



Cell differentiation, stem cell regulation and impact of the mutations: a stochastic approach

Celine Bonnet

► To cite this version:

Celine Bonnet. Cell differentiation, stem cell regulation and impact of the mutations: a stochastic approach. Probability [math.PR]. Institut Polytechnique de Paris, 2020. English. NNT: 2020IP-PAX016 . tel-02983173

HAL Id: tel-02983173

<https://theses.hal.science/tel-02983173>

Submitted on 29 Oct 2020

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.

Différentiation cellulaire, régulation des cellules souches et impact des mutations : une approche probabiliste

Thèse de doctorat de l'Institut Polytechnique de Paris
préparée à l'École polytechnique

École doctorale n°574 École Doctorale de Mathématiques Hadamard (EDMH)
Spécialité de doctorat : Mathématiques appliquées

Thèse présentée et soutenue à Antony en Visio conférence, le 27/05/2020, par

CÉLINE BONNET

Composition du Jury :

Marie Doumic Ingénieure en Chef des Ponts et Forêts, Inria et Sorbonne Universités (LJLL, UMR 7598)	Présidente
Léa Popovic Professeur, Concordia University (Dept of Mathematics and Statistics)	Rapporteuse
Fabien Crauste Directeur de Recherche, Université de Bordeaux (IMB, UMR 5251)	Rapporteur
Pierre-André Zitt Maître de conférences, Université Gustave Eiffel (LAMA, UMR 8050)	Examineur
Romain Yvinec Chargé de recherche, INRAE (PRC, UMR 85 / UMR 7247)	Examineur
Vincent Bansaye Professeur, École polytechnique (CMAP, UMR 7641)	Examineur
Sylvie Méléard Professeur, École polytechnique (CMAP, UMR 7641)	Directrice de thèse
Stéphane Giraudier Professeur, Hôpital Saint-Louis (INSERM Unité U1131)	Co-directeur de thèse

Remerciements

Je tiens à remercier Sylvie Méléard de m'avoir embarquée dans la grande aventure qu'a été cette thèse. Sans elle je ne pourrais remercier toutes ces personnes qui ont été à mes côtés pendant ces 3 années.

Telle une instructrice de voile, Sylvie, malgré mes quelques dérives, mes petits tours à la baille, mes pauses "admiration de la beauté du monde scientifique", tu as su m'aider à garder le cap et a été un support sans faille. Pour m'avoir fait grandir humainement et scientifiquement merci !

Merci à notre super équipe, les membres du projet M3PN, sans laquelle rien n'aurait été possible. Oui vous Evelyne, Stéphane et Simon pour nos échanges des plus dynamiques et acharnés.

Merci à Vincent Bansaye, d'avoir toujours été là quand j'en avais besoin.

Merci à mes rapporteurs Léa et Fabien, de m'avoir prodigué des conseils et surtout pour avoir lu avec tant d'attention mon manuscrit.

Merci à mon jury, Romain, Pierre-André et Marie d'avoir accepté de m'écouter et pour leur disponibilité et gentillesse durant nos quelques échanges.

Merci à Frédérique Clément, Marie Postel et Romain Yvinec pour nos collaborations des plus captivantes (en particulier en escalade) durant l'été 2018. Merci à Hugo et à Angélique, de belles rencontres de ce même été. Merci à Frédérique Robin pour nos échanges au resto bien sympathiques.

Merci aux membres de la chaire MMB avec lesquels j'ai partagé d'excellents moments à Aussois (je ne pense bien sûr pas qu'aux karaokés).

Merci à Charline et à Aline, pour leurs conseils avisés. Merci à Héléne pour nos discussions. J'ai hâte qu'on travaille ensemble.

Merci à tous les membres de l'équipe PEIPS. Je pense particulièrement à Amandine pour sa gentillesse sans pareil, à Lucas pour ses bêtises, à Carl pour ses interventions toujours intéressantes, à Milica pour ses conseils et à bien d'autres.

Merci à Anne Auger et Cheikh pour le temps qu'ils ont consacré à me parler du CMA-ES.

Merci à tous les membres du CMAP, les permanents et les doctorants. J'ai adoré passer du temps auprès de vous, nos discussions en salle café pour les permanents et au Magnan pour les doctorants. J'ai adoré rentrer dans le bureau de Mathilde et avant elle, celui de Tristan, pour extérioriser mes frustrations ou simplement discuter "du temps qui passe". Merci à Paul, je ne saurais dire exactement pourquoi ;). Il règne au CMAP une ambiance qu'on regrette d'avoir perdue pendant le confinement (surtout mardi dernier quand j'ai vidé mon bureau dans un laboratoire désert ^{sniff}). Vous allez tous me manquer (même si, soyons honnête, je suis très heureuse d'avoir fini cette étape de ma vie ;)).

Je souhaite remercier Nassera, et également Alexandra, et tout le personnel administratif sans lesquelles le CMAP ne serait pas ce qu'il est. J'ai adoré venir discuter avec vous pendant mes moments de pause et espère ne pas vous avoir trop dérangées ;).

Merci aux membres de mon bureau le 2010, Antoine pour tant d'excuses, Fedor pour tes moqueries sur mon anglais, Martin pour tes éblouissantes compétences aux échecs, Heytem pour tout et Juliette, toi que je poursuis depuis notre L3. Je pense particulièrement à vous 2 la DREAMteam du bureau 2010. Je n'oublie pas Josué mon petit frère de thèse (je te souhaite de t'épanouir dans la recherche) ni Adrien (je te souhaite la même chose), avec lesquels j'ai apprécié échanger sur tout et rien. Je retiens quand même, Josué, que tes conseils en matière de film d'horreur sont à prendre avec précautionsss. Et bien sûr bonne chance aux petits nouveaux ! (c'est pas si difficile une thèse, j'ai voulu me pendre juste une ou deux fois, pas plus... bon peut être 6 enfin disons 10 mais c'est tout).

Merci à Pierre mon tout premier co-bureau toujours prêt à rendre service et boire une bière.

Merci à Tran, Eugène, Kevish et Heythem. Je n'oublierai jamais notre collocation. Des moments importants dans ma vie.

Merci à mes amis de fac qui ne m'ont pas quitté depuis tant d'années : François, Robert, Diane (quoique récemment celle-ci on a réussi à l'envoyer au Gabon), Tran, Samuel, BG, Juliette (la best, il faut bien dans ce monde de chercheurs si masculin) et bien-sûr Pierre. Vous êtes vraiment les meilleurs !!

Merci à ma famille que j'aime énormément et qui est si génialissime. Du coup merci à mes deux grand mères et mes parents !! Gros bisous à Grandma, merci pour nos échanges réguliers.

Merci à Aurélie de m'avoir supportée pendant ces 3 ans et de m'avoir aidée à penser à d'autres choses qu'au travail. Sans toi je n'aurais pas survécu.

En parlant de penser à autre chose, merci Matthieu et merci Sophie, par ordre d'arrivée ;p. Et puis merci aussi aux filles de la "secte" et à Amine et Thomas pour les mêmes raisons.

Finalement je dédie ma thèse à mes deux piliers pendant ces trois années (désolée Juliette) : Geneviève Bonnet et Heythem Farhat.

Ps : j'ai eu envie d'ajouter Céline bonnet mais j'ai pas osé ^^".

Table des matières

Introduction	6
Mise en contexte biologique	7
Mise en contexte mathématique	8
Résultats du chapitre 1	10
Résultats du chapitre 2	18
Résultats du chapitre 3	24
Résultats du chapitre 4	30
 1 Régulation des cellules souches hématopoïétiques : Une approche macroscopique.	 31
1.1 Introduction	31
1.2 Erythropoïèse au repos : Un modèle mathématique	32
1.2.1 Le modèle probabiliste sous-jacent	32
1.2.2 Etude du modèle : un mécanisme d'amplification	35
1.2.3 Calibration du modèle à l'aide des données	37
1.3 Erythropoïèse de stress : Ajout d'une régulation	41
1.3.1 Des données par type cellulaire	41
1.3.2 Le modèle	43
1.4 Etude du modèle	46
1.4.1 Existence et unicité des points stationnaires	46
1.4.2 Le modèle probabiliste sous-jacent	49
1.5 Calibration du modèle régulé	53
1.6 Article associé (prochainement soumis)	56
 2 Large fluctuations in multi-scale modeling for rest erythropoiesis	 95
2.1 Introduction	95
2.2 Amplification mechanism : size-scale dynamics	100
2.2.1 The amplification mechanism	100
2.2.2 Asymptotic behavior on a finite time interval	102
2.3 Size-time multi-scale dynamics and asymptotic behavior	103
2.3.1 Asymptotic behavior at a time-scale of order K^{γ_2}	103
2.3.2 Asymptotic behavior at a time-scale of order K^{γ_3}	105
2.4 Amplified fluctuations	109
2.4.1 The large fluctuations of the second type	109
2.4.2 The large fluctuations of the third type	111
2.5 Appendix	116

3	Influence de la quiescence dans le développement d'un mutant : une approche déterministe par morceaux	117
3.1	Introduction	117
3.2	Convergence vers un PDMP	120
3.3	Comportement en temps long du processus limite.	125
3.4	Simulations et perspectives	134
3.5	Annexe	137
4	Multiscale population dynamics in reproductive biology.	138
4.1	Compartmental, ODE-based model	141
4.1.1	Initial model	141
4.1.2	Rescaled model	142
4.1.3	Limit model	143
4.1.4	Convergence in the linear case	144
4.1.5	Numerical convergence	145
4.2	PDE model	147
4.2.1	Rescaled model	147
4.2.2	Limit model	148
4.2.3	Convergence in the linear case	149
4.2.4	Numerical study	150
4.3	SDE model	156
4.3.1	Rescaled model	156
4.3.2	Limit model	157
4.3.3	Convergence in the linear case	159
4.3.4	Numerical convergence	161
4.4	Appendix - parameter values	164

Introduction

A l'origine de cette thèse se trouve une collaboration étroite avec des médecins de l'hôpital Saint Louis et une biologiste de l'Institut Cochin à Paris. Les mécanismes de l'érythropoïèse dans des conditions physiologiques normales (appelée érythropoïèse de repos) que l'on croyait avoir compris à partir d'observations sur des modèles de souris irradiées, ont été remis en cause en observant des souris in vivo. En particulier il a été remarqué, à travers différents mécanismes de régulation par le stress, que l'environnement a une importance capitale [13]. En s'inspirant des travaux de Busch et al [13], les échanges interdisciplinaires réguliers au sein de notre équipe ont mené à l'élaboration d'expériences biologiques sur des souris in vivo en condition normale et de stress. Les résultats de ces expériences ont montré le besoin de construire un nouveau modèle d'érythropoïèse permettant d'expliquer la dynamique cellulaire observée. La construction et l'étude de ce modèle d'érythropoïèse de repos et de stress sont développés dans le Chapitre 1 de cette thèse.

Plus particulièrement, nous nous intéressons à la dynamique de cellules ayant une capacité d'auto-renouvellement et aux conséquences dynamiques de cette capacité sur les cellules érythropoïétiques. Les expériences menées et le travail de modélisation associé ont mis en évidence un lien entre la capacité d'auto-renouvellement des cellules et le mécanisme d'amplification réalisé en érythropoïèse de repos. Ainsi en développant un modèle stochastique se concentrant sur l'érythropoïèse de repos, nous avons pu quantifier les grandes fluctuations observées au niveau des globules rouges, type cellulaire produit en fin de ce mécanisme d'amplification. Le modèle décrit la dynamique d'un processus de branchement de trois types cellulaires dont l'étude a fait apparaître une dynamique dite lente rapide nécessitant l'utilisation de méthodes de moyennisation développées par T. Kurtz [62]. Ce travail est présenté dans le Chapitre 2 de cette thèse.

Au delà du domaine de la recherche fondamentale, la compréhension des mécanismes dynamiques entourant cette capacité d'auto-renouvellement des cellules érythropoïétiques pourrait avoir des retombées dans la compréhension des causes de certains cancers. Nous verrons au Chapitre 3 l'influence du comportement dynamique particulier des cellules souches, à l'origine de l'érythropoïèse, dans l'émergence de mutations.

Mise en contexte biologique

Le terme cellule souche apparaît dans la littérature scientifique dès 1868, dans les travaux du biologiste allemand Ernst Haeckel. Partisan de la théorie de l'évolution de Darwin, Ernst Haeckel a dessiné un certain nombre d'arbres représentant l'évolution des divers organismes à partir d'ancêtres communs. Il a appelé ces arbres *Stammbaume* (*arbres souches* en allemand). Dans ce contexte, Ernst Haeckel a utilisé le terme *stammzelle* (*cellule souche* en allemand) pour décrire l'ancêtre unicellulaire à partir duquel auraient évolué tous les organismes multi-cellulaires [83].

L'hématopoïèse est le processus biologique permettant le renouvellement et le maintien des diverses cellules sanguines (globules rouges, plaquettes, globules blancs...) par différenciation successives d'une seule et même cellule multipotente appelée cellule souche hématopoïétique (HSC en anglais).

A chaque étape de différenciation la cellule initialement souche se transforme. La distinction majeure entre les cellules souches et les cellules plus matures appelées progéniteurs est leur capacité à reconstituer l'ensemble des cellules hématopoïétiques. Il a été montré que la transplantation de progéniteurs dans des souris irradiées, ayant perdu une bonne partie de leurs cellules hématopoïétiques, ne permet pas la survie de l'animal à la différence des cellules souches [31].

En effet, les cellules souches peuvent par division symétrique donner naissance à des cellules ayant le même potentiel multipotent qu'elles. On parle d'auto-renouvellement. Cette capacité importante d'auto-renouvellement permet aux cellules souches de produire les diverses cellules du sang tout au long de la vie de l'organisme.

Au fur et à mesure de leur différenciation les cellules initialement souches perdent progressivement leur capacité à s'auto-renouveler.

La dynamique des cellules hématopoïétiques est complexe et intéressante. Afin d'en comprendre de multiples aspects, nous nous intéressons, dans cette thèse, spécifiquement à la lignée des globules rouges (lignée des érythrocytes).

L'une des premières particularités de l'érythropoïèse est le mécanisme d'amplification cellulaire qu'elle représente. Les cellules souches dites «Long Term» (LT-HSC) initient l'érythropoïèse. Elles sont en grand nombre dans la moelle osseuse (10^4 cellules par souris, [12]) et produisent un encore plus grand nombre de globules rouges par jour (10^{10} globules rouges par souris) (voir Chapitre 1).

Les progrès scientifiques permettent aujourd'hui de quantifier le nombre de cellules de multiples types cellulaires constituant l'érythropoïèse. On observe ainsi une augmentation du nombre de cellules entre chaque étape de différenciation (voir Chapitre 1). Même si la description des étapes de l'érythropoïèse est encore en débat nous suivrons le schéma classique détaillé dans [13].

La communauté scientifique s'est également intéressée à l'estimation de paramètres cinétiques tels que la fréquence à laquelle les cellules se différencient, s'auto-renouvellent ou meurent dans des conditions normales. Nos travaux engagés initialement dans la même approche nous ont permis d'estimer les paramètres cinétiques en condition normale mais aussi en situation de stress. Nous obtenons en particulier la même estimation pour le taux de division des HSC au repos [13, 40]. Les cellules souches se divisent très rarement

(une division tous les 3-4 mois environ chez la souris). On dit qu'elles ont des phases de quiescence (inactivité) très longues [87].

En condition de repos les cellules semblent faire des divisions symétriques mais rien n'exclut la possibilité qu'elles se divisent de façon asymétrique [69]. Il existe des preuves expérimentales que les cellules peuvent se diviser symétriquement ou asymétriquement [100], mais les estimations de la probabilité relative de tels événements sont contradictoires [102, 64, 75, 88].

La dynamique de l'érythropoïèse de repos se résume alors ainsi :

- Les LT-HSCs se divisent rarement ;
- Les cellules érythropoïétiques présentes dans la moelle osseuse se divisent en deux cellules filles en donnant naissance soit à des cellules du même type cellulaire (par renouvellement), soit à des cellules différenciées (par différenciation). Ces divisions peuvent être symétriques comme asymétriques.
- Les LT-HSCs ont une probabilité de s'auto-renouveler à chaque division cellulaire plus élevée que celle des autres types cellulaires.

Une deuxième particularité de l'érythropoïèse est que le comportement dynamique des cellules qui la constituent semble soumis à des règles différentes lorsque le système est fortement perturbé. En effet, l'érythropoïèse dite de stress a été définie comme différente de l'érythropoïèse de repos. Les différences ne sont pas clairement établies mais il semble que les voies de différenciation ainsi que les règles de division cellulaire puissent être contournées par les cellules en situation de stress [46].

La compréhension de l'érythropoïèse de stress s'accompagne d'une compréhension des différents mécanismes de régulation. Les informations biologiques connues sur les régulations concernent majoritairement les cellules avancées dans la différenciation. Ces cellules ont des récepteurs à l'érythropoïétine (EPO). Les mécanismes de régulation en jeu dans l'action de l'EPO sont aujourd'hui bien connus et documentés. Or nous souhaitons comprendre les mécanismes de régulation agissant sur les cellules ayant une importante capacité à s'auto-renouveler. La régulation de ces cellules est bien moins documentée *in vivo*. Toutefois des hypothèses concernant cette régulation ont été formulées. En particulier, la dynamique de ces cellules présentes dans la moelle osseuse pourrait être influencée par la place disponible dans l'espace limité que constitue la moelle osseuse, comme envisagé dans [59]. De plus de multiples cytokines (enzymes de régulation) ont été identifiées *in vivo* chez la souris mais leurs effets et mécanismes sont loin d'être compris.

Afin d'aider à la compréhension de ces mécanismes de régulation, divers modèles mathématiques ont été introduits.

Mise en contexte mathématique

Dans le cadre d'une hématopoïèse de repos et de stress, des modèles déterministes multi-types, dont une confrontation aux données biologiques est réalisée uniquement au niveau des cellules matures, ont fait l'objet d'une analyse mathématique [75, 2]. Dans [75], les auteurs modélisent la régulation des différents types cellulaires en représentant les paramètres dynamiques des cellules par une fonction décroissante dépendant explicitement du nombre de cellules matures dans le système. Dans [2], les régulations sont modélisées par des cytokines ayant leur propre dynamique. Dans le Chapitre 1, nous présentons un modèle de l'érythropoïèse de repos et de stress permettant d'expliquer non seulement la

dynamique de retour à l'équilibre des cellules matures mais également celle de 5 autres types cellulaires.

Il existe également des modèles confrontés à un grand nombre de données *in vivo* dont la régulation est modélisée par une dépendance explicite des paramètres dynamiques en l'état du système [74, 59]. Nous verrons au Chapitre 1, qu'une telle modélisation de la régulation ne permet pas d'expliquer le retour à l'équilibre observé sur les données recueillies par nos collaborateurs. Nous verrons également que l'ajout d'équations modélisant la dynamique des régulations, nous a permis d'expliquer la dynamique cellulaire observée.

Des modèles prenant en compte la stochasticité de la dynamique cellulaire ont été développés. Certains modélisent la dynamique de mutation [60, 31, 25], d'autres modélisent la dynamique de l'érythropoïèse en confrontant le modèle à des données biologiques [1, 86]. Aucun d'eux ne s'intéresse au comportement asymptotique d'un processus de branchement multi-échelles comme on le fait dans le Chapitre 2.

Dans les travaux développés dans [60] et [86], les cellules souches basculent aléatoirement d'un état actif à un état quiescent, comme dans le modèle que nous développons au Chapitre 3. Toutefois les échelles de tailles des différentes populations et donc la vitesse d'évolution de chacune d'elles n'est pas prise en compte dans leur modèle. Nous avons choisi d'étudier le comportement dynamique de populations de cellules saines et mutantes ayant des échelles évolutives différentes en incluant une amplification entre le compartiment des cellules souches et celui des globules rouges. L'étude des changements d'état de l'unique cellule souche mutante a fait apparaître un processus de Markov déterministe par morceaux.

La capacité d'auto-renouvellement des cellules érythropoïétiques est au cœur des questions de cette thèse. Ainsi nous étudions au Chapitre 1, à l'aide d'une hiérarchie détaillée des étapes de différenciation, la régulation des cellules ayant une capacité d'auto-renouvellement plus ou moins grande (en fonction de leur type cellulaire). Cette étude se fait à l'aide de données biologiques *in vivo* sur des souris mises en situation de grand stress. Nous nous intéressons ensuite, dans le Chapitre 2, à une compréhension plus théorique du lien entre l'amplification observée entre les nombres de cellules souches et de globules rouges et les capacités d'auto-renouvellement et de différenciation des cellules. En supposant que le système est en grand nombre et à l'équilibre nous montrons que ce mécanisme d'amplification particulier a des effets sur la taille des populations mais également sur l'échelle temporelle à laquelle la dynamique des populations se déroule. Nous verrons que la prise en compte de l'aléa dans ce mécanisme conduit à l'apparition de fluctuations inattendues. Le troisième chapitre est consacré à l'étude de l'influence des phases de quiescence d'une cellule souche mutante dans l'invasion d'une population de cellules saines.

Dans chacun des chapitres, je développe un modèle adapté aux différentes questions posées sur la dynamique des cellules érythropoïétiques. L'étude de ces différents modèles nécessite une théorie mathématique spécifique. Dans chacune des sections suivantes je présente l'ensemble des résultats obtenus dans ma thèse ainsi que la théorie utilisée.

Le dernier chapitre est constitué d'un travail réalisé lors de ma participation au CEM-RACS 2018. Les différentes questions mathématiques apparues dans ce projet sont en lien avec les questions d'échelles apparues au chapitre 2 de cette thèse.

Résultats du chapitre 1

Ce chapitre présente les résultats de l'article intitulé " Modeling the behavior of hematopoietic compartments from stem to red cells in steady state and stress hematopoiesis", écrit en collaboration avec Panhong Gou, Simon Girel, Vincent Bansaye, Catherine Lacout, Karine Bailly, Marie-Hélène Schlagetter, Evelyne Lauret, Sylvie Méléard et Stéphane Giraudier dont la soumission est prévue dans le journal Cell Stem Cell.

Un modèle d'érythropoïèse de repos et de stress

Dans ce chapitre nous nous intéressons à l'étude d'un modèle décrivant la dynamique de k types cellulaires de l'érythropoïèse où le premier type représente les LT-HSC et le dernier type les globules rouges.

Nous supposons qu'une cellule de type i pour $i \in \{1, \dots, k-2\}$, se divise à taux τ_i en deux cellules filles. Les deux cellules filles sont de type i avec une probabilité d'auto-renouvellement p_i^R ou de types $i+1$ avec probabilité de différenciation érythroïde $\mu_i p_i^D = \mu_i (1 - p_i^R)$. Autrement dit à chaque division cellulaire les cellules s'auto-renouvellent ou se différencient. Lorsqu'elles se différencient les deux cellules filles sont de la lignée des érythrocytes avec une probabilité μ_i ou sont d'une autre lignée cellulaire avec probabilité $1 - \mu_i$.

On parle également du taux d'auto-renouvellement et du taux de différenciation érythroïde des cellules. En effet notons $\tau_i^R = \tau_i p_i^R$, $\tau_i^D = \tau_i \mu_i p_i^D$ et $s_i = \tau_i (1 - \mu_i) p_i^D$, alors une cellule de type i se divise en deux cellules de type i à taux τ_i^R , se divise en deux cellules de type $i+1$ à taux τ_i^D et sort du système à taux s_i .

Une cellule de type $k-1$ quant à elle se divise en deux cellules de type k à taux $\tau_{k-1}^R = \tau_{k-1} p_{k-1}^R$ et engendre 2^n cellules de type k à taux $\tau_{k-1}^D = p_{k-1}^D \tau_{k-1}$. Les probabilités de différenciation et d'auto-renouvellement vérifient $p_{k-1}^R + p_{k-1}^D = 1$.

Une cellule de type k meurt à taux constant $\tau_k > 0$.

Nous noterons $x_i(t)$ le nombre de cellules de type i au temps t et x_i^* le nombre de cellules de type i décrites par les données en condition normale, de repos.

La régulation du modèle sera modélisée par deux fonctions de régulations u_R et u_D . La régulation u_R agit sur le taux d'auto-renouvellement des cellules et est produite lorsque le système est en manque de cellules dans la moelle osseuse. La régulation u_D agit sur le taux de différenciation érythroïde des cellules et est produite lorsque la moelle osseuse est en surnombre. Les fonctions de régulation u_R et u_D décroissent selon un taux constant noté respectivement m_R et m_D .

L'objectif de notre modèle est d'expliquer le comportement dynamique des cellules érythropoïétiques présentes dans la moelle osseuse et ayant une forte capacité d'auto-renouvellement en introduisant un minimum de paramètres à calibrer à partir des données de nos collaborateurs biologistes.

Notre modèle est décrit par les fonctions continues $(x_1, \dots, x_k, u_R, u_D)$ solutions du système d'équations suivant pour lequel $x(0)$ est de première coordonnée strictement positive, $u_R(0) = \frac{1}{m_R(a+y(0))}$ et $u_D(0) = \frac{y(0)}{m_D(b+y(0))}$,

$$\left\{ \begin{array}{l} \frac{dx_1}{dt}(t) = (\tau_1^R(t) - \tau_1^D(t) - s_1) x_1(t) \\ \frac{dx_i}{dt}(t) = 2 \tau_{i-1}^D(t) x_{i-1}(t) + (\tau_i^R(t) - \tau_i^D(t) - s_i) x_i(t), \quad \forall i \in \llbracket 2, k-1 \rrbracket \\ \frac{dx_k}{dt}(t) = 2^{n(t)} \tau_{k-1}^D(t) x_{k-1}(t) - \tau_k x_k(t) \\ \frac{du_R}{dt}(t) = \frac{1}{a+y(t)} - m_R u_R(t) \\ \frac{du_D}{dt}(t) = \frac{y(t)}{b+y(t)} - m_D u_D(t) \end{array} \right. \quad (1)$$

où y représente la densité de cellules dans la moelle osseuse $y(t) = \sum_{j=1}^{k-1} \frac{x_j(t)}{x_j^*}$.

La régulation du système se décompose en trois parties :

- La plus connue de l'érythropoïèse, résumée au travers d'une fonction bornée et décroissante dépendant de la quantité de cellules matures et affectant les n dernières étapes de différenciation, pour tout t ,

$$n(t) = n^* \frac{1 + c_1}{1 + c_1 \left(\frac{x_k(t)}{x_k^*} \right)^{c_2}}.$$

- La régulation nommée u_R produite lorsque le système est en manque de cellules dans la moelle osseuse et affectant le taux de renouvellement des cellules

$$\tau_i^R(t) = \tau_i^* p_i^{R,*} \left(\frac{u_R(t)}{u_R^*} \right)^{r_i}.$$

- La régulation nommée u_D produite lorsque le nombre de cellules dans la moelle osseuse est trop important et affectant le taux de différenciation érythroïde des cellules

$$\tau_i^D(t) = \tau_i^* p_i^{D,*} \mu_i^* \left(\frac{u_D(t)}{u_D^*} \right)^{d_i}.$$

Les paramètres du modèle sont tous supposés positifs. On suppose de plus que

$$p_1^{R,*} = p_1^{D,*} \quad \text{et} \quad \forall i \geq 2, \quad p_i^{R,*} < p_i^{D,*}.$$

Rappelons que les paramètres désignés à l'aide d'une étoile en exposant font référence au modèle de l'érythropoïèse au repos.

Un tel modèle offre une vision nouvelle de la régulation d'un système biologique. En effet, ici l'homéostasie de l'érythropoïèse (autrement dit son activité au repos) ne nécessite pas l'intervention de la régulation. Elle n'apparaît dans les équations que lorsque le système est perturbé car dans le cas contraire $\frac{u_i}{u_i^*} = 1$ et les taux valent $\tau_i^D(t) = \tau_i^* p_i^{D,*} \mu_i^*$ et $\tau_i^R(t) = \tau_i^* p_i^{R,*}$.

Comportement en temps long des solutions du système (1)

Nous avons étudié le comportement asymptotique du système (1) et montré l'existence d'un unique point stationnaire pour lequel la première coordonnée est strictement positive. Nous avons également montré que ce point stationnaire est donné par le vecteur (x^*, u_R^*, u_D^*) .

Théorème 1. *Le système (1), pour lequel l'ensemble des paramètres sont strictement positifs, admet un unique point stationnaire, ayant toutes ses composantes strictement positives, donné par (x^*, u_R^*, u_D^*) .*

Bien que nous n'ayons pas prouvé théoriquement la convergence du modèle vers cet unique point stationnaire, nous avons illustré la convergence à l'aide de simulations numériques.

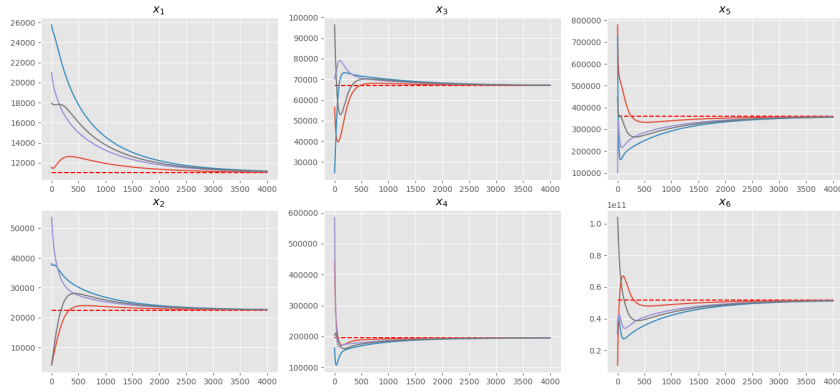


FIGURE 1 – Illustration de la convergence de l'unique solution du système (1) vers son unique point stationnaire non nul.

Nous observons les valeurs des solutions x_i au cours du temps pour des conditions initiales déterminées de façon aléatoire sur $[0, 3x_i^*]$. Pour cette simulation nous avons supposé les paramètres égaux aux valeurs présentées dans le tableau de la sous section 'Calibration du modèle : étape 1' et les autres paramètres sont supposés égaux à 1.

Cette particularité du modèle, ce lien qui existe entre l'érythropoïèse de repos et de stress dans notre modèle, sera utilisé pour calibrer par étapes le modèle à l'aide des données expérimentales. Avant cela, nous avons montré que le modèle ainsi créé est bien limite en grande population d'un modèle probabiliste.

