, $\epsilon > 0$, $\eta > 0$, there exists $\delta > 0$ and $N_0 \in \mathbb{N}^*$ such that for all sequence of stopping times $(\tau_N)_{N \in \mathbb{N}^*}$ (τ_N is a \mathcal{F}^N -stopping time) with $\tau_N \leq t$ we have

$$\sup_{N \geq N_0} \sup_{\theta \in [0, \delta]} P(|A_{\tau_N + \theta}^{f, N} - A_{\tau_N}^{f, N}| \geq \eta) \leq \epsilon, \quad (27)$$

$$\sup_{N \geq N_0} \sup_{\theta \in [0, \delta]} P(|\langle Z_{\tau_N + \theta}^{f, N} \rangle - \langle Z_{\tau_N}^{f, N} \rangle| \geq \eta) \leq \epsilon. \quad (28)$$

Proof of 1): Let $f \in \mathcal{C}^{1, \bullet}(\mathcal{X}, \mathbb{R})$ and $t \geq 0$. Under Hyp. 20, from Prop. 17.1 p.10, there exists an integer $p \geq 2$ and a constant $C(p, t)$ such that $\sup_{N \in \mathbb{N}^*} \mathbb{E} [\sup_{s \leq t} \langle \bar{\nu}_s^N, 1 \rangle^p] \leq C(p, t)$. Then, for every integer $q \leq p$,

$$\sup_{N \in \mathbb{N}^*} \mathbb{E} \left[\sup_{s \leq t} \langle \bar{\nu}_s^N, 1 \rangle^q \right] \leq C(p, t) + 1. \quad (29)$$

From Markov inequality, for any N ,

$$\begin{aligned} P(\langle \bar{\nu}_t^N, f \rangle \geq K) &\leq \frac{1}{K} \|f\|_\infty \mathbb{E} [\langle \bar{\nu}_t^N, 1 \rangle] \\ &\leq \frac{1}{K} \|f\|_\infty (C(p, t) + 1). \end{aligned}$$

Therefore, the sequence $(\langle \bar{\nu}_t^N, f \rangle)_{N \in \mathbb{N}^*}$ is tight.

Proof of 2): Let $t, \eta, \epsilon, \delta \geq 0$, and $\theta \leq \delta$, $\tau_N < t$. From proposition 18,

$$\begin{aligned} \mathbb{E} \left[|A_{\tau_N + \theta}^{f, N} - A_{\tau_N}^{f, N}| \right] &\leq \mathbb{E} \left[\left| \int_{\tau_N}^{\tau_N + \theta} \langle \bar{\nu}_s^N, \mathcal{F}(\cdot, \bar{\nu}_s^N * F) \nabla_z f \rangle ds \right| \right] \\ &+ \sum_{k \in \mathcal{E}} \mathbb{E} \left[\left| \int_{\tau_N}^{\tau_N + \theta} \int_{\mathcal{X}} q_k(x, \bar{\nu}_s^N * D(z), c_s^N * \psi_a(z)) \langle \mu_k(x), f \rangle \bar{\nu}_s^N(dx) ds \right| \right] \end{aligned}$$

A consequence of hypothesis 2 is that there exists a constant $C(\mathcal{F})$ such that, $\forall (z, y_1, y_2) \in \mathcal{Z} \times \mathbb{R}^d \times \mathbb{R}^d$,

$$|\mathcal{F}(z, y_1) - \mathcal{F}(z, y_2)| \leq C(\mathcal{F}) |y_1 - y_2|,$$

and so

$$|\mathcal{F}(z, y_1)| \leq C(\mathcal{F}) \times |y_1| + \mathcal{F}(z, 0) \leq C(\mathcal{F}) \times |y_1| + \|\mathcal{F}(\cdot, 0)\|_\infty < +\infty$$

as $z \mapsto \mathcal{F}(z, 0)$ is continuous on \mathcal{Z} compact. Then,

$$\begin{aligned} |\langle \bar{\nu}_s^N, \mathcal{F}(\cdot, \bar{\nu}_s^N * F) \nabla_z f \rangle| &\leq \|\nabla_z f\|_\infty \langle \bar{\nu}_s^N, |\mathcal{F}(\cdot, \bar{\nu}_s^N * F)| \rangle \\ &\leq \|\nabla_z f\|_\infty \langle \bar{\nu}_s^N, C(\mathcal{F}) |\bar{\nu}_s^N * F(\cdot)| + \|\mathcal{F}(\cdot, 0)\|_\infty \rangle \\ &\leq \|\nabla_z f\|_\infty (C(\mathcal{F}) \|F_{|\mathcal{Z}}\|_\infty \langle \bar{\nu}_s^N, 1 \rangle^2 + \|\mathcal{F}(\cdot, 0)\|_\infty \langle \bar{\nu}_s^N, 1 \rangle) \end{aligned}$$

as, $\forall z \in \mathcal{Z}$, $|\bar{\nu}_s^N * F(z)| \leq \|F_{|\mathcal{Z}}\|_\infty \langle \bar{\nu}_s^N, 1 \rangle$. Therefore,

$$\begin{aligned} \mathbb{E} \left[|A_{\tau_N+\theta}^{f,N} - A_{\tau_N}^{f,N}| \right] &\leq \|\nabla_z f\|_\infty C(\mathcal{F}) \|F_{|\mathcal{Z}}\|_\infty \mathbb{E} \left[\int_{\tau_N}^{\tau_N+\theta} \langle \bar{\nu}_s^N, 1 \rangle^2 ds \right] \\ &+ (\|\nabla_z f\|_\infty \|\mathcal{F}(\cdot, 0)\|_\infty + \sum_{k \in \mathcal{E}} 2q_k^\infty \|f\|_\infty) \mathbb{E} \left[\int_{\tau_N}^{\tau_N+\theta} \langle \bar{\nu}_s^N, 1 \rangle ds \right]. \end{aligned}$$

We have that $\int_{\tau_N}^{\tau_N+\theta} \langle \bar{\nu}_s^N, 1 \rangle^2 ds \leq \delta \sup_{s \leq t} \langle \bar{\nu}_s^N, 1 \rangle^2$. So, applying proposition 17.1 under Hyp. 20, we get that

$$\begin{aligned} \mathbb{E} \left[|A_{\tau_N+\theta}^{f,N} - A_{\tau_N}^{f,N}| \right] &\leq \delta \times \|\nabla_z f\|_\infty C(\mathcal{F}) \|F_{|\mathcal{Z}}\|_\infty (C(p, t + \delta) + 1) \\ &+ \delta \times (\|\nabla_z f\|_\infty \|\mathcal{F}(\cdot, 0)\|_\infty + \sum_{k \in \mathcal{E}} 2q_k^\infty \|f\|_\infty) (C(p, t + \delta) + 1). \end{aligned}$$

As δ can be made arbitrary small and the upper bound is uniform in N and θ , we get (27) by Markov inequality.

Following the same reasoning, we get that

$$\mathbb{E} \left[|\langle Z^N \rangle_{\tau_N+\theta} - \langle Z^N \rangle_{\tau_N}| \right] \leq \frac{\delta}{N} \times (C(p, t + \delta) + 1) \sum_{k \in \mathcal{T}} 2q_k^\infty \|f^2\|_\infty$$

and we deduce (28), which ends the proof of the lemma. \square

We now define two mappings. The first one is

$$\begin{aligned} \Phi : \mathcal{D}([0, T], \mathcal{M}_{F^+}(\mathcal{X})) &\rightarrow \left(W_{L^2(\mathcal{Z}_a)}^{2,1} \right)^{N_m} \cap \mathcal{C}([0, T], H^1(\mathcal{Z}_a)^{N_m}) \\ \nu &\mapsto \Phi(\nu) = c, \end{aligned} \quad (30)$$

which associates to ν the solution of the reaction-diffusion equation (9), with fixed c_0 and boundary conditions satisfying hypothesis 9. The second one is defined for a fixed $t \in [0, T]$, $f \in \mathcal{C}^{1,1,\bullet}([0, T] \times \mathcal{X}, \mathbb{R})$ by

$$\begin{aligned} \forall \nu \in \mathcal{D}([0, T], \mathcal{M}_{F^+}(\mathcal{X})), \\ \Psi_t(\nu) = \langle \nu_t, f_t \rangle - \langle \nu_0, f_0 \rangle - \int_0^t \langle \nu_s, \mathcal{F}(\cdot, \nu_s * F(\cdot)) \nabla_z f_s + \partial_s f_s \rangle ds \\ - \sum_{k \in \mathcal{E}} \int_0^t \int_{\mathcal{X}} \langle \mu_k(x), f_s \rangle q_k(x, \nu_s * D(z), \Phi(\nu)_s * \psi_a(z)) \nu_s(dx) ds. \end{aligned} \quad (31)$$

Let us establish two technical lemmas.

Lemma 23. Φ is continuous from $\mathcal{D}([0, T], \mathcal{M}_{F^+}(\mathcal{X}))$ to $\mathcal{C}([0, T], H^1(\mathcal{Z}_a)^{N_m})$.

Proof. Let $(\nu^N)_N$ be a sequence which converges to a given ξ in $\mathcal{D}([0, T], \mathcal{M}_{F^+}(\mathcal{X}))$. We denote $c^N = \Phi(\nu^N)$ and $\tilde{c} = \Phi(\xi) \in \mathcal{C}([0, T], H^1(\mathcal{Z}_a)^{N_m})$ the solutions of the reaction-diffusion equation (9), for fixed c_0 and boundary conditions. Let

$$G_N = \left(\sum_{l \in \mathcal{T}} \gamma_{1l}(c^N) \nu^{Nl} * \psi_a(\cdot), \dots, \sum_{l \in \mathcal{T}} \gamma_{N_m l}(c^N) \nu^{Nl} * \psi_a(\cdot) \right) \in \left(W_{L^2(\mathcal{Z}_a)}^{0,0} \right)^{N_m}$$

and

$$G = \left(\sum_{l \in \mathcal{T}} \gamma_{1l}(\tilde{c}) \xi^l * \psi_a(\cdot), \dots, \sum_{l \in \mathcal{T}} \gamma_{N_m l}(\tilde{c}) \xi^l * \psi_a(\cdot) \right) \in \left(W_{L^2(\mathcal{Z}_a)}^{0,0} \right)^{N_m}.$$

By the continuous injection of $\left(W_{L^2(\mathcal{Z}_a)}^{2,1}\right)^{N_m}$ in $\mathcal{C}([0, T], H^1(\mathcal{Z}_a)^{N_m})$ and the proposition 35 applied to $u = G_N - G$, we get that

$$\sup_{t \in [0, T]} \|c^N(t) - \tilde{c}(t)\|_{(L^2(\mathcal{Z}_a))^{N_m}} \leq \sup_{t \in [0, T]} \|c^N(t) - \tilde{c}(t)\|_{H^1(\mathcal{Z}_a)^{N_m}} \leq \beta_1 \|c^N - \tilde{c}\|_{\left(W_{L^2(\mathcal{Z}_a)}^{2,1}\right)^{N_m}} \quad (32)$$

and

$$\|c^N - \tilde{c}\|_{\left(W_{L^2(\mathcal{Z}_a)}^{2,1}\right)^{N_m}} \leq \beta_2 \|G_N - G\|_{\left(W_{L^2(\mathcal{Z}_a)}^{0,0}\right)^{N_m}}. \quad (33)$$

Then, using the same estimates as in Prop. 10 and lemma 11,

$$\begin{aligned} \|G_N - G\|_{\left(W_{L^2(\mathcal{Z}_a)}^{0,0}\right)^{N_m}} &\leq \sum_{m=1}^{N_m} \left(\int_0^T \left\| \sum_{l \in \mathcal{T}} (\gamma_{ml}(c^N(t)) \nu_t^{Nl} * \psi_a - \gamma_{ml}(\tilde{c}(t)) \xi^l * \psi_a) \right\|_{L^2(\mathcal{Z}_a)}^2 dt \right)^{1/2} \\ &\leq N_m \left[\gamma^\infty \sum_{l \in \mathcal{T}} \int_0^T \|(\nu_t^{lN} - \xi_t^l) * \psi_a\|_{L^2(\mathcal{Z}_a)} dt \right. \\ &\quad \left. + \|\psi_a\|_\infty \|\gamma\|_{Lip} C(T, \xi) \int_0^T \|c^N(t) - \tilde{c}(t)\|_{(L^2(\mathcal{Z}_a))^{N_m}} dt \right]. \end{aligned} \quad (34)$$

with, for all $\nu \in \mathcal{D}([0, T], \mathcal{M}_{F^+}(\mathcal{X}))$, $C(T, \nu) = \sup_{[0, T]} \|\nu_t\|_{TV} < \infty$. Combining (32), (33), (34) and applying Gronwall lemma, we obtain that $C := N_m \beta_1 \beta_2 \|\psi_a\|_\infty \|\gamma\|_{Lip} C(T, \xi)$ is a positive constant independent of N such that

$$\sup_{t \in [0, T]} \|c^N(t) - \tilde{c}(t)\|_{H^1(\mathcal{Z}_a)^{N_m}} \leq N_m \beta_1 \beta_2 \gamma^\infty e^{CT} \sum_{l \in \mathcal{T}} \int_0^T \|(\nu_t^{lN} - \xi_t^l) * \psi_a\|_{L^2(\mathcal{Z}_a)} dt. \quad (35)$$

From proposition 29, for almost every $t \in [0, T]$, $d_{PR}(\nu_t^N, \xi_t) \xrightarrow{n} 0$. So, for almost every t , $d_{PR}(\nu_t^{lN}, \xi_t^l) \xrightarrow{n} 0$. Then, from proposition 34 (to be perfectly rigorous: replace \mathcal{Z} by \mathcal{Z}_a in the statement of this proposition), for almost every t ,

$$\sup_{z \in \mathcal{Z}_a} |(\nu_t^{lN} - \xi_t^l) * \psi_a(z)| \xrightarrow{n} 0$$

and so

$$\|(\nu_t^{lN} - \xi_t^l) * \psi_a\|_{L^2(\mathcal{Z}_a)} \xrightarrow{n} 0.$$

From the dominated convergence theorem, we conclude that the right hand side of (35) goes to 0 when $N \rightarrow \infty$, which shows the desired continuity result. \square

From lemma 22 and Prokhorov theorem, we deduce that we can find a subsequence $(\bar{\nu}^{N(n)})_n$ of $(\bar{\nu}^N)_{N \in \mathbb{N}^*}$ which converges weakly in $\mathcal{D}([0, T], \mathcal{M}_{F^+}(\mathcal{X}))$ to a limit denoted $\tilde{\nu}$. Thanks to lemma 23 and the continuous mapping theorem, it is actually the couple $(\bar{\nu}^{N(n)}, c^{N(n)})$ which converges weakly in the product space to $(\tilde{\nu}, \tilde{c})$, \tilde{c} being the solution of (9) associated with $\tilde{\nu}$.

Before proving that the limit $\tilde{\nu}$ is solution of (25), we need a second technical lemma.