Le système (1) : une limite en grande population

Soit $K \in \mathbb{R}^+$ un paramètre d'échelle. Notons N^K le nombre de cellules des k populations au cours du temps. Supposons que le nombre de LT-HSC satisfait

$$N_1^K(0) \sim K.$$

Introduisons ensuite $(\mathcal{N}_i^j)_{\substack{1 \leq i \leq k \\ j \in \{-, +, s\}}}$, des mesures de Poisson ponctuelles indépendantes d'intensité $dvdu$ sur $\mathbb{R}_+^2$ ainsi que la filtration $(\mathcal{F}_t)_{t \geq 0}$ associée

$$\mathcal{F}_t = \sigma(\mathcal{N}_i^j([0, v] \times A); i \in \{1, \dots, k\}, j \in \{-, +, s\}, v \leq t, A \in \mathcal{B}(\mathbb{R})).$$

Nous pouvons alors définir le processus de Markov N^K comme l'unique solution des équations intégrales stochastiques suivantes,

$$\begin{aligned} N_1^K(t) = & K + \int_0^t \int_{\mathbb{R}_+} \mathbf{1}_{\{\mathbf{u} \leq \tau_1^R(\mathbf{u}_R^K(\mathbf{v}^-)) \mathbf{N}_1^K(\mathbf{v}^-)\}} \mathcal{N}_1^+(\mathbf{d}\mathbf{v}, \mathbf{d}\mathbf{u}) \\ & - \int_0^t \int_{\mathbb{R}_+} \mathbf{1}_{\{\mathbf{u} \leq \tau_1^D(\mathbf{u}_D^K(\mathbf{v}^-)) \mathbf{N}_1^K(\mathbf{v}^-)\}} \mathcal{N}_1^-(\mathbf{d}\mathbf{v}, \mathbf{d}\mathbf{u}) \\ & - \int_0^t \int_{\mathbb{R}_+} \mathbf{1}_{\{\mathbf{u} \leq \mathbf{s}_1 \mathbf{N}_1^K(\mathbf{v}^-)\}} \mathcal{N}_1^s(\mathbf{d}\mathbf{v}, \mathbf{d}\mathbf{u}) \end{aligned}$$

pour $i = 2, \dots, k-1$,

$$\begin{aligned} N_i^K(t) = & N_i^K(0) + 2 \int_0^t \int_{\mathbb{R}_+} \mathbf{1}_{\{\mathbf{u} \leq \tau_{i-1}^D(\mathbf{u}_D^K(\mathbf{v}^-)) \mathbf{N}_{i-1}^K(\mathbf{v}^-)\}} \mathcal{N}_{i-1}^-(\mathbf{d}\mathbf{v}, \mathbf{d}\mathbf{u}) \\ & + \int_0^t \int_{\mathbb{R}_+} \mathbf{1}_{\{\mathbf{u} \leq \tau_i^R(\mathbf{u}_R^K(\mathbf{v}^-)) \mathbf{N}_i^K(\mathbf{v}^-)\}} \mathcal{N}_i^+(\mathbf{d}\mathbf{v}, \mathbf{d}\mathbf{u}) \\ & - \int_0^t \int_{\mathbb{R}_+} \mathbf{1}_{\{\mathbf{u} \leq \tau_i^D(\mathbf{u}_D^K(\mathbf{v}^-)) \mathbf{N}_i^K(\mathbf{v}^-)\}} \mathcal{N}_i^-(\mathbf{d}\mathbf{v}, \mathbf{d}\mathbf{u}) \\ & - \int_0^t \int_{\mathbb{R}_+} \mathbf{1}_{\{\mathbf{u} \leq \mathbf{s}_i \mathbf{N}_i^K(\mathbf{v}^-)\}} \mathcal{N}_i^s(\mathbf{d}\mathbf{v}, \mathbf{d}\mathbf{u}) \end{aligned}$$

$$\begin{aligned} N_k^K(t) = & N_k^K(0) + \int_0^t \int_{\mathbb{R}_+} 2^{n^K(N_k^K(v^-))} \mathbf{1}_{\{\mathbf{u} \leq \tau_{k-1}^D(\mathbf{u}_D^K(\mathbf{v}^-)) \mathbf{N}_{k-1}^K(\mathbf{v}^-)\}} \mathcal{N}_{k-1}^-(\mathbf{d}\mathbf{v}, \mathbf{d}\mathbf{u}) \\ & - \int_0^t \int_{\mathbb{R}_+} \mathbf{1}_{\{\mathbf{u} \leq \tau_k \mathbf{N}_k^K(\mathbf{v}^-)\}} \mathcal{N}_k^-(\mathbf{d}\mathbf{v}, \mathbf{d}\mathbf{u}) \end{aligned} \quad (2)$$

avec

$$\left\{ \begin{aligned} \tau_i^R(u_1^K(t)) &= \tau_i^* p_i^{R,*} \left(\frac{u_1^K(t)}{u_R^*} \right)^{r_i} \\ \tau_i^D(u_2^K(t)) &= \tau_i^* p_i^{D,*} \mu_i^* \left(\frac{u_D^K(t)}{u_D^*} \right)^{d_i} \\ s_i &= \tau_i^* p_i^{D,*} (1 - \mu_i^*) \\ n^K(N_k^K(t)) &= n^* \frac{c_1}{1 + (c_1 - 1) \left(\frac{N_k^K(t)}{K x_k^*} \right)^{c_2}} \end{aligned} \right. \quad (3)$$

Les termes u_R^K et u_D^K sont des processus de Markov continus déterministes par morceaux définis entre chaque instant de sauts du processus N^K comme solution du système dyna-

mique

$$\left\{ \begin{array}{l} \frac{du_R^K(t)}{dt} = \frac{1}{a + \sum_{j=1}^{k-1} \frac{N_j^K(t)}{K x_j^*}} - m_R u_R^K(t) \\ \frac{du_D^K(t)}{dt} = \frac{\sum_{j=1}^{k-1} \frac{N_j^K(t)}{K x_j^*}}{b + \sum_{j=1}^{k-1} \frac{N_j^K(t)}{K x_j^*}} - m_D u_D^K(t) \end{array} \right. . \quad (4)$$

Les conditions initiales de u_1^K et u_D^K sont définies par

$$u_1^K(0) = \frac{1}{m_R(a + Y_0^K)} \quad \text{et} \quad u_2^K(0) = \frac{Y_0^K}{m_D(b + Y^K)}, \quad \text{avec} \quad Y_0^K = \sum_{j=1}^{k-1} \frac{N_j^K(0)}{K x_j^*}.$$

Notons le processus $X^K \in \mathbb{D}(\mathbb{R}_+, \mathbb{R}_+^{k-1})$ tel que

$$X^K = \frac{N^K}{K}.$$

Nous avons montré que le processus (u_1^K, u_2^K, X^K) converge en probabilité vers l'unique solution du système (1) (à condition initiale fixée).

Théorème 2. *Supposons que*

$$\sup_K \mathbb{E}[X^K(0)] < \infty$$

et que la variable aléatoire $X^K(0)$ converge en loi vers un vecteur x_0 à coordonnées strictement positives.

Alors pour tout $T > 0$, le processus (u_1^K, u_2^K, X^K) converge en loi (et en probabilité) dans $\mathbb{D}([0, T], \mathbb{R}_+^{k+2})$ lorsque K tend vers l'infini vers l'unique solution (u_R, u_D, x) du système (1) telle que $x(0) = x_0$, $u_R(0) = \frac{1}{m_R(a+y_0)}$ et $u_2(0) = \frac{y_0}{m_D(b+y_0)}$ avec $y_0 = \sum_{j=1}^{k-1} \frac{x_j(0)}{x_j^}$.*

Les fonctions u_1^K et u_2^K définissent une régulation du système sensible à la densité des cellules de type 1 à $k-1$.

Calibration du modèle : étape 1

On calibre dans un premier temps le modèle à l'aide de données biologiques en condition normale obtenues par nos collaborateurs en commençant par le nombre de compartiments k que nous choisissons égal à 6. Ce choix est fait à l'aide de différentes informations biologiques recueillies sur les paramètres du modèle et d'un raisonnement mathématique détaillé au Chapitre 1, § 1.2.3. En supposant que les paramètres dynamiques du modèle sont constants, on calibre le modèle relativement à une érythropoïèse de repos. Autrement dit on suppose que

$$\forall t, \quad \left\{ \begin{array}{l} \tau_i^R(t) = \tau_i^* p_i^{R,*} \\ \tau_i^D(t) = \tau_i^* p_i^{D,*} \mu_i^* \end{array} \right. . \quad (5)$$

Les paramètres désignés à l'aide d'une étoile en exposant font référence au modèle de l'érythropoïèse au repos.

En utilisant les données recueillies par nos collaborateurs dans des conditions de repos, ainsi que les données fournies dans [13], nous déduisons la valeur des paramètres dynamiques du modèle au repos.

type cellulaire*	1 : LT-HSC	2 : ST-HSC	3 : MPP	4 : CMP	5 : MEP	6 : RBC
taux de divisions (/jour)	$\tau_1^* = 0.01$	$\tau_2^* = 0.03$	$\tau_3^* = 0.07$	$\tau_4^* = 0.16$	$\tau_5^* = 0.35$	$\tau_6^* = 1/40$ (mort)
nombre de cellules (cellules/souris)	$x_1^* = 0.011$ ($\times 10^6$)	$x_2^* = 0.0225$ ($\times 10^6$)	$x_3^* = 0.067$ ($\times 10^6$)	$x_4^* = 0.196$ ($\times 10^6$)	$x_5^* = 0.359$ ($\times 10^6$)	$x_6^* = 1.48$ ($\times 10^{10}$)
facteur de différenciation	$D_1 = 0$	$D_2 = 0.16$	$D_3 = 0.167$	$D_4 = 0.174$	$D_5 = 0.176$	
probabilité du choix érythroïde	$\mu_1^* = 1$	$\mu_2^* = 1$	$\mu_3^* = 0.99$	$\mu_4^* = 3/5$		
nombre de mitoses terminales	$n^* = 12$					

où $D_i = p_i^{D,*} - p_i^{R,*}$, autrement dit $p_i^{D,*} = (1 + D_i)/2$ et $p_i^{R,*} = (1 - D_i)/2$.

* Les abréviations utilisées pour les différents types cellulaires correspondent respectivement à : LT-HSC, Long Term Hematopoietic Stem Cell, ST-HSC, Short Term Hematopoietic Stem Cell, MPP, Multipotent Progenitor, CMP, Common Myeloid Progenitor, MEP, Megakaryocyte-Erythroid Progenitor et RBC, Red Blood Cell.

Les paramètres du modèle de l'érythropoïèse au repos sont calculés à l'aide d'une formule explicite dépendant des données biologiques.

Le taux de division τ_i^* des cellules de chaque type cellulaire i se déduit directement des résultats d'une expérience réalisée *in vivo* par nos collaborateurs, informant sur la proportion de cellules de chaque type ayant fait au moins une division sur un temps donné. L'expérience en question est la suivante : une substance appelée BrdU (BromodeoxyUridine) est injectée dans une souris et laissée agir durant un temps $t_a \in [16h, 24h]$. Durant cette période de temps, toutes cellules ayant fait au moins une division se retrouve marquée par le BrdU. Notons $x_i^M(t)$ le nombre de cellules de type i marquées par le BrdU au temps t et p_i^{BrdU} la proportion de cellules de type i marquées sur la durée de temps $t_a \in [16h, 24h]$. Alors on obtient

$$p_i^{BrdU} = \frac{x_i^M(t_a)}{x_i^*} = 1 - e^{-\tau_i^* t_a}.$$

Les autres paramètres du modèle se déduisent de la littérature [13] et des expressions de l'unique équilibre du modèle (1) sous l'hypothèse (5).

Calibration du modèle : étape 2

Nous nous sommes finalement intéressés à confronter notre modèle à des données biologiques en situation de grand stress. L'objectif est de s'assurer que notre modèle peut

expliquer le comportement dynamique observé de retour à l'équilibre des cellules érythropoïétiques après un stress important.

Afin de comprendre les effets de la régulation sur chacun des 6 compartiments modélisés, nos collaborateurs ont choisi de réaliser l'expérience suivante sur des souris de laboratoire. Jour 0 : injection d'une dose de phénylhydrazine (PHZ). Cette dose de PHZ cause la destruction d'environ 45% des globules rouges du système en 3 jours. Aux jours 0, 1, 3, 5, 7, 10, 16 et 28, l'effectif cellulaire de chaque compartiment est relevé afin d'observer la réaction du système au stress. Chacun de ces relevés est présenté dans la figure 2.

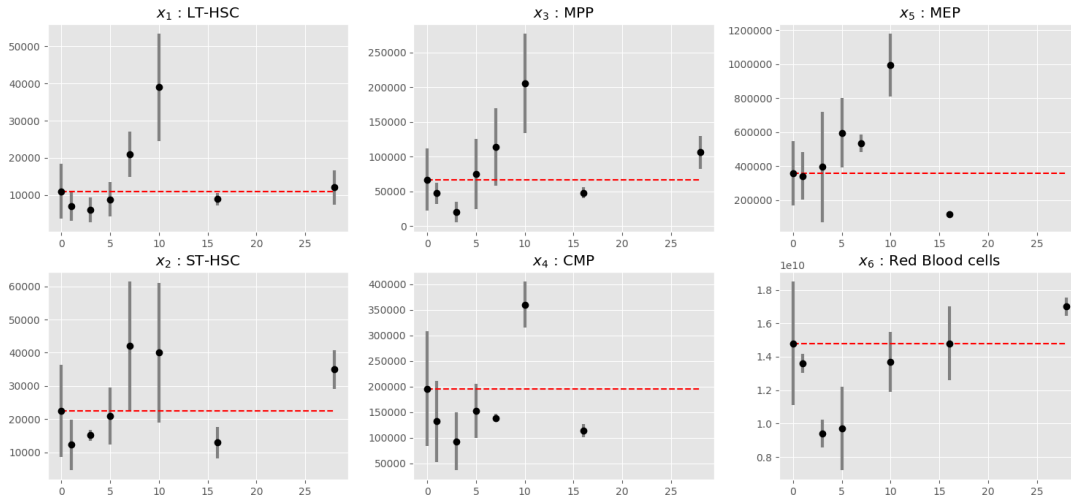


FIGURE 2 – Relevés des effectifs cellulaires par compartiment et par souris aux jours 0 (injection de la PHZ), 1, 3, 5, 7, 10, 16 et 28. La ligne en pointillés indique la valeur d'équilibre (Tableau 1) du nombre de cellules de chaque type.

La situation de stress rencontrée par les cellules est décrite par les données des tailles de populations au jour 3. De plus nous supposons qu'après le jour 3 les effets de la PHZ sont négligeables.

Afin de calibrer les 16 paramètres restants de notre modèle à l'aide des données présentées en figure 2, nous allons utiliser un algorithme d'optimisation stochastique nommé CMA-ES (Covariance Matrix Adaptation in Evolution Strategy). Cet algorithme a été implémenté en particulier sur le langage Python par l'équipe INRIA RandOpt. C'est un algorithme de stratégie évolutive. Dans sa version de base, un algorithme évolutionnaire (de stratégie évolutive) manipule itérativement un ensemble de vecteurs de variables réelles (de paramètres), à l'aide d'opérateurs de mutation et de sélection. La sélection s'effectue par un choix déterministe des meilleurs individus, en fonction des valeurs de la fonction de coût. L'étape de mutation est classiquement effectuée par l'ajout d'une valeur aléatoire, tirée au sein d'une distribution normale. La particularité du CMA-ES est en partie l'adaptation à chaque itération de la matrice de covariance de la distribution normale. Pour plus de détails sur le fonctionnement de l'algorithme se référer à l'article [48].

A l'aide d'une fonction de coût classique de type moindres carrés, nous obtenons le meilleur jeu de paramètres θ .

Un avantage lié à l'utilisation de cet algorithme d'optimisation est qu'il assure, lorsque comme nous différents runs indépendants de l'algorithme convergent vers la même solution, une identifiabilité dite numérique du résultat. En d'autres mots, la recherche du minimum

de la fonction coût est aléatoire. Ainsi la convergence vers le même minimum une dizaine de fois de suite assure que le jeu de paramètres obtenu est le meilleur sous deux critères : la minimisation de la fonction de coût en ce jeu de paramètres et la minimisation de la fonction de coût au voisinage de ce jeu de paramètres. Ainsi un jeu de paramètres pour lequel une légère variation autour de sa valeur induit une forte augmentation de la valeur de la fonction coût ne sera pas choisi par l'algorithme.

La solution $x(t, \theta)$ du système (1), où θ représente le jeu de paramètres obtenu à l'aide de l'algorithme, est présentée en figure 3.

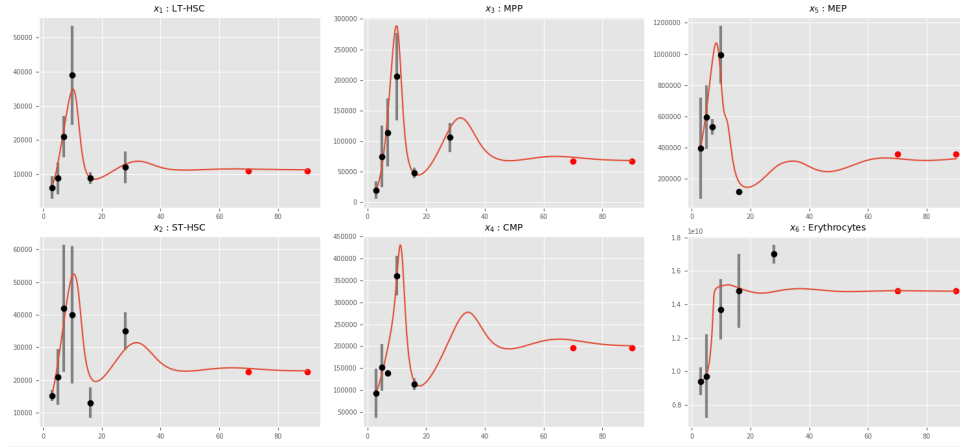


FIGURE 3 – Evolution au cours du temps du nombre de cellules de chaque type érythrocytaire modélisé pendant leur période de retour à l'équilibre suite à une injection de PHZ.

Le modèle présenté permet d'expliquer les résultats obtenus après la perturbation définie par les données au jour 3. Il est intéressant de remarquer que l'ajout de deux équations dans notre modèle pour décrire la dynamique de la régulation est un point clé permettant d'expliquer le retour à l'équilibre du système.

Un deuxième point intéressant concerne les données recueillies avant le jour 3. Nos collaborateurs ont remarqué une baisse des tailles de compartiments observés mais également une absence de division cellulaire durant ces trois jours. Ces résultats laissent à penser qu'en état de stress important les cellules peuvent se différencier sans se diviser, ce qui serait en accord avec l'idée selon laquelle les érythropoïèses de stress et de repos sont régies par des mécanismes cellulaires différents.

Résultats du chapitre 2

Ce chapitre est constitué de l'article [12], intitulé "Large fluctuations in multi-scale modeling for rest erythropoiesis", écrit en collaboration avec Sylvie Méléard, soumis à *Journal of Mathematical Biology*.

On s'intéresse ici au lien qui existe entre la probabilité d'auto-renouvellement des cellules précédemment évoquée et l'amplification observée entre les compartiments.

Nous savons, grâce au travail du Chapitre 1, d'une part que les équilibres du modèle de l'érythropoïèse de repos (1) vérifient

$$x_i^* = x_{i-1}^* \frac{2\mu_i^* p_i^{D*} \tau_i^*}{(p_i^{D*} - p_i^{R*}) \tau_{i-1}^*},$$

d'autre part que pour $i \geq 2$, $p_i^{D*} - p_i^{R*} \sim 10^{-1}$ (voir tableau présenté dans la sous section précédente 'Calibration du modèle : étape 1').

L'amplification entre deux compartiments est donc proportionnelle à la différence des probabilités de différenciation et d'auto-renouvellement des cellules

$$x_i^* \propto \frac{x_{i-1}^*}{p_i^{D*} - p_i^{R*}}. \quad (6)$$

En s'inspirant de cette réflexion issue du travail effectué au Chapitre 1, nous avons développé un modèle stochastique qui s'est avéré être un modèle multi-échelle.

Un modèle multi-échelle

Afin de comprendre plus précisément la dynamique induite par ce mécanisme d'amplification biologique particulier nous avons introduit un processus de branchement N^K de dimension 3. Le premier type cellulaire correspond aux LT-HSC. Le deuxième représente un type de cellules intermédiaires ayant une probabilité d'auto-renouvellement strictement positive. Le troisième type représente les globules rouges.

On suppose qu'une cellule de type 1 se divise à taux constant $\tau_1 > 0$ en deux cellules. Les deux cellules filles sont de type 1 avec probabilité 1/2 ou de type 2 avec la même probabilité. Une cellule de type 2 se divise à taux constant $\tau_2 > 0$ en deux cellules filles. Les deux cellules filles sont de type 2 avec probabilité p_2^R ou de type 3 avec probabilité $p_2^D = 1 - p_2^R$. Une cellule de type 3 meurt à taux constant $d_3 > 0$.

On suppose de plus

- L'existence de $K \in \mathbb{R}^+$ un paramètre d'échelle tel que

$$N_1^K(0) \sim K;$$

- L'existence de $1 > \gamma_3 > \gamma_2 > 0$ et $\tau_3 > 0$ telles que

$$p_2^D - p_2^R = K^{-\gamma_2} \quad \text{et} \quad d_3 = \tau_3 K^{-\gamma_3}.$$

Nous avons montré qu'une telle dynamique traduisait bien un mécanisme d'amplification. Pour cela nous avons étudié la convergence en loi lorsque K tend vers l'infini du processus N^K renormalisé. Un premier résultat nous a permis de réaliser la nécessité d'étudier

le processus sur une échelle de temps différente pour observer la dynamique des deux dernières composantes du processus. En effet nous avons montré que le processus N^K renormalisé en taille par K converge en loi, lorsque K tend vers l'infini, vers une fonction continue qui tend vers l'infini. Nous avons ensuite montré que lorsque le processus N^K est renormalisé en espace par les ordres de grandeurs qui semblent les plus appropriés celui-ci admet une limite dégénérée lorsque K tend vers l'infini.

Proposition 1. (i) Supposons que la suite $\left(\frac{N_1^K(0)}{K}, \frac{N_2^K(0)}{K}, \frac{N_3^K(0)}{K}\right)_{K \in \mathbb{N}^*}$ converge en loi, lorsque K tend vers l'infini, vers un vecteur $(x_1, 0, 0) \in \mathbb{R}_+^3$ et que

$$\sup_K \mathbb{E}\left[\frac{N_1^K(0)}{K}\right] < +\infty, \quad \sup_K \mathbb{E}\left[\frac{N_2^K(0)}{K}\right] < +\infty \quad \text{et} \quad \sup_K \mathbb{E}\left[\frac{N_3^K(0)}{K}\right] < +\infty.$$

Alors pour tout $T > 0$, la suite $\left(\left(\frac{N_1^K(t)}{K}, \frac{N_2^K(t)}{K}, \frac{N_3^K(t)}{K}\right), t \in [0, T]\right)_{K \in \mathbb{N}^*}$ converge en loi dans $\mathbb{D}([0, T], \mathbb{R}_+^3)$ vers $x_1(1, \tau_1 t, \frac{\tau_2}{2} t^2)$.

(ii) Soit le processus de sauts X^K défini pour tout $t \geq 0$ par

$$X^K(t) = \left(\frac{N_1^K(t)}{K}, \frac{N_2^K(t)}{K^{1+\gamma_2}}, \frac{N_3^K(t)}{K^{1+\gamma_2+\gamma_3}}\right).$$

Supposons que la suite $\left(\frac{N_1^K(0)}{K}, \frac{N_2^K(0)}{K^{1+\gamma_2}}, \frac{N_3^K(0)}{K^{1+\gamma_2+\gamma_3}}\right)_{K \in \mathbb{N}^*}$ converge en loi, lorsque K tend vers l'infini, vers un vecteur $(x_1, 0, 0) \in \mathbb{R}_+^3$ et que

$$\sup_K \mathbb{E}\left[\frac{N_1^K(0)}{K}\right] < +\infty, \quad \sup_K \mathbb{E}\left[\frac{N_2^K(0)}{K^{1+\gamma_2}}\right] < +\infty \quad \text{et} \quad \sup_K \mathbb{E}\left[\frac{N_3^K(0)}{K^{1+\gamma_2+\gamma_3}}\right] < +\infty.$$

Alors pour tout $T > 0$, la suite $\left(\left(\frac{N_1^K(t)}{K}, \frac{N_2^K(t)}{K^{1+\gamma_2}}, \frac{N_3^K(t)}{K^{1+\gamma_2+\gamma_3}}\right), t \in [0, T]\right)_{K \in \mathbb{N}^*}$ converge en loi dans $\mathbb{D}([0, T], \mathbb{R}_+^3)$ vers $(x_1, 0, 0)$.

Nous avons ensuite montré que chaque composante du processus N^K évolue selon ces échelles de taille et de temps typiques. Nous regroupons les résultats de ce chapitre décrivant la dynamique asymptotique pour K grand du processus N^K par composante et donc par échelle typique.

Comportement asymptotique de la 2ème coordonnée

Commençons par illustrer graphiquement (Figure 4) les résultats présentés dans cette sous section.

On observe sur ces simulations qu'à une échelle de temps d'ordre 1 (image de gauche), le processus N_2^K est loin d'avoir atteint l'ordre de grandeur de sa valeur moyenne à l'équilibre. On observe également (image au centre) que l'échelle de temps K^{γ_2} semble correspondre à l'échelle de temps typique pour observer la dynamique du processus N_2^K . Sur l'image de droite on peut observer que les fluctuations du processus N_2^K autour de sa valeur moyenne à l'équilibre semblent très grandes (comparées à celles observés pour des processus stochastiques de naissance et mort classiques).

Nous avons montré que l'échelle de temps K^{γ_2} était adapté à l'observation de la dynamique du deuxième type cellulaire lorsque celui-ci est renormalisé en espace par $K^{1+\gamma_2}$.

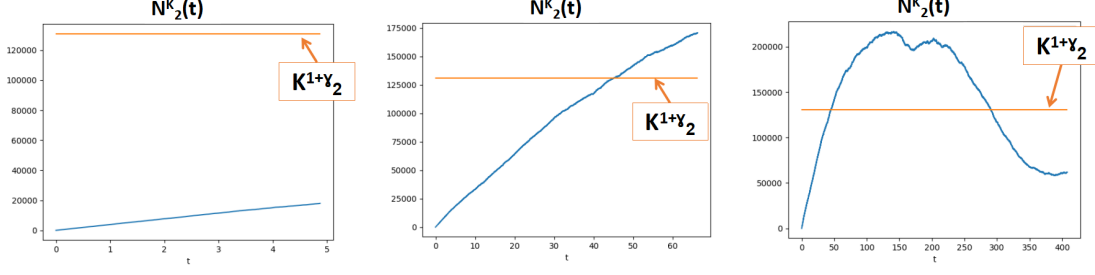


FIGURE 4 – Une trjetoire du processus N_2^K sur différentes échelles de temps $t \in [0, T]$ avec $T = O(1)$, $T = O(K^{\gamma_2})$ puis $T = O(K^{\gamma_3})$. Les paramètres sont supposés égaux à 1 à l'exception de $K = 2000$, $\gamma_2 = 0.55$ et $\gamma_3 = 0.8$.

En effet, notons pour tout $t \geq 0$,

$$Y^K(t) = \left(\frac{N_1^K(t K^{\gamma_2})}{K}, \frac{N_2^K(t K^{\gamma_2})}{K^{1+\gamma_2}} \right).$$

Nous avons dans un premier temps décrit la limite au premier ordre du processus Y^K en démontrant le théorème suivant.

Théorème 3. *Supposons l'existence d'un vecteur $(x_1, x_2) \in \mathbb{R}_+^2$ tel que la suite $(Y^K(0))_{K \in \mathbb{N}^*}$ converge en loi vers (x_1, x_2) lorsque K tend vers l'infini. Supposons de plus que*

$$\sup_K \mathbb{E}[Y_1^K(0)^2 + Y_2^K(0)^2] < \infty.$$

Alors pour tout $T > 0$, la suite $(Y^K)_{K \in \mathbb{N}^}$ converge en loi (et en probabilité) dans $\mathbb{D}([0, T], \mathbb{R}_+^2)$ vers la fonction continue $y = (y_1, y_2)$ telle que pour tout $t \geq 0$,*

$$\begin{cases} y_1(t) &= x_1 \\ y_2(t) &= \frac{\tau_1 x_1}{\tau_2} + \left(x_2 - \frac{\tau_1 x_1}{\tau_2}\right) e^{-\tau_2 t}. \end{cases} \quad (7)$$

Nous avons ensuite étudié la limite au deuxième et troisième ordre du processus Y^K en démontrant le théorème suivant.

Théorème 4. *Notons $(U^K)_K$ la suite de processus définie par*

$$\forall t \geq 0, \quad U^K(t) = K^{(1-\gamma_2)/2} (Y^K(t) - y(t)).$$

(i) Supposons l'existence de vecteur aléatoire $U_0 = (U_0^{(1)}, U_0^{(2)}) \in \mathbb{R}^2$ vers lequel la suite $(U^K(0))_{K \in \mathbb{N}^}$ converge en loi. Supposons de plus que*

$$\sup_K \mathbb{E}[U_1^K(0)^2 + U_2^K(0)^2] < +\infty. \quad (8)$$

Alors pour tout $T > 0$, la suite $(U^K)_{K \in \mathbb{N}^}$ converge en loi dans $\mathbb{D}([0, T], \mathbb{R}^2)$ vers le processus $U = (U_1, U_2)$ défini pour tout $t \geq 0$ par*

$$\begin{aligned} U_1(t) &= U_0^{(1)} + \sqrt{\tau_1 x_1} B_1(t), \\ U_2(t) &= U_0^{(2)} + \tau_1 \int_0^t U_1(s) ds - \tau_2 \int_0^t U_2(s) ds, \end{aligned}$$

où B_1 est un mouvement Brownien standard.

(ii) De plus, pour tout $T > 0$ la suite $(W_2^K)_{K \in \mathbb{N}^*}$ définie par

$$\forall t \in [0, T], \quad W_2^K(t) = K^{\gamma_2/2} [U_2^K(t) - U_0^{(2)} - \tau_1 \int_0^t U_1^K(s) ds + \tau_2 \int_0^t U_2^K(s) ds],$$

converge en loi dans $\mathbb{D}([0, T], \mathbb{R})$ vers le processus $(\sqrt{\tau_2 y_2(t)} B_2(t), \quad t \in [0, T])$ où B_2 est un mouvement Brownien standard indépendant du processus B_1 .

Supposons que U_0 vaut 0, l'approximation asymptotique en K du processus N_2^K s'écrit alors pour tout $t \geq 0$,

$$N_2^K(t) \sim K^{1+\gamma_2} y_2(t K^{-\gamma_2}) + K^{(1+3\gamma_2)/2} U_2(t K^{-\gamma_2}) + K^{(1+2\gamma_2)/2} \sqrt{\tau_2 y_2(t K^{-\gamma_2})} B_2(t K^{-\gamma_2})$$

où

$$\forall t, \quad U_2(t) = \tau_1 \sqrt{\tau_1 x_1} \int_0^t B_1(s) ds - \tau_2 \int_0^t U_2(s) ds$$

et B_1, B_2 sont des mouvements browniens indépendants.

Les preuves de ces résultats sont classiques toutefois l'obtention d'une telle asymptotique est peu commune. En effet on observe que les fluctuations de la première composante sont amplifiées par la dynamique du processus et apparaissent dans l'approximation au deuxième ordre de $N_2^K(t)$ par un terme en

$$K^{(1+3\gamma_2)/2} \int_0^t B_1(s) ds.$$

L'ordre de grandeur des fluctuations est donc plus grand que celui habituellement observé en racine de la taille de la population (i.e en $K^{(1+\gamma_2)/2}$). De plus les fluctuations browniennes de la première composante interviennent ici intégrées sous la forme d'un processus à variation finie.

Comportement asymptotique de la 3ème coordonnée

Nous nous sommes finalement intéressés à la dynamique de la troisième composante dont l'échelle de temps typique est donnée par K^{γ_3} et l'échelle de taille est donnée par $K^{1+\gamma_2+\gamma_3}$.

Soit le processus Z^K défini pour tout $t \geq 0$ par

$$Z^K(t) = \left(\frac{N_1^K(t K^{\gamma_3})}{K}, \frac{N_2^K(t K^{\gamma_3})}{K^{1+\gamma_2}}, \frac{N_3^K(t K^{\gamma_3})}{K^{1+\gamma_2+\gamma_3}} \right).$$

Nous avons montré que sous certaines hypothèses le processus Z^K admet la décomposition en semi-martingale suivante

$$\begin{aligned} Z_1^K(t) &= Z_1^K(0) + \widetilde{M}_1^K(t) \\ Z_2^K(t) &= Z_2^K(0) + \tau_1 K^{\gamma_3-\gamma_2} \int_0^t Z_1^K(s) ds - \tau_2 K^{\gamma_3-\gamma_2} \int_0^t Z_2^K(s) ds + \widetilde{M}_2^K(t) \\ Z_3^K(t) &= Z_3^K(0) + \left(1 + \frac{1}{K^{\gamma_2}}\right) \tau_2 \int_0^t Z_2^K(s) ds - \tau_3 \int_0^t Z_3^K(s) ds + \widetilde{M}_3^K(t) \end{aligned} \quad (9)$$

où $\widetilde{M}^K$ est une martingale qui converge en loi vers 0 lorsque K tend vers l'infini. On en a déduit à l'aide de la méthode de moyennisation développée par Thomas Kurtz le théorème suivant

Théorème 5. *Supposons l'existence d'un vecteur $(x_1, x_2, x_3) \in \mathbb{R}_+^3$ vers lequel converge en loi la suite $(Z^K(0))_{K \in \mathbb{N}^*}$ lorsque K tend vers l'infini. Supposons de plus que*

$$\sup_K \mathbb{E}[Z_1^K(0)] < +\infty, \quad \sup_K \mathbb{E}[Z_2^K(0)] < +\infty \quad \text{et} \quad \sup_K \mathbb{E}[Z_3^K(0)] < +\infty. \quad (10)$$

Notons Γ_2^K la mesure d'occupation associée au processus Z_2^K ,

$$\Gamma_2^K([0, t] \times B) = \int_0^t \mathbf{1}_{\{B\}}(\mathbf{Z}_2^K(s)) ds.$$

Alors pour tout $T > 0$, la suite $(Z_1^K, \Gamma_2^K, Z_3^K)_{K \in \mathbb{N}^*}$ converge en loi dans $\mathbb{D}([0, T], \mathbb{R}_+) \times l_m(\mathbb{R}_+) \times \mathbb{D}([0, T], \mathbb{R}_+)$ vers $(z_1, \delta_{z_2^*}(dz_2) ds, z_3)$. Les fonctions continues z_1 et z_3 sont définies pour tout $t \leq T$ par

$$\begin{cases} z_1(t) &= x_1 \\ z_3(t) &= \frac{\tau_2}{\tau_3} z_2^* + \left(x_3 - \frac{\tau_2}{\tau_3} z_2^*\right) e^{-\tau_3 t} \end{cases} \quad (11)$$

et z_2^* est la valeur vers laquelle la fonction y_2 tend à l'infini :

$$z_2^* = \frac{\tau_1 x_1}{\tau_2}.$$

À la différence des résultats précédents le Théorème 5 introduit un système lent rapide et nécessite l'utilisation de la mesure d'occupation associée à la seconde coordonnée. En effet, dans une telle échelle de temps la deuxième composante du processus a déjà atteint son équilibre. De plus la suite $(Z_2^K)_K$ ne peut être suffisamment contrôlée (voir l'expression (9)) pour vérifier des critères de tension à la différence de sa mesure d'occupation Γ_2^K , qui nécessite un contrôle plus faible,

$$\forall t > 0, \quad \sup_K \mathbb{E}\left[\int_0^t Z_2^K(s) ds\right] < +\infty.$$

La preuve de ce théorème utilise des méthodes de moyennisation introduite par Thomas Kurtz [62] et développées dans l'article [54].

Néanmoins, les échelles apparues dans notre modèle sont différentes. Cette différence implique que les variables lentes et rapides de notre modèle jouent des rôles différents au niveau des fluctuations du processus N_3^K . Nous avons ainsi décrit le comportement asymptotique lorsque K tend vers l'infini des fluctuations associées au processus N_3^K .

La démonstration de ce résultat nécessite d'identifier au niveau de la dynamique de la troisième coordonnée les fluctuations de la première. Pour cela nous avons introduit le processus

$$V_2^K(t) = \frac{K^{(1-\gamma_3)/2}}{K^{\gamma_3-\gamma_2}} (Z_2^K(t) - Z_2^K(0)).$$

La preuve de la convergence uniforme dans $\mathbb{L}^2$ sur tout intervalle de temps fini du processus V_2^K vers 0, a permis d'établir le théorème principal de ce chapitre.

Théorème 6. *Soient*

$$\forall t \geq 0, \quad \begin{cases} V_1^K(t) &= K^{(1-\gamma_3)/2} (Z_1^K(t) - x_1) \\ V_2^K(t) &= \frac{K^{(1-\gamma_3)/2}}{K^{\gamma_3-\gamma_2}} (Z_2^K(t) - Z_2^K(0)) \\ V_3^K(t) &= K^{(1-\gamma_3)/2} (Z_3^K(t) - z_3(t)) \end{cases}.$$

Supposons qu'il existe $V_0 = (V_0^{(1)}, V_0^{(3)})$ vecteur aléatoire de $\mathbb{R}^2$ vers lequel la suite $(V_1^K(0), V_3^K(0))_{K \in \mathbb{N}^}$ converge en loi. Supposons de plus que*

$$\sup_K \mathbb{E}[V_1^K(0)^4] < +\infty \quad ; \quad \sup_K \mathbb{E}[Z_2^K(0)^2] < +\infty. \quad (12)$$

$$\sup_K \mathbb{E}[|V_3^K(0)|] < +\infty. \quad (13)$$

Alors pour tout $T > 0$, la suite $(V_1^K, V_3^K)_{K \in \mathbb{N}^}$ converge en loi dans $\mathbb{D}([0, T], \mathbb{R}^2)$ vers les processus continus (V_1, V_3) tels que pour tout t ,*

$$V_1(t) = V_0^{(1)} + \sqrt{\tau_1 x_1} W_1(t)$$

$$V_3(t) = V_0^{(3)} + \tau_1 \int_0^t V_1(s) ds - \tau_3 \int_0^t V_3(s) ds,$$

où W_1 est un mouvement brownien standard.

En supposant que le vecteur V_0 est nul, on peut écrire l'approximation du processus $N_3^K(t)$ pour tout t et pour K grand,

$$N_3^K(t) \sim K^{1+\gamma_2+\gamma_3} z_3(t K^{-\gamma_3}) + K^{(1+2\gamma_2+3\gamma_3)/2} V_3(t K^{-\gamma_3}) \quad (14)$$

avec

$$\forall t, \quad V_3(t) = \tau_1 \sqrt{\tau_1 x_1} \int_0^t W_1(s) ds - \tau_3 \int_0^t V_3(s) ds$$

et W_1 un mouvement brownien standard.

Habituellement le théorème de la limite centrale décrit les fluctuations d'une population d'ordre $K^{1+\gamma_2+\gamma_3}$ comme étant d'ordre $K^{(1+\gamma_2+\gamma_3)/2}$. Ici les fluctuations de la première coordonnée sont prises dans le mécanisme d'amplification du processus et intègrent la dynamique de la troisième coordonnée par l'intermédiaire d'une terme d'ordre $K^{(1+2\gamma_2+3\gamma_3)/2} \gg K^{(1+\gamma_2+\gamma_3)/2}$.

Remarquons de plus que les fluctuations de l'hématocrite chez la souris en condition de repos ont été observées grâce à des souris contrôles lors des expériences effectuées dans [19]. On déduit de ces observations une variation d'environ 10% du nombre de globules rouges. Or en reprenant les valeurs obtenues au Chapitre 1 et le travail effectué ici on obtient une variation d'environ $K^{(\gamma_3-1)/2} \sim 6\%$.

Résultats du chapitre 3

Ce chapitre est constitué d'un travail en cours dans lequel nous étudions l'influence des phases de quiescence d'une unique cellule souche mutante sur l'évolution au cours du temps du nombre de globules rouges mutants qu'elle produit et sur l'évolution au cours du temps du nombre de cellules saines affectées par l'intermédiaire de régulations.

Nous nous intéressons plus particulièrement à un système à 4 populations de cellules : les cellules souches et les globules rouges sains et mutants. Nous modélisons l'amplification qui existe entre les cellules souches et les globules rouges à l'aide d'un taux de division asymétrique très élevé. Afin de quantifier ces ordres de grandeurs, nous noterons $K \in \mathbb{N}^*$ le paramètre d'échelle.

Nous supposons que les cellules souches saines sont en nombre constant, égal à K , au cours du temps et peuvent passer d'un état actif à un état quiescent et inversement. Ainsi une fois le nombre de cellules souches saines actives déterminé, on en déduit immédiatement le nombre de cellules souches quiescentes saines présentes dans le système.

Nous supposons de plus que l'amplification entre les cellules souches saines actives et les globules rouges sains se fait par l'intermédiaire d'un taux de division asymétrique d'ordre K^α où $\alpha > 0$. Une telle dynamique d'amplification permet de passer d'une population de cellules souches saines d'ordre K à une population de globules rouges sains d'ordre $K^{1+\alpha}$. De la même façon, nous supposons que les globules rouges mutants sont produits par division asymétrique de la cellule mutante saine lorsque celle-ci est active. Le taux de division asymétrique est proportionnel à K^β , avec $\beta > 0$. Ainsi bien que la population de globules rouges mutants ne bénéficie pas d'un réservoir de cellules souches mutantes en grand nombre, elle admet un facteur d'amplification différent de celui des saines (potentiellement plus élevé).

Notre modèle décrit la dynamique de trois types cellulaires dont les effectifs respectifs au cours du temps sont notés,

- N_1^K : le nombre de cellules souches actives ;
- N_2^K : le nombre de globules rouges sains ;
- N_3^K : le nombre de globules rouges mutants.

Chacune de ces trois populations est en grand nombre à la différence de l'unique cellule souche mutante dont l'état est désigné par le processus I^K ,

- $I^K = 1$: lorsque la cellule souche mutante est active ;
- $I^K = 0$: lorsque celle-ci est quiescente.

La dynamique des cellules est la suivante. Les cellules souches saines (respectivement mutantes) basculent de l'état actif à l'état quiescent à taux constant $a > 0$ (respectivement $a_M > 0$). Elles basculent de l'état quiescent à l'état actif à un taux q^K (respectivement q_M^K). Les taux q^K et q_M^K sont des fonctions continues croissantes dépendantes du nombre de globules rouges sains et du nombre de globules rouges mutants. Elles dépendent de K de la façon suivante,

$$\begin{cases} q^K(N_2^K, N_3^K) = q\left(\frac{N_2^K}{K^{1+\alpha}}, \frac{N_3^K}{K^\beta}\right) = q_1 + q_2 \frac{N_2^K}{K^{1+\alpha}} + q_3 \frac{N_3^K}{K^\beta}, \\ q_M^K(N_2^K, N_3^K) = q_M\left(\frac{N_2^K}{K^{1+\alpha}}, \frac{N_3^K}{K^\beta}\right) = q_{1,M} + q_{2,M} \frac{N_2^K}{K^{1+\alpha}} + q_{3,M} \frac{N_3^K}{K^\beta} \end{cases},$$

avec $q_1 > 0$, $q_{1,M} > 0$, $q_2 \geq 0$, $q_3 \geq 0$, $q_{2,M} \geq 0$ et $q_{3,M} \geq 0$.

Une cellule souche active se divise de manière asymétrique à taux τ pour les saines et τ_M pour la mutante. A chaque division elle produit une cellule souche active et un globule rouge. Ces taux de division sont régulés par les nombres de globules rouges sains et mutants. Ils dépendent de K de la façon suivante,

$$\tau(N^K) = K^\alpha r\left(\frac{N_2^K}{K^{1+\alpha}}, \frac{N_3^K}{K^\beta}\right), \quad \tau_M(N^K) = K^\beta r_M\left(\frac{N_2^K}{K^{1+\alpha}}, \frac{N_3^K}{K^\beta}\right)$$

avec $\alpha > 0$ et $\beta > 0$.

Les fonctions r et r_M s'écrivent pour tout $(y_2, y_3) \in \mathbb{R}_+^2$,

$$r(y_2, y_3) = \frac{c_1}{1 + c_2 y_2 + c_3 y_3} \quad r_M(y_2, y_3) = \frac{c_{1,M}}{1 + c_{2,M} y_2 + c_{3,M} y_3},$$

avec $c_1 > 0$, $c_{1,M} > 0$, $c_2 \geq 0$, $c_3 \geq 0$, $c_{2,M} \geq 0$ et $c_{3,M} \geq 0$.

Elles modélisent la régulation de la production des globules rouges sains, respectivement mutants en fonction du nombre à la fois de globules rouges sains et de globules rouges mutants dans le système. Les puissances α et β résument l'amplification entre le nombre de cellules souches et celui des globules rouges, due à une succession d'étapes de différenciation intermédiaires ([69]).

Les globules rouges ont quant-à-eux un taux de mort individuel constant noté $d > 0$ pour les sains et $d_M > 0$ pour les mutants.

La dynamique du modèle peut alors être résumée ainsi,

$$\begin{array}{llll} N_1^K & \longrightarrow & N_1^K + 1 & \text{à taux } a(K - N_1^K) \\ N^K & \longrightarrow & (N_1^K - 1, N_2^K, N_3^K) & \text{à taux } q^K(N_2^K, N_3^K) N_1^K \\ N^K & \longrightarrow & (N_1^K, N_2^K + 1, N_3^K) & \text{à taux } K^\alpha r\left(\frac{N_2^K}{K^{1+\alpha}}, \frac{N_3^K}{K^\beta}\right) N_1^K \\ N_2^K & \longrightarrow & N_2^K - 1 & \text{à taux } d N_2^K \\ (N^K, I^K) & \longrightarrow & (N_2^K, N_3^K + 1, I^K) & \text{à taux } K^\beta r_M\left(\frac{N_2^K}{K^{1+\alpha}}, \frac{N_3^K}{K^\beta}\right) I^K \\ N_3^K & \longrightarrow & N_3^K - 1 & \text{à taux } d_M N_3^K \\ I^K & \longrightarrow & I^K + 1 & \text{à taux } a_M(1 - I^K) \\ (N_2^K, N_3^K, I^K) & \longrightarrow & (N_2^K, N_3^K, I^K - 1) & \text{à taux } q_M^K(N_2^K, N_3^K) I^K. \end{array}$$

Rappelons que K représente le nombre total de cellules souches saines au cours du temps. Ainsi $K - N_1^K$ représente le nombre de cellules souches quiescentes saines.

Notation. Soit $E = [0, 1] \times \mathbb{R}_+^2$ muni de la norme suivante

$$\forall x \in E, \|x\|^2 = \sum_i x_i^2.$$

Soit $\mathcal{E}$ le sous-ensemble de $\mathbb{R}_+^4$ défini par $\mathcal{E} = E \times \{0, 1\}$.

L'ensemble $C_b^1(\mathcal{E})$ est défini par $C_b^1(\mathcal{E}) = \{f : \mathcal{E} \rightarrow \mathbb{R} \text{ bornée dont la restriction à } E \text{ est } C^1\}$.

A l'aide d'une renormalisation adéquate du processus N^K nous avons montré la convergence en loi du processus, lorsque K tend vers l'infini, vers un processus de Markov déterministe par morceaux (PDMP).

Théorème 7. Soient $T > 0$ et X^K le processus stochastique à valeurs dans $\mathbb{D}([0, T], E)$ défini par

$$\forall t \in [0, T], \quad X^K(t) = \left(\frac{N_1^K(t)}{K}, \frac{N_2^K(t)}{K^{1+\alpha}}, \frac{N_3^K(t)}{K^\beta} \right).$$

Supposons que le vecteur aléatoire $(X^K(0), I^K(0))$ converge en loi vers $(x_0, i_0) \in \mathcal{E}$ et vérifie

$$\sup_K \mathbb{E}[\|X^K(0)\|^2] < \infty. \quad (15)$$

Alors la suite $((X^K(t), I^K(t)), t \in [0, T])_K$ converge en loi dans $\mathbb{D}([0, T], \mathcal{E})$, lorsque K tend vers l'infini, vers l'unique processus stochastique (X, I) à valeurs dans $\mathcal{E}$ de condition initiale (x_0, i_0) et de générateur infinitésimal $\mathcal{L}$ défini pour $f \in C_b^1(\mathcal{E})$ par

$\forall x \in E, i \in \{0, 1\},$

$$\mathcal{L}f(x, i) = \sum_{j=1}^3 \left(\frac{\partial f}{\partial x_j}(x, i) g_j(x, i) \right) + a_M(1-i)(f(x, i+1) - f(x, i)) + q_M(x_2, x_3)i(f(x, i-1) - f(x, i))$$

où

$$g(x, i) = (a - (a + q(x_2, x_3))x_1, r(x_2, x_3)x_1 - dx_2, r_M(x_2, x_3)i - d_Mx_3)^T. \quad (16)$$

Bien que ce résultat aurait pu être prouvé en utilisant les arguments de [20], nous avons préféré démontrer ce théorème par une preuve directe utilisant des arguments de tension pour extraire une sous-suite convergente puis en identifiant la limite à l'aide du Théorème de Cauchy-Lipschitz et d'un résultat de représentation des trajectoires de processus de saut pur ([66], Th 22).

Comportement en temps long du PDMP obtenu.

Les processus de Markov déterministes par morceaux ont été définis par Mark Davis ([21]) à l'aide de trois caractéristiques :

- un ensemble de champs de vecteurs $(g(\cdot, i))_{i \in \{0,1\}}$ dont le flot associé partant de $x \in E$ est noté $t \mapsto \phi_t^i(x)$.

Pour le processus limite (X, I) , les champs de vecteurs sont donnés par (16).

- un taux de saut $\tau : E \times \{0, 1\} \rightarrow \mathbb{R}_+$ défini pour (X, I) par

$$\forall (x, i) \in \mathcal{E}, \quad \tau(x, i) = \tau(i) = a_M(1-i) + q_M(x_2, x_3)i.$$

- une mesure de transition $Q : E \times \{0, 1\} \rightarrow \mathcal{P}(E)$ telle que pour toute fonction $f : E \times \{0, 1\}$ continue sur E , pour tout $(x, i) \in E \times \{0, 1\}$

$$Qf(x, i) = \int_{E \times \{0,1\}} f(z) Q(dz; (x, i)) = \frac{1}{\tau(x, i)} ((f(x, i+1) a_M(1-i) + f(x, i-1) q_M(x_2, x_3)i).$$

On notera pour $(x, i) \in \mathcal{E}$, $Q(x, i, 1-i) = \frac{a_M(1-i) + q_M(x_2, x_3)i}{\tau(x, i)}$ et $Q(x, i, i) = 1 - Q(x, i, 1-i)$

Le processus (X, I) est un PDMP particulier dont les trajectoires sont presque sûrement enfermées dans un compact que nous noterons B . Cette catégorie de processus a été étudiée dans [7] où les auteurs détaillent une méthode permettant de construire le processus (X, I) à l'aide de sa chaîne incluse $(Z_n)_n$. Rappelons cette construction afin de nous familiariser avec la dynamique du processus limite (X, I) dont l'étude du comportement en temps long est l'objet des résultats de ce chapitre.

Soit (U_n) une suite de variables aléatoires de loi exponentielle de paramètre

$$\lambda = \max_{x \in B} (Q(x, 0, 1) + Q(x, 1, 0)).$$

Le processus (X, I) s'écrit pour tout $t \in [T_n, T_n + U_{n+1}]$ avec $T_n = \sum_{j=0}^n U_j$,

$$(x_t, I_t) = \left(\phi_{t-T_n}^{\tilde{I}_n}(\tilde{X}_n), \tilde{I}_n \right)$$

où

- $Z_0 = (\tilde{X}_0, \tilde{I}_0)$ est une variable aléatoire appartenant à $B \times \{0, 1\}$ indépendant de $(U_n)_n$.
- Pour tout $n \in \mathbb{N}$,

$$Z_{n+1} = (\tilde{X}_{n+1}, \tilde{I}_{n+1})$$

avec

$$\tilde{X}_{n+1} = \phi_{U_{n+1}}^{\tilde{I}_n}(\tilde{X}_n) \quad \text{et} \quad \mathbb{P}(\tilde{I}_{n+1} = j | \tilde{X}_{n+1}, \tilde{I}_n = i) = \frac{Q(x, i, j)}{\lambda}, \quad i, j \in \{0, 1\}.$$

Nous avons dans un premier temps décrit le compact dans lequel le processus (X, I) est confiné.

Proposition 1. *Soit l'ensemble compact $B \subset E$ tel que*

$$B = [b_1, B_1] \times [b_2, \frac{c_1}{d}] \times [0, \frac{c_{1,M}}{d_M}]$$

avec

$$b_1 = \frac{a}{a + q_1 + q_2 \frac{c_1}{d} + q_3 \frac{c_{1,M}}{d_M}} \quad , \quad b_2 = \frac{c_1 b_1}{d (c_2 \frac{c_1}{d} + c_3 \frac{c_{1,M}}{d_M})}$$

et

$$B_1 = \frac{a}{a + q_1}.$$

Alors B est positivement invariant par rapport aux flots ϕ^i , $i \in \{0, 1\}$, autrement dit

$$\forall i \in \{0, 1\}, \forall t \geq 0, \quad \phi_t^i(B) \subset B.$$

De plus pour toute condition initiale du processus (X, I) , appartenant à $E \times \{0, 1\}$, il existe $t \geq 0$ tel que

$$(X(t), I(t)) \in B \times \{0, 1\}.$$

Nous avons ensuite décrit de façon rigoureuse l'ensemble des trajectoires de probabilité non nulle du processus (X, I) . Pour cela nous avons introduit la notion de point accessible pour le processus (X, I) .

Posons pour tout $n \in \mathbb{N}^*$,

$$\mathbb{T}_n = \{(\bar{i}, \bar{u}) = (i_1, \dots, i_n), (u_1, \dots, u_n) \in \{0, 1\}^n \times \mathbb{R}_+^n\}.$$

Alors les trajectoires de (X, I) peuvent s'écrire à l'aide des flots $(\phi^i)_i$ de la façon suivante, pour $x \in B$ et $(\bar{i}, \bar{u}) \in \mathbb{T}_n$,

$$\Phi_{\bar{u}}^{\bar{i}}(x) = \phi_{u_n}^{i_n} \circ \dots \circ \phi_{u_1}^{i_1}(x).$$

Une trajectoire de X se construit en suivant le flot associé au champ de vecteur $g(\cdot, i_j)$ sur un laps de temps donné par le réel u_j .

Soit $x \in B$, on définit l'orbite positive de x par

$$\gamma^+(x) = \{\Phi_{\bar{u}}^{\bar{i}}(x) : (\bar{i}, \bar{u}) \in \bigcup_{n \in \mathbb{N}^*} \mathbb{T}_n\}.$$

On dit alors qu'un point x est accessible d'un singleton $\{y\}$ si $x \in \overline{\gamma^+(y)}$. De façon plus générale, on définit l'ensemble Γ des points accessibles pour le processus (X, I) par

$$\Gamma = \bigcap_{x \in B} \overline{\gamma^+(x)}.$$

Nous avons montré l'existence d'un unique point stationnaire asymptotiquement globalement stable pour le champ de vecteurs $g(\cdot, 0)$, noté p . Nous avons alors pu décrire l'ensemble Γ à l'aide de ce point stationnaire p dont la dernière coordonnée est nulle

$$\Gamma = \overline{\gamma^+(p)}.$$

L'ensemble de ces résultats nous a ensuite permis de montrer que le processus (X, I) admet une unique probabilité invariante, qu'elle est absolument continue par rapport à la mesure de Lebesgue, que son support est donné par $\Gamma \times \{0, 1\}$ et que le processus converge vers elle en variation totale.

Nous avons prouvé ce résultat en vérifiant les conditions (i) et (ii) suivantes (voir [6] et [67]).

Notons $\mathcal{F}_0 = \{F^i\}_{i \in \{0,1\}}$ puis $\mathcal{F}_{j+1} = F^j \cup \{[F^i, V], V \in \mathcal{F}_k\}$ où

$$[V, W](x) = DW(x)V(x) - DV(x)W(x), \quad x \in \mathbb{R}^3.$$

On dit que la condition (i) est vérifiée s'il existe $s \in \mathbb{R}$ tel que

$$s g(p, 0) + (1 - s) g(p, 1) = 0.$$

On dit de plus que la condition (ii) est vérifiée s'il existe $x \in B$ accessible de $\{p\}$ tel que

$$Vect(V(x) : V \in \bigcup_{j \geq 0} \mathcal{F}_j) = \mathbb{R}^3.$$

Nous avons finalement identifié la mesure invariante π , à l'aide du générateur de (X, I) et de la solution d'un système d'équations aux dérivées partielles.

Théorème 8. *Sous certaines hypothèses, il existe un couple de fonctions intégrables et positives (h_0, h_1) solution du système d'équations aux dérivées partielles suivant sur $\overset{\circ}{B} =]b_1, B_1[\times]b_2, \frac{c_1}{d}[\times]0, \frac{c_{1,M}}{d_M}[$,*

$$\begin{cases} g(x, 0)^T \nabla h_0(x) + h_0(x) \left[\sum_{j=1}^3 \partial_j g_j(x, 0) - a_M \right] & = -q_M(x_2, x_3) h_1(x) \\ g(x, 1)^T \nabla h_1(x) + h_1(x) \left[\sum_{j=1}^3 \partial_j g_j(x, 1) - q_M(x_2, x_3) \right] & = -a_M h_0(x) \end{cases} \quad (17)$$

telles que pour tout $i \in \{0, 1\}$,

$$\pi(dx, \{i\}) = \sum_{j=0}^1 \delta_{ij} h_j(x) dx \quad \text{et} \quad \int_B (h_0(x) + h_1(x)) dx = 1.$$

Nous avons ainsi démontré l'existence d'un couple de fonctions positives et intégrables, solution non triviale du système (17). La preuve de l'unicité des solutions de ce système semble pouvoir se déduire du résultat d'unicité obtenu pour la mesure invariante, en utilisant le générateur infinitésimal du PDMP (X, I) . (Ce dernier point reste à éclaircir.)

Résultats du chapitre 4 : CEMRACS 2018

Durant l'été 2018, j'ai eu l'occasion de participer au CEMRACS, une école d'été de 6 semaines au CIRM, à Marseille. Lors de ces 6 semaines, j'ai travaillé avec Romain Yvinec, Frédérique Clément et Marie Postel sur des questions de dynamique dite lente/rapide, présentes dans des modèles de croissance de follicules ovariens. Dans notre modèle, nous avons distingué deux types de follicules ovariens, primordiaux et croissants (de densités respectives ρ_0 et ρ). Les follicules primordiaux peuvent devenir des follicules croissants ou/et mourir, selon des taux très faibles (de l'ordre de $\epsilon \ll 1$), reflétant une dynamique très lente. De plus à l'instant initial considéré (la puberté), ces follicules sont en grand nombre (de l'ordre de $1/\epsilon$). Les follicules croissants, quant-à-eux, sont caractérisés par un trait x représentant, par exemple, la taille du follicule (en fonction des données, il peut être un nombre entier ou un réel). Ces hypothèses nous ont conduits à étudier un système dit lent-rapide selon trois modèles différents : le premier décrit par des équations différentielles ordinaires (EDO), le second par des équations aux dérivées partielles (EDP) et le dernier par des équations différentielles stochastiques. Le second modèle décrit à l'aide d'une EDP est constitué d'un couplage entre une EDO, décrivant une décroissance non linéaire de la densité ρ des follicules primordiaux ρ_0 en fonction du temps, et une équation de transport non linéaire sur la densité des follicules croissants. La fonction ρ dépend alors du temps et du trait continu x . La condition de bord associée à l'équation de transport est donnée par l'interaction entre les follicules primordiaux et croissants.

Nous avons étudié la convergence de ces trois modèles, lorsque ϵ tend vers 0, numériquement par simulations sur le langage Python et théoriquement dans le cas linéaire.

Ce travail a été accepté et sera publié dans ESAIM : Proceedings and Surveys ([11]).

Chapitre 1

Régulation des cellules souches hématopoïétiques : Une approche macroscopique.

Ce chapitre présente les résultats de l'article intitulé " Modeling the behavior of hematopoietic compartments from stem to red cells in steady state and stress hematopoiesis", écrit en collaboration avec Panhong Gou, Simon Girel, Vincent Bansaye, Catherine Lacout, Karine Bailly, Marie-Hélène Schlagetter, Evelyne Lauret, Sylvie Méléard et Stéphane Giraudier dont la soumission est prévue prochainement. Vous trouverez dans ce chapitre le travail mathématique entourant les résultats, plus approfondi que celui présenté dans l'article.

1.1 Introduction

L'hématopoïèse est un mécanisme cellulaire permettant la production de cellules sanguines (globules rouges, globules blancs, plaquettes...) par différenciations successives de cellules souches dites cellules souches hématopoïétiques. Dans ce chapitre nous nous intéresserons à la modélisation de ce mécanisme en nous focalisant exclusivement sur la différenciation en globules rouges. On parle alors d'érythropoïèse. La modélisation mathématique de ces mécanismes est très variée. En effet il existe des modèles déterministes, mais également probabilistes, certains intégrant l'apparition de mutations [25], d'autres l'effet de traitements [89], décrivant les stades de différenciation cellulaire selon un continuum [29] ou un nombre de compartiments fixé [74]. Une description plus détaillée des différents modèles mathématiques en hématopoïèse a été présentée en introduction de la thèse.

Bien que de nombreuses études s'intéressent à la dynamique cellulaire hématopoïétique, aucune n'a encore construit de modèle où 5 types cellulaires, connus pour leur capacité à s'auto-renouveler, sont étudiés, en situation de stress et de repos et où chacune des populations est confrontée à des données réelles in vivo en différents instants.

Notre travail s'inscrit dans la même veine que le travail de Busch et al [13]. Dans cette article de 2015, il a été décrit le comportement au repos des cellules au début de l'hématopoïèse à l'aide de données biologiques in vivo sur chaque population et d'un modèle mathématique se limitant aux phases précoces de l'hématopoïèse de repos. Nous avons choisi de concentrer nos recherches sur la lignée érythrocytaire afin d'expliquer la dynamique des cellules au repos mais également en situation de stress et ce en s'appuyant sur des données biologiques in vivo par type cellulaire.