Lemma 24. *Let $\nu \in \mathcal{C}([0, T], \mathcal{M}_{F^+}(\mathcal{X}))$, then for any sequence $(\nu^n)_{n \in \mathbb{N}}$ of $\mathcal{D}([0, T], \mathcal{M}_{F^+}(\mathcal{X}))$ such that $\lim_{n \rightarrow \infty} \nu^n = \nu$, we have that*

$$\lim_{n \rightarrow \infty} \Psi_t(\nu^n) = \Psi_t(\nu).$$

Proof. Let $C(T, \nu) = \sup_{[0, T]} \langle \nu_t, 1 \rangle$. Under the hypothesis that $\lim_{n \rightarrow \infty} \nu^n = \nu$ with $\nu \in \mathcal{C}([0, T], \mathcal{M}_{F^+}(\mathcal{X}))$, we know that the convergence is uniform in time (Prop. 28). So, for all $\epsilon > 0$, there exists $N_\epsilon \in \mathbb{N}$ such that for all $n \geq N_\epsilon$,

$$\sup_{t \in [0, T]} d_{PR}(\nu_t^n, \nu_t) \leq \epsilon \quad (36)$$

As a consequence of the definition of d_{PR} (see (62)), we have

$$\sup_{t \in [0, T]} \langle \nu_t^n, 1 \rangle \leq \sup_{t \in [0, T]} \langle \nu_t, 1 \rangle + \epsilon \leq C(T, \nu) + \epsilon \quad (37)$$

Therefore, we restrict ourselves without loss of generality to sequences such that, for all $n \in \mathbb{N}$,

$$\sup_{t \in [0, T]} \langle \nu_t^n, 1 \rangle \leq 2C(T, \nu). \quad (38)$$

Moreover, uniform convergence of $(\nu^n)_{n \in \mathbb{N}}$ to ν implies that for all $t \leq T$, the projection at t $\pi_t : \nu \mapsto \nu_t$ from $\mathcal{D}([0, T], \mathcal{M}_{F^+}(\mathcal{X}))$ to $\mathcal{M}_{F^+}(\mathcal{X})$ at every element ν continuous in time. Prokhorov distance metrizes the weak topology on $\mathcal{M}_{F^+}(\mathcal{X})$, that is to say that for all continuous bounded function $f \in \mathcal{C}_b(\mathcal{X}, \mathbb{R})$, the mapping $\nu \mapsto \langle \nu, f \rangle$ is continuous on $\mathcal{M}_{F^+}(\mathcal{X})$. So $\nu \mapsto \langle \pi_t(\nu), f \rangle$ is continuous at each element of $\mathcal{C}([0, T], \mathcal{M}_{F^+}(\mathcal{X}))$.

We can now show the continuity of Ψ_t at ν . We denote $c^n = \Phi(\nu^n)$ and $c = \Phi(\nu)$, then

$$\begin{aligned} \Psi_t(\nu^n) - \Psi_t(\nu) &= \underbrace{\langle \nu_t^n - \nu_t, f_t \rangle - \langle \nu_0^n - \nu_0, f_0 \rangle - \int_0^t \langle \nu_s^n - \nu_s, \partial_s f_s \rangle ds}_{(A_n)} \\ &\quad - \underbrace{\left(\int_0^t \langle \nu_s^n, \mathcal{F}(\cdot, \nu_s^n * F(\cdot)) \nabla_z f_s \rangle ds - \int_0^t \langle \nu_s, \mathcal{F}(\cdot, \nu_s * F(\cdot)) \nabla_z f_s \rangle ds \right)}_{(B_n)} \\ &\quad - \underbrace{\sum_{k \in \mathcal{E}} \int_0^t \left(\int_{\mathcal{X}} \langle \mu_k(x), f_s \rangle q_k(x, \nu_s^n * D(z), c_s^n * \psi_a(z)) \nu_s^n(dx) ds - \int_{\mathcal{X}} \langle \mu_k(x), f_s \rangle q_k(x, \nu_s * D(z), c_s * \psi_a(z)) \nu_s(dx) \right) ds}_{(C_n)}. \end{aligned}$$

According to the foregoing, regularity of f and compactness of \mathcal{X} , we have $\langle \nu_t^n - \nu_t, f_t \rangle \xrightarrow[n \rightarrow +\infty]{} 0$, $\langle \nu_0^n - \nu_0, f_0 \rangle \xrightarrow[n \rightarrow +\infty]{} 0$ and $s \mapsto \langle \nu_s^n - \nu_s, \partial_s f_s \rangle$ converges pointwise to 0. Moreover,

$$\begin{aligned} |\langle \nu_s^n - \nu_s, \partial_s f_s \rangle| &\leq \sup_{(s, x)} \{ \partial_s f_s(x) \} \times (\langle \nu_s^n, 1 \rangle + \langle \nu_s, 1 \rangle) \\ &\leq 3C(T, \nu) \sup_{(s, x)} \{ \partial_s f_s(x) \} < +\infty. \end{aligned}$$

By dominated convergence, $\int_0^t \langle \nu_s^n - \nu_s, \partial_s f_s \rangle ds \xrightarrow[n \rightarrow +\infty]{} 0$. Therefore, $\lim_{n \rightarrow +\infty} A_n = 0$. We have $B_n = B_n^1 + B_n^2$ with

$$B_n^1 = \int_0^t \langle \nu_s^n - \nu_s, \mathcal{F}(\cdot, \nu_s^n * F(\cdot)) \nabla_z f_s \rangle ds$$

and

$$B_n^2 = \int_0^t \langle \nu_s^n, (\mathcal{F}(\cdot, \nu_s^n * F(\cdot)) - \mathcal{F}(\cdot, \nu_s * F(\cdot))) \nabla_z f_s \rangle ds.$$

$\mathcal{F}(\cdot, \nu_s * F(\cdot)) \nabla_z f_s$ is continuous bounded as \mathcal{X} is compact, so the integrand in B_n^1 converges pointwise to 0. As for A_n , by the dominated convergence theorem, $B_n^1 \xrightarrow[n \rightarrow +\infty]{} 0$. Moreover, \mathcal{F} is Lipschitz according to Hyp. 2.1, therefore

$$B_n^2 \leq \int_0^t \langle \nu_s^n, \|\mathcal{F}\|_{Lip} \times |\nu_s^n * F(\cdot) - \nu_s * F(\cdot)| \times |\nabla_z f_s| \rangle ds.$$

As F is Lipschitz and bounded on the compact set $\hat{\mathcal{Z}}$ under Hyp. 2.1, we can apply proposition 34: $\forall t \in [0, T]$:

$$\sup_{z \in \hat{\mathcal{Z}}} |(\nu_t^n - \nu_t) * F(z)| \xrightarrow[n \rightarrow \infty]{} 0.$$

So $B_n^2 \xrightarrow[n \rightarrow +\infty]{} 0$.

Finally, let $C_n = C_n^1 + C_n^2$ with

$$C_n^1 = \sum_{k \in \mathcal{E}} \int_0^t \int_{\mathcal{X}} \langle \mu_k(x), f_s \rangle q_k(x, \nu_s^n * D(z), c_s^n * \psi_a(z)) (\nu_s^n(dx) - \nu_s(dx)) ds$$

and

$$C_n^2 = \sum_{k \in \mathcal{E}} \int_0^t \int_{\mathcal{X}} \langle \mu_k(x), f_s \rangle (q_k(x, \nu_s * D(z), c_s * \psi_a(z)) - q_k(x, \nu_s^n * D(z), c_s^n * \psi_a(z))) \nu_s(dx) ds$$

The term $C_n^1 \xrightarrow{n \rightarrow +\infty} 0$ as q_k are continuous bounded according to Hyp. 14. To tackle the term C_n^2 , we observe that for all $t \in [0, T]$ and all $z \in \mathcal{Z}$,

$$\begin{aligned} |(c_t^n - c_t) * \psi_a(z)| &\leq \|\psi_a\|_{L^2} \|c_t^n - c_t\|_{L^2(\mathcal{Z}_a)^{N_m}} \\ &\leq \|\psi_a\|_{L^2} \sup_{t \in [0, T]} \|\Phi(\nu^n)_t - \Phi(\nu)_t\|_{H^1(\mathcal{Z}_a)^{N_m}} \xrightarrow{n \rightarrow +\infty} 0 \end{aligned}$$

where we used lemma 23. Due to Hyp. 8 on ψ_a , we use proposition 34 and get

$$\sup_{z \in \mathcal{Z}} |(\nu_t^n - \nu_t) * D(z)| \xrightarrow{n \rightarrow \infty} 0.$$

Finally, using that q_k is Lipschitz with respect to its second and third variable (Hyp. 14), proposition 34 for the term involving D (Lipschitz and bounded on $\hat{\mathcal{Z}}$ according to Hyp. 14) and the dominated convergence theorem, we get that $C_n^2 \xrightarrow{n} 0$.

This ends the proof of the continuity of Ψ_t at ν . \square

Lemma 25. $\tilde{\nu}$ is almost surely a solution of (25).

Proof. We observe that $\tilde{\nu}$ is solution of (25) if and only if $\Psi_t(\tilde{\nu}) = 0$ a.s. for all t and f . We therefore prove the latter. We first show that

$$\tilde{\nu} \in \mathcal{C}([0, T], \mathcal{M}_{F^+}(\mathcal{X})) \text{ a.s.} \quad (39)$$

$\tilde{\nu}$ is right continuous by definition. Let us prove the left continuity. We use total variation norm on $\mathcal{M}_{F^+}(\mathcal{X})$ defined by (see A p.29)

$$\|\mu - \nu\|_{TV} = \sup \{ \langle \mu - \nu, f \rangle; f \in \mathcal{C}(\mathcal{X}), \|f\|_{\infty} \leq 1 \}.$$

Let $f \in \mathcal{C}(\mathcal{X})$, $\|f\|_{\infty} \leq 1$. For all n , random jumps of $\bar{\nu}^{N(n)}$ happen one at a time a.s. so, for all $t \leq T$,

$$\begin{aligned} |\langle \bar{\nu}_t^{N(n)}, f \rangle - \langle \bar{\nu}_{t^-}^{N(n)}, f \rangle| &\leq 2 \frac{\|f\|_{\infty}}{N(n)} \leq \frac{2}{N(n)} \text{ a.s.} \\ \Rightarrow \sup_{0 \leq t \leq T} \|\bar{\nu}_t^{N(n)} - \bar{\nu}_{t^-}^{N(n)}\|_{TV} &\leq \frac{2}{N(n)} \text{ a.s..} \end{aligned}$$

Total variation distance dominates over Prokhorov distance on $\mathcal{M}_{F^+}(\mathcal{X})$ (30 p.30). Therefore,

$$\lim_{n \rightarrow \infty} \sup_{0 \leq t \leq T} d_{PR}(\bar{\nu}_t^{N(n)}, \bar{\nu}_{t^-}^{N(n)}) = 0 \text{ a.s.}$$

and, by theorem 10.2 p.148 in [EK86],

$$\sup_{0 \leq t \leq T} d_{PR}(\tilde{\nu}_t, \tilde{\nu}_{t^-}) = 0 \text{ a.s.}$$

which proves (39).

We now observe that

$$\Psi_t(\bar{\nu}^N) = Z_t^{f, N}$$

with $Z_t^{f, N}$ the martingale defined in Prop. 18. Then, under Hyp. 20:

$$\mathbb{E} \left[\Psi_t(\bar{\nu}^{N(n)})^2 \right] = \mathbb{E} \left[\langle Z^{f, N(n)} \rangle_t \right] \leq \frac{1}{N(n)} \|f\|_{\infty}^2 t (C(p, t) + 1) \sum_{k \in \mathcal{E}} q_k^{\infty} \xrightarrow{n \rightarrow \infty} 0 \quad (40)$$

with $C(p, t)$ a constant defined in Prop. 17. Therefore, $(\Psi_t(\bar{\nu}^{N(n)}))_n$ converges in L^2 , and so in L^1 by Cauchy-Schwartz, to 0.

Given $\bar{\nu}^{N(n)}$ converges weakly to $\tilde{\nu}$ and Ψ_t is continuous at $\tilde{\nu}$, we get from lemma 24 and the continuous mapping theorem that $|\Psi_t(\bar{\nu}^{N(n)})|$ converges weakly to $|\Psi_t(\tilde{\nu})|$. Moreover, under Hyp. 20 and from Prop. 17.1, the sequence $(\Psi_t(\bar{\nu}^{N(n)}))_n$ is bounded in L^p with $p \geq 2$ and so uniformly integrable. Then, we have (see theorem 3.5 in [Bil99]) that, $\forall q \leq p$,

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[|\Psi_t(\bar{\nu}^{N(n)})|^q \right] = \mathbb{E} [|\Psi_t(\tilde{\nu})|^q].$$

In particular,

$$\begin{aligned} 0 &= \lim_{n \rightarrow \infty} \mathbb{E} \left[|\Psi_t(\bar{\nu}^{N(n)})| \right] = \mathbb{E} [|\Psi_t(\tilde{\nu})|] \\ \Leftrightarrow \Psi_t(\tilde{\nu}) &= 0 \text{ as.} \end{aligned}$$

Note that $\tilde{\nu}$ is continuous in time a.s., and therefore $\bar{\nu}_0^N$ converges weakly to $\tilde{\nu}_0$. But, under Hyp. 20, $\bar{\nu}_0^N$ converges in distribution to ξ_0 so ξ_0 and $\tilde{\nu}_0$ have the same law.

This is true for any $t \geq 0$, we deduce that $\tilde{\nu}$ is solution of (25), and the sequence of concentrations $(c^{N(n)})_n$ converges as well to the solution of (26) in $\mathcal{C}([0, T], H^1(\mathcal{Z}_a)^{N_m})$. \square

So far, we showed that all accumulation points of $(\bar{\nu}^N)_{N \in \mathbb{N}^*}$ are positive measures, solutions of (25), and related accumulation points in the concentrations' space are solutions of (26) in $\mathcal{C}([0, T], H^1(\mathcal{Z}_a)^{N_m})$. To complete the proof of theorem 21, we tackle uniqueness of the solution.

Lemma 26. *If (25-26) has a solution (ξ, c) such that $\xi \in \mathcal{C}([0, T], \mathcal{M}_{F^+}(\mathcal{X}))$ then this solution is unique.*

Proof. We proceed in two steps.

Step 1 : on a finite time horizon, positive solutions of (25) have finite mass. Let $(c_0, \xi_0) \in H^1(\mathcal{Z}_a)^{N_m} \times \mathcal{M}_{F^+}(\mathcal{X})$ be a deterministic initial condition. We suppose that there exists a solution $(c_t, \xi_t)_{t \geq 0} \in \mathcal{C}([0, T], H^1(\mathcal{Z}_a)^{N_m}) \times \mathcal{D}([0, T], \mathcal{M}_{F^+}(\mathcal{X}))$. Let show that the mass of ξ_t is bounded at finite time by a constant depending only of ξ_0 and t . From (25), for $t \leq T$,

$$\begin{aligned} \langle \xi_t, 1 \rangle &= \langle \xi_0, 1 \rangle + \sum_{k \in \mathcal{E}} \int_0^t \sum_{l \in \mathcal{T}} \int_{\mathcal{Z}} \langle \mu_k(z, l), 1 \rangle q_k^\infty q_k((z, l), \xi_s * D(z), c_s * \psi_a(z)) \xi_s(dz, l) ds \\ &\leq \langle \xi_0, 1 \rangle + \sum_{k \in \mathcal{E}} q_k^\infty \int_0^t \sum_{l \in \mathcal{T}} \int_{\mathcal{Z}} \xi_s(dz, l) ds = \langle \xi_0, 1 \rangle + \sum_{k \in \mathcal{E}} q_k^\infty \int_0^t \langle \xi_s, 1 \rangle ds. \end{aligned} \quad (41)$$

By Gronwall lemma, we see that

$$\langle \xi_t, 1 \rangle \leq \langle \xi_0, 1 \rangle \exp(t \sum_{k \in \mathcal{E}} q_k^\infty) < +\infty,$$

which is the desired result.