Ainsi, tout au long de ce chapitre nous développons un modèle déterministe de 6 types cellulaires traduisant la dynamique des cellules érythropoïétiques à la fois au repos et suite à un stress biologique par hémolyse. Notre travail se décompose en cinq sections. La première correspond à la description mathématique du modèle en érythropoïèse de repos. Nous mettrons en lumière le mécanisme d'amplification observé biologiquement à l'aide d'une étude mathématique du comportement en temps long du modèle. En nous appuyant sur cette première étude et sur une exploitation des données biologiques pour chaque population, nous calibrerons entièrement notre modèle au repos. Cette calibration passera en premier lieu par le choix du nombre de compartiments. Ce choix se fera de sorte à minimiser le nombre de compartiments considérés tout en détaillant la dynamique des cellules ayant une forte capacité à s'auto-renouveler, cette phase constituant le début de l'érythropoïèse [69]. La section 1.2 constitue donc une base solide destinée à être étoffée dans la section suivante avec l'ajout de la régulation érythropoïétique.

La régulation de la dynamique des cellules érythropoïétiques a été étudiée de façon approfondie en ce qui concerne les dernières étapes de l'érythropoïèse et est très documentée ([61, 71, 70, 98, 101, 19, 18, 2]). En effet ces cellules avec une faible capacité à s'auto-renouveler sont soumises à des régulations telles que l'érythropoïétine. Nous avons choisi de modéliser cette partie terminale de l'érythropoïèse par un unique paramètre traduisant l'amplification "finale". En ce qui concerne la régulation des cellules qui nous intéressent, dans les premiers stades de l'érythropoïèse, les informations sont bien moins affinées. Prenons l'exemple de la dynamique des cellules souches hématopoïétiques. Elle a pendant longtemps été soumise à un dogme relatif à leur prétendue indépendance à l'effet des régulations. Nous allons voir que cette idée communément acceptée, peu à peu remise en question, ne permet absolument pas d'expliquer les observations biologiques de notre équipe portant sur la dynamique des cellules suite à un stress. Nous allons voir en Section 1.3 quel type de régulation permet d'expliquer les oscillations observées lors du retour à l'équilibre du nombre de cellules de chaque type après une perturbation biologique par hémolyse. A cette fin, nous construirons un modèle à 8 équations différentielles ordinaires décrivant la dynamique en situation de stress des 6 types cellulaires précédemment introduits. Nous montrerons en Section 1.4 que le modèle ainsi construit est limite (en un certain sens) d'un modèle probabiliste et admet un unique point stationnaire strictement positif. Nous montrerons que ce point stationnaire correspond la dynamique de l'érythropoïèse de repos précédemment modélisée. Nous calibrerons, en Section 1.5, les 16 paramètres du modèle régulé à l'aide des données expérimentales in vivo recueillies et montrerons qu'il permet d'expliquer les observations biologiques. Finalement, vous trouverez en Section 1.6 l'article de biologie correspondant à ce travail.

1.2 Erythropoïèse au repos : Un modèle mathématique

Notre premier objectif est de construire un modèle réaliste, avec un minimum de paramètres à déterminer, décrivant la dynamique des cellules érythropoïétiques au repos.

1.2.1 Le modèle probabiliste sous-jacent

Notre modèle décrit une communauté de k types cellulaires évoluant aléatoirement dans le temps ($k \in \mathbb{N}^*$ est un paramètre à déterminer). Le premier type cellulaire représente les cellules souches hématopoïétiques dites Long Term (LT-HSC) tandis que le k -ème

type représente les globules rouges. Nous considérons un processus de branchement décomposable multi-types pour modéliser ce mécanisme de différenciation cellulaire qu'est l'érythropoïèse au repos (cf. [94]). Plus précisément, nous supposons que chaque cellule de type i se divise à taux constant τ_i , indépendamment des autres cellules. A chaque division d'une cellule de type $i \in \{1, \dots, k-2\}$, deux cellules filles sont créées. Elles sont toutes deux de type i avec une probabilité dite de renouvellement $p_i^R \in [0, 1]$ ou toutes deux d'un type plus différencié avec une probabilité dite de différenciation $p_i^D = 1 - p_i^R$. Comme nous l'avons vu précédemment, il existe différentes lignées cellulaires en hématopoïèse et nous nous intéressons uniquement à la lignée des globules rouges. Ainsi nous noterons μ_i la probabilité que les deux cellules filles issues de la différenciation d'une cellule de type i soient toutes deux de type $i+1$. Autrement dit, avec probabilité $p_i^D \mu_i$, lorsqu'une cellule de type i , se divise elle engendre deux cellules filles de type $i+1$.

Lorsqu'une cellule de type $k-1$ se divise, elle se renouvelle avec une probabilité p_{k-1}^R et se différencie avec une probabilité $p_{k-1}^D = 1 - p_{k-1}^R$. A la différence des précédents types cellulaires lorsqu'une telle cellule se différencie elle engendre 2^n cellules de type k . Ainsi le paramètre n représente le nombre de mitoses successives faites par les cellules incapables de se renouveler constituant la fin de l'érythropoïèse. Finalement les cellules de type k , représentant les globules rouges sont dépourvues de noyaux. Elles meurent indépendamment des autres cellules à taux constant τ_k sans se diviser.

Le type d'une cellule est caractérisé par sa dynamique, c'est-à-dire son taux de division, ses probabilités de différenciation, de différenciation érythroïde et de renouvellement et son taux de mortalité (qui concerne les globules rouges).

La dynamique peut donc s'écrire comme suit

1. Une cellule de type $i \in \{1, \dots, k-2\}$ se divise à taux $\tau_i > 0$, en deux cellules de type i avec probabilité $p_i^R > 0$ ou en deux cellules de type $i+1$ avec probabilité $p_i^D \mu_i = (1 - p_i^R) \mu_i > 0$.
- $k-1$. Une cellule de type $k-1$ se "divise" à taux $\tau_{k-1} > 0$, en deux cellules de type $k-1$ avec une probabilité p_{k-1}^R ou en 2^n cellules de type k avec probabilité $p_{k-1}^D (= 1 - p_{k-1}^R)$ ($n \in \mathbb{N}^*$).
- k . Une cellule de type k meurt à taux constant $\tau_k > 0$.

Remarque 1. (i) Les probabilités $(p_i^D (1 - \mu_i))_i$ peuvent être considérées comme des probabilités de sortie du système à chaque étape de différenciation. Cette probabilité de sortie n'apparaît pas entre le type $k-1$ et le type k . En effet, cette sortie de système peut facilement être prise en compte dans le paramètre n décrivant les 2^n globules rouges engendrés par différenciation d'une cellule de type $k-1$.

(ii) Nous pourrions également supposer les divisions asymétriques (donnant deux cellules filles de types différents) avec une certaine probabilité strictement positive. Toutefois cela complique le modèle et ne change pas l'analyse mathématique présentée dans cette section ni le calibrage de chacun des paramètres du modèle de l'érythropoïèse au repos. Ce travail n'exclut pas un étoffement ultérieur du modèle lorsque des données futures le permettront.

(iii) Lorsque nous parlons de division d'une cellule à taux τ_i cela signifie qu'au bout d'un temps aléatoire T la cellule se divise en deux. La loi de ce temps est donnée par une loi exponentielle de paramètre τ_i . En d'autres mots, la valeur moyenne que prendra la variable aléatoire T sera donnée par $1/\tau_i$ (valeur qui dépend du type de la cellule).

(iv) Le modèle décrit ici est stochastique. Toutefois nous allons voir qu'il peut être approché en grande population par un modèle déterministe. Bien que l'aléa de la dynamique des cellules nous intéresse particulièrement dans notre réflexion sur le long terme, le travail effectué et présenté dans ce chapitre utilise l'approximation déterministe du modèle.

Nous pouvons prouver que le processus stochastique $(N_t, t \geq 0) = ((N_1(t), \dots, N_k(t)), t \geq 0)$ appartenant à $\mathbb{D}(\mathbb{R}_+, (\mathbb{R}_+)^k)$ et décrivant la dynamique du nombre de cellules de chaque type au temps t , est bien défini et peut être obtenu comme solution d'un système différentiel stochastique dirigé par des mesures de Poisson ponctuelles indépendantes bien choisies. Il s'agit donc d'une semi-martingale k -dimensionnelle dont le drift et la variation quadratique sont faciles à obtenir.

$\mathbb{D}(\mathbb{R}_+, (\mathbb{R}_+)^k)$ représente l'espace des fonctions càdlàg (continues à droite et admettant une limite à gauche) à valeur dans $(\mathbb{R}_+)^k$ et définies sur $\mathbb{R}_+$.

Introduisons un paramètre d'échelle $K \geq 1$ qui représente l'ordre de grandeur de la taille de la population de LT-HSC. Plus précisément, nous supposons que $N_0 = (x_1^* K, 0, \dots, 0)$. Alors tout le processus N^K dépend de K . Nous considérons le processus $(X_t^K = K^{-1} N_t, t \geq 0)$. En utilisant la décomposition en semi-martingale des processus $(X^K, K \geq 1)$ et un argument de compacité-unicité, on peut prouver la convergence de X^K vers l'unique solution d'un système déterministe (cf. [41]).

Théorème 9. Soit $T > 0$, supposons que $X_0^K = (x_1^*, 0, \dots, 0) \in \mathbb{R}_+^k$ alors la suite de processus $(X_t^K, t \in [0, T])_K$ converge en loi (et en probabilité) dans $\mathbb{D}([0, T], \mathbb{R}_+^k)$, lorsque K tend vers l'infini, vers la fonction continue $(x(t), t \in [0, T])$ à valeur dans $\mathbb{R}_+^k$ telle que $x(0) = (x_1^*, 0, \dots, 0)$ et qui est solution du système différentiel suivant :

$$\left\{ \begin{array}{l} \frac{dx_1}{dt}(t) = (p_1^R - p_1^D) \tau_1 x_1(t) \\ \forall i \in \llbracket 2, k-1 \rrbracket, \frac{dx_i}{dt}(t) = 2\mu_{i-1} p_{i-1}^D \tau_{i-1} x_{i-1}(t) - (p_i^D - p_i^R) \tau_i x_i(t) \\ \frac{dx_k}{dt}(t) = 2^n p_{k-1}^D \tau_{k-1} x_{k-1}(t) - \tau_k x_k(t) \end{array} \right. \quad (1.1)$$

Afin de s'assurer de l'existence d'un équilibre stable pour le nombre de cellules LT-HSC qui initient le mécanisme de différenciation cellulaire modélisé, nous supposons que les probabilités de différenciation et de renouvellement de ce premier type (LT-HSC) sont égales

$$p_1^D = p_1^R = 1/2. \quad (1.2)$$

Ainsi une approximation du processus N^K pour K suffisamment grand est donnée par

$$\forall t, \quad N_1^K \sim Kx_1^*, N_2^K \sim Kx_2, \dots, N_k^K \sim Kx_k.$$

Dans la sous-section suivante (1.2.2), nous allons étudier le comportement en temps long de l'unique solution x du système (1.1).

1.2.2 Etude du modèle : un mécanisme d'amplification

Supposons dans un premier temps que pour $i \in \llbracket 2, k-1 \rrbracket$, la probabilité de différenciation est différente de celle de renouvellement

$$\forall i \in \llbracket 2, k-1 \rrbracket, \quad p_i^D \neq p_i^R.$$

On obtient alors aisément l'existence d'un unique (à condition initiale de la 1ère coordonnée près) point stationnaire $x^* = (x_1^*, \dots, x_k^*)$ décrit par

$$\left\{ \begin{array}{l} x_2^* = \frac{\mu_1 \tau_1}{(p_2^D - p_2^R) \tau_2} x_1^* \\ \forall i \in \llbracket 3, k-1 \rrbracket, x_i^* = \frac{2 \mu_{i-1} p_{i-1}^D \tau_{i-1}}{(p_i^D - p_i^R) \tau_i} x_{i-1}^* \\ x_k^* = \frac{2^n p_{k-1}^D \tau_{k-1}}{\tau_k} x_{k-1}^* \end{array} \right. \quad (1.3)$$

Remarquons que l'existence d'un type cellulaire $i \in \llbracket 2, k-1 \rrbracket$ tel que $p_i^R = p_i^D$ implique que le modèle linéaire n'admet aucun point stationnaire tel que $x_1^* > 0$ puisque chacun des paramètres dynamiques du modèle est strictement positif.

Désormais nous désignerons par D_i la différence entre les probabilités de se différencier et de se renouveler des cellules de type i , appelé ensuite facteurs de différenciation,

$$D_i := p_i^D - p_i^R.$$

On en déduit par l'hypothèse (1.2) que $D_1 = 0$ et par l'hypothèse $p_i^D + p_i^R = 1$, que

$$p_i^D = \frac{1 + D_i}{2} \quad \text{et} \quad p_i^R = \frac{1 - D_i}{2}.$$

Le système (1.1) étant linéaire, il se réécrit : $\forall t, x'(t) = M x(t)$ avec

$$M = \begin{pmatrix} 0 & 0 & \dots & \dots & \dots & 0 \\ \mu_1 \tau_1 & -D_2 \tau_2 & \ddots & & & \vdots \\ 0 & \mu_2 (1 + D_2) \tau_2 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & -D_{k-2} \tau_{k-2} & \ddots & \vdots \\ \vdots & & \ddots & \mu_{k-2} (1 + D_{k-2}) \tau_{k-2} & -D_{k-1} \tau_{k-1} & 0 \\ 0 & \dots & \dots & 0 & 2^n p_{k-1}^D \tau_{k-1} & -\tau_k \end{pmatrix}.$$

Le comportement en temps long d'un tel modèle linéaire se déduit aisément.

- Théorème 10.** 1. Si pour tout $i \in \llbracket 2, k-1 \rrbracket$, $D_i > 0$,
alors l'unique solution x du problème de Cauchy $\begin{cases} \forall t \in \mathbb{R}_+, x'(t) = Mx(t) \\ x(0) = (x_1^*, 0, \dots, 0) \end{cases} (*)$
est une solution globale qui, quand t tend vers l'infini, converge vers x^* , définie précédemment par (1.3).
2. S'il existe $i \in \llbracket 2, k-1 \rrbracket$ tel que $D_i < 0$, alors la solution x peut avoir deux comportements différents en temps long, suivant les valeurs de x_0 : soit elle converge vers l'unique point stationnaire (1.3) qui a au moins une coordonnée négative, soit toutes ses coordonnées à partir de la i -ème, autrement dit x_j pour $j \in \{i, \dots, k\}$, tendent vers $\pm\infty$.

Remarque 2. Bien que ce résultat soit standard et énoncé dans un cas particulier, j'ai souhaité l'énoncer ainsi car il est loin d'être évident pour la communauté médicale ou biologique. En effet, ce théorème signifie que quelles que soient les conditions potentiellement critiques du système, il n'a pas besoin de boost de régulation pour revenir à son équilibre. La régulation permet simplement un retour plus rapide vers son équilibre comme le montre la figure 1.1 représentant le retour à l'équilibre du nombre de globules rouges suite à une baisse de 45% pour le modèle théorique présenté contre des données expérimentales.

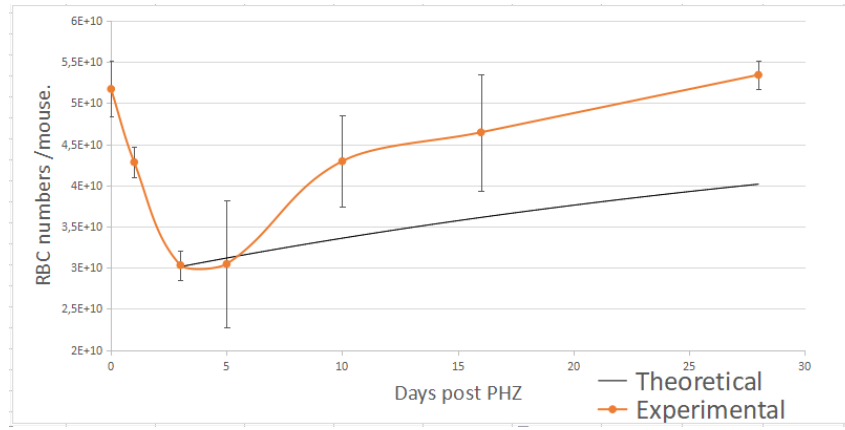


FIGURE 1.1 – Observation biologique (Experimental) du retour à l'équilibre du nombre de globules rouges suite à une baisse de 45% en 3 jours. Comparaison avec la dynamique du modèle sans régulation (Theoretical)

Nous supposons donc dans toute la suite que pour $i \in \llbracket 2, k-1 \rrbracket$, la probabilité de se différencier des cellules de type i est plus élevée que celle de s'auto-renouveler

$$\forall i \in \llbracket 2, k-1 \rrbracket, \quad D_i > 0.$$

La taille de chaque compartiment peut s'écrire de la façon suivante en fonction des quantités D_i ,

$$\begin{aligned} x_2^* &= \frac{\mu_1 \tau_1 x_1^*}{D_2 \tau_2} ; \dots ; x_{k-1}^* = \frac{\mu_{k-1}(1 + D_{k-1}) \dots \mu_2(1 + D_2) \mu_1 \tau_1 x_1^*}{D_{k-1} \dots D_2 \tau_{k-1}} ; \\ x_k^* &= \frac{2^n \mu_{k-1}(1 + D_{k-1}) \dots \mu_2(1 + D_2) \mu_1 \tau_1 x_1^*}{D_{k-1} \dots D_2 \tau_k}. \end{aligned} \quad (1.4)$$

En remarquant que $D_i \in [0, 1]$, on déduit de cette dernière formule que l'amplification entre les populations des cellules de type 1, x_1^* et de type k , x_k^* est principalement due à deux termes :

$$2^n \quad \text{et} \quad \prod_{i=2}^{k-1} \left(1 + \frac{1}{D_i}\right)$$

La différence entre ces deux termes est donnée par la capacité des types cellulaires de 1 à $k-1$ à s'auto-renouveler. En effet, en l'absence de cette capacité, on aurait pour tout i , $p_i^R = 0$, autrement dit $D_i = 1$. Dans ce cas les deux termes seraient semblables.

Ce qui nous intéresse dans ce travail de modélisation est de comprendre l'évolution dynamique de ces cellules capables d'auto-renouvellement autrement dit pour lesquelles D_i est faible.

Terminons cette sous-section en soulignant le lien qui existe entre les paramètres k , n et D_i et l'amplification du nombre de cellules observées, à l'aide de la simplification suivante. Supposons que pour tout $i \in \{2, \dots, k-1\}$, $D_i \geq D \in]0, 1[$. Alors, sachant que $\forall i, \mu_i \leq 1$, on obtient à l'aide de la formule (1.4), les inégalités suivantes

$$\frac{x_k^* \tau_k}{x_1^* \tau_1} \leq \frac{x_k^* \tau_k}{x_1^* \tau_1 \mu_1 \dots \mu_{k-2}} \leq \frac{2^n (1 + D)^{k-2}}{D^{k-2}}. \quad (1.5)$$

Nous allons maintenant à l'aide de cette formule et de différentes données biologiques calibrer l'ensemble des paramètres de notre modèle.

1.2.3 Calibration du modèle à l'aide des données

La première étape consiste à déterminer le nombre de compartiments k , autrement dit le nombre de types cellulaires, ayant la capacité de s'auto-renouveler, que nous souhaitons considérer.

Le nombre de compartiments k

Nous allons ici fixer le nombre k de compartiments selon deux critères :

1. Simplicité : k doit être le plus petit possible pour avoir le moins de paramètres à calibrer dans notre modèle que l'on souhaite être simple.
2. Réalité : nous voulons que notre modèle prenne en compte les informations biologiques à notre disposition.

Bien qu'encore incomplètes nous disposons de certaines informations sur les cellules souches. En effet, les expériences de nos collaborateurs ont montré que

$$x_1^* = 11\,000 \quad (\text{cellules/souris}), \quad \tau_1 = 1/100 \quad (/ \text{jour}),$$

$$x_k^* = 1.48 \cdot 10^{10} \quad (\text{cellules/souris}), \quad \tau_k = 1/40 \quad (/ \text{jour}).$$

Ces valeurs sont en accord avec la littérature ([13], [19]).

Nous savons, de plus que les cellules perdent leur capacité à s'auto-renouveler au fur et à mesure de leur différenciation [13]. Autrement dit nous savons que la suite $(D_i)_{i \in \{2, \dots, k-1\}}$ est croissante. En suivant la convention prise par Katrin Busch et son équipe [13], nous supposons que les cellules de type 2 de notre modèle représentent les cellules souches dites Short-Term (ST-HSC). Alors en utilisant la formule (1.3) ainsi que les informations

recueillies sur les ST-HSC détaillées dans la suite de cette section, nous déduisons une estimation de la valeur minimale de $D_2 \geq 0.16$.

Dans un article de 2019 ([68]), les auteurs étudient un type cellulaire nommé BFU-E, dont la capacité d'auto-renouvellement est considérée négligeable dans notre modèle. Ils ont montré par des expériences in vitro que ces cellules nécessitent jusqu'à 11 divisions cellulaires pour générer des érythrocytes. Nous supposons donc que le paramètre n représentant le nombre de mitoses entre les deux derniers compartiments est supérieur ou égal à 12.

En reprenant la formule (1.5), nous pouvons alors déduire des informations présentées une estimation du nombre minimal de types cellulaires, ayant une capacité à s'auto-renouveler non négligeable, que nous souhaitons considérer dans notre modèle,

$$k \geq \frac{\log(\frac{x_k^* \tau_k}{2^n \tau_1 x_1^*})}{\log(1 + \frac{1}{D})} + 2 \sim 6.$$

Bien-sûr, augmenter fortement la valeur de n diminuerait celle de k puisque l'amplification se fera alors principalement par différenciation et non par auto-renouvellement des cellules.

Dans la suite nous fixerons donc le nombre de compartiments de notre modèle à $k = 6$.

Nos collaborateurs distinguent physiologiquement 6 compartiments, ainsi nous appellerons ces compartiments LT-HSC (Long Term Hematopoietic Stem Cell), ST-HSC (Short Term Hematopoietic Stem Cell), MPP (Multipotent Progenitor), CMP (Common Myeloid Progenitor), MEP (Megakaryocyte-Erythroid Progenitor) et Globules rouges.

Nous allons maintenant calibrer les 16 paramètres de notre modèle pour finaliser la modélisation de l'érythropoïèse de repos. En prévision du travail qui sera effectué en section suivante, nous noterons ces paramètres à l'aide d'une étoile en exposant pour faire référence à l'état de repos de l'érythropoïèse qu'ils servent à modéliser. Ainsi il nous faut calibrer les paramètres suivants : $(\tau_i^*)_{i \in \{1, \dots, 6\}}$, $(p_i^{D,*})_{i \in \{1, \dots, 5\}} = (1 - p_i^{R,*})_{i \in \{1, \dots, 5\}}$, $(\mu_i^*)_{i \in \{1, \dots, 4\}}$ et n^* .

Les taux de division de chaque type cellulaire

Nos collaborateurs ont, lors d'expériences réalisées in vivo en condition normale chez la souris, recueilli des informations sur la proportion de cellules de chaque type qui se sont divisées dans un laps de temps donné. Nous allons voir comment ces informations nous donnent accès aux taux de division de chaque type cellulaire : $(\tau_i^*)_{i \in \{1, \dots, 5\}}$.

La BromodésoxyUridine (BrdU) est utilisée pour déterminer la phase du cycle cellulaire dans laquelle se trouve une cellule et notamment si cette dernière est en phase de répliation de l'ADN.

Commençons par décrire l'expérience réalisée, étape par étape, afin d'en déduire le lien qui existe entre les taux de division des cellules dans notre modèle et la proportion de cellules observées marquées par le BrdU de chaque type cellulaire.

Les expériences suivantes sont faites sur différentes souris.

- Avant l'injection de BrdU, le système est à l'équilibre et toutes les cellules sont non marquées par le BrdU.

- Le BrdU agit ensuite durant 16 à 24 heures. Nous noterons t_a cette période de temps.
- Durant la période d'action du BrdU, toutes les cellules qui se sont divisées donnent naissance à des cellules marquées par le BrdU.
En d'autres termes lorsqu'une cellule de type i non marquée se divise (ce qui arrive à taux τ_i^*) elle donne naissance à deux cellules de type i marquées avec probabilité $p_i^{R,*}$ ou à deux cellules de type $i + 1$ marquées avec probabilité $p_i^{D,*} = 1 - p_i^{R,*}$.

La dynamique des cellules non marquées est alors très simple.

En effet, notons $x_i^{NM}(t)$ et $x_i^M(t)$ respectivement l'approximation du nombre de cellules non marquées et marquées par le BrdU de type i au temps t après injection. On déduit de l'expérience que pour tout t et pour tout $i = 1, \dots, 5$,

$$\frac{d x_i^{NM}}{dt}(t) = -\tau_i^* x_i^{NM}(t).$$

Ainsi on obtient la dynamique des cellules non marquées au cours du temps.

$$\forall t, \quad x_i^{NM}(t) = x_i^{NM}(0) e^{-\tau_i^* t}.$$

La dynamique des cellules marquées se déduit alors aisément puisque l'expérience a lieu à l'équilibre, autrement dit le nombre de cellules de chaque type reste constant au cours du temps. De plus au début de l'expérience aucune cellule n'est marquée par le BrdU.

$$\forall t, \quad x_i^M(t) + x_i^{NM}(t) = x_i^* \quad \text{et} \quad x_i^{NM}(0) = x_i^*.$$

On en déduit ainsi la dynamique des cellules marquées

$$\forall t, \quad x_i^M(t) = x_i^* - x_i^{NM}(t) = x_i^* - x_i^* e^{-\tau_i^* t}.$$

Notons alors p_i^{BrdU} la proportion de cellules de type i marquées par le BrdU à la fin de l'expérience. En conséquence, au temps t_a on obtient finalement la formule suivante

$$p_i^{BrdU} = \frac{x_i^M(t_a)}{x_i^*} = 1 - e^{-\tau_i^* t_a}.$$

Les taux de division s'écrivent donc clairement en fonction de la proportion de cellules marquées par le BrdU et le temps d'action de celle-ci.

$$\forall i \in \{1, \dots, 5\}, \quad \tau_i^* = -\frac{1}{t_a} \log(1 - p_i^{BrdU}). \quad (1.6)$$

A partir des données suivantes sur la proportion de cellules marquées par le BrdU,

$$\left\{ \begin{array}{l} p_1^{BrdU} \in [0, 3.85] \quad (\text{en } \%) \\ p_2^{BrdU} \in [0.42, 1.86] \quad (\text{en } \%) \\ p_3^{BrdU} \in [2.04, 11.36] \quad (\text{en } \%) \\ p_4^{BrdU} \in [6.48, 15.18] \quad (\text{en } \%) \\ p_5^{BrdU} \in [17.92, 33.92] \quad (\text{en } \%) \end{array} \right. . \quad (1.7)$$

et des données concernant le temps d'action du BrdU,

$$t_a \in [16h, 24h]$$

on déduit des bornes de valeurs pour les taux de division de chaque type cellulaire.

$$\left\{ \begin{array}{l} \tau_1^* \in [0, 0.06] \quad (/j) \\ \tau_2^* \in [0.004, 0.03] \quad (/j) \\ \tau_3^* \in [0.02, 0.18] \quad (/j) \\ \tau_4^* \in [0.067, 0.25] \quad (/j) \\ \tau_5^* \in [0.2, 0.62] \quad (/j) \end{array} \right. . \quad (1.8)$$

Remarque 3. La formule (1.6) est obtenue à l'aide d'un raisonnement macroscopique où l'on suppose que le nombre de cellules est constant au cours du temps. Un raisonnement probabiliste conduisant au même résultat est présenté dans l'article [104] où les auteurs s'intéressent à la dynamique de cellules souches neuronales soumises à l'effet d'un traitement.

Ensemble des valeurs des paramètres du modèle linéaire

Grâce aux expériences de nos collaborateurs, nous connaissons la taille de chaque compartiment en érythropoïèse de repos :

$$\left\{ \begin{array}{l} x_1^* = 0.011 \quad (\times 10^6 \text{ cellules/souris}) \\ x_2^* = 0.0225 \quad (\times 10^6 \text{ cellules/souris}) \\ x_3^* = 0.067 \quad (\times 10^6 \text{ cellules/souris}) \\ x_4^* = 0.196 \quad (\times 10^6 \text{ cellules/souris}) \\ x_5^* = 0.359 \quad (\times 10^6 \text{ cellules/souris}) \\ x_6^* = 1.48 \quad (\times 10^{10} \text{ cellules/souris}) \end{array} \right. \quad (1.9)$$

De plus les probabilités μ_i qu'après différenciation une cellule de type i choisisse la lignée érythroïde (au détriment des autres lignées cellulaires) sont publiées dans l'article [13] :

$$\text{pour } i = 1, 2, \quad \mu_i^* = 1, \quad \mu_3^* = 3.992/4.014 \text{ (MPP)} \quad \text{et} \quad \mu_4^* = 3/5 \text{ (CMP)}.$$

Nous pouvons alors déduire de ces informations et des bornes sur le taux de division de chaque type cellulaire (1.6), l'ordre de grandeur de chacun des paramètres de modèle à l'aide de la formule (1.3).

Pour des raisons algorithmiques nous avons besoin de définir des valeurs précises pour chacun des paramètres du modèle. Ces valeurs sont répertoriées dans le tableau suivant :

type cellulaire	1 : LT-HSC	2 : ST-HSC	3 : MPP	4 : CMP	5 : MEP	6 : RBC
taux de divisions (/jour)	$\tau_1^* = 0.01$	$\tau_2^* = 0.03$	$\tau_3^* = 0.07$	$\tau_4^* = 0.16$	$\tau_5^* = 0.35$	$\tau_6^* = 1/40$ (mort)
nombre de cellules (cellules/souris)	$x_1^* = 0.011$ ($\times 10^6$)	$x_2^* = 0.0225$ ($\times 10^6$)	$x_3^* = 0.067$ ($\times 10^6$)	$x_4^* = 0.196$ ($\times 10^6$)	$x_5^* = 0.359$ ($\times 10^6$)	$x_6^* = 1.48$ ($\times 10^{10}$)
facteur de différenciation	$D_1 = 0$	$D_2 = 0.16$	$D_3 = 0.167$	$D_4 = 0.174$	$D_5 = 0.176$	
probabilité du choix érythroïde	$\mu_1^* = 1$	$\mu_2^* = 1$	$\mu_3^* = 0.99$	$\mu_4^* = 3/5$		
nombre de mitoses terminales	$n^* \sim 12$					

Ces valeurs sont choisies de sorte à prendre en compte deux informations biologiques supplémentaires : la capacité des cellules à s’auto-renouveler diminue au fur et à mesure de leur différenciation, le taux de division des ST-HSC est plus élevé que celui des LT-HSC ([13],[69]).

Bien que nos modèles soient différents, les valeurs obtenues par l’équipe de K.Busch [13] ont des ordres de grandeurs en accord avec nos valeurs, à l’exception des taux de division des MPP.

1.3 Erythropoïèse de stress : Ajout d’une régulation

Nous allons voir maintenant que la réaction des LT-HSC, après une perturbation du système au niveau des érythrocytes appelée hémolyse, traduit l’effet d’un boost de la régulation sur la dynamique de ces cellules.

1.3.1 Des données par type cellulaire

Afin de comprendre les effets de la régulation sur chacun des 6 compartiments modélisés, nos collaborateurs Stéphane Giraudier et Evelyne Lauret ont choisi de réaliser l’expérience suivante sur des souris de laboratoire. Jour 0 : injection d’une dose de phénylhydrazine (PHZ). Cette dose de PHZ va causer la destruction d’environ 45% des globules rouges du système en 3 jours. Aux jours 0, 1, 3, 5, 7, 10, 16 et 28 l’effectif cellulaire de chaque compartiment est relevé afin d’observer la réaction du système au stress. Chacun de ces relevés est présenté dans la figure 1.2.

On observe une destruction effective du nombre de globules rouges d’environ 45% jusqu’au jour 3 puis une remontée de ceux-ci vers leur valeur dite de repos (Jour 28), référencée par la ligne rouge en pointillés.

Lors de cette expérience seuls les globules rouges sont directement affectés par la PHZ. En effet des expériences supplémentaires ont montré que la baisse observée des autres types cellulaires n’est pas due à un effet direct de la PHZ.

La figure 1.2 montre une dynamique semblable pour les populations de type 1 à 4. On observe une baisse du nombre de cellules entre les jours 0 et 3 puis une augmentation massive jusqu’au jour 10. Les cellules de type 5, les MEP, ne présentent pas de baisse en amont de leur expansion massive au jour 10.

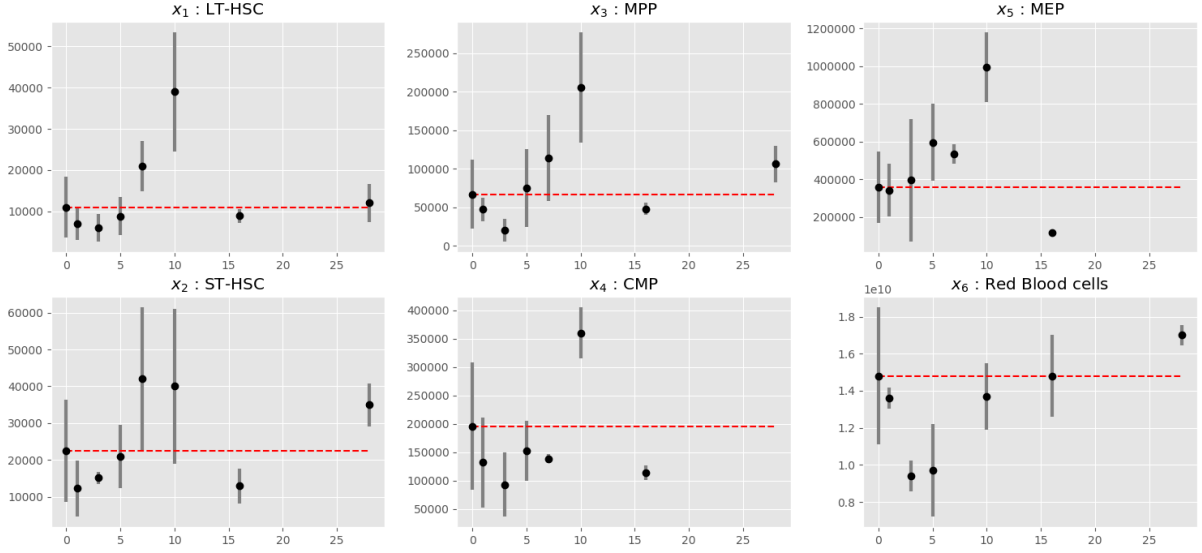


FIGURE 1.2 – Relevés des effectifs cellulaires par compartiment et par souris aux jours 0 (injection de la PHZ), 1, 3, 5, 7, 10, 16 et 28. La ligne en pointillés indique la valeur d'équilibre (1.9) du nombre de cellules de chaque type. Chaque point noir correspond à une valeur moyenne sur un effectif variant de 3 à 10 souris selon les points.

Les lignes en pointillés rouges représentent les valeurs des tailles de population de l'érythropoïèse de repos précédemment modélisée. Ces valeurs correspondent aux données au jour 0 de l'expérience (jour où la PHZ est injectée).

Ces résultats montrent une variation importante du nombre de cellules de type 1 lors de l'expérience (voir Figure 1.2). Or la dynamique des LT-HSC est simple,

$$\forall t, \quad \frac{dx_1(t)}{dt} = \left(p_1^R(t) - p_1^D(t) \right) \tau_1(t) x_1(t).$$

On en déduit que les probabilités de se renouveler ou de se différencier des cellules de type 1 ne peuvent rester constantes au cours de l'expérience. Il nous faut donc ajouter de la régulation dans notre modèle au niveau des probabilités de se renouveler et donc de se différencier des LT-HSC. Remarquons de plus qu'il faut également réguler le taux de division des LT-HSC pour pouvoir expliquer ces résultats. En effet, on passe de 20 000 LT-HSC à 50 000 en 3 jours. Or si l'on suppose que nous sommes dans un cas extrême ie, $p_1^R = 1$ et que le taux de division des LT-HSC n'est pas régulé, donc égal à sa valeur au repos $\tau_1^* = 1/100$, on obtient

$$N_1(J10) = N_1(J7) e^{(10-7)\tau_1^*} = 20\,000 e^{3\tau_1^*} \sim 20600 \quad (\text{cellules/souris}).$$

Pour expliquer la croissance observée entre J7 et J10, nous devons donc ajouter une régulation sur les probabilités ainsi que sur les taux de division des cellules.

Nous allons voir quels types de régulations ajoutées au modèle (1.1) permettent d'observer une telle dynamique oscillante, à la fois au niveau du dépassement de l'équilibre au jour 10, alors qu'au jour 7 toutes les populations ont déjà dépassé leurs valeurs d'équilibres ou en sont proches, et également au niveau de la baisse observée au jour 16.

1.3.2 Le modèle

A ce stade, il est évident qu'il faut ajouter une régulation à notre modèle.

Présentons dans un premier temps notre modèle puis nous reviendrons sur les choix de modélisation effectuée.

Notre modèle est décrit par les fonctions continues $(x_1, \dots, x_6, u_R, u_D)$ solutions du système d'équations suivant pour lequel $u_R(0) = \frac{1}{m_R(a+y(0))}$ et $u_D(0) = \frac{y(0)}{m_D(b+y(0))}$,

$$\left\{ \begin{array}{l} \frac{dx_1}{dt}(t) = (\tau_1^R(t) - \tau_1^D(t) - s_1) x_1(t) \\ \frac{dx_i}{dt}(t) = 2 \tau_{i-1}^D(t) x_{i-1}(t) + (\tau_i^R(t) - \tau_i^D(t) - s_i) x_i(t), \quad \forall i \in \llbracket 2, 5 \rrbracket \\ \frac{dx_6}{dt}(t) = 2^{n(t)} \tau_5^D(t) x_5(t) - \tau_6 x_6(t) \\ \frac{du_R}{dt}(t) = \frac{1}{a + y(t)} - m_R u_R(t) \\ \frac{du_D}{dt}(t) = \frac{y(t)}{b + y(t)} - m_D u_D(t) \end{array} \right. \quad (1.10)$$

où le terme $y(t)$ représente l'environnement du système au temps t dans lequel chaque type cellulaire vivant dans la moelle osseuse influence à part égale la production de u_R et u_D ,

$$y(t) = \sum_{j=1}^5 \frac{x_j(t)}{x_j^*}.$$

Les informations biologiques dont nous disposons sur la régulation des 5 premiers types cellulaires sont faibles. En revanche, la régulation des derniers types cellulaires, modélisée par la fonction n , est plus connue. Cette régulation a fait l'objet de nombreuses études à la fois biologiques et mathématiques, en particulier concernant l'action d'une cytokine nommée érythropoïétine (EPO), ([71, 70, 98, 101, 61, 19, 2]).

Nous avons, par conséquent, choisi de décomposer la régulation du modèle en trois parties :

- La régulation la plus connue affecte le nombre de cellules de type 6 (globules rouges) produits à partir d'une cellule de type 5 (MEP), autrement dit affecte la dynamique des n dernières mitoses. Nous modélisons cette régulation par une fonction avec saturation décroissante avec le nombre de globules rouges,

$$n(t) = n^* \frac{1 + c_1}{1 + c_1 \left(\frac{x_6(t)}{x_6^*} \right)^{c_2}}.$$

Les paramètres c_1 et c_2 sont strictement positifs. Leur effet sur les n dernières mitoses en fonction du nombre de globules rouges dans le système est décrit dans la figure 1.3.

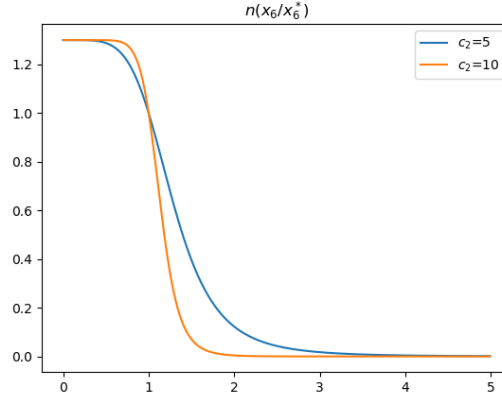


FIGURE 1.3 – Allure du paramètre n en fonction de la valeur de la fraction $\frac{x_6(t)}{x_6^*}$ pour $c_1 = 0.3$. La valeur $1 + c_1$ représente le maximum de la fonction n .

- La deuxième régulation est décrite par le régulateur u_R . Celui-ci est produit dans le système lorsqu'il y a un manque de cellules dans la moelle osseuse et décroît selon une constante de dégradation notée m_R . Nous supposons qu'il affecte le taux d'auto-renouvellement des cellules

$$\forall i \in \{1, \dots, 5\}, \quad \tau_i^R(t) = \tau_i^* p_i^{R,*} \left(\frac{u_R(t)}{u_R^*} \right)^{r_i}.$$

Les paramètres $(r_i)_i$ sont positifs. Ils représentent la sensibilité des cellules de type i à la présence du régulateur u_R . La constante u_R^* représente la valeur de u_R à l'équilibre,

$$u_R^* = \frac{1}{m_R(a+5)}.$$

- La troisième régulation est décrite par le régulateur u_D . Ce dernier est produit lorsque la moelle osseuse est en surnombre et décroît selon une constante de dégradation notée m_D . Nous supposons qu'il affecte le taux de différenciation érythroïde des cellules

$$\forall i \in \{1, \dots, 5\}, \quad \tau_i^D(t) = \tau_i^* p_i^{D,*} \mu_i^* \left(\frac{u_D(t)}{u_D^*} \right)^{d_i}.$$

Les paramètres $(d_i)_i$ sont positifs. Ils représentent la sensibilité des cellules de type i à la présence du régulateur u_D . La constante u_D^* représente la valeur de u_D à l'équilibre,

$$u_D^* = \frac{5}{m_D(b+5)}.$$

Nous avons remarqué en Section 1.3.1 que la régulation des cellules de type 1 devait se faire à la fois sur leur probabilité d'auto-renouvellement (et donc de différenciation) et sur leur taux de division. Nous avons choisi de faire agir les régulateurs u_R et u_D non pas sur le taux de division des cellules et sur la probabilité d'auto-renouvellement directement mais sur ce qu'on a appelé le taux d'auto-renouvellement et le taux de différenciation érythroïde plus facilement reliés à la dynamique du système. En effet, ces deux taux agissent comme des taux de naissance (auto-renouvellement) et de mort (différenciation) sur le système.

Les fonctions communément choisies pour modéliser la régulation des taux de division sont les fonctions de Michaelis-Menten ou de Hill. Ces fonctions modélisent une forme de saturation associée à l'effet de la régulation. Nous avons choisi de modéliser cette saturation au niveau de la production des régulateurs u_R et u_D et non au niveau des taux directement afin de minimiser le nombre de paramètres du modèle. Précisons également que la forme des fonctions choisies pour la régulation des taux induit à la fois l'augmentation des taux d'auto-renouvellement lorsque $u_R(t)$ est au-dessus de sa valeur d'équilibre u_R^* et une inhibition de celui-ci lorsque sa valeur est en dessous (la même dynamique est observée pour u_D).

Le terme s_i désigne, quant-à-lui, le taux de sortie du système par différenciation non érythroïde pour tout $i \in \{1, \dots, 4\}$

$$\forall t, \quad s_i = \tau_i^* p_i^{D,*} (1 - \mu_i^*).$$

N'ayant aucune information sur la dynamique des autres lignées cellulaires, nous avons supposé ce taux constant. Pour simplifier les notations nous supposons que $s_5 = 0$.

Remarque 4. Ici le nombre de globules rouges n'influence pas la production des régulateurs u_R et u_D . L'ajout des globules rouges ne change ni le résultat final ni la complexité du modèle. Par contre mathématiquement, l'absence de cette dépendance en les globules rouges permet de s'assurer l'unicité d'un point stationnaire tel que le nombre de cellules de chaque type cellulaire soit strictement positif. De plus cette hypothèse possède un certain sens biologique, puisque les globules rouges vivent dans le sang tandis que les autres types cellulaires considérés dans notre modèle sont présents dans la moelle osseuse.

Récapitulons la liste des paramètres qu'il nous faut calibrer à l'aide des données.

Les 16 paramètres de notre modèle à calibrer sont :

$(c_i)_{i \in \{1,2\}}$	constantes modélisant la régulation induite par le nombre de globules rouges sur n , indice d'amplification entre les MEP et les globules rouges
$(r_i)_{i \in \{1, \dots, 5\}}$	constantes modélisant l'impact de la régulation u_R sur le taux de prolifération des cellules de type $i = 1, \dots, 5$
$(d_i)_{i \in \{1, \dots, 5\}}$	constantes modélisant l'impact de la régulation u_D sur le taux de différenciation érythroïde des cellules de type $i = 1, \dots, 5$
a	constante intervenant dans la fonction de Michaelis Menten modélisant la production de u_R dans le système
b	constante intervenant dans la fonction de Michaelis Menten modélisant la production de u_D dans le système
$(m_i)_{i \in \{R,D\}}$	constantes de dégradation de u_i .

Les paramètres déjà calibrés en Section 1.2 sont :

$(\tau_i^*)_{i \in \{1, \dots, 6\}}$	taux de divisions des cellules de type i au repos (taux de mortalité pour les globules rouges)
$(p_i^{R,*})_{i \in \{1, \dots, 5\}}$	probabilité de s'auto-renouveler des cellules de type i au repos
$(p_i^{D,*} = 1 - p_i^{R,*})_{i \in \{1, \dots, 6\}}$	probabilité de se différencier des cellules de type i au repos
$(\mu_i^*)_{i \in \{1, \dots, 5\}}$	probabilité que lorsqu'une cellule de type i se différencie, ses deux cellules filles soient de la lignée des globules rouges.
n^*	indice d'amplification entre les MEP et les globules rouges.

Nous allons maintenant étudier le modèle (1.10) afin de montrer que notre modèle de l'érythropoïèse de stress pris à l'équilibre correspond bien au modèle de l'érythropoïèse de repos.

1.4 Etude du modèle

Dans cette section, nous étudions le modèle (1.10). Nous allons démontrer qu'il admet un unique point stationnaire dont toutes les composantes sont strictement positives et que celui-ci correspond au point stationnaire du modèle sur l'érythropoïèse de repos étudié en section 1.2. Nous allons également observer numériquement la convergence de l'unique solution décrite par le modèle vers ce point pour différentes conditions initiales choisies de façon aléatoire. Finalement nous montrerons que notre modèle régulé peut être décrit comme limite (en un certain sens) d'un modèle probabiliste.

1.4.1 Existence et unicité des points stationnaires

Notre modèle est décrit par le système dynamique à 8 équations suivant

$$\left\{ \begin{array}{l} \frac{du_R(t)}{dt} = \frac{1}{a+y(t)} - m_R u_R(t) \\ \frac{du_D(t)}{dt} = \frac{y(t)}{b+y(t)} - m_D u_D(t) \\ \frac{dx_1}{dt}(t) = (p_1^{R,*} \left(\frac{u_R(t)}{u_R^*}\right)^{r_1} - p_1^{D,*} \left(\frac{u_D(t)}{u_D^*}\right)^{d_1}) \frac{\tau_1^*}{2} x_1(t) \\ \forall i \in \llbracket 2, 5 \rrbracket, \frac{dx_i}{dt}(t) = 2 \mu_{i-1}^* p_{i-1}^{D,*} \tau_{i-1}^* \left(\frac{u_D(t)}{u_D^*}\right)^{d_{i-1}} x_{i-1}(t) \\ \quad - (p_i^{D,*} \left(\frac{u_D(t)}{u_D^*}\right)^{d_i} - p_i^{R,*} \left(\frac{u_R(t)}{u_R^*}\right)^{r_i}) \tau_i^* x_i(t) \\ \frac{dx_6}{dt}(t) = 2^{n(t)} p_5^{D,*} \tau_5^* \left(\frac{u_D(t)}{u_D^*}\right)^{d_5} x_5(t) - \tau_6 x_6(t) \end{array} \right. , \quad (1.11)$$

avec

$$y(t) = \sum_{j=1}^5 \frac{x_j(t)}{x_j^*} \quad \text{et} \quad n(t) = n^* \frac{1 + c_1}{1 + c_1 \left(\frac{x_6(t)}{x_6^*}\right)^{c_2}}.$$

Théorème 11. *Le système (1.11), pour lequel l'ensemble des paramètres sont strictement positifs, admet un unique point stationnaire, ayant toutes ses composantes strictement positives, donné par (u_R^*, u_D^*, x^*) .*

Démonstration. Supposons l'existence d'un point stationnaire $(\bar{u}_R, \bar{u}_D, \bar{x}_1, \dots, \bar{x}_6)$ pour (1.11) dont toutes les composantes sont strictement positives.

Alors on sait que pour $i = 1, 2$, $\frac{d\bar{u}_i}{dt} = 0$, autrement dit que

$$\bar{u}_R = \frac{1}{m_R(a + \bar{y})}, \quad \text{et} \quad \bar{u}_D = \frac{\bar{y}}{m_D(b + \bar{y})} \quad (1.12)$$

$$\text{où } \bar{y} = \sum_{j=1}^5 \frac{\bar{x}_j}{x_j^*}.$$

On sait également que $\frac{d\bar{x}_1}{dt} = 0$ et que

$$u_R^* = \frac{1}{m_R(a + 5)} \quad \text{et} \quad u_D^* = \frac{5}{m_D(b + 5)}.$$

On en déduit la condition suivante sur $\bar{y}$,

$$\left(\frac{a+5}{a+\bar{y}}\right)^{r_1} = \left(\frac{\bar{y}(b+5)}{5(b+\bar{y})}\right)^{d_1}, \quad (1.13)$$

avec $r_1 > 0$, $d_1 > 0$. Par décroissance stricte de la fonction $\bar{y} \in \mathbb{R}^+ \mapsto \left(\frac{a+5}{a+\bar{y}}\right)^{r_1} \in [0, (\frac{a+5}{a})^{r_1}]$ et croissance stricte de la fonction $\bar{y} \in \mathbb{R}^+ \mapsto \left(\frac{\bar{y}(b+5)}{5(b+\bar{y})}\right)^{d_1} \in [0, (\frac{b+5}{b})^{d_1}]$, on obtient l'unicité d'un tel $\bar{y}$.

Au vu des expressions (1.12), l'unicité de $\bar{y}$ implique l'unicité des $\bar{u}_i$.

Nous savons de plus que pour tout $i \in \{2, \dots, 5\}$,

$$2\mu_{i-1}^* p_{i-1}^{D,*} \tau_{i-1}^* \left(\frac{\bar{u}_D}{u_D^*}\right)^{d_{i-1}} \bar{x}_{i-1} = (p_i^{D,*} \left(\frac{\bar{u}_D}{u_D^*}\right)^{d_i} - p_i^{R,*} \left(\frac{\bar{u}_R}{u_R^*}\right)^{r_i}) \tau_i^* \bar{x}_i \quad (1.14)$$

Montrons par l'absurde que pour tout $i \in \{2, \dots, 5\}$,

$$p_i^{D,*} \left(\frac{\bar{u}_D}{u_D^*}\right)^{d_i} \neq p_i^{R,*} \left(\frac{\bar{u}_R}{u_R^*}\right)^{r_i}. \quad (1.15)$$

Pour cela, supposons qu'il existe $i_0 \in \{2, \dots, 5\}$ tel que

$$p_{i_0}^{D,*} \left(\frac{\bar{u}_D}{u_D^*}\right)^{d_{i_0}} = p_{i_0}^{R,*} \left(\frac{\bar{u}_R}{u_R^*}\right)^{r_{i_0}}.$$

Alors par (1.14) et par la stricte positivité de l'ensemble des paramètres, on déduit que $\bar{x}_{i_0-1} = 0$. Cela implique une contradiction avec l'hypothèse de stricte positivité des composantes du point stationnaire. On en déduit donc que pour tout $i \in \{2, \dots, 5\}$,

$$p_i^{D,*} \left(\frac{\bar{u}_D}{u_D^*}\right)^{d_i} \neq p_i^{R,*} \left(\frac{\bar{u}_R}{u_R^*}\right)^{r_i}.$$

Alors, on obtient pour tout $i \in \{2, \dots, 5\}$,

$$\bar{x}_i = \frac{2\mu_{i-1}^* p_{i-1}^{D,*} \tau_{i-1}^* \left(\frac{\bar{u}_D}{u_D^*}\right)^{d_{i-1}}}{(p_i^{D,*} \left(\frac{\bar{u}_D}{u_D^*}\right)^{d_i} - p_i^{R,*} \left(\frac{\bar{u}_R}{u_R^*}\right)^{r_i}) \tau_i^*} \bar{x}_{i-1}. \quad (1.16)$$

On en déduit que les $\bar{x}_i$ pour $i \in \{2, \dots, 5\}$ sont tous proportionnels à $\bar{x}_1$. Par définition de $\bar{y}$, on obtient

$$\bar{y} = \sum_{j=1}^5 \bar{x}_j = \bar{x}_1 \sum_{j=1}^5 C_j.$$

De plus $\sum_{j=1}^5 C_j \neq 0$. Dans le cas contraire $\bar{y} = 0$ et la condition (1.13) ne serait pas vérifiée.

On déduit alors de l'unicité de $\bar{y}$ l'unicité de $\bar{x}_1$ puis celle des $\bar{x}_i$.

Finalement, nous savons que $\frac{d\bar{x}_6}{dt} = 0$, c'est à dire que $\bar{x}_6$ vérifie

$$\bar{x}_6 = \frac{2^{\bar{n}(\bar{x}_6)} p_5^{D,*} \tau_5^* \left(\frac{\bar{u}_D}{u_D^*}\right)^{d_5}}{\tau_6^*} \bar{x}_5$$

avec

$$\bar{n}(\bar{x}_6) = \frac{c_1}{1 + (1 - c_1) \left(\frac{\bar{x}_6}{x_6^*}\right)^{c_2}}.$$

Par stricte croissance de la fonction identité, et stricte décroissance de la fonction $\bar{x}_6 \mapsto$

$\frac{2^{\bar{n}(\bar{x}_6)} p_5^{D,*} \tau_5^* \left(\frac{\bar{u}_D}{u_D^*}\right)^{d_5}}{\tau_6^*} \bar{x}_5$, on obtient l'unicité d'un tel $\bar{x}_6$.

Nous avons ainsi démontré que si un point stationnaire pour (1.11) dont toutes les composantes sont strictement positives existe, alors il est unique. La preuve de l'existence d'un tel point est simple. Il suffit de constater par calcul que le point $(u_R^*, u_D^*, x_1^*, \dots, x_6^*)$ définit un tel point.

□

La preuve de la convergence en temps long de la solution du système (1.11) vers cette unique point stationnaire strictement positif n'est pas évidente. Numériquement, la stabilité de ce point stationnaire semble être vraie (voir figure 1.4). En effet, on observe sur la figure 1.4, les trajectoires de notre modèle pour différentes conditions initiales, choisies aléatoirement selon une loi uniforme sur $[0, 3x^*]$. Théoriquement, la convergence en temps long vers ce point est difficile à démontrer car le système est en grande dimension. Il est donc difficile d'exclure l'existence d'un cycle limite ou d'autres comportements dynamiques. Ainsi dans un premier temps il serait intéressant de démontrer la stabilité de ce point dans un cas plus simple où seule la population des cellules de type 1, les LT-HSC, est considérée.

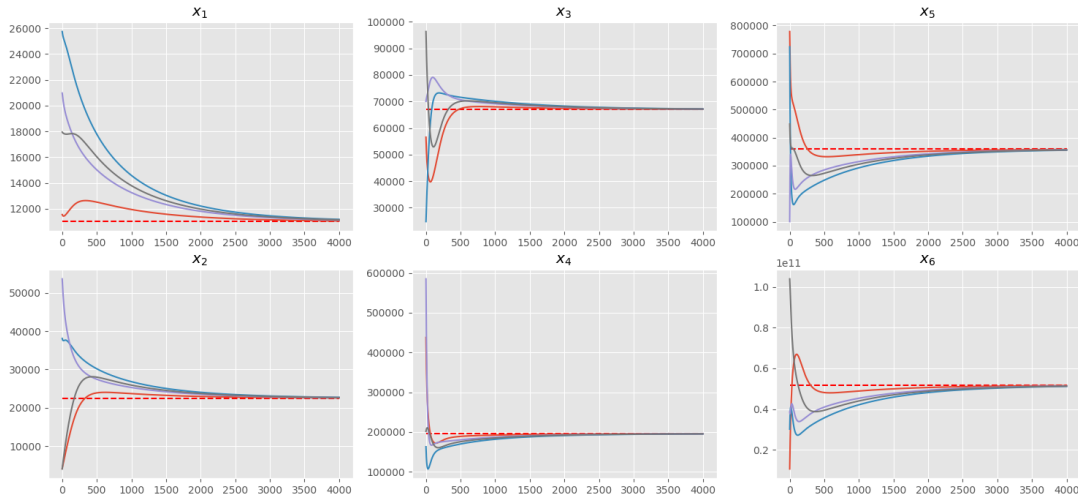


FIGURE 1.4 – Quatre trajectoires de la fonction x sur $[0, 4000]$ pour des conditions initiales choisies aléatoirement sur $[0, 3x^*]$. La ligne en pointillés représente l'unique équilibre stationnaire du système (1.10) dont les composantes sont strictement positives.

1.4.2 Le modèle probabiliste sous-jacent

Nous allons voir que le modèle défini par (1.11) s'écrit comme limite macroscopique d'un modèle probabiliste. Pour cela, notons $N^K(t)$, le vecteur représentant le nombre de cellules de chaque type cellulaire au temps t . Supposons que pour tout K , $N^K(0)$ soit un vecteur aléatoire intégrable de $\mathbb{N}$ tel que

$$N_1^K(0) \sim K.$$

Introduisons ensuite $(\mathcal{N}_i^j)_{\substack{1 \leq i \leq 6 \\ j \in \{-, +, s\}}}$, 16 mesures de Poisson ponctuelles indépendantes d'intensité $dvdu$ sur $\mathbb{R}_+^2$ ainsi que la filtration $(\mathcal{F}_t)_{t \geq 0}$ associée

$$\mathcal{F}_t = \sigma(\mathcal{N}_i^j([0, v] \times A); i \in \{1, \dots, 6\}, j \in \{-, +, s\}, v \leq t, A \in \mathcal{B}(\mathbb{R})).$$

Nous pouvons alors définir le processus de Markov N^K comme l'unique solution des équations intégrales stochastiques suivantes,

$$\begin{aligned}
N_1^K(t) &= K + \int_0^t \int_{\mathbb{R}_+} \mathbf{1}_{\{u \leq \tau_1^R(u_R^K(v^-)) N_1^K(v^-)\}} \mathcal{N}_1^+(dv, du) \\
&\quad - \int_0^t \int_{\mathbb{R}_+} \mathbf{1}_{\{u \leq \tau_1^D(u_D^K(v^-)) N_1^K(v^-)\}} \mathcal{N}_1^-(dv, du) - \int_0^t \int_{\mathbb{R}_+} \mathbf{1}_{\{u \leq s_1 N_1^K(v^-)\}} \mathcal{N}_1^s(dv, du) \\
\text{pour } i &= 2, \dots, 5, \\
N_i^K(t) &= N_i^K(0) + 2 \int_0^t \int_{\mathbb{R}_+} \mathbf{1}_{\{u \leq \tau_{i-1}^D(u_D^K(v^-)) N_{i-1}^K(v^-)\}} \mathcal{N}_{i-1}^-(dv, du) \\
&\quad + \int_0^t \int_{\mathbb{R}_+} \mathbf{1}_{\{u \leq \tau_i^R(u_R^K(v^-)) N_i^K(v^-)\}} \mathcal{N}_i^+(dv, du) \\
&\quad - \int_0^t \int_{\mathbb{R}_+} \mathbf{1}_{\{u \leq \tau_i^D(u_D^K(v^-)) N_i^K(v^-)\}} \mathcal{N}_i^-(dv, du) - \int_0^t \int_{\mathbb{R}_+} \mathbf{1}_{\{u \leq s_i N_i^K(v^-)\}} \mathcal{N}_i^s(dv, du) \\
N_6^K(t) &= N_6^K(0) + \int_0^t \int_{\mathbb{R}_+} 2^{n^K(N_6^K(v^-))} \mathbf{1}_{\{u \leq \tau_5^D(u_D^K(v^-)) N_5^K(v^-)\}} \mathcal{N}_5^-(dv, du) \\
&\quad - \int_0^t \int_{\mathbb{R}_+} \mathbf{1}_{\{u \leq \tau_6 N_6^K(v^-)\}} \mathcal{N}_6^-(dv, du)
\end{aligned} \tag{1.17}$$

avec

$$\left\{ \begin{aligned} \tau_i^R(u_1^K(t)) &= \tau_i^* p_i^{R,*} \left(\frac{u_1^K(t)}{u_R^*} \right)^{r_i} \\ \tau_i^D(u_2^K(t)) &= \tau_i^* p_i^{D,*} \mu_i^* \left(\frac{u_D^K(t)}{u_D^*} \right)^{d_i} \\ s_i &= \tau_i^* p_i^{D,*} (1 - \mu_i^*) \\ n^K(N_6^K(t)) &= n^* \frac{1 + c_1}{1 + c_1 \left(\frac{N_6^K(t)}{K x_6^*} \right)^{c_2}} \end{aligned} \right.$$

où u_R^K , u_D^K sont des processus de Markov continus déterministes par morceaux définis entre chaque instant de sauts du processus N^K comme solution du système dynamique

$$\left\{ \begin{aligned} \frac{du_R^K(t)}{dt} &= \frac{1}{a + \sum_{j=1}^5 \frac{N_j^K(t)}{K x_j^*}} - m_R u_R^K(t) \\ \frac{du_D^K(t)}{dt} &= \frac{\sum_{j=1}^5 \frac{N_j^K(t)}{K x_j^*}}{b + \sum_{j=1}^5 \frac{N_j^K(t)}{K x_j^*}} - m_D u_D^K(t) \end{aligned} \right. \tag{1.18}$$

Les conditions initiales de u_1^K et u_D^K sont définies par

$$u_1^K(0) = \frac{1}{m_R(a + Y_0^K)} \quad \text{et} \quad u_2^K(0) = \frac{Y_0^K}{m_D(b + Y_0^K)}, \quad \text{avec} \quad Y_0^K = \sum_{j=1}^5 \frac{N_j^K(0)}{K x_j^*}.$$

Notons le processus $X^K \in \mathbb{D}(\mathbb{R}_+, \mathbb{R}_+^6)$ tel que

$$X^K = \frac{N^K}{K}.$$

Nous allons montrer que le processus (u_1^K, u_2^K, X^K) converge en probabilité vers l'unique solution du système (1.11) (à condition initiale fixée).