Step 2 : a positive solution of (25-26) is unique. Let (c_t^1, ξ_t^1) and (c_t^2, ξ_t^2) be two solutions of (25-26) on $[0, T]$, with ξ_t^i positive as in the previous step. We use the flat metric d_F on $\mathcal{M}_{F^+}(\mathcal{X})$ defined as (more details are given in A)

$$d_F(\nu, \xi) = \sup_{f \in \mathcal{L}ip_{1,1}(\mathcal{X}, \mathbb{R})} \langle \nu - \xi, f \rangle,$$

and derive a bound for $\|c_t^1 - c_t^2\|_{L^2} + d_F(\xi_t^1, \xi_t^2)$, for $t \in [0, T]$.

From step 1, we can define

$$C(t, \xi^1 + \xi^2) := \sup_{s \leq t} \langle \xi_s^1 + \xi_s^2, 1 \rangle < +\infty.$$

Moreover, ξ_t^1, ξ_t^2 are finite positive measures so, from lemma 11, there exists a constant C such that

$$\|c_t^1 - c_t^2\|_{L^2} \leq C \int_0^t d_F(\xi_s^1, \xi_s^2) ds. \quad (42)$$

We now bound $d_F(\xi_t^1, \xi_t^2)$ using formula (25). As done in [Tra06], the idea is to cancel in equation (25) the integral term accounting for the deterministic movement of cells between jumps. Let $u \in \mathcal{L}ip_{1,1} \cap \mathcal{C}^{1,\bullet}(\mathcal{X}, \mathbb{R})$. From lemma 38, there exists a unique function

$$\begin{aligned} u^1 : \quad [0, T] \times \mathcal{X} &\rightarrow \mathbb{R} \\ (s, z, l) &\mapsto u_s^1(z, l) \end{aligned}$$

such that, $\forall l \in \mathcal{T}$, $u_s^1(z, l)$ is solution of (74), with $\xi_s = \xi_s^1$ and the condition $u_t^1(z, l) = u(z, l)$, that is to say

$$\partial_s u_s^1(z, l) + \mathcal{F}(z, \xi_s^1 * F(z)) \nabla_z u_s^1(z, l) = 0 \quad \text{and} \quad u_t^1(x) = u(x).$$

We also know that $u_s^1(z, l) = u(Z^1(s; t, z), l)$ where $Z^1(\cdot)$ are the characteristics associated to (74) when $\xi_s = \xi_s^1$. Similarly, let $u_s^2(z, l) = u(Z^2(s; t, z), l)$, with $\xi_s = \xi_s^2$ in (74). Moreover, still from lemma 38,

$$u^i(\cdot, l) \in \mathcal{C}^1([0, T] \times \mathcal{Z}, \mathbb{R}).$$

We then apply formula (25), with $\xi_0^1 = \xi_0^2 = \xi_0$.

$$\begin{aligned} & | \langle \xi_t^1 - \xi_t^2, u \rangle | = | \langle \xi_0, u_0^1 - u_0^2 \rangle | \\ & + \underbrace{\sum_{k \in \mathcal{E}} \int_0^t \sum_{l \in \mathcal{T}} \int_{\mathcal{Z}} \langle \mu_k(z, l), u_s^1 \rangle q_k((z, l), \xi_s^1 * D(z), c_s^1 * \psi_a(z)) \xi_s^1(dz, l) ds}_{(A)} \\ & - \sum_{k \in \mathcal{E}} \int_0^t \sum_{l \in \mathcal{T}} \int_{\mathcal{Z}} \langle \mu_k(z, l), u_s^2 \rangle q_k((z, l), \xi_s^2 * D(z), c_s^2 * \psi_a(z)) \xi_s^2(dz, l) ds \\ & = \underbrace{| \langle \xi_0, u_0^1 - u_0^2 \rangle |}_{(A)} \\ & + \underbrace{\sum_{k \in \mathcal{E}} \int_0^t \sum_{l \in \mathcal{T}} \int_{\mathcal{Z}} \langle \mu_k(z, l), u_s^1 - u_s^2 \rangle q_k((z, l), \xi_s^1 * D(z), c_s^1 * \psi_a(z)) \xi_s^1(dz, l) ds}_{(B)} \\ & + \underbrace{\sum_{k \in \mathcal{E}} \int_0^t \sum_{l \in \mathcal{T}} \int_{\mathcal{Z}} \langle \mu_k(z, l), u_s^2 \rangle (q_k((z, l), \xi_s^1 * D(z), c_s^1 * \psi_a(z)) - q_k((z, l), \xi_s^2 * D(z), c_s^2 * \psi_a(z))) \xi_s^1(dz, l) ds}_{(C)} \\ & + \underbrace{\sum_{k \in \mathcal{E}} \int_0^t \sum_{l \in \mathcal{T}} \int_{\mathcal{Z}} \langle \mu_k(z, l), u_s^2 \rangle (q_k((z, l), \xi_s^2 * D(z), c_s^1 * \psi_a(z)) - q_k((z, l), \xi_s^2 * D(z), c_s^2 * \psi_a(z))) \xi_s^1(dz, l) ds}_{(D)} \\ & + \underbrace{\sum_{k \in \mathcal{E}} \int_0^t \sum_{l \in \mathcal{T}} \int_{\mathcal{Z}} \langle \mu_k(z, l), u_s^2 \rangle q_k((z, l), \xi_s^2 * D(z), c_s^2 * \psi_a(z)) (\xi_s^1 - \xi_s^2)(dz, l) ds}_{(E)}. \end{aligned}$$

Before proceeding, we recall the following simple properties that we will use extensively. Let f be Lipschitz on \mathcal{X} , then $\|\frac{f}{\|f\|_{\mathcal{V}}}\|_{\infty} \leq 1$, $\|\frac{f}{\|f\|_{\mathcal{V}}}\|_{Lip} \leq 1$ and, for $\nu_1, \nu_2 \in \mathcal{M}_{F^+}(\mathcal{X})$

$$\langle \nu_1 - \nu_2, f \rangle \leq \|f\|_{\mathcal{V}} \times d_F(\nu_1, \nu_2).$$

We first derive a bound for (C). Remember that $\langle \mu_k(z, l), u_s^2 \rangle$ is the sum of at most two punctual evaluations of u_s^2 . So $\|\langle \mu_k(z, l), u_s^2 \rangle\|_{\infty} \leq 2$. We supposed (Hyp. 14) that q_k is Lipschitz with respect to its second variable, therefore there exists a constant $C(q_k)$ such that

$$(C) \leq 2 \sum_{k \in \mathcal{E}} \int_0^t \sum_{l \in \mathcal{T}} \int_{\mathcal{Z}} C(q_k) |\xi_s^1 * D(z) - \xi_s^2 * D(z)| \xi_s^1(dz, l) ds.$$

D is locally Lipschitz in z under Hyp. 14 on the compact set $\hat{\mathcal{Z}} := \{z - y : (z, y) \in \mathcal{Z}^2\}$. Therefore, for all $z \in \mathcal{Z}$, $|\xi_s^1 * D(z) - \xi_s^2 * D(z)| \leq \|D_{a|\hat{\mathcal{Z}}}\|_{\mathcal{V}} d_F(\xi_s^1, \xi_s^2)$. Let $C_{(C)} := 2 \sum_{k \in \mathcal{E}} C(q_k) C(t, \xi_1 + \xi_2) \|D_{a|\hat{\mathcal{Z}}}\|_{\mathcal{V}}$. Then

$$(C) \leq C_{(C)} \int_0^t d_F(\xi_s^1, \xi_s^2) ds.$$

We find an upper bound for the term (D) using again the fact that q_k is Lipschitz with respect to its

third variable:

$$\begin{aligned}
(D) &\leq 2 \sum_{k \in \mathcal{E}} C(q_k) \int_0^t \sum_{l \in \mathcal{T}} \int_{\mathcal{Z}} |(c_s^1 - c_s^2) * \psi_a(z)| \xi_s^1(dz, l) ds \\
&\leq 2 \sum_{k \in \mathcal{E}} C(q_k) \|\psi_a\|_{L^2} C(t, \xi_1 + \xi_2) \int_0^t \|c_s^1 - c_s^2\|_{L^2(\mathcal{Z}_a)^{N_m}} ds \\
&\leq C_{(D)} \int_0^t \|c_s^1 - c_s^2\|_{L^2(\mathcal{Z}_a)^{N_m}} ds
\end{aligned}$$

setting $C_{(D)} := 2 \sum_{k \in \mathcal{E}} C(q_k) \|\psi_a\|_{L^2(\mathbb{R}^d)} C(t, \xi_1 + \xi_2)$.

We move on to term (E). For $s \leq t$, let $f_s : (z, l) \mapsto \langle \mu_k(z, l), u_s^2 \rangle q_k(z, l, \xi_s^2 * D(z), c_s^2 * \psi_a(z))$ be defined on \mathcal{X} . Let us show that f_s is uniformly bounded and uniformly Lipschitz for $s \leq t$.

On the one hand, q_k and u^2 are bounded, so $\sup_{s \leq t} \|f_s\|_\infty < +\infty$.

On the other hand, $\forall l \in \mathcal{T}$, $z \mapsto q_k(z, l, \xi_s^2 * D(z), c_s^2 * \psi_a(z))$ is Lipschitz uniformly for $s \leq t$. Indeed, $z \mapsto \xi_s^2 * D(z)$ is Lipschitz uniformly for $s \leq t$ under Hyp. 14. It is the case of $c_s^2 * \psi_a(z)$ as well because $\sup_{s \leq t} \|c_s^2\|_{L^2(\mathcal{Z}_a)^{N_m}} < +\infty$ from Prop. 10 and $\psi_a(z)$ is Lipschitz on $\hat{\mathcal{Z}}$ under Hyp. 8. Eventually, $\forall l \in \mathcal{T}$, $q_k(\cdot, l, \cdot)$ is Lipschitz under Hyp. 14 so $z \mapsto q_k(z, l, \xi_s^2 * D(z), c_s^2 * \psi_a(z))$ is Lipschitz uniformly for $s \leq t$ from Prop. 1.

Moreover, for all l , $z \mapsto u_s^2(z, l)$ is Lipschitz uniformly in time as it is continuously derivable, so u_s^2 is Lipschitz on \mathcal{X} uniformly in time. Then, when the jump type k is a differentiation or an extrusion, $\langle \mu_k(z, l), u_s^2 \rangle$ is also Lipschitz uniformly in time. If k is a division, it is the case as well because $\lambda(z)$ is Lipschitz according to Hyp. 12. Eventually, the product of two Lipschitz bounded functions is Lipschitz. Therefore, f_s is Lipschitz uniformly for $s \leq t$. Let $C_{(E)} := \sup_{s \leq t} \|f_s\|_v < +\infty$. Then,

$$(E) \leq C_{(E)} \int_0^t d_F(\xi_s^1, \xi_s^2) ds.$$

In order to bound the remaining terms (A) and (B), we want an upper bound on $|u_s^1(z) - u_s^2(z)|$ depending of $d_F(\xi_s^1, \xi_s^2)$. We have

$$\begin{aligned}
|u_s^1(z) - u_s^2(z)| &= |u(Z^1(t; s, z)) - u(Z^2(t; s, z))| \\
&\leq \|u\|_v |Z^1(t; s, z) - Z^2(t; s, z)| \leq |Z^1(t; s, z) - Z^2(t; s, z)|
\end{aligned}$$

since $u \in \mathcal{L}ip_{1,1}(\mathcal{X}, \mathbb{R})$.

Using proposition 39, there exists a constant $C(t)$ such that

$$\begin{aligned}
(A) &= \sum_{l \in \mathcal{T}} \int_{\mathcal{Z}} u(Z^1(t; 0, z), l) - u(Z^2(t; 0, z), l) \xi_0(dz, l) \\
&\leq \sum_{l \in \mathcal{T}} \int_{\mathcal{Z}} |Z^1(t; 0, z) - Z^2(t; 0, z)| \xi_0(dz, l) \\
&\leq \langle \xi_0, 1 \rangle C(t) \int_0^t d_F(\xi_s^1, \xi_s^2) ds \\
&\leq C_{(A)} \int_0^t d_F(\xi_s^1, \xi_s^2) ds
\end{aligned}$$

with $C_{(A)} := \langle \xi_0, 1 \rangle C(t)$. Note that the first inequality comes from the assumption that u is 1-Lipschitz, which motivates the use of the flat metric d_F .

With proposition 39 and as u is 1-Lipschitz, we can also tackle term (B):

$$\begin{aligned}
(B) &\leq \sum_{k \in \mathcal{E}} 2q_k^\infty C(t, \xi^1 + \xi^2) C(t) \int_0^t \int_s^t d_F(\xi_r^1, \xi_r^2) dr ds \\
&\leq \sum_{k \in \mathcal{E}} 2q_k^\infty C(t, \xi^1 + \xi^2) C(t) t \int_0^t d_F(\xi_s^1, \xi_s^2) ds \\
&\leq C_{(B)} \int_0^t d_F(\xi_s^1, \xi_s^2) ds
\end{aligned}$$

with $C_{(B)} := \sum_{k \in \mathcal{E}} 2q_k^\infty C(t, \xi^1 + \xi^2 C(t)t$.

Therefore, we showed that there exists a constant C such that, $\forall u \in \mathcal{L}ip_{1,1}(\mathcal{X}, \mathbb{R}) \cap \mathcal{C}^{1,\bullet}(\mathcal{X}, \mathbb{R})$,

$$|\langle \xi_t^1 - \xi_t^2, u \rangle| \leq C \int_0^t \|c_s^1 - c_s^2\|_{L^2(\mathcal{Z})} + d_F(\xi_s^1, \xi_s^2) ds.$$

As \mathcal{X} is compact, according to Stone-Weierstrass theorem, the set of functions $\mathcal{C}^{1,\bullet}(\mathcal{X}, \mathbb{R})$ is dense in $\mathcal{C}(\mathcal{X}, \mathbb{R})$ for the uniform norm. So

$$\sup_{u \in \mathcal{L}ip_{1,1} \cap \mathcal{C}^{1,\bullet}} |\langle \xi_t^1 - \xi_t^2, u \rangle| = \sup_{u \in \mathcal{L}ip_{1,1}} |\langle \xi_t^1 - \xi_t^2, u \rangle| = d_F(\xi_t^1, \xi_t^2)$$

and

$$d_F(\xi_t^1, \xi_t^2) \leq C \int_0^t d_F(\xi_s^1, \xi_s^2) + \|c_s^1 - c_s^2\|_{L^2(\mathcal{Z})} ds. \quad (43)$$

Using (42), we obtain that there exists a constant C such that

$$\|c_t^1 - c_t^2\|_{L^2} + d_F(\xi_t^1, \xi_t^2) \leq C \int_0^t \|c_s^1 - c_s^2\|_{L^2(\mathcal{Z})} + d_F(\xi_s^1, \xi_s^2) ds. \quad (44)$$

From Gronwall lemma,

$$\|c_t^1 - c_t^2\|_{L^2} + d_F(\xi_t^1, \xi_t^2) = 0$$

which proves the uniqueness of a positive solution. \square

This shows that the sequence $(\bar{\nu}^N)_{N \in \mathbb{N}^*}$ is tight with a unique accumulation point, which is solution of (25), so it is convergent and the sequence of concentrations converges to the corresponding solution of (26) in $\mathcal{C}([0, T], H^1(\mathcal{Z}_a)^{N_m})$. This ends the proof of theorem 21.