Théorème 12. *Supposons que l'ensemble des paramètres du modèle soient strictement positifs, que*

$$\sup_K \mathbb{E}[X^K(0)] < \infty$$

et que la variable aléatoire $X^K(0)$ converge en loi vers x_0 , un réel positif.

Alors pour tout $T > 0$, le processus (u_1^K, u_2^K, X^K) converge en loi (et en probabilité) dans $\mathbb{D}([0, T], \mathbb{R}_+^8)$ lorsque K tend vers l'infini vers l'unique solution (u_R, u_D, x) du système (1.11) telle que $x(0) = x_0$,

$$u_R(0) = \frac{1}{m_R(a + y_0)} \quad \text{et} \quad u_D(0) = \frac{y_0}{m_D(b + y_0)} \quad \text{avec} \quad y_0 = \sum_{j=1}^5 \frac{x_j(0)}{x_j^*}.$$

Démonstration. La dynamique du processus u_R^K consiste en des "switch" aléatoires entre différents systèmes dynamiques (1.18). Une sur-solution commune à chacun de ces systèmes est donnée par l'unique solution du système dynamique

$$\forall t, \quad \frac{du}{dt} = \frac{1}{a} - m_R u(t),$$

telle que $u(0) = u_1^K(0)$. Par conséquent, on obtient la majoration presque sûre suivante

$$\forall t, \quad u_R^K(t) \leq \frac{1}{a m_R} + \left(u_1^K(0) - \frac{1}{a m_R} \right) \leq \frac{1}{a m_R}.$$

De la même façon, presque sûrement

$$\forall t, \quad u_D^K(t) \leq \frac{1}{m_D} + \left(u_2^K(0) - \frac{1}{m_D} \right) \leq \frac{1}{m_D}.$$

Par un argument standard de localisation, le lemme de Gronwall ainsi que l'inégalité de Doob, on déduit des majorations précédentes et de celle de n^K par la constante $n^* c_1$ que pour tout $T > 0$,

$$\sup_K \mathbb{E} \left[\sup_{t \in [0, T]} (\|X^K(t)\|^2 + u_1^K(t)^2 + u_2^K(t)^2) \right] < \infty. \quad (1.19)$$

Nous déduisons ensuite des équations (1.17), la forme semi-martingale suivante pour X^K ,

$$\begin{aligned}
X_1^K(t) &= X_1^K(0) + \int_0^t (\tau_1^R(v) - \tau_1^D(v) - s_1) X_1^K(v) dv + M_1^K \\
\text{pour } i &= 2, \dots, 5, \\
X_i^K(t) &= X_i^K(0) + 2 \int_0^t \tau_{i-1}^D(v) X_{i-1}^K(v) dv + \int_0^t (\tau_i^R(v) - \tau_i^D(v) - s_i) X_i^K(v) dv + M_i^K(t) \\
X_6^K(t) &= X_6^K(0) + \int_0^t 2^{n^K(K X_6^K(v))} \tau_5^D(v) X_5^K(v) dv - \int_0^t \tau_6 X_6^K(v) dv + M_6^K(t)
\end{aligned} \tag{1.20}$$

où les variations quadratiques des martingales de carrés intégrables M_i^K s'écrivent

$$\begin{aligned}
\langle M_1^K \rangle_t &= \frac{1}{K} \int_0^t (\tau_1^R(v) - \tau_1^D(v) - s_1) X_1^K(v) dv, \\
\text{pour } i &= 2, \dots, 5, \\
\langle M_i^K \rangle_t &= \frac{1}{K} 2 \int_0^t \tau_{i-1}^D(v) X_{i-1}^K(v) dv + \int_0^t (\tau_i^R(v) - \tau_i^D(v) - s_i) X_i^K(v) dv, \\
\langle M_6^K \rangle_t &= \frac{1}{K} \left(\int_0^t 2^{n^K(K X_6^K(v))} \tau_5^D(v) X_5^K(v) dv - \int_0^t \tau_6 X_6^K(v) dv \right) \\
\text{pour } i &= 2, \dots, 5, \\
\langle M_{i-1}^K, M_i^K \rangle_t &= -\frac{1}{K^2} \int_0^t \tau_{i-1}^D(v) X_{i-1}^K(v) dv \\
\langle M_5^K, M_6^K \rangle_t &= -\frac{1}{K^2} \int_0^t 2^{n^K(K X_6^K(v))} \tau_5^D(v) X_5^K(v) dv
\end{aligned} \tag{1.21}$$

et

$$n^K(K X_6^K(t)) = n^* \frac{c_1}{1 + (c_1 - 1) \left(\frac{X_6^K(t)}{x_6^*} \right)^{c_2}}.$$

On déduit, des contrôles de moment (1.19), la tension de la suite de loi de $(X^K, u_1^K, u_2^K)_K$ puis la convergence en probabilité des martingales vers 0. La preuve du théorème s'obtient finalement en suivant le raisonnement détaillé par exemple dans le livre [5].

Sans détailler la preuve nous donnons ici quelques éléments.

On déduit de la tension du processus l'existence d'une sous-suite convergente par le Théorème de Prohorov. En observant que la taille des sauts du processus X^K est proportionnelle à $\frac{1}{K}$ et que u_R^K et u_D^K sont à trajectoires continues, on peut montrer que la loi limite ne charge que l'ensemble des fonctions continues dans $\mathbb{D}([0, T], \mathbb{R}_+^8)$.

Définissons les fonctions $\phi_t : \mathbb{D}([0, T], \mathbb{R}_+^8) \rightarrow \mathbb{R}_+^8$ par

$$\phi_t(u_R, u_D, x) = \begin{pmatrix} u_R(t) - u_R(0) - \int_0^t \left(\frac{1}{a + y(s)} - m_R u_R(s) \right) ds \\ u_D(t) - u_D(0) - \int_0^t \left(\frac{y(s)}{b + y(s)} - m_D u_D(s) \right) ds \\ x_1(t) - x_1(0) - \int_0^t \left(\tau_1^R(u_R(s)) - \tau_1^D(u_D(s)) \left(\frac{u_D(s)}{u_D^*} \right)^{d_1} \right) x_1(s) ds \\ x_i(t) - x_i(0) - 2 \int_0^t \tau_{i-1}^D(u_D(s)) x_{i-1}(s) ds \\ \quad - \int_0^t \left(\tau_i^R(u_R(s)) - \tau_i^D(u_D(s)) - s_i \right) x_i(s) ds \\ x_6(t) - x_6(0) - \int_0^t \left(2^{n(x_6(s))} \tau_5^D(u_D(s)) x_5(s) - \tau_6 x_6(s) \right) ds \end{pmatrix},$$

avec

$$y(s) = \sum_{j=1}^5 \frac{x_j(s)}{x_j^*}.$$

On peut alors d  duire du contr  le (1.19), l  uniforme int  grabilit   de la sous-suite de martingales de $(\phi_t(u_R^K, u_D^K, X^K))_K$.

Le Th  or  me de Cauchy-Lipschitz donne l  unicit   de la solution (u_R, u_D, x) du syst  me (1.11) telle que $x(0) = x_0$, $u_R(0) = \frac{1}{m_R(a+y_0)}$ et $u_D(0) = \frac{y_0}{m_D(b+y_0)}$ avec $y_0 = \sum_{j=1}^5 \frac{x_j(0)}{x_j^*}$.

De plus on peut d  duire de l  uniforme int  grabilit   de la sous-suite de $(\phi_t(u_R^K, u_D^K, X^K))_K$ que la loi limite charge uniquement l  ensemble des solutions du syst  me

$$\forall t \geq 0, \quad \phi_t(u_R, u_D, x) = 0.$$

En remarquant que les solutions de ce syst  me correspondent    celle du syst  me (1.11), on conclut la preuve. □

1.5 Calibration du mod  le r  gul  

L  objectif ici est de mod  liser la r  gulation de l  rythropo  se de stress suite    une perturbation. Plus particuli  rement nous mod  lisons le retour    l   quilibre du syst  me apr  s la perturbation d  finie par l   tat du syst  me au jour 3. En effet, nous choisissons de mod  liser la dynamique des cellules apr  s le jour 3. La PHZ perturbe fortement le syst  me durant 3 jours (baisse des globules rouges) et nous souhaitons non pas mod  liser son action mais la r  gulation du syst  me une fois son effet dissip  .

Nous avons de plus choisi de mod  liser les conditions initiales des r  gulateurs u_R et u_D    l  aide des donn  es des tailles de populations au jour 3 de l  exp  rience de la fa  on suivante,

$$u_R(3) = \frac{1}{m_R(a + y(3))},$$

$$u_D(3) = \frac{y(3)}{m_D(b + y(3))}.$$

Ce choix correspond à l'annulation de la dérivée de u_R et u_D au jour 3 ($\frac{du_i(3)}{dt} = 0$). Il est inspiré du rôle d'instant charnière que représente ce jour : on passe d'une dynamique de "perturbation" à une dynamique de retour à l'équilibre.

Afin de calibrer notre modèle, nous allons utiliser les données présentées à la sous section 1.3.1 mais également le travail effectué en section 1.2. En effet, le système sera revenu à son équilibre 2 mois après l'injection de PHZ. Ainsi nous avons choisi d'ajouter deux points supplémentaires au jour 70 et au jour 90 correspondant aux valeurs à l'équilibre des tailles de population vues en section 1.2.

Plus précisément, les données décrites nous permettent de calibrer les paramètres du modèle en utilisant la fonction coût suivante,

$$f_{cost}(\theta) = \sum_{i=1}^6 \sum_{j \in \{J5, J10, J16, J28, J70, J90\}} \delta_{i,j} \left(\frac{x_i(t_j, \theta) - S_i(t_j)}{x_i(t_j, \theta)} \right)^2. \quad (1.22)$$

Ici, $\delta_{i,j}$ est égale à 0 pour $(i, j) = (4, J28), (5, J28), (6, J7)$, et est égale à 1 pour tous les autres couples (i, j) . Tandis que $\theta \in \Theta$ représente un jeu de paramètres parmi l'ensemble Θ (décrit ensuite). Le terme $x_i(t_j, \theta)$ représente la i -ème coordonnée de la solution du système (1.10) au tps t_j , lorsque l'ensemble des paramètres sont donnés par θ et la condition initiale par les données au jour 3. Le terme $S_i(t_j)$ représente la valeur des données relatives au nombre de cellules de type i au temps t_j .

L'ensemble Θ dans lequel les paramètres sont recherchés se décrit comme suit,

$$m_R \in [0, 1], \quad m_D \in [0, 1], \quad a \in [10^{-4}, 10], \quad b \in [0, 100], \quad c_1 \in [0, 0.3], \quad c_2 \in [0, 50], \\ r_i \in [0, 10], \quad d_i \in [0, 10].$$

Nous avons choisi ces bornes pour différentes raisons :

- Les constantes m_i sont prises inférieures à 1 pour envisager une dégradation progressive des régulateurs.
- Les constantes a et b jouent un rôle de saturation concernant la production de quantité u_R et u_D . Leurs bornes sont prises suffisamment grandes pour ne pas influencer la recherche de l'algorithme d'optimisation utilisé pour la calibration. En effet, nous allons voir que le rôle de ces bornes sur le résultat final est négligeable.
- Les constantes c_1 et c_2 sont liées à la dynamique de n . Nous avons précédemment mentionné les travaux [19], [2] concernant cette partie de l'érythropoïèse. En particulier, nous savons que L'EPO agit sur la survie des cellules ([61]). De telles bornes permettent d'une part d'expliquer la dynamique des globules rouges et sont d'autre part en accord avec les travaux précédemment effectués sur cette partie de l'érythropoïèse.
- Ces bornes prises pour r_i et d_i respectent les ordres de grandeur estimés pour les valeurs maximales que peuvent prendre les taux de division des cellules en cas de stress ([91], [49]).

L'utilisation d'un algorithme d'optimisation stochastique basé sur une méthode appelée CMA-ES (Covariance Matrix Adaptation in Evolution Strategy, [48]), nous a alors permis de minimiser la fonction de coût (1.22). Nous avons ainsi obtenu les meilleurs paramètres (relativement à cette fonction de coût) permettant aux courbes des tailles de populations décrites par notre modèle d'expliquer les données biologiques. L'algorithme stochastique

utilisé, dont le fonctionnement a été présenté en Introduction du présent manuscrit, permet d'assurer une identifiabilité dite numérique de notre modèle régulé.

La solution $x(t, \theta)$ du système (1.10), où θ représente le jeu de paramètres obtenu à l'aide de l'algorithme, est présentée en figure 1.5.

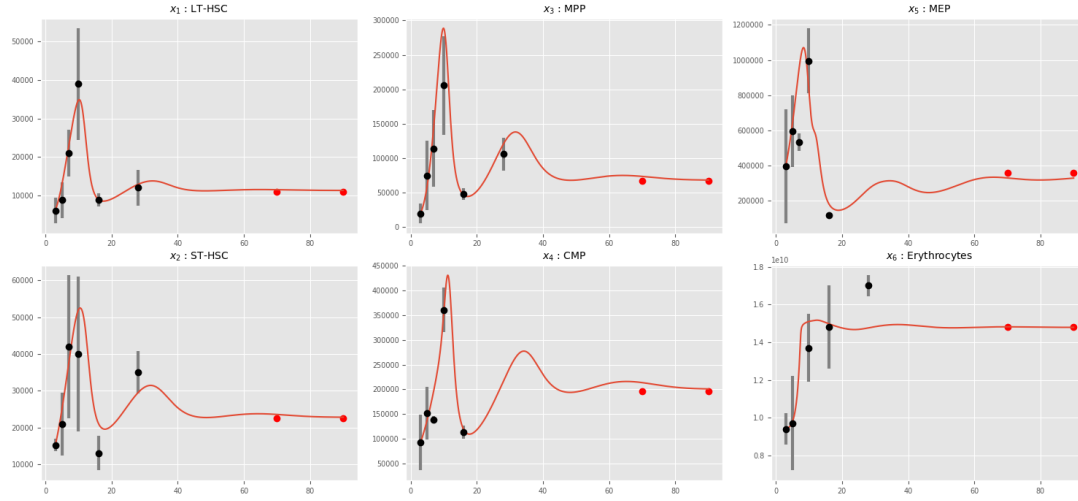


FIGURE 1.5 – Evolution au cours du temps du nombre de cellules de chaque type cellulaire de la lignée rouge dans leur période de retour à l'équilibre suite à une injection de PHZ.

La courbe en ligne continue représente la solution x tandis que les points représentent l'ensemble des données utilisées pour calibrer l'algorithme.

La figure 1.5 montre que notre modèle est en adéquation avec les données biologiques obtenues par notre équipe. Toutefois nous avons conscience qu'un travail plus détaillé des mécanismes de division cellulaire entre les MEP et les globules rouges nous permettrait sûrement d'améliorer le "fit" concernant la dynamique des globules rouges.

A l'aide de l'algorithme, nous avons obtenu le jeu de paramètres optimal suivant (ces paramètres sont sans unité à l'exception de m_r et m_d),

$$\begin{aligned} m_R &= 0.03 \text{ (/j)}, \quad m_D = 0.2 \text{ (/j)}, \quad a = 10^{-4}, \quad b = 100, \quad c_1 = 0.15, \quad c_2 = 50, \\ r_1 &= 9.2, \quad r_2 = 6.35, \quad r_3 = 6.38, \quad r_4 = 2.76, \quad r_5 = 0.33, \\ d_1 &= 10, \quad d_2 = 8.4, \quad d_3 = 7.72, \quad d_4 = 7.4, \quad d_5 = 4.3. \end{aligned} \tag{1.23}$$

Remarquons plusieurs points concernant la valeur de ces paramètres

- Les paramètres a et b atteignent leurs bornes. En effet le paramètre a est pris de telle sorte qu'il est négligeable devant la somme y au cours de l'expérience (voir (1.10)). Le paramètre b quant-à-lui est choisi par l'algorithme de sorte à rendre la somme y au dénominateur négligeable devant lui. Leurs valeurs exactes ne changent pas le résultat obtenu.
- Le paramètre c_2 atteint également sa borne. L'effet de ce paramètre sur la dynamique de n est décrite par la figure 1.3. Exactement comme précédemment, l'influence de ce paramètre est limitée. Ainsi passée une certaine valeur, le résultat obtenu ne dépend plus de ce paramètre.

- Le paramètre d_1 atteint lui aussi sa borne. De façon générale, le choix de l'intervalle $[0, 10]$ pour les paramètres r_i et d_i ne se fait pas au hasard. En effet, j'ai mené différents tests qui ont révélé qu'une borne maximale inférieure à 10 ne permet pas aux cellules (dont la dynamique est décrite par le modèle) de proliférer suffisamment pour atteindre les valeurs des données au jour 10. De plus lorsque cette borne est prise trop grande, on perd ce que j'appelle en Introduction de la thèse l'identifiabilité numérique du modèle. En effet, la dynamique des u_i et la valeur des puissances associées aux différents taux forment alors un couple dont les effets de l'un peuvent compenser ceux de l'autre.
- On observe une décroissance des paramètres r_i et d_i avec le type cellulaire i . Cette décroissance suggère que les cellules les plus avancées dans la différenciation de notre modèle (CMP, MEP) ont moins besoin d'être régulées pour observer la dynamique de retour à l'équilibre attendu. Bien que moins détaillé, ce comportement a été remarqué par les auteurs de l'article [13], lors d'un stress différent (au 5-Fluoruracil).

Une discussion concernant le modèle pourra être trouvée dans l'article joint au chapitre. Ajoutons simplement ici deux remarques concernant la dynamique des cellules.

(1) Notre modélisation, nous donne également accès aux valeurs des taux de division des cellules au cours de l'expérience. Ces valeurs sont inférieures à une division par jour pour les LT-HSC, ce qui est en accord avec la littérature [91]. De plus les taux de division des autres types cellulaires peuvent quant à eux atteindre une valeur maximale d'un peu moins de 3 divisions par jour. Ce qui est cohérent avec les capacités dynamiques qu'ont montré ces cellules sous certaines conditions [49].

(2) On observe que notre modèle ainsi calibré induit une décroissance du nombre de cellules de type 5 un peu avant le jour 10. Cette décroissance s'explique par la valeur proche de 0 de la constante r_5 à l'opposée de celle plus élevée de la constante d_5 . En effet, la dynamique de la 5ème population est donnée par

$$\frac{dx_5(t)}{dt} = \tau_4^* p_4^{D,*} \mu_4^* \left(\frac{u_D(t)}{u_D^*} \right)^{d_4} x_4(t) - \left(p_5^{D,*} \left(\frac{u_D(t)}{u_D^*} \right)^{d_5} - p_5^{R,*} \left(\frac{u_R(t)}{u_R^*} \right)^{r_5} \right) \tau_5^* x_5(t).$$

Cette dynamique se décompose en deux parties : l'une corrélée à celle du compartiment en amont et l'autre due à la régulation du compartiment lui-même (Régulation, ayant ici une sensibilité r_5 à s'auto-renouveler faible). Ce comportement est intéressant car il est une conséquence du caractère "en cascade" ou décomposable du processus.

1.6 Article associé (prochainement soumis)

Modeling the behavior of Hematopoietic Compartments from Stem to Red Cells in steady state and stress hematopoiesis.

Céline Bonnet^{*3}, Panhong Gou^{*1}, Simon Girel³, Vincent Bansaye³, Catherine Lacout¹, Karine Bailly², Marie-Hélène Schlagetter¹, Evelyne Lauret^{✧2} Sylvie Méléard^{✧3,4}, and Stéphane Giraudier^{✧1}.

*These two authors contribute equally to this work.

✧ These three authors are co-last authors.

¹Université de Paris, Hôpital Saint Louis, INSERM U1131, F-750 Paris, France

² Université de Paris, Institut Cochin, INSERM U1016, CNRS UMR8104, F-75014 PARIS, France

³ CMAP, CNRS, Ecole polytechnique, Institut Polytechnique de Paris, 91128, Palaiseau.

⁴ Institut universitaire de France

Short title: Modelling hematopoiesis from stem to red cells

Corresponding author: Stéphane Giraudier

Address: INSERM U1131, Centre Hayem, Hôpital Saint Louis, 1 avenue Claude Vellefaux, 75010 Paris, France Phone: +33 (0)1 Fax: +33 (0)1 e-mail: stephane.giraudier@aphp.fr

Supported by grants from the INSERM and INCa 2018 and fellowships from The Research ministry of China.

Abstract

Mathematical modelling integrating both unperturbed and stress hematopoiesis is a powerful approach to address key questions in hematopoiesis. Here we propose an erythropoiesis model leaning on a minimum of 6 cell-amplification compartments able to reproduce the steady state approach, and further complexified by integrating stress regulatory processes. Such mathematical model mimics well what is observed *in vivo* with the LT-HSC, ST-HSC, MPP, CMP, MEP and mature red blood cell compartments. After a chemical hemolytic anemia, we observed a severe but rapidly recovered anemia and a reduction of all progenitor compartments followed by a rebound before normalization. By adding to the stochastic model regulatory processes, we faithfully recapitulated the different progenitor compartment evolutions observed after stress as well as in steady state. To our knowledge, this is the first multi-step modelling of hematopoiesis, based on multi-step *in vivo* data, taking into consideration the time-dependent effects of regulation able to recapitulate and reconcile steady state as well as stressed erythropoiesis.

INTRODUCTION

Hematopoiesis provides the lifelong supply of large number of mature blood cells derived from a rare population of multipotent hematopoietic stem cells (HSCs) ¹. The lineage relationship between HSCs and mature cells was proposed as a multistep process in which generations of diverse blood cells are coupled with movement through consecutive amplification motors, from multi-, to oligo-, and lastly unipotent progenitors to mature blood cells ²⁻⁴. HSC division and differentiation are controlled by extracellular signals originating from the microenvironment, and by intracellular networks. In the classical model, successive progenitor compartments, Long-Term HSC (LT-HSC), Short-Term HSC (ST-HSC), MultiPotent Progenitor (MPP), Common Myeloid progenitor (CMP), etc.. of the hematopoietic differentiation process are defined on their immunophenotypic level, typically by the expression of cell surface marker combinations ⁵. In this context, since the 1960s, HSCs are typically functionally defined by their ability to sustain multi-lineage engraftment for an extended period of time upon serial transplantation into irradiated recipient mice. Recently, new generation of experimental tools allowing *in situ* analysis of HSC output has challenged the classical model by pointing out fundamental differences between physiological unperturbed hematopoiesis and stress hematopoiesis such as the reconstitution of the hematopoietic system after transplantation. In particular, in native hematopoiesis, progenitors directly downstream of LT-HSCs (ST-HSC, MPP...) serve as a major, nearly self-renewing source of day-to-day hematopoiesis, rendering the blood and immune system less dependent on LT-HSC, in agreement with the quiescence of adult LT-HSCs ⁶⁻¹⁶. In addition, recent single cell profiling analysis has unveiled the tremendous heterogeneity of each previously defined compartment. Thus, developmental stages downstream of LT-HSCs, such as MPPs or CMPs should rather be considered as transitory states within a hematopoietic stem and progenitor (HSPC) continuum ¹⁷⁻¹⁹.

Mathematical modelling integrating both unperturbed and stress hematopoiesis should be a powerful approach to address key questions in hematopoiesis, able to provide qualitative and quantitative insights into stem cell dynamics and fate commitment. To date, most of hematopoiesis modelling are based on the existence of few cell compartments (stem cells, progenitors and mature cells), and do not take into consideration all of the hematopoietic immature compartments described above ^{20,21}. Rodewald's group has introduced a major breakthrough by modelling hematopoiesis based upon all immature compartments in unperturbed hematopoiesis (Busch et al., 2015). Nevertheless, they notice that their steady state model cannot recapitulate stress hematopoiesis such as 5-fluorouracile (5-FU) treatment, suggesting that unperturbed and stress hematopoiesis do not follow the same biological rules.

Consequences of acute stress on hematopoiesis are generally studied through administration of myelotoxic agents (such as 5-FU or irradiation), known to impact all proliferating cells. Regarding stress impacting solely mature cells, most of the studies have focused their attention on committed progenitors, but the effects on more immature compartments have been poorly addressed.

Here, we hypothesized that stress and unperturbed hematopoiesis can be combined in a single system. We conducted *in silico* modelling in parallel with *in vivo* experiments to integrate homeostasis and stress hematopoiesis (such as phenylhydrazine treatment destroying mature red blood cells, RBC) ^{22,23} and define minimal number of hematopoietic compartments allowing to recapitulate the feedback effects of acute stress on immature progenitor compartments.

RESULTS

We first developed a steady-state mathematical model able to recapitulate hematopoiesis from quiescent LT-HSC to mature RBC. Modelling quiescence/kinetics of proliferation/differentiation processes requires a theoretical number of compartments between LT-HSC and mature cells. However, deciphering an exact compartment number remains elusive since currently, nature and number of compartments are not standardized from one study to another (when a continuum is not considered) ^{17,24}. We assumed that the theoretical number of compartments should rely upon at least 3 criteria for each compartment: 1- size, 2- duration of proliferative steps and 3-self-renewal versus differentiation capacities.

We hypothesized that the model relies upon cell divisions, and integrates stochasticity and amplification in each compartment as well as plasticity in the number of intermediate compartments. Based on this assumption, a stochastic multi-type branching process was designed to recapitulate native (unperturbed) hematopoiesis, which led to the derivation of a macroscopic deterministic system. *In vivo* steady-state parameters applied to this model help us to determine the minimal number of amplification motors (k cell compartments) necessary to obtain a lifelong stable RBC production.

Modelling of unperturbed early erythropoiesis

Our model describes a hierarchy of k cell compartments evolving over time. We assumed that each cell of a given compartment divides with a constant rate independently from the others. A cell in a compartment is characterized by its division rate, its self-renewal and differentiation capacity, and for the last compartment (corresponding to RBC), its death rate.

The entire dynamics of hematopoiesis presented in **Figure 1A and 1B** can be summarized as follows:

1- For any $i = 1, \dots, k - 2$, a cell belonging to the compartment i divides at rate $\tau_i > 0$ in two cells of the same compartment with a probability $0 < p_r^i < 1$ (self-renewal) or in two cells of compartment $i + 1$ with a probability $\mu_i p_d^i (= \mu_i (1 - p_r^i))$ (erythroid differentiation). The probabilities p_d^i correspond to differentiation, and μ_i to commitment towards erythroid rather than non-erythroid lineages. Note that the $i=1$ compartment corresponds to LT-HSC and $i=k$ to mature RBC.

2- A $(k-1)$ -type cell “divides” at rate $\tau_{k-1} > 0$. Upon division, the cell renews with probability $0 < p_r^{k-1} < 1$, and gives rise to 2^n k -type cells with probability $p_d^{k-1} = 1 - p_r^{k-1}$. This dynamic summarizes the n mitosis with negligible self-renewal observed at the end of the erythroid differentiation process. We simplified this amplification as 2^n RBC issued from 1 cell of the previous compartment since terminal erythroid differentiation has been largely modeled before and is not our main concern.

3 - A k -type cell corresponding to a mature RBC dies at rate $\tau_k > 0$.

We are aware that stem cell division could be asymmetric (a cell of type i gives birth to two daughter cells, one of type i and the other of type $i+1$). However, this feature would not change the mathematical analysis of the system; we therefore neglected it to simplify the model. Furthermore, we presumed that the LT-HSC compartment is at equilibrium meaning that the number of cells that self-renew is equal to the the number of cells that differentiate, *i.e* $p_d^1 = p_r^1$. Moreover we assumed that for any $i \in \{2, \dots, k-1\}$, $p_d^i - p_r^i > 0$ in order to ensure the stability of the system. For any type $i = 2, \dots, k - 1$, the i -compartment differentiation factor is defined as $D_i = p_d^i - p_r^i$. Since $p_d^i + p_r^i = 1$, $D_i = 1 - 2 p_r^i = 2 p_d^i - 1$.

According to the literature ²⁵, the number of LT-HSC is large enough to approximate the stochastic model by a system of differential equations (**Figure 1C**) and then, compartment sizes at equilibrium can be expressed (**Figure 1D**). In particular, the i compartment size relies upon the one of compartment $i-1$ multiplied by $1/D_i$, meaning that the smaller the D_i , the bigger the i compartment size.

As a first step, we assumed that the D_i are equal to a positive value D and the probabilities μ_i are equal to 1. Then a direct relationship can be computed between the number of LT-HSC, their division rate, the D parameter, the n number of final mitoses, the number of mature RBC, and their death rate (**Figure 1E**). The minimal number of amplification cell compartment can then be computed based on this formula.

Determination of the minimal compartment number

The determination of the theoretical minimum number of compartments is based on **Figure 1E**, and then depends on the number of RBC, their death rate, the number of LT-HSC, their division rate, the D parameter and the number of mitosis between the two last compartments.

The number of RBC in mouse, and their death rate are known (ie $\approx 10^{10}$ cells/ mouse, and 40-day life expectancy). The number of LT-HSC was evaluated on the basis of the proportion of LT-HSC (Lin- Sca-1⁺ c-kit⁺, LSK CD150⁺CD48⁻) estimated through a gating strategy of 2.5×10^6 cytometry analyzed bone marrow cells as described in **Figure S1**. Knowing that a C57Bl6 mouse displays $\approx 250 \times 10^6$ bone marrow cells ²⁶⁻²⁸, the LT-HSC number was estimated to be 15248 +/- 2200 cells /mouse.

The division rate of LT-HSC can be computed from the proportion of LT-HSC retaining their phenotype after 20 hours of BrdU incorporation at steady state (1.82 in [0-4,7] % of cycling

LT-HSC) (**Figure 1F** and **Figure S1**). Hence the division rate of LT-HSC is approximated by 1/100 per day.

Assuming now that the D_i must increase as a function of i and since $D_k=1$ (no self-renewal), D_i belongs to $[D_2;1]$. As in Busch's report, we considered that the second compartment corresponds to the ST-HSC compartment. Then a minimal value for D_2 ($D_2 \geq 0.16$) was computed using **Figure 1D**, with bounds for division rate obtained from **Figure 1F** and data from **Table 1**.

The n number of mitosis between the two last compartments has been assessed as follows: we assume that the $k-1$ compartment is upstream of BFU-E (Burst-Forming-Unit Erythroid)-enriched cells known to be able to perform 11 cell divisions to generate mature RBC *in vitro* (Li et al., 2019). We therefore assume that “ n ” should be over 12.