6 Numerical convergence of the PDMP towards a deterministic limit

6.1 Description of the PDMP model

We illustrate our result on a simple model inspired from a more complex intestinal crypt model previously introduced [Dar+22]. We first set the precise numerical experiments that are conducted, considering two populations of different types and stochastic regulation processes including cell division, extrusion and differentiation. Then, we introduce the corresponding deterministic limit in large population for this model.

We consider a one dimensional spatial model on $\mathcal{Z} = [0, z_{max}]$, $0 < z_{max} < \infty$, with two different type of cells: progenitor cell (pc) and enterocyte cells (ent) that is to say $\mathcal{T} = \{\text{pc}, \text{ent}\}$, and two metabolites: butyrate and oxygen ($N_m = 2$). We denote $c_b(z, t)$ and $c_o(z, t)$ the respective concentration of butyrate and oxygen. The size of the cells is modelled through a parameter $a > 0$.

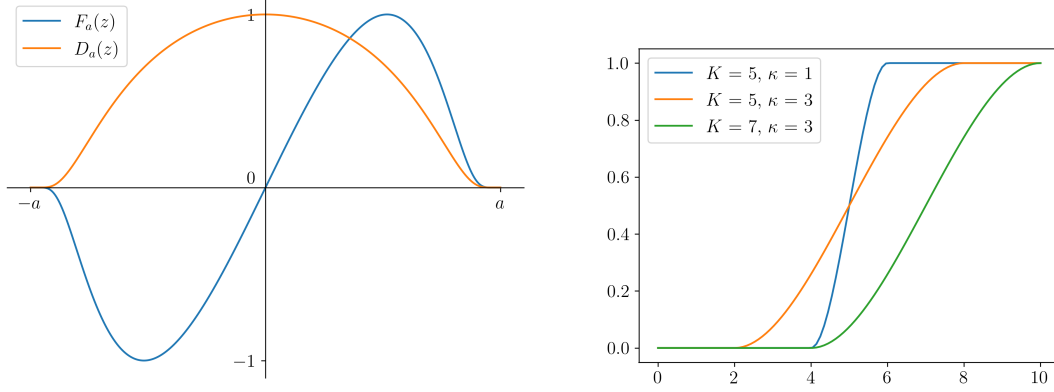
Mechanical interactions for cell migration. We set

$$\frac{dz^i}{dt}(t) = \mathcal{F}(z^i(t), \nu_t * F_a(z^i(t))) := \phi(z^i(t)) \nu_t * F_a(z^i(t)) \quad (45)$$

with

$$F_a(z) = k_a \frac{f_a(z)}{\|f_a\|_\infty} \quad \text{where} \quad f_a(z) = \begin{cases} \frac{z}{a} \exp\left(-\frac{|z|}{1-\frac{|z|}{a}}\right) & \text{if } |z| \leq a, \\ 0 & \text{otherwise.} \end{cases} \quad (46)$$

k_a is a scaling parameter driving the maximal intensity of the force. F_a is plotted in Fig. 1a with $k_a = 1$. $\phi(z)$ is a positive continuously derivable function, such that $0 \leq \phi(z) \leq 1$, $\phi(0) = \phi(z_{max}) = 0$ and $\phi(z) = 1$ if $z \in [\epsilon, z_{max} - \epsilon]$. The function ϕ models the vanishing of interaction forces near the boundaries, induced by geometrical constraints (see [Dar+22] for more details). The exact expression of ϕ is given in equation (78).



(a) Interaction kernels F_a (with $k_a = 1$, blue) and D_a (with $d_a = 1$, orange). The function F_a describes repulsive forces, while D_a renders density dependant interactions in jump rates.

(b) Regulation function $y \mapsto R(y, K, \kappa)$ for various values of K and κ . The parameter K tunes the location of the transition and κ its slope.

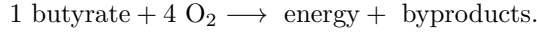
Figure 1: **Interactions and regulations.** Shapes of interaction kernels and regulation functions used in the model

Reaction-diffusion of chemicals. The oxygen and butyrate concentrations c_o and c_b follow the reaction-diffusion equation (9), leading to the system:

$$\partial_t c_o - \sigma \Delta c_o = \sum_{l \in \{\text{pc}, \text{ent}\}} \gamma_{ml}(c_o, c_b) \nu_t^l * \psi_a(z), \quad (47)$$

$$\partial_t c_b - \sigma \Delta c_b = \sum_{l \in \{\text{pc}, \text{ent}\}} \gamma_{ml}(c_o, c_b) \nu_t^l * \psi_a(z). \quad (48)$$

Progenitor cells do not use oxygen and butyrate for their metabolism, while enterocytes catalyze them with the following chemical reaction



Therefore we set for the reaction speed in equation (47)-(48), $\gamma_{m,l}(c_o, c_b) = 0$ if $l = \text{pc}$, and, if $l = \text{ent}$,

$$\gamma_{m,l}(c_o, c_b) = -s_m \gamma^\infty \frac{c_o^4 c_b}{c_o^4 c_b + K^5} \quad (49)$$

where s_m is the stoichiometric coefficient of metabolite m . γ^∞ is the maximal reaction speed of the reaction and $K > 0$ is a constant. We define $\psi_a(z)$ as follow. Let

$$D_a(z) := d_a D\left(\frac{z}{a}\right), \quad \text{where} \quad D(z) := \begin{cases} \exp\left(-\frac{z^2}{1-z^2}\right) & \text{if } 0 \leq |z| \leq 1 \\ 0 & \text{otherwise,} \end{cases} \quad (50)$$

and d_a is a scaling parameter (see Fig. 1a for a plot with $d_a = 1$). We then derive a normalized version of the density kernel D_a by writing

$$\psi_a(z) := \frac{D_{a/2}(z)}{\int_{-a/2}^{a/2} D_{a/2}(u) du}. \quad (51)$$

Then, for all $z \in [0, z_{max}]$, $\int_{\mathcal{Z}_a} \psi_a(z-y) dy = 1$. We see with this kernel definition that the extended domain \mathcal{Z}_a has been introduced for compatibility reasons between the reaction-diffusion domain and the epithelial cell domain $[0, z_{max}]$.

This system of equations is supplemented with boundary conditions modeling the chemical fluxes at the crypt extremities. For butyrate:

$$\begin{cases} c_b = c_{b,\text{lum}} & \text{at } z = z_{max} + \frac{a}{2} \\ \partial_z c_b = 0 & \text{at } z = -\frac{a}{2} \end{cases} \quad (52)$$

and for oxygen:

$$\begin{cases} \partial_z c_o = 0 & \text{at } z = z_{max} + \frac{a}{2} \\ c_o = c_{o,bot} & \text{at } z = -\frac{a}{2} \end{cases} \quad (53)$$

Constant parameters $c_{i,lum}$, $c_{i,bot}$ represent the chemical concentrations in the gut lumen or at the bottom of the crypt, respectively reflecting microbial or host physiology.

Stochastic jumps modelling cell fate. We consider three types of stochastic events. Progenitor cells can divide symmetrically or differentiate into enterocytes, and enterocytes can die. As explained in section 3.2, these stochastic events are regulated in various ways, which are summarized in table 1:

- the rate of progenitor cell division $q_{pc,pc}(\cdot)$ is a decreasing function of local butyrate concentration, height in the crypt ($z \in [0, z_{max}]$) and local cell density. Upon division, the spatial shift of the daughter cell λ is

$$\lambda(z) = \begin{cases} \frac{a}{2} & \text{if } z \leq z_{max}/2 - a, \\ \frac{-a}{2} & \text{if } z \geq z_{max}/2 + a, \\ \frac{a}{2} - a \frac{z - (z_{max}/2 - a)}{2a} & \text{otherwise,} \end{cases}$$

- The rate of progenitor cell differentiation $q_{pc,ent}(\cdot)$ is positively correlated to height in the crypt.
- The rate of enterocyte death $q_{ent,\emptyset}(\cdot)$ is positively correlated to height in the crypt and local cell density.

More precisely, we define the local density of the population \mathbf{d} as

$$\mathbf{d}(\nu_t, z) = d(z)\nu_t * D_a(z) \quad (54)$$

where D_a is defined in Eq.(50) and

$$d(z) = \frac{\int_{-a}^a D_a(u) du}{\int_{\max(-a, -z)}^{\min(a, z_{max}-z)} D_a(u) du} \quad (55)$$

is a rescaling function ensuring that the density kernel sums to 1 even near the boundaries. As $a \ll z_{max}$, $1 \leq d(z) \leq 2$.

We also define the local concentration of butyrate \mathbf{c}_b as

$$\mathbf{c}_b(c_b, z) = c_b * \psi_a(z) \quad (56)$$

where $\psi_a(z)$ is introduced in Eq. (51). Eventually, for a cell at $x := (z, l) \in \mathcal{X}$, in a population $\nu \in \mathcal{M}_{F^+}(\mathcal{X})$ and for butyrate and oxygen concentration $c := (c_o, c_b)$, we define the individual jump rate of progenitor cell division by

$$\begin{aligned} q_{pc,pc}(x, \nu * D(z), c * \psi_a(z)) = & \underbrace{0 \text{ if not prog. cell}}_{\mathbf{1}\{l = pc\}} \times \underbrace{(1 - R(z, K_{pc,pc}[z], \kappa_{pc,pc}[z]))}_{\text{inhibition by } z} \\ & \times \underbrace{(1 - R(\mathbf{d}(\nu_t, z), K_{pc,pc}[dens], \kappa_{pc,pc}[dens]))}_{\text{inhibition by density}} \\ & \times \underbrace{(1 - R(\mathbf{c}_b(c_b, z), K_{pc,pc}[but], \kappa_{pc,pc}[but]))}_{\text{inhibition by butyrate}} \end{aligned} \quad (57)$$

with $K_{pc,pc}[\cdot]$ and $\kappa_{pc,pc}[\cdot]$ strictly positive constants and $R(y, K, \kappa)$ is the piecewise polynomial function (plotted on Fig. 1b for several values of K , and κ) defined by

$$R(y, K, \kappa) := \begin{cases} 0 & \text{if } y \leq K - \kappa, \\ -\frac{1}{4\kappa^3}y^3 + \frac{3K}{4\kappa^3}y^2 - \frac{3K^2 - 3\kappa^2}{4\kappa^3}y + \frac{K^3 + 2\kappa^3 - 3K\kappa^2}{4\kappa^3} & \text{if } K - \kappa \leq y \leq K + \kappa, \\ 1 & \text{if } K + \kappa \geq y. \end{cases}$$

Note that R connects continuously the values 0 and 1 for $K - \kappa < z < K + \kappa$ and has value 1/2 in $y = K$.

We see that the jump rate depends on a product of three regulation terms involving the function R that corresponds to the three biological regulation of progenitor cell division given in table 1. Using the same construction, we define functions $q_{pc,ent}$ and $q_{ent,\emptyset}$ based on biological regulations given in table 1. For a positive regulation, we use a term $R(\cdot)$ instead of $1 - R(\cdot)$. And no regulation correspond to a constant term equal to 1.

The values of parameters used for simulation are given in table 2.

Jump type	Jump index ($k \in \mathcal{E}$)	Regulation pathway (j)		
		butyrate [but]	position [z]	density [$dens$]
Progenitor cell division	pc, pc	−	−	−
Progenitor cell differentiation	pc, ent	+	+	\emptyset
Enterocyte death	ent, \emptyset	\emptyset	+	+

Table 1: **Enumeration of stochastic jump types and their corresponding regulation.** For the jump k and a given regulation pathway j , a negative regulation (−) is represented by the term $(1 - R(y, K_k[j], \kappa_k[j]))$, a positive regulation (+) by the term $R(y, K_k[j], \kappa_k[j])$ and a neutral regulation (\emptyset) by a constant value equal to 1, where $y = z$, $y = \mathbf{d}(\nu_t, z)$ or $y = \mathbf{c}_b(c_b, z)$ for respectively position, density or butyrate regulation. The values of $K_k[j]$ and $\kappa_k[j]$ can be found in table 2.

Initial conditions Initial conditions for the position of the cells in the stochastic process are sampled in distributions ρ_{pc} and ρ_{ent} defined in eqs. (79)-(80), in D with a Monte-Carlo sampling, hence guarantying the weak convergence of the rescaled initial condition towards ρ_{pc} and ρ_{ent} (see next section).

Furthermore, initial conditions for chemicals are defined : $c_b(0, z) = c_{b0}(z)$ and $c_o(0, z) = c_{o0}(z)$, for smooth c_{b0} and c_{o0} functions defined in eq. (81).

Rescaled parameters. To illustrate the convergence of the stochastic model, we used the following rescaled parameters:

$$F_a^N = \frac{F_a}{N}, \quad D_a^N = \frac{D_a}{N}, \quad \gamma^N = \frac{\gamma}{N}.$$

where N is a multiplicative factor applied to the progenitor and enterocyte cell populations included in the initial condition.

6.2 Deterministic model obtained as a limit of the PDMP: SA equation

Since the PDMP model definition complies with the hypothesis 2, 7, 8, 9, 12, 14 and 20, according to Th. 21 and formula 25, the deterministic limit $\xi_t^{(SA)}$ of the PDMP model is the weak solution of the following system, here written in the strong form:

$$\partial_t \xi_t^{(SA)} + \nabla_z \left(\mathcal{F}(\cdot, \xi_t^{(SA)} * F_a) \xi_t^{(SA)} \right) = \sum_{k \in \mathcal{E}} \sum_{l \in \mathcal{T}} \int_{\mathcal{Z}} \mu_k(z, l) q_k^\infty q_k((z, l), \xi_t^{(SA)} * D_a(z), c_t * \psi_a(z)) \xi_t^{(SA)}(dz, l). \quad (58)$$

This equation can be interpreted as a non-linear advection-reaction equation, that will be termed self-aggregation (SA) equation in the sequel. Note that the non linear transport term $\nabla_z (\mathcal{F}(\cdot, \xi_t * F_a) \xi_t)$ comes from the cell interaction model of the deterministic flow of the PDMP model. The reaction term is a deterministic approximation of the stochastic terms of the PDMP.

For the metabolites $m \in \{\text{but, oxy}\}$, the same PDE equations as in the PDMP model are kept, except for the reaction term that now includes the density $\xi_t^{(SA)}$

$$\partial_t c_m^{(SA)} - \sigma \Delta c_m^{(SA)} = \sum_{l \in \mathcal{T}} \gamma_{ml} (c^{(SA)}) \xi_t^{(SA)l} * \psi_a(z). \quad (59)$$

This model is supplemented by a deterministic version of the boundary conditions

$$\begin{aligned} \mathcal{F}(\cdot, \xi_t^{(SA)} * F) \xi_t^{(SA)}(0, t) &= 0 \\ c_m^{(SA)}(z, t) &= f_m(z, t) \text{ for } z \in \Gamma, \\ \xi_0^{(SA)}(\cdot, 0) &= (\rho_{pc}, \rho_{ent}) \\ c_m^{(SA)}(\cdot, 0) &= c_{0m}, \end{aligned}$$

6.3 Numerical implementation

Simulations are performed with dedicated Python scripts available here <https://forgemia.inra.fr/leo.darrigade/cvg>. To solve the PDMP, we use a standard Gillespie algorithm with an acceptance-reject method for stochastic events, while the deterministic flows are solved with an explicit Euler method for the cell movement, and with a semi-implicit Euler scheme with operator splitting for the reaction-diffusion equations on butyrate and oxygen. To solve the deterministic limit, we use a semi-implicit scheme for the non-linear diffusion term and a Runge-Kutta4 scheme for time integration. Scripts use NumPy[Har+20], Scipy[Vir+20] sparse algebra and Numba [LPS15] packages for additional code acceleration.

6.4 Numerical assessment of the convergence

As the deterministic limit of cell densities $\xi_t^{(SA)}$ is actually continuous in time (see proof of lemma 25), the weak convergence of $\bar{\nu}_t^N$ to $\xi_t^{(SA)}$ at a fixed time t is true for almost every t . Then, for a continuous and bounded function $f : \mathcal{X} \rightarrow \mathbb{R}$, $\langle \bar{\nu}_t^N, f \rangle$ converges weakly to $\langle \xi_t^{(SA)}, f \rangle$. From uniform integrability of the sequence $(\langle \bar{\nu}_t^N, f \rangle)_N$, as in the proof of lemma 25, we get that

$$\mathbb{E} [\langle \bar{\nu}_t^N, f \rangle] \rightarrow_N \langle \xi_t^{(SA)}, f \rangle. \quad (60)$$

We now illustrate this convergence using some specific functions f .

6.4.1 Total populations

First, taking $f = 1$, the quantity $\mathbb{E} [\langle \bar{\nu}_t^N, f \rangle]$ and $\langle \xi_t^{(SA)}, f \rangle$ represent the total cell population at time t in the PDMP and the SA equation, respectively. Fig. 2 displays the evolution of the total populations in the PDMP (mean and quantiles 2.5% and 97.5% for $n_r = 30$ replicates) when the population scaling $N = 60$. We can observe that the progenitor cells of the SA equation is in the range of the PDMP outputs, while the total number of enterocytes is over-estimated by the SA deterministic limit by 10% at $t = 29h$. This discrepancy could be due to the effect of death events occurring in the upper part of the crypt near z_{max} . This results in lower cell populations in this region, for which the convergence towards the deterministic limit is very slow.

The inaccuracies in total enterocyte populations are propagated to the evaluation of the total number of cells (progenitor and enterocyte cells).

6.4.2 Convergence for growing N

Let us first introduce, for $(z_a, z_2) \in (0, z_{max})^2$ and a scaling parameter $\alpha > 0$ the kernel

$$\varphi(z_1, z_2 | \alpha) = \begin{cases} \exp\left(\frac{-1}{1 - \left(\frac{z_1 - z_2}{\alpha}\right)^2}\right) & \text{if } |z_1 - z_2| < \alpha \\ 0 & \text{otherwise} \end{cases}$$

and, for $z \in (0, z_{max})$, we define the function ϕ_z indexed by z so that

$$\phi_z(x) := \varphi(z, x) \quad \text{for any } x \in (0, z_{max}).$$

The functions ϕ_z allow to test the PDMP and SA solutions in different zones of the crypt.

We finally define the residual for cell populations with

$$\mathcal{R}^N(t, z) := \frac{1}{2} \sum_{l \in \{\text{pc}, \text{ent}\}} \left| \mathbb{E} \left[\langle \bar{\nu}_{l,t}^N - \xi_{l,t}^{(SA)}, \phi_z \rangle \right] \right| \quad (61)$$

where the true expectation $\mathbb{E}[\cdot]$ is approximated by the mean over n_r replicates of the PDMP, and $\bar{\nu}_{l,t}^N = \frac{1}{N} \nu_{l,t}^N$ is the rescaled solution of the PDMP for the rescaling parameter N and the population $l \in \{\text{pc}, \text{ent}\}$ at time $t > 0$.

According to the convergence theorem, we must have for any $t > 0$ and $z \in (0, z_{max})$

$$\mathcal{R}^N(z, t) \rightarrow_N 0$$

We display the curves $N \mapsto \mathcal{R}^N(t, z)$ ($N \in \{15, 20, 25, 30, 35, 40, 45, 50, 55, 60\}$) for several t (8 h, 16 h, 24 h, 29 h) and several z (34 μm , 42 μm , 59 μm , 68 μm), either grouped by t (Figure 3) or

by z (Figure 4). In all curves, we observe a consistent reduction in residuals between the PDMP and SA equation as N increases, despite some local fluctuations. The convergence rate is relatively slow : an average calculation shows that, with a fourfold increase in the population scaling parameter (from $N = 15$ to $N = 60$), residuals decrease by an average factor of 0.86.

We note that at a given time, this decrease follows a common pattern for every $z \in \{34, 42, 59, 68\}$, since the curves are nearly parallel in Figure 3. However, convergence seems to be faster for higher z , since red curves ($z = 68$) and green curves ($z = 59$) have the steepest decrease in the four panels of Figure 3. At a given z (Figure 4), the decrease of residuals tends to be more stochastic than for the curves grouped by times (Figure 3). The convergence rate is lower at the end of the simulation, for $t = 29$ h, but no general trend can be deciphered for earlier times.

The residual $\mathcal{R}^N(z, t)$ is a metric that only focuses on cell populations. We now look at the convergence of the PDMP towards the SA equation for the metabolite concentrations. In Figure 5, the concentration distributions $z \mapsto c_o^N$ and $z \mapsto c_b^N$ obtained with the PDMP for different scaling parameters $N \in \{15, 20, 50, 60\}$ are displayed at $t = 29$ h together with the functions $z \mapsto c_o^{(SA)}$ and $z \mapsto c_b^{(SA)}$ computed with the deterministic SA equation. We can see that the concentration maps are very similar whatever the value of N . The SA equation renders very well the butyrate distribution, which has the main impact on the cell population dynamics through butyrate regulation of growth rates (see eq.(57)). Comparatively, discrepancies are higher for oxygen distribution, specially at the top of the crypt ($z = 100 \mu\text{m}$), probably for the same reasons as above: the very slow convergence towards the asymptotic cell density in this region, with an underestimation of the enterocytes, is reflected in the underestimation of the oxygen consumption. Note that the Dirichlet condition at the top of the crypt renders this effect invisible for the butyrate concentration.

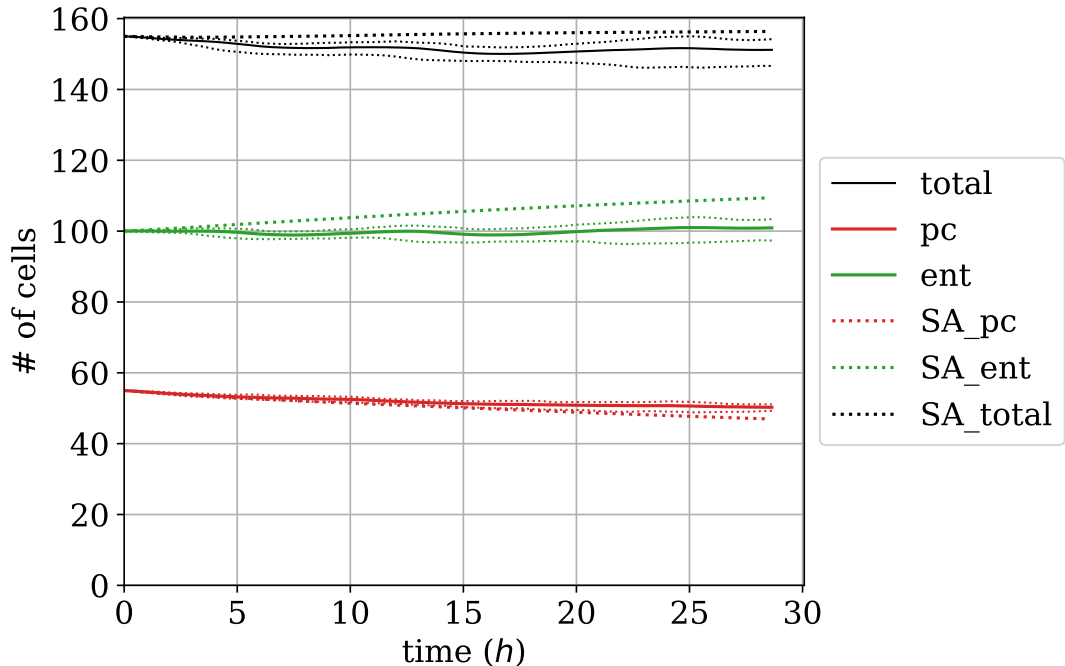


Figure 2: **Total populations over time** : The total population of progenitor cell and enterocytes are plotted over time for the PDMP (with population scaling = 60) and the SA equation. The averaged over 30 PDMP repetitions are plotted (continuous line) together with quantiles 2.5% and 97.5% (thin dots). The deterministic self-aggregation limit is plotted with thick dots.

7 Conclusive discussion

In this study, we introduced a generic model of epithelial tissues interacting with diffusive compounds, in a rigorous PDMP framework. We determined the deterministic limit of the PDMP model, which is a self-aggregation like equation. We then proved the convergence of the IBM model towards this limit

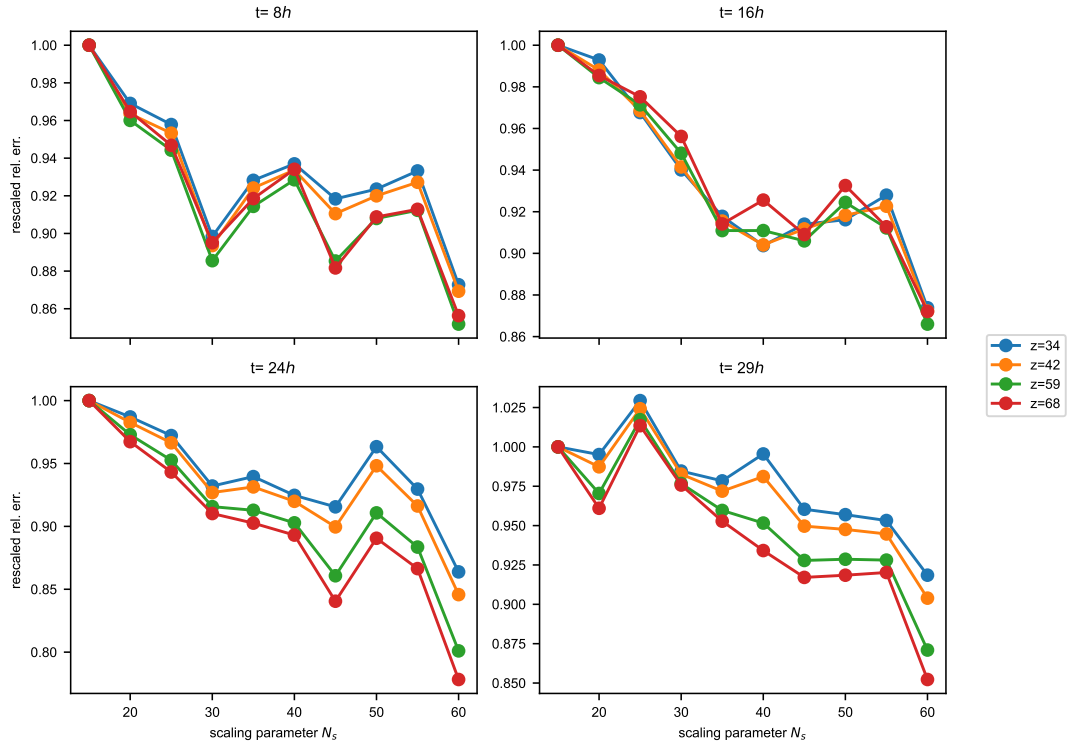


Figure 3: **Evolution of the residual $\mathcal{R}^N(z, t)$ for increasing N at different t and z , grouped by times:** we display for different times (8h, upper left; 16h, upper right; 24h, lower left and 29h, lower right) and different z (34 μm , 42 μm , 59 μm , 68 μm) the evolution of the residual for growing N ($N \in \{15, 20, 25, 30, 35, 40, 45, 50, 55, 60\}$), grouped by times. To visualize all the graphs in a common plot, we normalized the residuals by the value of the residual for $N = 15$: we then plot the functions $N \mapsto \mathcal{R}^N(z, t)/\mathcal{R}^{15}(z, t)$.

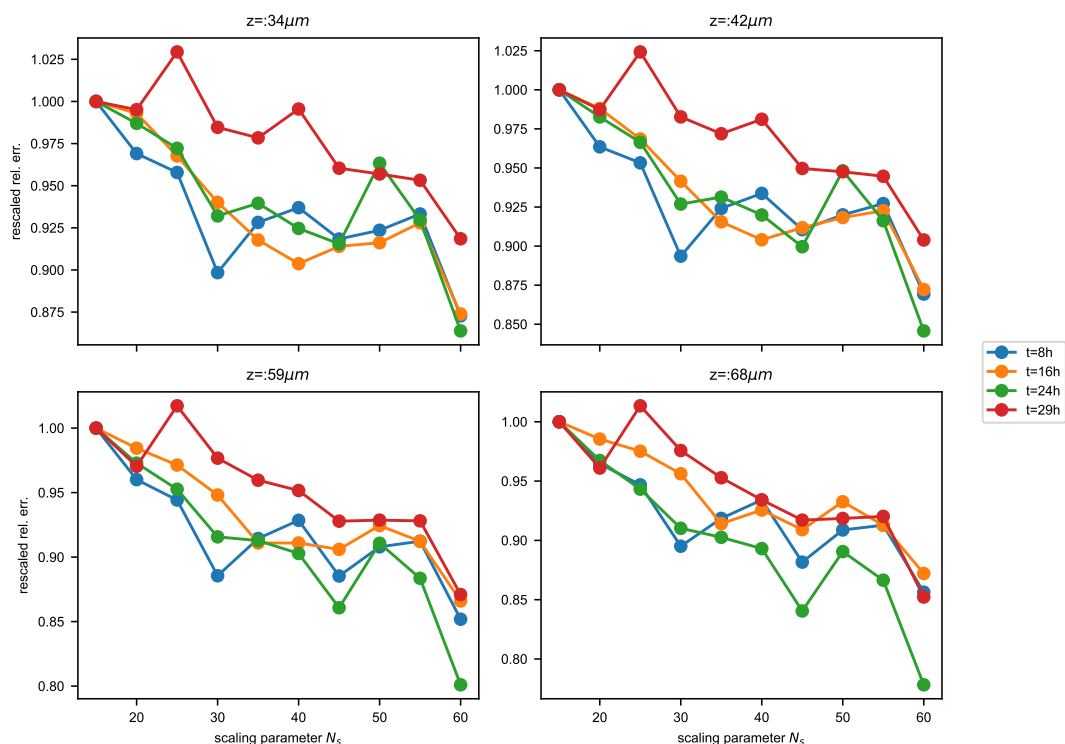


Figure 4: **Evolution of the residual $\mathcal{R}^N(z, t)$ for increasing N at different t and z , grouped by z :** we display for different times (8h, 16h, 24h, 29h) and different z ($34\mu\text{m}$, upper left; $42\mu\text{m}$, upper right; $59\mu\text{m}$, lower left; $68\mu\text{m}$ lower right) the evolution of the residual for growing N ($N \in \{15, 20, 25, 30, 35, 40, 45, 50, 55, 60\}$), grouped by z . To visualize all the graphs in a common plot, we normalized the residuals by the value of the residual for $N = 15$: we then plot the functions $N \mapsto \mathcal{R}^N(z, t)/\mathcal{R}^{15}(z, t)$