Taking into consideration all these parameters with n equal 12 and **Figure 1E**, we can compute that if $k=5$, then $D=0.12$; if $k=6$, then $D=0.23$; if $k=7$, then $D=0.35$. Therefore, the minimum compartment number to obtain $D \in [0.16;1]$ is 6. This compartment number mostly fulfills those generally assessed as LT-HSC, ST-HSC, MPP, CMP, MEP, and RBC. As previously assessed, the last steps (MEP to RBC) in our mathematical modelling, is simplified as 2^n RBC from 1 committed MEP. Based on cytometric analysis, we evaluated the number of cells in each compartment at steady state and their division rates (**Table 2**).

Parameter calibration of the mathematical model for steady-state erythropoiesis.

The 6-compartment model allows us to calibrate all parameters. We first gave bounds to the division rate per day. Taking into account variation due to the experimental procedure on BrdU we used bounds rather than median values to delineate division rates (see Figure 1F). The bounds gave us the opportunity to include standard deviation from one experiment to another and a variability of the time interval between BrdU injection and analysis (20h +/-4h).

We then computed appropriate values of all parameters (**Table 2**). To this aim, we took into consideration 1- the population size presented in **Table 1**, 2- the probabilities μ_i (based on Busch data) of differentiation toward the erythroid lineage, 3- the decreased probability of self-renewal along the differentiation process, and 4- the bounds of the division rate of each cell type. The erythro-myeloid commitment probabilities μ_i were obtained considering Busch et al. data and dividing the differentiation rate toward erythro-myeloid lineages by the sum of differentiation rates toward all lineages. Based on a decreasing probability of self-renewal in the different compartments along the differentiation process, the D_i were computed allowing us to calibrate the model using **Figure 1D**. Our results presented in Table 2 showed that this steady state erythropoiesis model was able to reproduce unperturbed native hematopoiesis for at least two years, the normal mouse life time.

Steady-state model applied to acute erythroid stress.

In order to validate the 6-amplification compartments model of hematopoiesis in stress conditions, we applied the highly reproducible phenylhydrazine administration (**Figure S2A**, protocol), which induces an acute hemolytic anemia after administration of a single dose of 60 mg/Kg IP. After the 40%-hemolysis stress observed at day 3, a rapid recovery of RBC was

evidenced after 10-16 days (**Figure 2A**). Our 6-step model of erythropoiesis was applied to recapitulate the RBC recovery after day 3. Based on this steady-state model, the recovery time appeared more than one month, in total discrepancy with the biological data observed (**Figure 2A**). This indicated that, although able to fulfill hematopoiesis at steady state, this model was unable to recapitulate the hematopoietic stress response, and required additional regulations.

Analysis of compartment kinetics during acute hematopoietic stress in mice.

In order to better decrypt the biological mechanisms responsible for this fast RBC recovery, we analyzed all compartment modifications after PHZ injection in blood, bone marrow, and spleen. Briefly, during the first 3 days after injection, we observed a drastic fall of RBC, and an increase in platelets and leukocytes counts (**Figure 2A-C**). While the total number of bone marrow cells was slightly reduced at day 3 (-25%) and enhanced at day 10 (+33%) (**Figure 2G**), the spleen weight strongly increased 24h after PHZ injection (**Figure 2D-F**), becoming 3 times larger than normal at day 3, as previously observed after erythroid stress^{22,29} and social stress in mice³⁰. The three hematopoietic tissues (bone marrow, spleen and blood) were analyzed to assess the proportion and absolute numbers of cells in the different compartments (LT-HSC, ST-HSC, MPP, CMP, and MEP) as well as their percentages into cell cycle (BrdU incorporation), and apoptosis (Annexin V labelling) during the hemolytic phase. In the bone marrow, a decrease in the size of all compartments (from LT-HSC to CMP) was evidenced. Bone marrow MEP population remained relatively stable during this phase, while a huge increase in the number of splenic MEP was observed 3 and 5 days after PHZ injection (**Figure 2G**). No difference in BrdU incorporation was noticed at day 3 in the different bone marrow compartments except for MEP, which displayed a 2-fold increased cell proliferation, as compared to D0 (**Figure S2F**). To further get insight into mechanisms involved in the drastic decrease of compartment cell

numbers, we first analyzed apoptosis. No significant difference in the proportion of Annexin V-labelled cells was observed between PHZ-treated and control mice whatever the progenitor compartment examined, allowing us to exclude any excess of cell death in the bone marrow progenitor compartments, from LT-HSC to MEP (**Figure S3B**). Blood cell analysis performed at day 1 and 3 after PHZ injection showed a 15-fold increase in the number of mononuclear cells in the blood (**Figure 2B**) yet with a neglectable proportion of lin⁻ cells, precluding an egress of LSK from the BM to the blood during these 3 days (**Figure S3A**). In the spleen, a stress erythropoietic organ in mouse, the number of mature cells increased (3-fold) during these 3 first days (**Figure 2E**) whereas LSK, CMP, and MEP compartments displayed no significant variations in cell number ruling out any role of splenic immature progenitors at day 3 (**Figure 2G**). Lastly, to confirm the stemness properties of LT-HSC assessed by cytometry and exclude any change in the expression of CD48 or CD150 markers after exposure to PHZ, competitive BMT with 50% of CD45.1 bone marrow cells from untreated animals and 50% of total bone marrow cells from day 3-PHZ-treated mice were performed in CD45.2 9.5 Gy-irradiated recipients. Blood chimerism analysis performed 3 months post-transplantation showed a decreased proportion of PHZ-treated BM cells, confirming the decrease at day 3 in the number of LT-HSC in bone marrow of PHZ-treated mice (**Figure S3C**).

The decrease of the bone marrow progenitor compartment sizes, without increased proliferation, or apoptosis or egress of these cells to the spleen or the blood, together with a stable number of MEP displaying a two-fold increased proliferation in bone marrow and spleen appears to derive from a rapid differentiation process, like a flush, to quickly compensate the loss of mature RBC through a versatile mechanism, as previously highlighted for platelet production in stress circumstances ³¹.

Recovery of progenitor compartment after PHZ stress in mice.

To assess the dynamics of progenitor compartments during recovery after PHZ treatment, we analyzed at days 3, 5, 7, 10, 16 and 28 the different compartment sizes and proliferation rates as previously described. The recovery was almost effective in all compartments at day 7-10 (**Figure 4A**). Surprisingly, at day 10, all compartments displayed a 2-3-fold increase in cell numbers before returning to normal values at day 28. In parallel, BrdU analysis showed that all progenitor compartments exhibited a drastic increase in cell proliferation at day 10 in the bone marrow, at the time point of the hematocrit normalization. Thereafter, BrdU incorporation returned to normal values at days 16 and 28 (**Figure S2C-G**).

Modelling steady-state and stress hematopoiesis by adding compartment regulations

The steady state model previously introduced cannot explain the fast return of RBC to equilibrium. Therefore, to model all at once unperturbed and stress hematopoiesis, we introduced regulations by assuming that the previously defined parameters depend indirectly on the size of the compartments through dynamics of two regulators (u_r and u_d) modulating self-renewal and erythroid differentiation bringing positive and negative effects on the compartment sizes, respectively (**Figure 3A**).

The dynamics of the two regulators followed a double system of production and clearance. The production is given by Michaelis-Menten type functions depending on the bone marrow cell numbers. The clearance of the regulator after its production is given by a linear degradation term (**Figure 3B**). This modelling induced a delay between the dynamics of the regulators and

the changes in the compartment sizes, explaining the huge oscillations of cell numbers after PHZ injection.

Assuming that the proliferation rate during late erythropoiesis is regulated in a specific manner (Erythropoietin, Stem Cell Factor, etc...), the production dynamics of RBC from MEP is controlled with a third type of regulation. MEP differentiation gives birth to $2^{n(t)}$ RBC summarizing the $n(t)$ last mitosis with a Hill function (Michaelis-Menten's type) of the RBC number (**Figure 3C**). Finally, we hypothesized that the differentiation towards a non-erythroid lineage was not affected by these regulators.

As shown in **Figure 4B and Table 3**, this regulated model led us to fulfill the experimental data in all compartments in stress condition as well as in steady state.

DISCUSSION

Our present work provides a mathematical modelling able to reconcile steady state and stressed hematopoiesis. Furthermore, it highlights a mechanism by which transient stress exposure leads to a protracted influence on all hematopoietic compartments, leading to a fast repopulation of mature compartment.

In order to determine whether our model could integrate steady-state as well as stress hematopoiesis, we analyzed the consequences of stress in the theoretical compartment behavior. We here focused on erythroid response since erythroid cells are the most abundant cells in the blood (approximately 1000-times more than white blood cells and 10-times more than platelets). To decipher the mechanisms involved in the normalization of RBC number at steady

state and after an acute stress, we first developed a compartmental model of hematopoiesis based on the classical hematopoietic compartmentalization established upon phenotypic, functional and molecular characteristics. The theoretical number of compartments sufficient to fulfill stable erythropoiesis was relied upon the number of LT-HSC and mature RBC in mouse. Our modelling proposes a minimum of 6 amplification compartments sufficient to ensure steady state erythropoiesis. We hypothesized that the last compartment could be resumed from MEP to RBC as a simple amplification factor according to the n number of mitosis required from 1 differentiated MEP to RBC. Based on the division number between BFU-E and erythrocytes, this n number could vary from 11 to 14 ³². We neglected megakaryocytic differentiation that theoretically only represents 20% of the MEP differentiation ³³ as well as other myelo-lymphoid cells that represents a quantitative minority of cells. Then, this 6-compartment model parallels the generally assessed compartment hierarchy with LT-HSC, ST-HSC, MPP, CMP, MEP and mature RBC.

The initial stress assessed in our “stress model” was a destruction of red blood cells by chemical hemolysis but this hemolysis process induces also a huge inflammatory response leading to changes in all blood cell types. In agreement with this drastic inflammatory process a huge cytokine storm in the next few days after hemolysis was noticed (data not shown). Stress hematopoiesis has been defined as different from unperturbed hematopoiesis. The differentiation process could be accelerated or inversely the step number during differentiation could be multiplied as illustrated by the megakaryocytic differentiation process under stress conditions ³¹. More recently, the step number of the granulocytic differentiation has also been demonstrated as varying according to stress (Simultaneous tracking of division and differentiation from individual hematopoietic stem and progenitor cells reveals within-family homogeneity despite population heterogeneity ³⁴).

Based on our data and modelling, we demonstrate here that regulations of compartment proliferation and differentiation (u_r and u_d) are required to reproduce the stress response while not necessary for steady-state erythropoiesis, the regulators remaining constant in the latter.

Our model allows us also to compute the self-renewing rate affecting the different compartments at steady state. BrdU is considered as a marker of cell proliferation and allow us to assess self-renewal in the LT-HSC compartment (since differentiating LT-HSC becomes a ST-HSC). In contrast, in the following compartments, BrdU labeling results from self-renewal and cell differentiation from the upstream compartment. Probabilities of self-renewal and differentiation toward the different lineages were then computed based on the BrdU data in one hand and Busch derived data for the differentiation steps observed toward non-erythroid lineages (μ_i in our model) in the other hand.

Our model also led us to analyze division rates in each compartment after stress. The division rate of each progenitor compartment reached its maximum value around day 11 but still remained lower than 3 divisions per day that is in accordance with biological data. (**Table 4**).

Previous mathematical models based on *in vivo* experiments mostly described hematopoiesis dynamics at steady state. These studies rely upon parameters such as cell division rates or probabilities of self-renewing^{7,10,35–40}. Models integrating steady state as well as *in vivo* stress hematopoiesis are less frequent in the literature^{20,21,37,39,41–44} and did not reproduce the oscillations observed in the size of the progenitor compartments apart from the modelling of^{20,41}. These authors modeled the regulation dynamics by a differential equation with a delay, enabling to predict the oscillating behavior of progenitor compartments characterized by a regulated self-renewal. Their model integrates a theoretical 3-compartment model (stem cells/progenitors/mature erythrocytes) based only on RBC recovery data after stress. Other

regulation-based models have been developed but do not integrated *in vivo* data ^{45,46}. In contrast, our present model was designed to reproduce the *in vivo* dynamics with 6 compartments observed throughout unperturbed and stress erythropoiesis, using 8 differential equations, including two for regulator dynamics and naturally creates a delay in the oscillation phenomenon due to the duration needed for regulator production and clearance.

Little is known regarding *in vivo* consequences of a “peripheral” stress on upstream progenitor compartments such as LT-, ST-HSC, MPP, CMP, and MEP. Our *in vivo* experiments demonstrate that after an acute stress destroying mature (red) blood cells, the first compensatory mechanism entails all compartments, from LT-HSC to CMP, and surprisingly, requires as a first step, differentiation without enhanced proliferation just like a flush. This kind of flush previously described for other cell lineage regeneration than erythrocytes ^{31,47} appears as a common response to acute stress. Thereafter, all depleted compartments enter into proliferation during a recovery phase suggesting a similar mechanism of regulation whatever the compartment. This recovery phase not only replenishes the different “flushed” compartments but exceeds steady-state values. Normalization of all compartment sizes occurs in a third phase just like spring oscillation around the standard values.

These oscillations in cell number from each compartment are in accordance with *in vivo* data previously reported in post-chemotherapy stressed hematopoiesis ⁷. In contrast, the MEP compartment was not subjected to such a high regulation process in the bone marrow probably because of the well-known spleen stress erythropoiesis process that follows a different kind of kinetic and regulation as reported here and in previous reports ⁴⁸.

In conclusion, we model here the plasticity of hematopoiesis recapitulating steady-state as well as stress hematopoiesis, no compartment being “protected” from stress consequences. We demonstrated that regulators are the main factors affecting cell proliferation and differentiation.

Thus, our model opens the way to better understand malignant clones development and invasion: Stem cell disorders just like leukemias and more generally cancers could arise from successive stress if pathological clonal cells react differently from normal cells in response to aggression.

MATERIALS AND METHODS

This section is described in details in the supplemental information section.

ACKNOWLEDGMENT

The authors are grateful to C. Chomienne for helpful discussions and improving the English manuscript. The authors greatly acknowledge the Cochin Cytometry and Immunobiology Facility.

AUTHORSHIP

CB, SiG, VB, SM conducted modeling of the paper, PG, CL, KB, MHS, EL and StG conducted in vivo experimental procedures, EL and SG performed the analysis of the data. CB, SiG, EL, SM and StG wrote the paper.

DISCLOSURE OF CONFLICT INTEREST. The authors declare no competing financial interests.

REFERENCES

1. Orkin SH, Zon LI. Hematopoiesis: an evolving paradigm for stem cell biology. *Cell*. 2008;132(4):631–644.
2. Cabezas-Wallscheid N, Klimmeck D, Hansson J, et al. Identification of regulatory networks in HSCs and their immediate progeny via integrated proteome, transcriptome, and DNA methylome analysis. *Cell Stem Cell*. 2014;15(4):507–522.
3. Seita J, Weissman IL. Hematopoietic stem cell: self-renewal versus differentiation. *Wiley Interdiscip Rev Syst Biol Med*. 2010;2(6):640–653.
4. Till JE, McCULLOCH EA. A direct measurement of the radiation sensitivity of normal mouse bone marrow cells. *Radiat. Res*. 1961;14:213–222.
5. Eich M, Trumpp A, Schmitt S. OMIP-059: Identification of Mouse Hematopoietic Stem and Progenitor Cells with Simultaneous Detection of CD45.1/2 and Controllable Green Fluorescent Protein Expression by a Single Staining Panel. *Cytometry A*. 2019;95(10):1049–1052.
6. Bernitz JM, Kim HS, MacArthur B, Sieburg H, Moore K. Hematopoietic Stem Cells Count and Remember Self-Renewal Divisions. *Cell*. 2016;167(5):1296–1309.e10.
7. Busch K, Klapproth K, Barile M, et al. Fundamental properties of unperturbed haematopoiesis from stem cells in vivo. *Nature*. 2015;518(7540):542–546.
8. Carrelha J, Meng Y, Kettyle LM, et al. Hierarchically related lineage-restricted fates of multipotent haematopoietic stem cells. *Nature*. 2018;554(7690):106–111.
9. Chapple RH, Tseng Y-J, Hu T, et al. Lineage tracing of murine adult hematopoietic stem cells reveals active contribution to steady-state hematopoiesis. *Blood Adv*. 2018;2(11):1220–1228.
10. Höfer T, Barile M, Flossdorf M. Stem-cell dynamics and lineage topology from in vivo fate mapping in the hematopoietic system. *Curr. Opin. Biotechnol*. 2016;39:150–156.
11. Pei W, Feyerabend TB, Rössler J, et al. Polylox barcoding reveals haematopoietic stem cell fates realized in vivo. *Nature*. 2017;548(7668):456–460.
12. Pei W, Shang F, Wang X, et al. Resolving fate and transcriptome of hematopoietic stem cell clones. *bioRxiv*. 2020;2020.03.25.008433.
13. Rodriguez-Fraticelli A, Wolock S, Weinreb C, et al. Clonal Analysis of Lineage Fate in Native Haematopoiesis. *Nature*. 2018;553(7687):.
14. Säwen P, Eldeeb M, Erlandsson E, et al. Murine HSCs contribute actively to native hematopoiesis but with reduced differentiation capacity upon aging. *Elife*. 2018;7:.
15. Sun J, Ramos A, Chapman B, et al. Clonal dynamics of native haematopoiesis. *Nature*. 2014;514(7522):322–327.
16. Yu VWC, Yusuf RZ, Oki T, et al. Epigenetic Memory Underlies Cell-Autonomous Heterogeneous Behavior of Hematopoietic Stem Cells. *Cell*. 2016;167(5):1310–1322.e17.
17. Laurenti E, Göttgens B. From haematopoietic stem cells to complex differentiation landscapes. *Nature*. 2018;553(7689):418–426.
18. Paul F, Arkin Y, Giladi A, et al. Transcriptional Heterogeneity and Lineage Commitment in Myeloid Progenitors. *Cell*. 2015;163(7):1663–1677.
19. Velten L, Haas SF, Raffel S, et al. Human haematopoietic stem cell lineage commitment is a continuous process. *Nat. Cell Biol*. 2017;19(4):271–281.
20. Crauste F, Pujo-Menjouet L, Génieys S, Molina C, Gandrillon O. Adding self-renewal in committed erythroid progenitors improves the biological relevance of a mathematical model of erythropoiesis. *J. Theor. Biol*. 2008;250(2):322–338.
21. Crauste F, Demin I, Gandrillon O, Volpert V. Mathematical study of feedback control roles and relevance in stress erythropoiesis. *J. Theor. Biol*. 2010;263(3):303–316.
22. Hara H, Ogawa M. Erythropoietic precursors in mice with phenylhydrazine-induced anemia. *Am. J. Hematol*. 1976;1(4):453–458.

23. Klinken SP, Holmes KL, Fredrickson TN, Erner SM, Morse HC. Phenylhydrazine stimulates lymphopoiesis and accelerates Abelson murine leukemia virus-induced pre-B cell lymphomas. *J. Immunol.* 1987;139(9):3091–3098.
24. Karamitros D, Stoilova B, Aboukhalil Z, et al. Single-cell analysis reveals the continuum of human lympho-myeloid progenitor cells. *Nat. Immunol.* 2018;19(1):85–97.
25. Ethier S, Kurtz T. Markov Processes: Characterization and Convergence. *Markov Processes: Characterization and Convergence.* 1986;
26. Boggs DR. The total marrow mass of the mouse: a simplified method of measurement. *Am. J. Hematol.* 1984;16(3):277–286.
27. Chervenick PA, Boggs DR, Marsh JC, Cartwright GE, Wintrobe MM. Quantitative studies of blood and bone marrow neutrophils in normal mice. *Am. J. Physiol.* 1968;215(2):353–360.
28. Mahajan MM, Cheng B, Beyer AI, et al. A quantitative assessment of the content of hematopoietic stem cells in mouse and human endosteal-bone marrow: a simple and rapid method for the isolation of mouse central bone marrow. *BMC Hematol.* 2015;15:9.
29. Sanchez M, Weissman IL, Pallavicini M, et al. Differential amplification of murine bipotent megakaryocytic/erythroid progenitor and precursor cells during recovery from acute and chronic erythroid stress. *Stem Cells.* 2006;24(2):337–348.
30. McKim DB, Yin W, Wang Y, et al. Social Stress Mobilizes Hematopoietic Stem Cells to Establish Persistent Splenic Myelopoiesis. *Cell Rep.* 2018;25(9):2552-2562.e3.
31. Haas S, Hansson J, Klimmeck D, et al. Inflammation-Induced Emergency Megakaryopoiesis Driven by Hematopoietic Stem Cell-like Megakaryocyte Progenitors. *Cell Stem Cell.* 2015;17(4):422–434.
32. Li H, Natarajan A, Ezike J, et al. Rate of Progression through a Continuum of Transit-Amplifying Progenitor Cell States Regulates Blood Cell Production. *Dev. Cell.* 2019;49(1):118-129.e7.
33. Mori Y, Chen JY, Pluvinaige JV, Seita J, Weissman IL. Prospective isolation of human erythroid lineage-committed progenitors. *PNAS.* 2015;112(31):9638–9643.
34. Tak T, Prevedello G, Simon G, et al. Simultaneous tracking of division and differentiation from individual hematopoietic stem and progenitor cells reveals within-family homogeneity despite population heterogeneity. *bioRxiv.* 2019;586354.
35. Abkowitz JL, Golinelli D, Harrison DE, Gutter P. In vivo kinetics of murine hemopoietic stem cells. *Blood.* 2000;96(10):3399–3405.
36. Foudi A, Hochedlinger K, Van Buren D, et al. Analysis of histone 2B-GFP retention reveals slowly cycling hematopoietic stem cells. *Nature Biotechnology.* 2009;27(1):84–90.
37. Klose M, Florian MC, Gerbaulet A, Geiger H, Glauche I. Hematopoietic Stem Cell Dynamics Are Regulated by Progenitor Demand: Lessons from a Quantitative Modeling Approach. *Stem Cells.* 2019;37(7):948–957.
38. Mackey MC. Cell kinetic status of haematopoietic stem cells. *Cell proliferation.* 2001;34(2):71–83.
39. Manesso E, Teles J, Bryder D, Peterson C. Dynamical modelling of haematopoiesis: an integrated view over the system in homeostasis and under perturbation. *J R Soc Interface.* 2013;10(80):20120817.
40. Sawai CM, Babovic S, Upadhaya S, et al. Hematopoietic Stem Cells Are the Major Source of Multilineage Hematopoiesis in Adult Animals. *Immunity.* 2016;45(3):597–609.
41. Angulo O, Crauste F, Lopez-Marcos JC. Numerical integration of an erythropoiesis model with explicit growth factor dynamics. *Journal of Computational and Applied Mathematics.* 2018;770.

42. Loeffler M, Pantel K, Wulff H, Wichmann HE. A mathematical model of erythropoiesis in mice and rats. Part 1: Structure of the model. *Cell Tissue Kinet.* 1989;22(1):13–30.
43. Marciniak-Czochra A, Stiehl T, Ho AD, Jäger W, Wagner W. Modeling of asymmetric cell division in hematopoietic stem cells--regulation of self-renewal is essential for efficient repopulation. *Stem Cells Dev.* 2009;18(3):377–385.
44. Roeder I, Kamminga LM, Braesel K, et al. Competitive clonal hematopoiesis in mouse chimeras explained by a stochastic model of stem cell organization. *Blood.* 2005;105(2):609–616.
45. Kirouac DC, Ito C, Csaszar E, et al. Dynamic interaction networks in a hierarchically organized tissue. *Molecular systems biology.* 2010;6(1):.
46. Mahadik B, Hannon B, Harley BA. A computational model of feedback-mediated hematopoietic stem cell differentiation in vitro. *PloS one.* 2019;14(3):.
47. Kang Y-A, Pietras EM, Passegué E. Deregulated Notch and Wnt signaling activates early-stage myeloid regeneration pathways in leukemia. *J Exp Med.* 2020;217(3):.
48. Perry JM, Harandi OF, Porayette P, et al. Maintenance of the BMP4-dependent stress erythropoiesis pathway in the murine spleen requires hedgehog signaling. *Blood.* 2009;113(4):911–918.

TABLES

<i>Cell Type</i>	<i>n (10⁶/mouse)</i>	<i>BrdU (%)</i>
<i>LT-HSC</i>	0.011+/-0.070	2.35+/-2.71
<i>ST-HSC</i>	0.022+/-0.013	0.90+/-0.75
<i>MPP</i>	0.066+/-0.045	6.32+/-3.58
<i>CMP</i>	0.196+/-0.111	12.42+/-2.31
<i>MEP</i>	0.358+/-0.189	31.75+/-1.15
<i>GMP</i>	0.677+/-0.430	19.50+/-7.50

Table 1: Determination of quantitative parameters in steady state hematopoiesis and

BrdU analysis. Each value is the mean of at least 6 experiments.

<i>Compartment</i>	1 LT-HSC	2 ST-HSC	3 MPP	4 CMP	5 MEP	6 RBC
<i>Division rate/ day</i>	$\tau_1^* = 0.01$	$\tau_2^* = 0.03$	$\tau_3^* = 0.07$	$\tau_4^* = 0.16$	$\tau_5^* = 0.35$	$\tau_6^* = 1/40$
<i>Cell number (10³/mouse)</i>	$x_1^* = 11$	$x_2^* = 22.5$	$x_3^* = 67$	$x_4^* = 196$	$x_5^* = 359$	$x_6^* = 5.17 \cdot 10^7$
<i>D (Differentiation factor)</i>	$D_1 = 0$	$D_2 = 0.16$	$D_3 = 0.167$	$D_4 = 0.174$	$D_5 = 0.176$	

Table 2: Determination of cell number, division rate and D factor for each compartment at steady state.

<i>Proliferation rate</i>	$r_1=9.2$	$r_2=6.4$	$r_3=6.4$	$r_4=2.8$	$r_5=0.33$
<i>Erythroid Differentiation rate</i>	$d_1=10$	$d_2=8.4$	$d_3=7.7$	$d_4=7.4$	$d_5=4.3$
<i>Regulator clearance (/day)</i>	$m_1=0.03$	$m_2=0.2$			
<i>Constants of the regulators production functions</i>	$a=10^{-4}$	$b=100$			
<i>Last mitosis number</i>	$c_1=1.18$	$c_2=31.3$			
<i>Steady state value of regulator u_1</i>	$u^*_1=1/(m_1(a+5))$				
<i>Steady state value of regulator u_2</i>	$u^*_2=5/(m_2(b+5))$				

Table 3: Proliferation rate, erythroid differentiation rate, regulators parameters values and mitosis number from MEP to mature RBC constants after stress.

<i>Cell type division rate value (/day)</i>	<i>LT-HSC</i>	<i>ST-HSC</i>	<i>MPP</i>	<i>CMP</i>	<i>MEP</i>
<i>Day 0</i>	$\tau_1^* = 0.01$	$\tau_2^* = 0.03$	$\tau_3^* = 0.07$	$\tau_4^* = 0.16$	$\tau_5^* = 0.35$
<i>Day 5</i>	0.37	0.24	0.57	0.28	0.21
<i>Day 11</i>	0.97	1.4	2.4	2.7	2
<i>Day 16</i>	0.12	0.14	0.34	0.33	0.48

Table 4: Division rates at different time points in the different progenitor compartments.

FIGURE LEGENDS

Figure 1 : Mathematical parameters required to Model steady state Erythropoiesis

(A) Scheme of the cell dynamics for type $i = 1$ to $k-2$. C_i : cell of i -type. τ_i : division rate of i -type cells. (B) Scheme of the two last cell dynamics. τ_{k-1} : division rate of $(k-1)$ -type cells. τ_k : death rate of RBC (k -type cells). (C) Differential system modeling the steady-state hematopoiesis. τ_i : division rate of i -type cells. τ_k : death rate of RBC (k -type cells). $dx_i/dt(t)$:

derivative of the function x_i at time t . 2^n : number of RBC (type k cells) created after one differentiation of a $(k-1)$ -type cell. **(D) Value of the unique steady state in the system Figure 1C depending on the differentiation factors $D_i = p_d^i - p_r^i$.** **(E) Number of RBC at equilibrium, for the differentiation factors D_i , $i=2, \dots, k-1$, equal to D and the probabilities μ_i equal to 1.** **(F) Estimation from BrdU data of the division rate at equilibrium for each cell type.** p_i^{BrdU} : proportion at steady-state of i -type cells BrdU+ 20 hours after BrdU incorporation. (see Figure S1 and supplemental information in supplemental data).

Figure 2: Analysis of s Blood, Spleen and bone marrow parameters after PHZ-mediated chemical hemolysis.

Mice were injected with PHZ (60 mg/kg). Blood, bone marrow and spleen were harvested at the indicated times. **(A) Red blood cell counts (10^9 /mouse)** (orange line, biological values, black line, computed values). **(B) White blood cells (10^6 /mouse)** and **(C) platelets (10^6 /mouse).** **(D) Spleen morphological changes**, Spleen photography at day 0,1,3 and 5. **(E) total spleen cell number per mouse (10^6 /mouse):**, and **(F) total spleen weight variations.** **(G) Comparison of Bone Marrow (blue line) and Spleen cell (black line) numbers and MEP numbers kinetics during time** (Bone marrow MEP: Green line, Spleen MEP: orange line).

Figure 3. Mathematical parameters required to Model stress Erythropoiesis

(A) Effects of regulators u_r and u_d on self-renewal and differentiation rates. $p_r^i(t)$: self-renewal probability at time t of i -type cells. $p_d^i(t)$: differentiation probability at time t of i -type cells. $u_r(t)$ (resp. $u_d(t)$) : self-renewal (resp. differentiation) regulation effect at time t . r_i :

sensitivity of the i -type cell to the regulator u_r (resp. d_i). **(B) Dynamics of the regulators u_r and u_d .** $du_r/dt(t)$: derivative of the function u_r at time t (similar notation for u_d). $u_r(t)$ (resp. $u_d(t)$) : self-renewal (resp. differentiation) regulation effect at time t . $a > 0$ (resp. $b > 0$): Michaelis-Menten's type constant of the regulator u_r (resp. u_d). m_r (resp. m_d) : clearance rate of the regulators u_r (resp. u_d). **(C). Regulation of the terminal erythropoiesis production (from MEP to RBC).** $2^{n(t)}$: number of RBC (k -type cells) created after one differentiation of a ($k-1$)-type cell at time t . $c_1 > 1$ and $c_2 > 0$: Hill constants of the regulated terminal mitosis of erythropoiesis.

τ_i^* : steady-state division rate of i -type cells. p_r^i : steady-state self-renewal probability of i -type cells. p_d^i : steady-state differentiation probability of i -type cells. μ_i : steady-state erythromyeloid commitment probability of the i -type cells. $x_i(t)$: number of i -type cells at time t . x_i^* : steady-state number of i -type cells. 2^{n^*} : steady-state number of RBC (type k cells) created after one differentiation of a ($k-1$)-type cell. k : number of compartments considered.