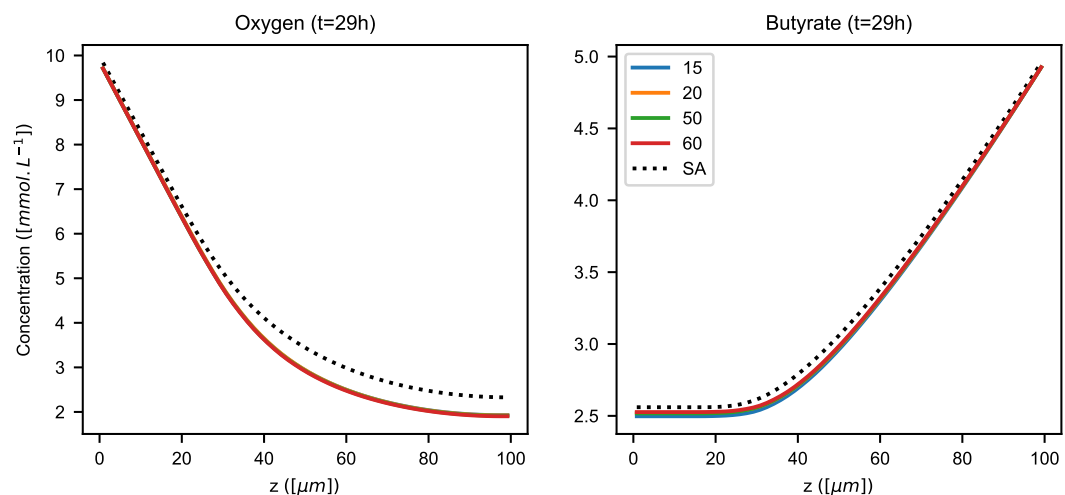


Figure 5: **Metabolite concentrations :** the distribution of oxygen (left) and butyrate (right) concentrations at $t = 29h$ is plotted for the PDMP (with population scalings = 15,20,50 and 60) and the self-aggregation equation (dotted lines). The mean over 30 PDMP repetitions are plotted.

under a large population assumption. This convergence is finally checked numerically by solving both the deterministic limit model and the PDMP model for several population scaling parameters.

This work is of special interest for modelers of epithelial tissues. It effectively links a PDMP IBM and its deterministic, continuous population counterpart, combining the strength of both modeling frameworks for epithelial tissues modeling. Micro-scale data like microscopy imaging can be used to calibrate the IBM model, while the deterministic limit allows upscaling this calibration to the continuous population model. Conversely, continuous population model parameters inferred from macro-scale data like metabolomics can be down-scaled to the IBM. Following the same idea, a PDMP IBM model can be used to model accurately intra-specific distribution in cell populations, but the deterministic limit would be more appropriate for intensive numerical exploration such as sensitivity analysis. Such an approach has been previously used to analyse a PDMP model of the cell turn-over in the intestinal crypt [Dar+22].

The main achievement of this work is the rigorous derivation of the deterministic limit and the proof of the convergence of the PDMP model towards this limit. Although the general functional analysis framework is well established, in particular in the theoretical ecology and evolution community [FM04; CFM07; Tra06], technical issues had to be overcome in the specific context of epithelial tissues. Notably, dealing with cell-to-cell interactions introduced non-linearities in the deterministic limit, and complex regulatory processes in epithelial cell turnover added technical difficulties in proofs.

The main limitation of our study is the lack of convergence rate estimates. We could check numerically that in our test case, this convergence rate is relatively slow. Consequently, we have no guarantee that the number of cells observed in real tissues is in the convergence regime. Most probably, at the crypt scale, cell numbers might be too low for stochastic events to approximate the "average process" accurately, suggesting the importance to retain stochasticity at this scale to model population distribution around the mean. However, at the tissue scale, with numerous crypts, the large population assumption could be more valid, making the deterministic limit accurate. In conclusion, modelers should not avoid a discussion of the number of cells in the tissue of interest before choosing between a PDMP or a continuous population model, bearing in mind that connections can be made between these formalisms.

A Metrics on $\mathcal{M}_{F^+}(\mathcal{X})$, properties of $\mathcal{D}([0, T], (\mathcal{M}_{F^+}(\mathcal{X}), d_{PR}))$

We remind the reader that (\mathcal{X}, d) , with d defined at (2) is a complete separable metric space (Polish space). For a closed set F of \mathcal{X} and $\epsilon \geq 0$, let $F_\epsilon = \{x \in \mathcal{X} : d(x, F) < \epsilon\}$. The Prokhorov distance d_{PR} on $\mathcal{M}_{F^+}(\mathcal{X})$ is defined by (Appendix A2.5 of [DVon])

$$d_{PR}(\mu, \nu) = \inf \{ \epsilon \geq 0 : \text{for all closed } F \subset \mathcal{X}, \mu(F) \leq \nu(F_\epsilon) + \epsilon \text{ and } \nu(F) \leq \mu(F_\epsilon) + \epsilon \}. \quad (62)$$

The first important property of the Prokhorov distance is that it metrizes the weak convergence topology on $\mathcal{M}_{F^+}(\mathcal{X})$ (Corollary A2.5.II in [DVon]) and this is why we use it. Another useful property is that, if (X, d) is Polish, the space $(\mathcal{M}_{F^+}(\mathcal{X}), d_{PR})$ is Polish too (Proposition A.2.5.III of [DVon]).

On $\mathcal{M}_F(\mathcal{X})$, we also define the total variation norm:

$$\begin{aligned} \|\mu\|_{TV} &= \mu^+(\mathcal{X}) + \mu^-(\mathcal{X}) \\ &= \sup_{f \text{ measurable}, \|f\|_\infty \leq 1} \left\{ \int_{\mathcal{X}} f \, d\mu \right\} \end{aligned} \quad (63)$$

where the couple μ^+, μ^- is the Hahn-Jordan decomposition of μ .

Eventually, we will use the flat distance (or Kantorovitch-Rubinstein distance) defined on $\mathcal{M}_{F^+}(\mathcal{X})$ as

$$d_F(\nu, \xi) = \sup_{f \in \mathcal{L}ip_{1,1}(\mathcal{X}, \mathbb{R})} \langle \nu - \xi, f \rangle. \quad (64)$$

We now remind (and prove some) useful properties on these metrics.

Proposition 27 (Theorem 1.2 p.96 in [EK86]). *Let μ and ν be two probability measures on \mathcal{X} . Let $\Pi(\mu, \nu)$ be the set of probability measures on $\mathcal{X} \times \mathcal{X}$ with marginal distributions μ and ν . Then*

$$d_{PR}(\mu, \nu) = \inf_{\gamma \in \Pi(\mu, \nu)} \inf \{ \epsilon > 0 : \gamma\{(x, y) \in \mathcal{X}^2 : d(x, y) \geq \epsilon\} \leq \epsilon \}. \quad (65)$$

Once a metric is defined on $\mathcal{M}_{F^+}(\mathcal{X})$, we can define the Skorokhod metric (see [Bil99]) on the space $\mathcal{D}([0, T], (\mathcal{M}_{F^+}(\mathcal{X}), d_{PR}))$ of càdlàg functions of $[0, T]$ to $\mathcal{M}_{F^+}(\mathcal{X})$. The space $\mathcal{D}([0, T], (\mathcal{M}_{F^+}(\mathcal{X}), d_{PR}))$ equipped with the Skorokhod metric inherits the Polish space property from $(\mathcal{M}_{F^+}(\mathcal{X}), d_{PR})$. Polish spaces are well suited to study weak convergence of random variable.

Proposition 28. *Let $(\nu^n)_{n \in \mathbb{N}}$ be a sequence which converges to ν in $\mathcal{D}([0, T], \mathcal{M}_{F^+}(\mathcal{X}))$. We suppose that $\nu \in \mathcal{C}([0, T], \mathcal{M}_{F^+}(\mathcal{X}))$. Then, $(\nu^n)_n$ converges to ν uniformly in time:*

$$\sup_{0 \leq t \leq T} d_{PR}(\nu_t^n, \nu_t) \xrightarrow{n} 0.$$

Proof. From the definition of the Skorokhod metric (Prop. 5.3 p.119 in [EK86]), $\nu_n \rightarrow \nu$ iff

1. there exists a family of functions $(\lambda_n)_n$ strictly increasing from $[0, T]$ to $[0, T]$, with $\lambda_n(0) = 0$ and $\lambda_n(T) = T$, and such that

$$\lim_{n \rightarrow \infty} \sup_{0 \leq t \leq T} |t - \lambda_n(t)| = 0.$$

2. And

$$\lim_{n \rightarrow \infty} \sup_{0 \leq t \leq T} d_{PR}(\nu_t^n, \nu_{\lambda_n(t)}) = 0.$$

Therefore:

$$\sup_{0 \leq t \leq T} d_{PR}(\nu_t^n, \nu_t) \leq \sup_{0 \leq t \leq T} d_{PR}(\nu_t^n, \nu_{\lambda_n(t)}) + \sup_{0 \leq t \leq T} d_{PR}(\nu_{\lambda_n(t)}, \nu_t).$$

But, when $n \rightarrow \infty$, the first term of the right hand side goes to 0 from the definition of convergence in the Skorokhod space, and the second term goes also to 0 from the uniform continuity in time of ν on $[0, T]$. \square

Proposition 29. *Let $(\nu^n)_{n \in \mathbb{N}}$ be a sequence which converges to ν in $\mathcal{D}([0, T], \mathcal{M}_{F^+}(\mathcal{X}))$. Then, $(\nu_t^n)_n$ converges to ν_t for almost every $t \in [0, T]$, that is to say*

$$d_{PR}(\nu_t^n, \nu_t) \xrightarrow{n} 0 \quad t - a.e..$$

Proof. ν has at most a countable set of time discontinuities, written A (lemma 5.1 in [EK86]). Then, by definition of the Skorokhod metric (see Prop. 28), $\forall t_0 \in [0, T] \setminus A$,

$$d_{PR}(\nu_{t_0}^n, \nu_{t_0}) \leq d_{PR}(\nu_{t_0}^n, \nu_{\lambda_n(t_0)}) + d_{PR}(\nu_{\lambda_n(t_0)}, \nu_{t_0}) \xrightarrow{n} 0.$$

Indeed, the first term of the right hand side goes to 0 by definition of the convergence in Skorokhod distance, and the second as well as ν_t is continuous at t_0 . \square

Proposition 30. *The total variation distance is greater than the Prokhorov distance on $\mathcal{M}_{F^+}(\mathcal{X})$.*

Proof. For all $\mu \in \mathcal{M}_F(\mathcal{X})$, we admit that (and this can be simply proven using the Hahn-Jordan decomposition of μ):

$$\|\mu\|_{TV} \leq 2 \sup_{B \in \mathcal{B}(\mathcal{X})} \{|\mu(B)|\} \leq 2\|\mu\|_{TV}, \quad (66)$$

with $\mathcal{B}(\mathcal{X})$ the Borel sets. Let F be a closed set of \mathcal{X} , μ and $\nu \in \mathcal{M}_{F^+}(\mathcal{X})$,

$$\begin{aligned} \mu(F) &= \mu(F) - \nu(F) + \nu(F) \leq \sup_{B \in \mathcal{B}(\mathcal{X})} \{|\mu(B) - \nu(B)|\} + \nu(F) \\ &\leq \frac{1}{2} \|\mu - \nu\|_{TV} + \nu(F_{\frac{1}{2}\|\mu - \nu\|_{TV}}) \end{aligned}$$

where we used the fact that ν is a positive measure. We can switch μ and ν in those expressions and we obtain that

$$d_{PR}(\mu, \nu) \leq \frac{1}{2} \|\mu - \nu\|_{TV}. \quad \square$$

Proposition 31. *Let $\nu \in \mathcal{M}_F(\mathcal{X})$,*

$$\|\nu\|_{TV} = \sup \{ \langle \nu, f \rangle; f \in \mathcal{C}(\mathcal{X}), \|f\|_\infty \leq 1 \}. \quad (67)$$

Proof. As ν is finite $L^\infty(\mathcal{X}, |\nu|) \subset L^1(\mathcal{X}, |\nu|)$, and continuous functions are dense in $L^1(\mathcal{X}, |\nu|)$. \square

Proposition 32. *Let $\alpha \in \mathbb{R}^+$, $\nu \in \mathcal{M}_{F^+}(\mathcal{X})$,*

$$d_{PR}(\alpha\nu, \nu) \leq |\alpha - 1| \times \|\nu\|_{TV}.$$

Proof. Suppose that $0 \leq \alpha \leq 1$. Let $\epsilon \geq 0$, F a closed set of \mathcal{X} . Then

$$\alpha\nu(F) \leq \nu(F^\epsilon) + \epsilon$$

as ν is a positive measure. Moreover

$$\nu(F) - \alpha\nu(F^\epsilon) \leq \nu(F) - \alpha\nu(F) \leq (1 - \alpha)\|\nu\|_{TV}.$$

Therefore

$$d_{PR}(\alpha\nu, \nu) \leq (1 - \alpha)\|\nu\|_{TV}.$$

Now suppose that $\alpha \geq 1$. Then $d_{PR}(\alpha\nu, \nu) = d_{PR}(\mu, \frac{1}{\alpha}\mu)$ with $\mu := \alpha\nu$. Using what we have shown above:

$$d_{PR}(\alpha\nu, \nu) \leq (1 - \frac{1}{\alpha})\|\mu\|_{TV} = (1 - \frac{1}{\alpha})\alpha\|\nu\|_{TV} = (\alpha - 1)\|\nu\|_{TV}.$$

This completes the proof. \square

Proposition 33. *Let $\alpha > 0$ and $\nu, \mu \in \mathcal{M}_{F^+}(\mathcal{X})$.*

$$d_{PR}(\alpha\mu, \alpha\nu) \leq \max(\alpha, \frac{1}{\alpha})d_{PR}(\mu, \nu).$$

Proof. Let F be a closed set of \mathcal{X} .

$$\begin{aligned} \nu(F) &\leq \mu(F^{d_{PR}(\mu, \nu)}) + d_{PR}(\mu, \nu) \\ \Leftrightarrow \alpha\nu(F) &\leq \alpha\mu(F^{d_{PR}(\mu, \nu)}) + \alpha d_{PR}(\mu, \nu). \end{aligned}$$

The second inequality implies that

$$\alpha\nu(F) \leq \alpha\mu(F^{\max(\alpha, \frac{1}{\alpha})d_{PR}(\mu, \nu)}) + \max(\alpha, \frac{1}{\alpha})d_{PR}(\mu, \nu).$$

From the definition of the Prokhorov metric, this concludes the proof. \square

Proposition 34. *Let $\nu \in \mathcal{M}_{F^+}(\mathcal{X})$ and $(\nu^n)_{n \in \mathbb{N}}$ a sequence of $\mathcal{M}_{F^+}(\mathcal{X})$ such that $d_{PR}(\nu, \nu^n) \xrightarrow{n} 0$. Let $G : \mathbb{R}^d \rightarrow \mathbb{R}$ Lipschitz bounded. Then,*

$$\sup_{z \in \mathcal{Z}} |(\nu - \nu^n) * G(z)| \xrightarrow{n} 0.$$

Proof. We remind the reader that if $d_{PR}(\nu^n, \nu) \xrightarrow{n} 0$ then $\|\nu^n\|_{TV} \xrightarrow{n} \|\nu\|_{TV}$, with $\|\nu\|_{TV} = \nu(\mathcal{X})$ when $\nu \in \mathcal{M}_{F^+}(\mathcal{X})$. Indeed, if $d_{PR}(\nu, \nu^n) \leq \epsilon$, from the definition (62), $|\nu(\mathcal{X}) - \nu^n(\mathcal{X})| \leq \epsilon$.

We start by treating the case when ν is the null measure. Then, $\forall z \in \mathcal{Z}$,

$$\begin{aligned} |(\nu - \nu^n) * G(z)| &\leq |\nu * G(z)| + |\nu^n * G(z)| \leq \|G\|_\infty \|\nu^n\|_{TV} \\ &\xrightarrow{n \rightarrow \infty} 0. \end{aligned}$$

Let us suppose now that $\|\nu\|_{TV} > 0$. As $\|\nu^n\|_{TV} \rightarrow \|\nu\|_{TV}$, we also suppose that $\|\nu^n\|_{TV} > 0 \forall n \in \mathbb{N}$. We want to use proposition 27 p.29 which is valid for probability measures. We have that

$$d_{PR}(\nu, \nu^n) \xrightarrow{n} 0 \quad \Rightarrow \quad d_{PR}\left(\frac{\nu}{\|\nu\|_{TV}}, \frac{\nu^n}{\|\nu^n\|_{TV}}\right) \xrightarrow{n} 0. \quad (68)$$

Indeed:

$$d_{PR}\left(\frac{\nu}{\|\nu\|_{TV}}, \frac{\nu^n}{\|\nu^n\|_{TV}}\right) \leq d_{PR}\left(\frac{\nu}{\|\nu\|_{TV}}, \frac{\nu^n}{\|\nu\|_{TV}}\right) + d_{PR}\left(\frac{\nu^n}{\|\nu\|_{TV}}, \frac{\nu^n}{\|\nu^n\|_{TV}}\right).$$

The first term of the right hand side goes to 0 by proposition 33. The second term of the right hand side goes to 0 by proposition 32. Let ϵ such that $0 < \epsilon$. From what was shown above, there exists N_ϵ such that, if $n \geq N_\epsilon$,

$$d_{PR} \left(\frac{\nu}{\|\nu\|_{TV}}, \frac{\nu^n}{\|\nu^n\|_{TV}} \right) \leq \epsilon \quad \text{and} \quad \left| 1 - \frac{\|\nu^n\|_{TV}}{\|\nu\|_{TV}} \right| \leq \epsilon.$$

Let $n \geq N_\epsilon$. Let $G^+(z)$ and $G^-(z)$ be two positive functions such that $G(z) = G^+(z) - G^-(z)$.

$$|(\nu^n - \nu) * G(z)| \leq |(\nu^n - \nu) * G^+(z)| + |(\nu^n - \nu) * G^-(z)|.$$

Then,

$$\begin{aligned} |(\nu - \nu^n) * G^+(z)| &= \|\nu\|_{TV} \times \left| \left(\frac{\nu}{\|\nu\|_{TV}} - \frac{\|\nu^n\|_{TV}}{\|\nu\|_{TV}} \times \frac{\nu^n}{\|\nu^n\|_{TV}} \right) * G^+(z) \right| \\ &\leq \|\nu\|_{TV} \times \left| \left(\frac{\nu}{\|\nu\|_{TV}} - \frac{\nu^n}{\|\nu^n\|_{TV}} \right) * G^+(z) \right| + \|\nu\|_{TV} \times \left| \left(1 - \frac{\|\nu^n\|_{TV}}{\|\nu\|_{TV}} \right) \times \frac{\nu^n}{\|\nu^n\|_{TV}} * G^+(z) \right| \\ &\leq \|\nu\|_{TV} \times \left| \left(\frac{\nu}{\|\nu\|_{TV}} - \frac{\nu^n}{\|\nu^n\|_{TV}} \right) * G^+(z) \right| + \|\nu\|_{TV} \epsilon \|G^+\|_\infty \end{aligned}$$

Let $\hat{\nu}^n := \frac{\nu^n}{\|\nu^n\|_{TV}}$ and $\hat{\nu} := \frac{\nu}{\|\nu\|_{TV}}$, which are probability measures on \mathcal{X} . From proposition 27 p.29, $\forall \delta$ such that $\epsilon < \delta$, there exists a probability measure on $\mathcal{X} \times \mathcal{X}$ $\gamma \in \Pi(\hat{\nu}^n, \hat{\nu})$ and a set $E \supset \{(x, y) \in \mathcal{X}^2 : d(x, y) \geq \delta\}$ such that $\gamma(E) \leq \delta$. Then

$$\begin{aligned} (\hat{\nu}^n - \hat{\nu}) * G^+(z) &= \sum_{l \in \mathcal{T}} \int_{\mathcal{Z}} G^+(z - y) (\hat{\nu}^n - \hat{\nu})(dy, dl) \\ &= \sum_{l \in \mathcal{T}} \int_{\mathcal{Z}} G^+(z - y) \hat{\nu}^n(dy, dl) - \sum_{k \in \mathcal{T}} \int_{\mathcal{Z}} G^+(z - u) \hat{\nu}(du, dk) \\ &= \sum_{l \in \mathcal{T}} \int_{\mathcal{Z}} \sum_{k \in \mathcal{T}} \int_{\mathcal{Z}} G^+(z - y) - G^+(z - u) \gamma(dy, dl, du, dk) \\ &\leq 2\delta \|G^+\|_\infty + \int_{\mathcal{X} \setminus E} \|G^+\|_{Lip} \delta \gamma(dy, dl, du, dk) \\ &\leq \delta(2\|G^+\|_\infty + \|G^+\|_{Lip}). \end{aligned}$$

We can moreover assume that $\epsilon < \delta < \text{diam}(\mathcal{Z})$. Then, for all $x^1 = (z^1, l^1)$, $x^2 = (z^2, l^2) \in \mathcal{X}^2$, if $l^1 \neq l^2$ then $(x^1, x^2) \in E$. Thus, we can use the fact that G^+ is Lipschitz in space to find an upper bound for the last integral on the domain $\mathcal{X} \setminus E$:

$$\begin{aligned} &\sum_{l \in \mathcal{T}} \int_{\mathcal{Z}} \sum_{k \in \mathcal{T}} \int_{\mathcal{Z}} G^+(z - y) - G^+(z - u) \gamma(dy, dl, du, dk) \\ &\leq 2\delta \|G^+\|_\infty + \int_{\mathcal{X} \setminus E} \|G^+\|_{Lip} \delta \gamma(dy, dl, du, dk) \\ &\leq \delta(2\|G^+\|_\infty + \|G^+\|_{Lip}). \end{aligned}$$

As this is true for any $\delta > \epsilon$, we get that

$$(\hat{\nu}^n - \hat{\nu}) * G^+(z) \leq \epsilon(2\|G^+\|_\infty + \|G^+\|_{Lip}). \quad (69)$$

The same kind of upper bound can be found for $|(\nu^n - \nu) * G^-(z)|$. And so, $\forall z \in \mathcal{Z}$,

$$|(\nu^n - \nu) * G(z)| \leq 2\epsilon \|\nu\|_{TV} (3\|G\|_\infty + \|G\|_{Lip}).$$

This concludes the proof as $\epsilon \rightarrow 0$. □

B Reaction-diffusion equations

In this section we recall some classical results on linear reaction-diffusion equations with non-homogeneous boundary conditions. The main result is proposition 35 which states existence and uniqueness of solution

for the reaction-diffusion problem (73). We use this result to obtain the existence and uniqueness of solutions to the problem governing the evolution of chemical concentrations in our model (proposition 10 in the main text).

Let $U \subset \mathbb{R}^d$ be a bounded open set of regularity C^∞ . We denote ∂U the boundary of U and consider the homogeneous Robin boundary condition

$$\partial_n u|_{\partial U} + bu|_{\partial U} = 0 \quad (70)$$

where b is a function in $C^\infty(\bar{U})$ and ∂_n is the normal derivative on ∂U . It is a standard result (see [Paz12; Ama93]) that the unbounded operator $A = \sigma\Delta$ whose domain $\mathcal{D}(A)$ is the closure in $H^2(U)$ of the set of function in $C^2(\bar{U})$ which satisfy (70) generates a strongly continuous semi-group on $L^2(U)$, denoted S . Moreover, this semi-group is positive and analytical.

Let us consider the non-homogeneous elliptic problem defined on U :

$$-\lambda c + Ac = 0 \quad (71)$$

$$\partial_n c|_{\partial U} + bc|_{\partial U} = f \quad (72)$$

where $\lambda \geq 0$ is a real number greater than the largest spectral value of A . For all $\alpha \geq 0$, the problem (71) has a unique solution c in $H^{2\alpha}(U)$ when $f \in H^{2\alpha-\frac{3}{2}}(\partial U)$, denoted $c = D(f)$. Operator D is bounded from $H^{2\alpha-\frac{3}{2}}(\partial U)$ to $H^{2\alpha}(U)$. These standard results can be found in [LM72] or [Ama93].

Using notations taken from [Las80], we define the following functional spaces, for all $\alpha \geq 0$ and for $L^2(U) = H^0(U)$,

$$W_{L^2(U)}^{2\alpha, [\alpha]} = \left\{ f \in L^2([0, T], H^{2\alpha}(U)); \forall j \in \{1, \dots, [\alpha]\} \frac{\partial^j f}{\partial t^j} \in L^2([0, T], H^{2\alpha-2j}(U)) \right\}$$

where $[\cdot]$ is the ceiling function. $W_{L^2(U)}^{2\alpha, [\alpha]}$ is a Hilbert space endowed with the norm

$$\left(\|f\|_{L^2([0, T], H^{2\alpha}(U))}^2 + \sum_{j=1}^{[\alpha]} \left\| \frac{\partial^j f}{\partial t^j} \right\|_{L^2([0, T], H^{2\alpha-2j}(U))}^2 \right)^{\frac{1}{2}}.$$

Note that with these notations, $W_{L^2(U)}^{0,0} = L^2([0, T] \times U)$. We define in the same way the space $W_{L^2(\partial U)}^{2\alpha, [\alpha]}$. From [LM72] theorem 3.1 p. 19, we know that:

- if $f \in W_{L^2(\partial U)}^{\frac{3}{2}, 1}$, then $f \in L^2([0, T], H^{\frac{3}{2}}(\partial U))$ and $\frac{\partial f}{\partial t} \in L^2([0, T], H^{-\frac{1}{2}}(\partial U))$, so $f \in \mathcal{C}([0, T], H^{\frac{1}{2}}(\partial U))$ and the injection is continuous.
- Similarly, if $c \in W_{L^2(U)}^{2,1}$ then $c \in \mathcal{C}([0, T], H^1(U))$ and the injection is continuous.

With these definitions and properties, we have the following result

Proposition 35. *If $b \in C^\infty(\bar{U})$, $c_0 \in H^1(U)$, $f \in W_{L^2(\partial U)}^{\frac{1}{2}, 1}$, and $u \in W_{L^2(U)}^{0,0}$, then*

$$c(\cdot, t) = S(t)c_0 + (\lambda I - \sigma\Delta) \left(\int_0^t S(t-s)D(f)(\cdot, s)ds \right) + \int_0^t S(t-s)u(\cdot, s)ds$$

is the unique weak solution in $W_{L^2(U)}^{2,1}$ of the following evolution problem

$$\begin{aligned} \partial_t c - \sigma\Delta c &= u, \\ \partial_n c|_{[0, T] \times \partial U} + bc|_{[0, T] \times \partial U} &= f \\ c(\cdot, 0) &= c_0 \end{aligned} \quad (73)$$

This solution belongs to $W_{L^2(U)}^{2,1} \cap \mathcal{C}([0, T], H^1(U))$. Moreover, the operator $(c_0, f, u) \mapsto c$ is continuous from $\left(H^1(U) \times W_{L^2(\partial U)}^{\frac{1}{2}, 1} \times W_{L^2(U)}^{0,0} \right)$ to $W_{L^2(U)}^{2,1}$.

Proof. It is a consequence of theorem 6.1 from [Las80] for problems with mixed (Robin) boundary conditions in the case $\alpha = 1/4$ and $k = 0$. \square

Remark 36. The same type of results hold for Dirichlet boundary conditions with $f \in W_{L^2(\partial U)}^{\frac{3}{2}, 1}$ if the compatibility condition $c_0|_{\partial U} = f(0)$ is satisfied (see [Las80]).

Remark 37. It is difficult to state general results when U is non smooth or in the case of mixed Dirichlet-Neumann boundary conditions, where the analysis of each specific case is required.

C Technical results for the proof of theorem 21

Lemma 38. *Under hypotheses 2, let $(c_s, \xi_s)_{s \geq 0}$ be a solution of (25-26) on $[0, T]$, let $t \in [0, T]$ and $u^*(z) \in \mathcal{C}^1(\mathcal{Z}, \mathbb{R})$. The transport equation on $[0, T]$*

$$\begin{cases} \partial_s u(s, z) + \mathcal{F}(\cdot, \xi_s * F) \nabla_z u(s, z) = 0, \\ u(t, z) = u^*(z), \end{cases} \quad (74)$$

admits a unique solution constant along characteristics $Z(s; t, z)$:

$$u(s, z) = u^*(Z(t; s, z)), \quad (75)$$

where characteristics $Z(s; t, z)$ are solutions of

$$\begin{cases} d_s Z(s; t, z) = \mathcal{F}(Z(s; t, z), \xi_s * F(Z(s; t, z))), \\ Z(t; t, z) = z \quad \text{for } z \in \mathcal{Z}. \end{cases} \quad (76)$$

Moreover, the solution $u \in \mathcal{C}^1([0, T] \times \mathcal{Z}, \mathbb{R})$.

Proof. This is an application of theorem 6.1 in [Per07], which relies on the continuous derivability of the flow associated to the ODE (76). To prove the latter, one must show that the vector field $(t, z) \mapsto \mathcal{F}(z, \xi_t * F(z))$ is continuously derivable on $[0, T] \times \mathcal{Z}$, which is a standard consequence of the \mathcal{C}^1 regularity assumption (1) in hypothesis 2. The detailed proof can be found in A.3 p.119 of [Dar20]. \square

Let (c_t^1, ξ_t^1) and (c_t^2, ξ_t^2) be two solutions of (25-26) on $[0, T]$ with ξ_t^i positive, and let $Z^1(s; t, z)$ and $Z^2(s; t, z)$ be the associated characteristics defined in (76). Then, the following proposition is true.

Proposition 39. *Let $z \in \mathcal{Z}$ and $0 \leq s \leq t$,*

$$|Z^1(t; s, z) - Z^2(t; s, z)| \leq C(t) \int_s^t d_F(\xi_r^1, \xi_r^2) dr \quad (77)$$

with $C(t)$ a constant which depends on t .

Proof. Proof can be found in A.4 p.125 of [Dar20]. \square

D Exact expression and parameters values for the numerical implementation

Shape of the domain The function ϕ appearing in equation (45) is defined as

$$\phi(z) := \begin{cases} \frac{\frac{f(z) - f(0)}{r_0} - \frac{f(0)}{r_0}}{1 - \frac{f(0)}{r_0}} & \text{if } z \leq r_0 - \epsilon \\ 1 & \text{if } r_0 - \epsilon < z < z_{max} - r_0 + \epsilon \\ \frac{\frac{f(z_{max}) - f(z)}{r_0} - \frac{f(z)}{r_0}}{\frac{f(z_{max}) - f(z)}{r_0} - 1} & \text{if } z \geq z_{max} - r_0 + \epsilon \end{cases} \quad (78)$$

with

$$f(z) = \begin{cases} r_0 \sqrt{\frac{z+\epsilon}{r_0} \left(2 - \frac{z+\epsilon}{r_0}\right)} & \text{if } z \leq r_0 - \epsilon \\ r_0 & \text{if } r_0 - \epsilon < z < z_{max} - r_0 + \epsilon \\ r_0 \left(2 - \sqrt{\frac{z_{max}-z+\epsilon}{r_0} \left(2 - \frac{z_{max}-z+\epsilon}{r_0}\right)}\right) & \text{if } z \geq z_{max} - r_0 + \epsilon \end{cases}$$

where we assume that r_0 is small compared to z_{max} , and ϵ is small compared to r_0 .

Initial conditions Let

$$\rho_{\text{pc}}(z) = \begin{cases} \frac{1}{20} & \text{if } z \leq 20, \\ \frac{1}{20} \times \frac{30-z}{10} & \text{if } 20 < z < 30, \\ 0 & \text{if } 30 \leq z, \end{cases} \quad (79)$$

and

$$\rho_{\text{ent}}(z) = \begin{cases} 0 & \text{if } z \leq 20, \\ \frac{1}{70} \times \frac{z-20}{10} & \text{if } 20 < z < 30, \\ \frac{1}{70} & \text{if } 30 \leq z. \end{cases} \quad (80)$$

To initialize the simulations of the PDMP, we sample randomly the position of $55 \times N$ progenitor cells in ρ_{pc} and $100 \times N$ enterocytes in ρ_{ent} .

The initial conditions c_{b0} and c_{o0} are defined with

$$c_{b0}(z) = \frac{z}{z_{\text{max}}} c_{b,\text{lum}} \quad c_{o0}(z) = \frac{z_{\text{max}} - z}{z_{\text{max}}} c_{o,\text{bot}} \quad (81)$$

where $c_{b,\text{lum}}$ and $c_{o,\text{bot}}$ are the Dirichlet boundary conditions for respectively the butyrate and the oxygen (see Table 2).

Table 2: **Parameters for Sect. 6.1.** For references, see [Dar+22].

	Value	Unit	Description
z_{max}	100	μm	height of the crypt
r_0	12.5	μm	shape parameter of the crypt
ϵ	$= 0.1 \times r_0$	μm	shape parameter of the crypt
a	10	μm	cell diameter
k_a	15	$\mu\text{m} \times h^{-1}$	maximal intensity of \mathcal{F}
d_a	1	\emptyset	maximal value of D
$q_{\text{pc,pc}}^\infty$	0.11	h^{-1}	maximal rate of progenitor cell division
$K_{\text{pc,pc}}[z]$	20	μm	space regulation for progenitor cell div.
$\kappa_{\text{pc,pc}}[z]$	7.5	μm	space regulation for progenitor cell div.
$K_{\text{pc,pc}}[\text{dens}]$	20	<i>cell</i>	cell density regulation for progenitor cell div.
$\kappa_{\text{pc,pc}}[\text{dens}]$	6	<i>cell</i>	space regulation for progenitor cell div.
$K_{\text{pc,pc}}[\text{but}]$	2.5		butyrate regulation for progenitor cell div.
$\kappa_{\text{pc,pc}}[\text{but}]$	5		butyrate regulation for progenitor cell div.
$q_{\text{pc,ent}}^\infty$	0.125	h^{-1}	max. rate of progenitor cell to enterocyte
$K_{\text{pc,ent}}[z]$	$= K_{\text{pc,pc}}[z]$	μm	space regulation of progenitor cell to diff.
$\kappa_{\text{pc,ent}}[z]$	$= \kappa_{\text{pc,pc}}[z]$	μm	space regulation of progenitor cell to diff.
$K_{\text{pc,ent}}[\text{but}]$	$= K_{\text{pc,pc}}[\text{but}]$	μm	butyrate regulation of progenitor cell to diff.
$\kappa_{\text{pc,ent}}[\text{but}]$	$= \kappa_{\text{pc,pc}}[\text{but}]$	μm	butyrate regulation of progenitor cell to diff.
$q_{\text{ent},\emptyset}^\infty$	0.17	h^{-1}	maximal rate of diff. extrusion
$K_{\text{ent},\emptyset}[z]$	90	μm	space regulation for diff. extrusion
$\kappa_{\text{ent},\emptyset}[z]$	10	μm	space regulation for diff. extrusion
$K_{\text{ent},\emptyset}[\text{dens}]$	20	<i>cell</i>	cell density regulation for diff. extrusion
$\kappa_{\text{ent},\emptyset}[\text{dens}]$	$= \kappa_{\text{pc,pc}}[\text{dens}]$	<i>cell</i>	cell density regulation for diff. extrusion
s_o	-4	\emptyset	stoichiometric coefficient for O_2
s_b	-1	\emptyset	stoichiometric coefficient for butyrate
σ_o	3.6×10^4	$\mu\text{m}^2/h$	diffusion coefficient for O_2
σ_{but}	$= \sigma_o$	$\mu\text{m}^2/h$	diffusion coefficient for butyrate
γ^∞	1.6×10^2	$10^{-15} \times \frac{\text{mmol}}{h \times \mu\text{m}^2}$	max. reaction speed of β -oxydation
K	$1840^{1/5}$	<i>mM</i>	affinity of β -oxydation
$c_{b,\text{lum}}$	5	<i>mM</i>	luminal concentration of butyrate
$c_{o,\text{bot}}$	10	a.u.	O_2 concentration at the bottom of the crypt

References

- [Ama93] Herbert Amann. “Nonhomogeneous linear and quasilinear elliptic and parabolic boundary value problems”. In: *Function spaces, differential operators and nonlinear analysis*. Springer, 1993, pp. 9–126.
- [Bar14] Nick Barker. “Adult intestinal stem cells: critical drivers of epithelial homeostasis and regeneration”. In: *Nature reviews Molecular cell biology* 15.1 (2014), pp. 19–33.
- [Bat+02] Eduard Batlle et al. “ β -Catenin and TCF mediate cell positioning in the intestinal epithelium by controlling the expression of EphB/ephrinB”. In: *Cell* 111.2 (2002), pp. 251–263.
- [BNL19] Martine Ben Amar, Pierre Nassoy, and Loic LeGoff. “Physics of growing biological tissues: the complex cross-talk between cell activity, growth and resistance”. In: *Philosophical Transactions of the Royal Society A* 377.2144 (2019), p. 20180070.
- [Ben+15] Michel Benaïm et al. “Qualitative properties of certain piecewise deterministic Markov processes”. In: *Annales de l’IHP Probabilités et statistiques*. Vol. 51. 2015, pp. 1040–1075.
- [Bil99] Patrick Billingsley. *Convergence of probability measures*. New York: Wiley, 1999.
- [Bus+12] Peter Buske et al. “On the biomechanics of stem cell niche formation in the gut - Modelling growing organoids”. In: *FEBS Journal* 279.18 (2012), pp. 3475–3487. DOI: 10.1111/j.1742-4658.2012.08646.x. eprint: 84865978356.
- [CF14] Fabien Campillo and Coralie Fritsch. “Weak Convergence of a Mass-Structured Individual-Based Model”. In: *Applied Mathematics & Optimization* 72.1 (2014), pp. 37–73. DOI: 10.1007/s00245-014-9271-3.
- [CFM07] Nicolas Champagnat, Régis Ferrière, and Sylvie Méléard. “Individual-based probabilistic models of adaptive evolution and various scaling approximations”. In: *Seminar on stochastic analysis, random fields and applications V*. Springer. 2007, pp. 75–113.
- [Cop86] Sylvie Roelly- Copoletta. “A criterion of convergence of measure-valued processes: application to measure branching processes”. In: *Stochastics* 17.1-2 (1986), pp. 43–65. DOI: 10.1080/17442508608833382. eprint: <https://doi.org/10.1080/17442508608833382>.
- [CZ12] Ruth F Curtain and Hans Zwart. *An introduction to infinite-dimensional linear systems theory*. Vol. 21. Springer Science & Business Media, 2012.
- [DVon] D. J. Daley and D. Vere-Jones. *An Introduction to the Theory of Point Processes. Volume 1: Elementary Theory and Methods*. Probability and Its Applications. Springer, 2003, second edition.
- [Dar20] Léo Darrigade. “Modélisation du dialogue hôte-microbiote au voisinage de l’épithélium de l’intestin distal”. PhD thesis. Ecole Doctorale de Mathématiques Hadamard, Université Paris-Saclay, 2020.
- [Dar+22] Léo Darrigade et al. “A PDMP model of the epithelial cell turn-over in the intestinal crypt including microbiota-derived regulations”. In: *Journal of Mathematical Biology* 84.7 (2022), pp. 1–67. DOI: 10.1007/s00285-022-01766-8.
- [DNO13] Sara Jane Dunn, Inke S. Näthke, and James M. Osborne. “Computational models reveal a passive mechanism for cell migration in the crypt”. In: *PLoS ONE* 8.11 (2013), pp. 1–18. DOI: 10.1371/journal.pone.0080516.
- [Eis+12] George T Eisenhoffer et al. “Crowding induces live cell extrusion to maintain homeostatic cell numbers in epithelia”. In: *Nature* 484.7395 (2012), pp. 546–549.
- [EK86] Stewart N. Ethier and Thomas G. Kurtz. *Markov Processes: Characterization and Convergence*. Wiley Series in Probability and Its Applications. Wiley, 1986.
- [FS08] Nele Famaey and Jos Vander Sloten. “Soft tissue modelling for applications in virtual surgery and surgical robotics”. In: *Computer methods in biomechanics and biomedical engineering* 11.4 (2008), pp. 351–366.
- [FO22] Alexander G Fletcher and James M Osborne. “Seven challenges in the multiscale modeling of multicellular tissues”. In: *WIREs mechanisms of disease* 14.1 (2022), e1527.
- [FM04] Nicolas Fournier and Sylvie Méléard. “A microscopic probabilistic description of a locally regulated population and macroscopic approximations”. In: *Ann. Appl. Probab.* 14.4 (Nov. 2004), pp. 1880–1919. DOI: 10.1214/105051604000000882.

- [Har+20] Charles R. Harris et al. “Array programming with NumPy”. In: *Nature* 585.7825 (Sept. 2020), pp. 357–362. DOI: 10.1038/s41586-020-2649-2.
- [IW89] Nobuyuki Ikeda and Shinzo Watanabe. *Stochastic differential equations and diffusion processes*. North-Holland Publishing Company, 1989.
- [JM86] A. Joffe and M. Metivier. “Weak Convergence of Sequences of Semimartingales with Applications to Multitype Branching Processes”. In: *Advances in Applied Probability* 18.1 (1986), pp. 20–65.
- [Joh+07] Matthew D Johnston et al. “Mathematical modeling of cell population dynamics in the colonic crypt and in colorectal cancer”. In: *Proceedings of the National Academy of Sciences* 104.10 (Mar. 2007), pp. 4008–4013. DOI: 10.1073/pnas.0611179104.
- [Kai+16] Gerard E. Kaiko et al. “The Colonic Crypt Protects Stem Cells from Microbiota-Derived Metabolites”. In: *Cell* 165.7 (2016), pp. 1708–1720. DOI: 10.1016/j.cell.2016.05.018.
- [LPS15] Siu Kwan Lam, Antoine Pitrou, and Stanley Seibert. “Numba: A LLVM-based Python JIT Compiler”. In: *Proceedings of the Second Workshop on the LLVM Compiler Infrastructure in HPC*. Austin Texas: ACM, Nov. 2015, pp. 1–6. DOI: 10.1145/2833157.2833162.
- [Las80] Irena Lasiecka. “Unified theory for abstract parabolic boundary problems—a semigroup approach”. In: *Applied Mathematics and Optimization* 6.1 (1980), pp. 287–333.
- [LM72] Jacques Louis Lions and Enrico Magenes. *Non-homogeneous boundary value problems and applications*. Vol. 1. Die Grundlehren der mathematischen Wissenschaften. Berlin: Springer-Verlag, 1972.
- [LD08] Richard Losick and Claude Desplan. “Stochasticity and Cell Fate”. In: *Science* 320.5872 (Apr. 4, 2008), pp. 65–68. DOI: 10.1126/science.1147888.
- [MFK09] Martin Meier-Schellersheim, Iain DC Fraser, and Frederick Klauschen. “Multiscale modeling for biologists”. In: *Wiley Interdisciplinary Reviews: Systems Biology and Medicine* 1.1 (2009), pp. 4–14.
- [Mic17] Philippe Michel. “Large population and size scale limit of a stochastic particle model”. In: *Mathematical Models and Methods in Applied Sciences* 27.03 (2017), pp. 581–615. DOI: 10.1142/S0218202517500105. eprint: <https://doi.org/10.1142/S0218202517500105>.
- [Mur+11] Philip J Murray et al. “Comparing a discrete and continuum model of the intestinal crypt”. In: *Physical Biology* 8.2 (Apr. 2011), p. 026011. DOI: 10.1088/1478-3975/8/2/026011.
- [Osb+17] James M Osborne et al. “Comparing individual-based approaches to modelling the self-organization of multicellular tissues”. In: *PLoS computational biology* 13.2 (2017), e1005387.
- [Paz12] Amnon Pazy. *Semigroups of linear operators and applications to partial differential equations*. Vol. 44. Springer Science & Business Media, 2012.
- [Per07] Benoît Perthame. *Transport Equations in Biology*. Basel: Birkhäuser Basel, 2007.
- [Roc+19] Russell C Rockne et al. “The 2019 mathematical oncology roadmap”. In: *Physical biology* 16.4 (2019), p. 041005.
- [Ros+20] Jakob Rosenbauer et al. “Modeling of Wnt-mediated tissue patterning in vertebrate embryogenesis”. In: *PLOS Computational Biology* 16.6 (2020), e1007417.
- [Smi+12] Aaron M Smith et al. “Incorporating chemical signalling factors into cell-based models of growing epithelial tissues”. In: *Journal of mathematical biology* 65.3 (2012), pp. 441–463.
- [Tha+18] Torsten Thalheim et al. “Linking stem cell function and growth pattern of intestinal organoids”. In: *Developmental Biology* 433.2 (Jan. 2018), pp. 254–261. DOI: 10.1016/j.ydbio.2017.10.013.
- [Tra08] Viet Chi Tran. “Large population limit and time behaviour of a stochastic particle model describing an age-structured population”. In: *ESAIM: Probability and Statistics* 12 (2008), pp. 345–386.
- [Tra06] Viet Chi Tran. “Stochastic particle models for problems of adaptive evolution and for the approximations of statistical solutions”. Theses. Université de Nanterre - Paris X, Dec. 2006.
- [Vir+20] Pauli Virtanen et al. “SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python”. In: *Nature Methods* 17 (2020), pp. 261–272. DOI: 10.1038/s41592-019-0686-2.

- [Wan+18] Yuli Wang et al. “Bioengineered systems and designer matrices that recapitulate the intestinal stem cell niche”. In: *Cellular and molecular gastroenterology and hepatology* 5.3 (2018), pp. 440–453.