Figure 4: Experimental and model parameters from different hematopoietic parameters after PHZ-mediated chemical hemolysis. **(A) Bone marrow cell numbers per mouse** (10^6 /mouse) (blue line: Total BM cells, green line: MEP, dark blue line: ST-HSC, brown line: LT-HSC, orange line : MPP, grey line: CMP). **B. Evolution of the number of each type cells for the regulated model** (red curve) **compared to the data** (black point). The red points represent the number of cells at steady state and correspond to the x_i^* given in Table 2.(See supplemental informations).

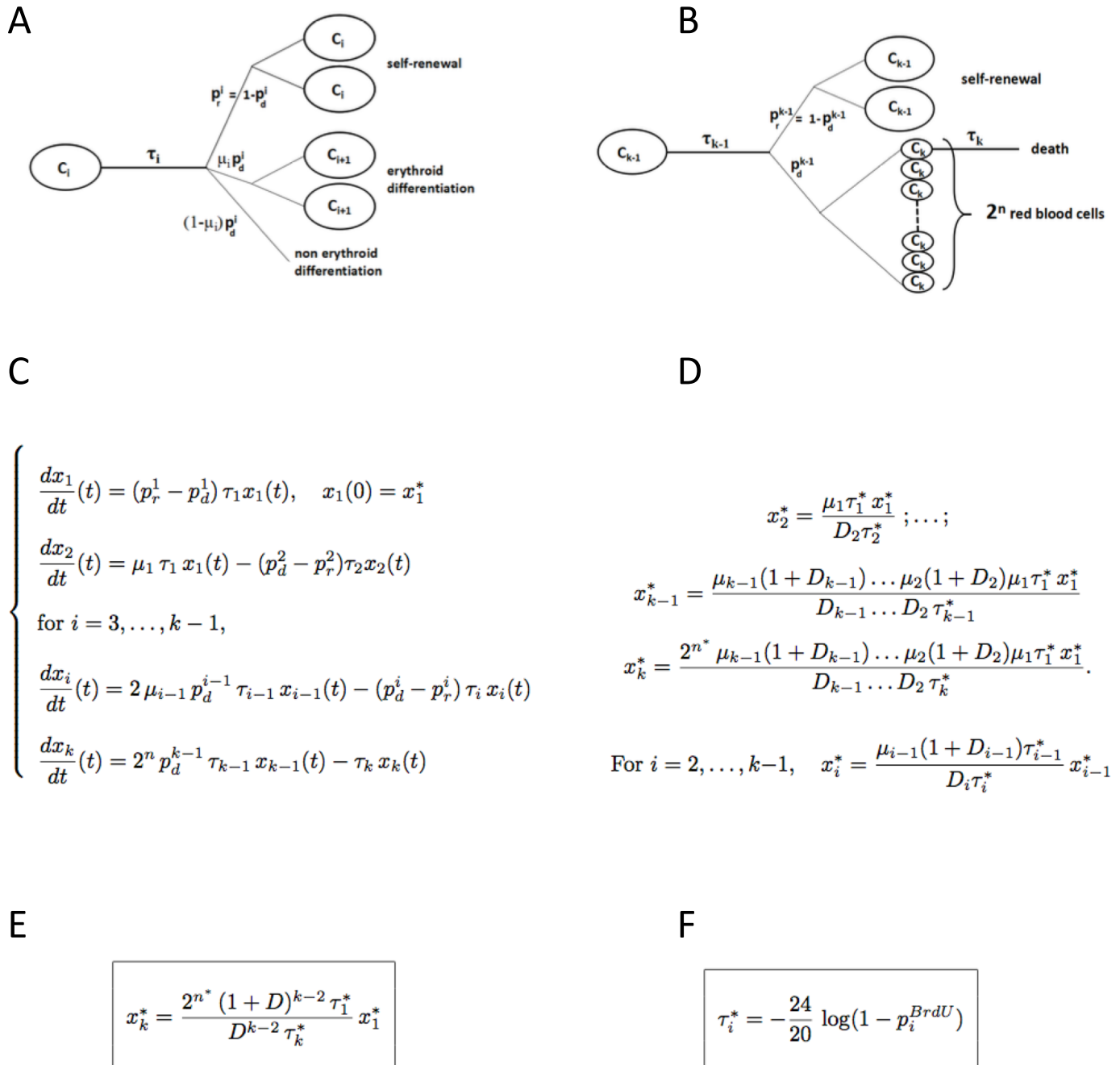


Figure 1, Bonnet et al.

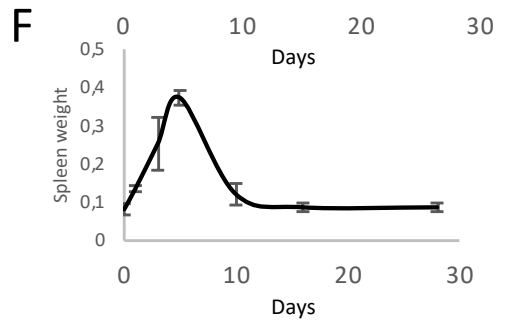
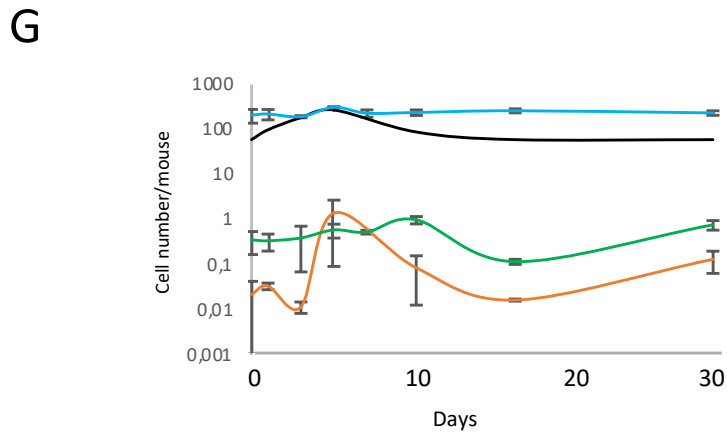
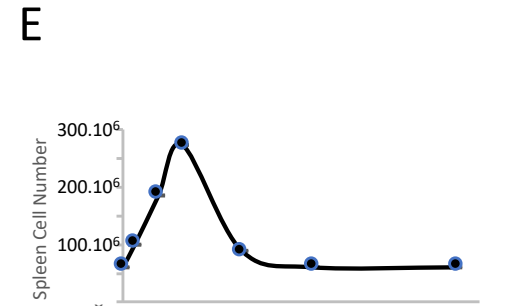
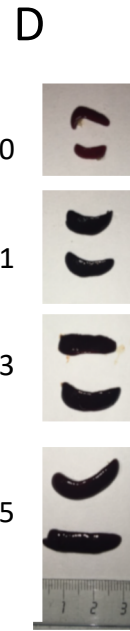
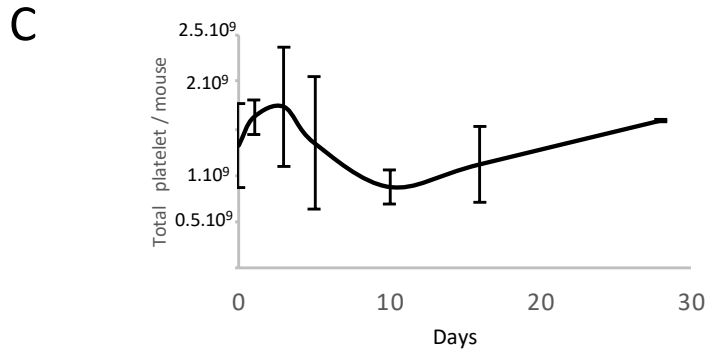
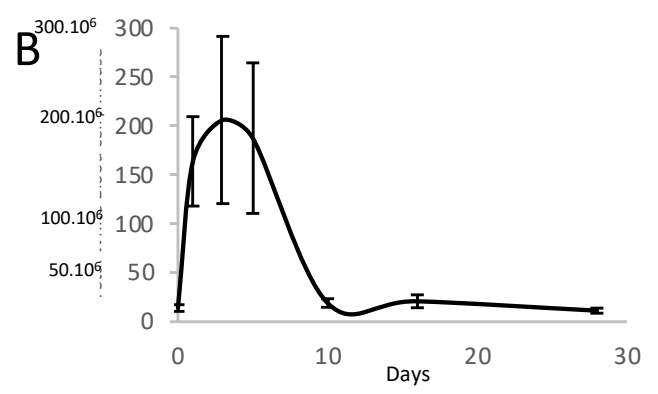
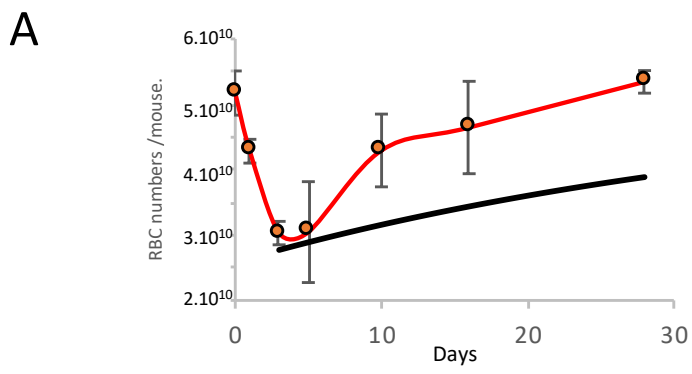


Figure 2, Bonnet et al.

A

$$\begin{cases} p_r^i(t) \tau_i(t) = p_r^i \tau_i^* \left(\frac{u_r(t)}{u_r^*} \right)^{r_i} \\ \mu_i(t) p_d^i(t) \tau_i(t) = \mu_i p_d^i \tau_i^* \left(\frac{u_d(t)}{u_d^*} \right)^{d_i} \end{cases}$$

where $(r_i)_i$, $(d_i)_i$, u_r^* and u_d^* are positive parameters.

C

$$n(t) = n^* \frac{c_1}{1 + (c_1 - 1) \left(\frac{x_6(t)}{x_6^*} \right)^{c_2}}.$$

B

$$\begin{cases} \frac{du_r(t)}{dt} = \frac{1}{a + y(t)} - m_r u_r(t) \\ \frac{du_d(t)}{dt} = \frac{y(t)}{b + y(t)} - m_d u_d(t) \end{cases}$$

with $y(t) = \frac{x_1(t)}{x_1^*} + \frac{x_2(t)}{x_2^*} + \frac{x_3(t)}{x_3^*} + \frac{x_4(t)}{x_4^*} + \frac{x_5(t)}{x_5^*}$

Figure 3. Bonnet et al.

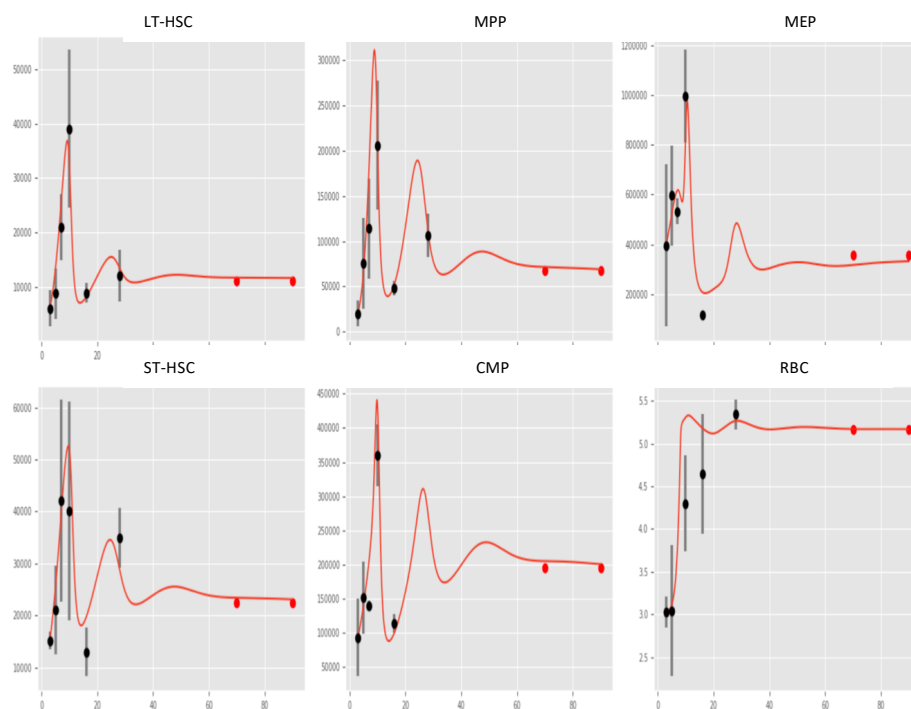
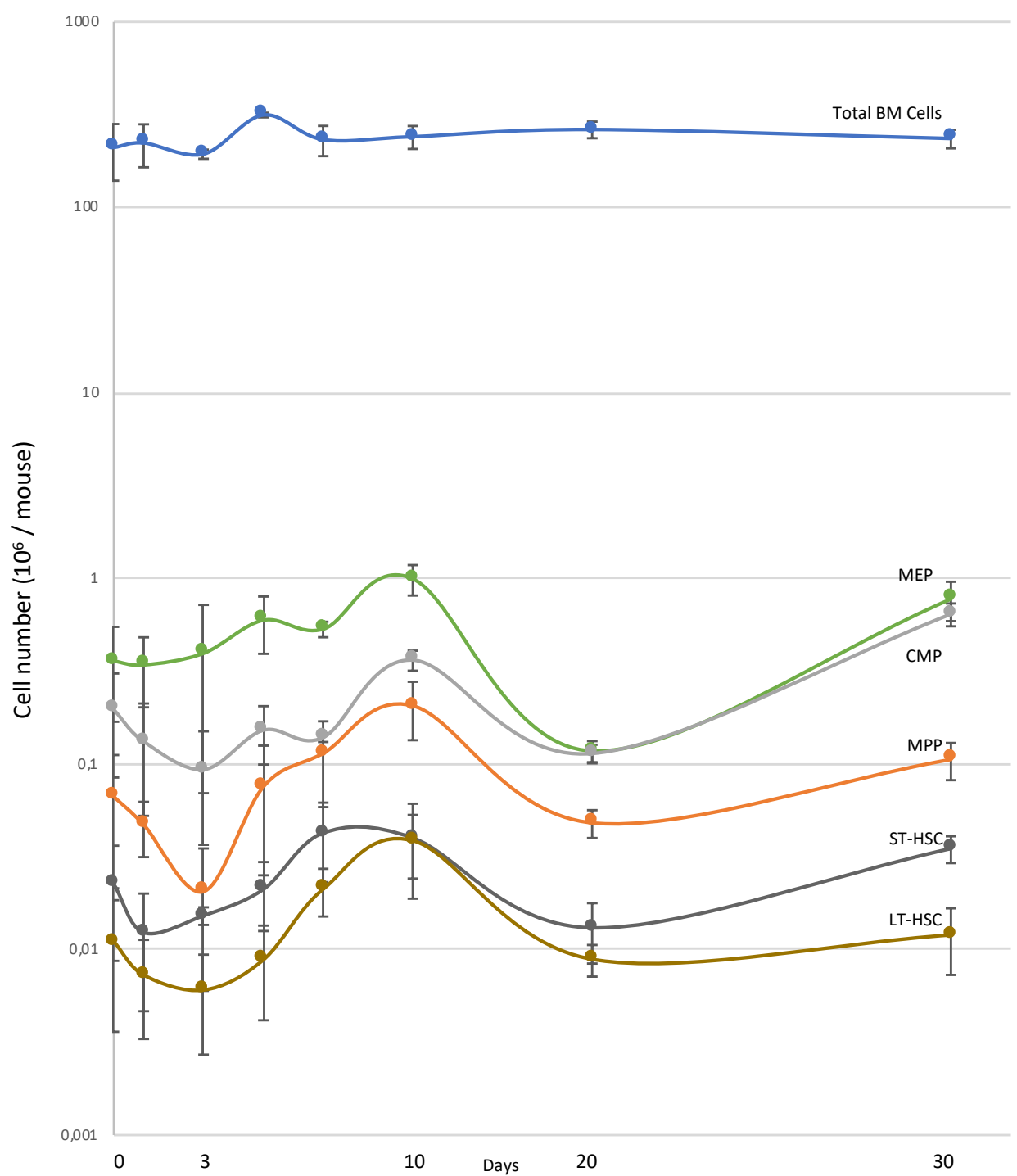


Figure 4, Bonnet et al.

SUPPLEMENTAL MATERIALS AND METHODS

Animals

We used adult 6 to 12 weeks old C57BL/6 Ly5.2 and Ly5.1 mice obtained from Janvier, (Le Genest, France). Mice were bred and maintained in pathogen-free conditions in our Institute Animal facilities Unit for Laboratory Animal Medicine at the Université de Paris, Centre Hayem, Saint Louis hospital). All procedures performed were approved by the local Committee on the Use and Care of Animals and by the Institutional Animal Care and Use Committee.

Hematologic evaluations

Orbital plexus blood was collected in EDTA tubes from the anesthetized mice. Nucleated blood cells, hematocrit, and platelet counts were determined using an automated blood coulter calibrated for mouse blood (MS9, Schloessing Melet, Cergy-Pontoise, France). Blood samples were also used for cytometry analysis and serum were collected, and stored at -20°C and used for determination of cytokines levels.

After experiments, mice were killed by cervical dislocation. Bones (femurs, tibias, pelvis and humerus) and spleen were harvested from mice, muscle and tendon tissue were removed using a scalpel and kimwipes. Femurs and tibias were used for cell count, and cytometry. The same protocols were used for spleens which were also weighted.

Peripheral Blood and BM Analysis

The bone marrow fraction was flushed out using a syringe containing 1x PBS complemented with 2% Fetal Bovine Serum, and spleen. The resulting cell suspension were filtered through a 40 µM cell strainer and pelleted by centrifugation.

Spleen cell Analysis

For mechanical grinding, spleens were smashed and ground between the rough sides of frosted glass slides, and cells were collected in DMEM containing 2% FBS. After incubation in a 24-well plate for 30 min at 37°C in a humidified incubator, cell suspensions were passed through a 100 µm Falcon nylon cell strainer (Corning, NY, USA) and cells were re-suspended in DMEM containing 10% FBS. Suspensions were subjected to cytometry analysis after red cell lysis.

Cytometry analysis

For flow cytometry analysis, erythrocytes were lysed using ACK Lysing Buffer (Lonza, Basel, Switzerland) before flow cytometry stainings. Total BM and spleen cells were stained with the Zombie Viability kit for 15 min at room temperature, and thereafter stained with the biotinylated anti-mouse Lineage Cocktail of antibodies [anti-CD4 (clone GK1.5), anti-CD8a (53-6.7), anti-CD11b (M1/70), anti- B220 (RA3-6B2), anti-GR1 (RB6-8C5) and anti-TER119 (Ter-119)] for 30 min (Biolegend). After washing, cells were stained for 30 min using the following monoclonal antibodies in the BD Horizon Brilliant staining buffer: anti-CD117/c-Kit (2B8)-BB700 and anti-CD34 (RAM34)-AF647 (BD Pharmingen); anti-Ly6a/Sca-1 (D7)-BV510; anti-CD150 (TC15-12F12.2)-PE-Cy7; anti-CD48 (HM48-1)-BV711 ; Streptavidin APC-Cy7 (Biolegend).

For proliferation analysis, surface stainings were performed as described above, and cells were proceeded for cell proliferation according to manufacturer's recommendations (BD Pharmingen BrdU Flow kits) with BrdU-AF488 (BD Biosciences, Franklin Lakes, NJ, USA). Data acquisition and/or data analysis were performed on the Cochin Cytometry and Immunobiology Facility Cytometry analyses on a Fortessa cytometer (Becton Dickinson), and analyses were done on Kaluza software.

To assess cell viability, surface labelled cells were resuspended in Annexin V binding buffer, and Annexin V (FITC) was added 15 min before analysis (FITC-conjugated Annexin V labeling detection kit, BD Pharmingen).

Reconstitution assays

Donor cells were isolated from 6-12-week-old B6.SJL-Ptprc^aPepc^b/BoyCrl Congenic mice (*Ly5.1*) and *Ly5.2* mice after Phenylhydrazine (PHZ) (60 mg/Kg, IP) injection or not. From these mice, hind limbs were extracted and cleaned. Total bone marrow were flushed from the bone cavity separately with ice cold PBS using a 21 G x 1.5 needle, then passed through a 70 μ M cell strainer to obtain a single cell suspension. A mix of cells from PHZ-treated mouse and untreated mouse was prepare in a 1:1 ratio. A mix of 5 x 10⁶ cells were resuspended in a total volume of 0.2 mL and transplanted into 5 irradiated (9 Gy) CD45.1 C57/BL6 recipient mice. Chimerism analysis were performed three months post reconstitution using cytometry.

Statistical analysis

The results are presented as mean $\pm$ SD. The data were analyzed with the 2-tailed Student t-test.

Reagent or Resource	Source	Identifier
Antibodies		
Biotin anti-mouse Lin Panel	Biolegend	133307
Ter-119, M1/70, RB6-8C5, 145-2C11, RA3-6B2		
APC/Cy7 Streptavidin	Biolegend	405208
BV 510 TM anti-mouse Ly-6A/E (Sca-1-) D7	Biolegend	108129
BV421 anti-mouse CD135 A2F10	Biolegend	135315
BV711 anti-mouse CD48 HM48-1	Biolegend	103439

PE/Cy7 anti-mouse CD150 TC15-12F12.2	Biolegend	115914
BB700 Anti-Mouse CD117 (PerCP) 2B8	BD Pharmingen	566414
BV786 Anti-Mouse CD16/CD32 2.4G2	BD Pharmingen	740851
Alexa Fluor® 647 anti-Mouse CD34 Ram34	BD Pharmingen	560230
Alexa Fluor 488 anti-BrdU Antibody	Biolegend	364105
FITC anti-mouse CD45.1 A20	Biolegend	110705
PE anti mouse CD45.2 104	Biolegend	109808

Chemicals, Peptides, and Recombinant Proteins

PHZ (Phenylhydrazine)	Sigma-Aldrich	P26252
-----------------------	---------------	--------

Critical Commercial Assays

Zombie UV fixable viability kit	Ozyme	BLE423107
BD Horizon Brilliant staining buffer	BD Pharmingen	563794
Ultracomp ebeads	Life Technologies	01-2222-42
Rainbow Calibration Particles (8 - PEaks), 3.0-3.4 mm SPHERO Particles	BD Pharmingen	559123
BD Pharmingen BrdU Flow Kit	BD Pharmingen	557892
FITC Annexin V Apoptosis Detection Kit	BD Pharmingen	559763
ACK Lysing Buffer, Quality Biological	VWR	10128-802

Experimental Models: Organisms/Strains

C57BL/6J	Janvier (Le Genest, France)
B6.SJL-Ptprc ^a Pepc ^b /BoyCrl Congenic (Ly5.1 mice)	Janvier (Le Genest, France)

Software and Algorithms

Diva software	Beckman
Kakuza software	Beckman
Covariance matrix adaptative in Evolution algorithm (cmaes)	Python language

SUPPLEMENTAL FIGURES

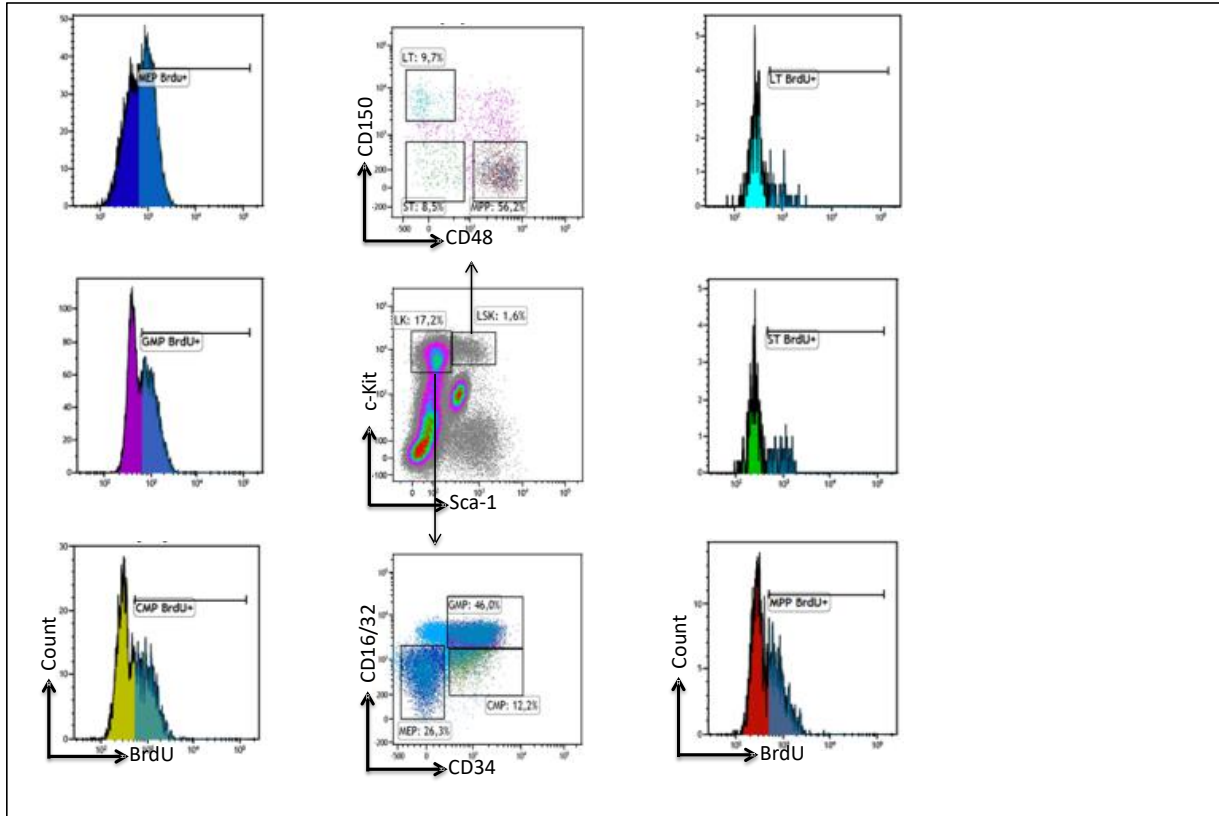


Figure S1: Single staining 9-parameter analysis of mouse bone marrow cells. After lysis of the erythrocytes, cellular debris or doublets were discriminated from single cells by analyzing the side scatter area and width signals. Single cells were then stained with the life/dead discriminator (Zombie), and alive cells were then plotted for lineage markers (CD3e, CD11b, B220, Gr1, and Ter119). Lineage negative cells were analyzed for cKit and Sca1 to define the LSK and LS-K cell populations. LS-K cells were discriminated by expression of CD16/32 and CD34 to identify GMP, CMP and MEP populations. LSK cells were discriminated by expression of CD150 and CD48 to identify LT-HSC, ST-HSC and MPP populations. BrdU incorporation was assessed in all these populations.

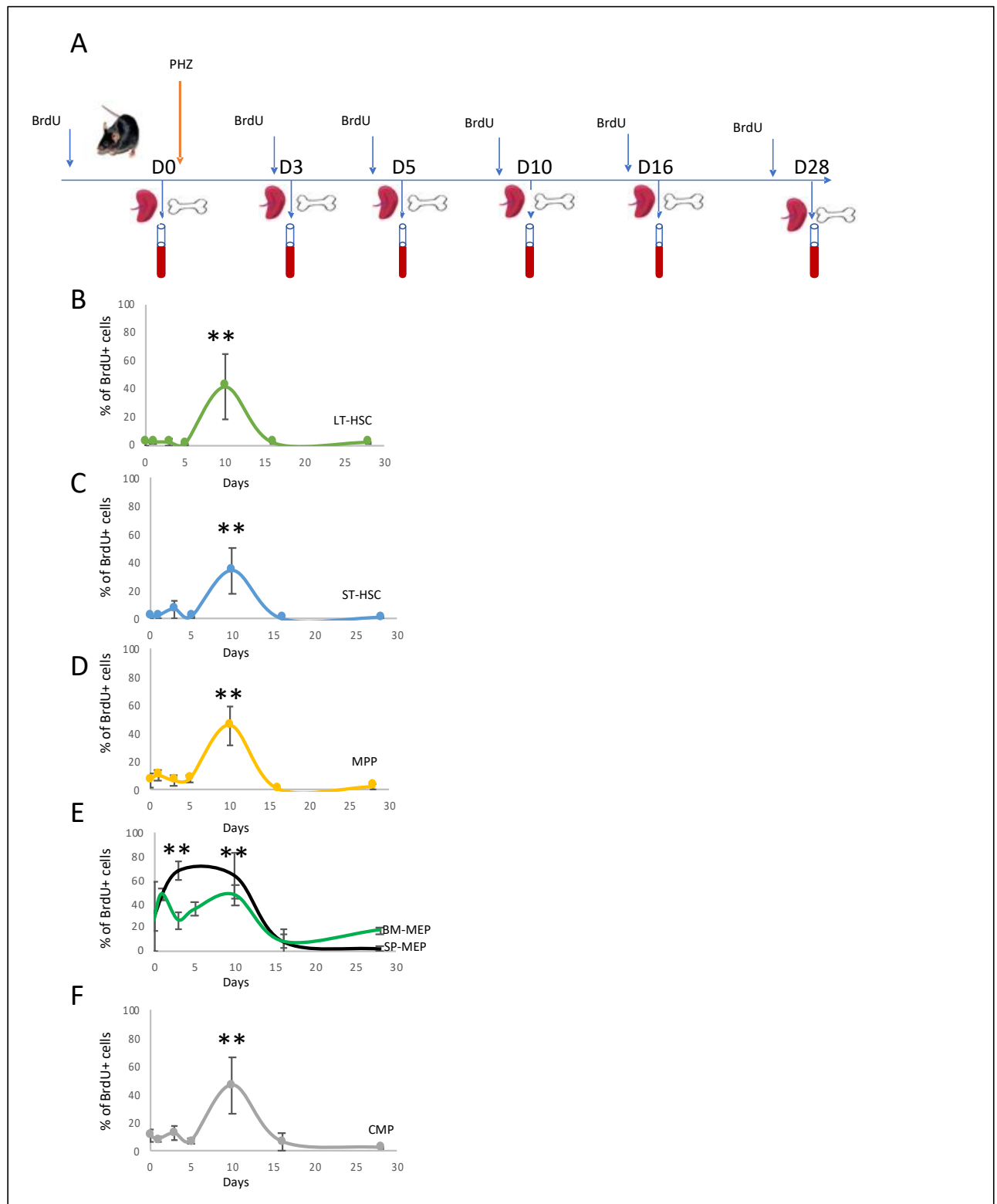


Figure S2: BrdU modifications during stress. Summary scheme of PHZ treatment in mouse. *In vivo* cell cycle analysis of bone marrow cells after chemical hemolysis: Percentage of BrdU⁺ cells 16h after intraperitoneal BrdU injection at different time points after PHZ

administration.C: BM LT-HSC, D: BM ST-HSC, E: BM MPP, F: MEP from BM (red)and
spleen (purple)and G: CMP

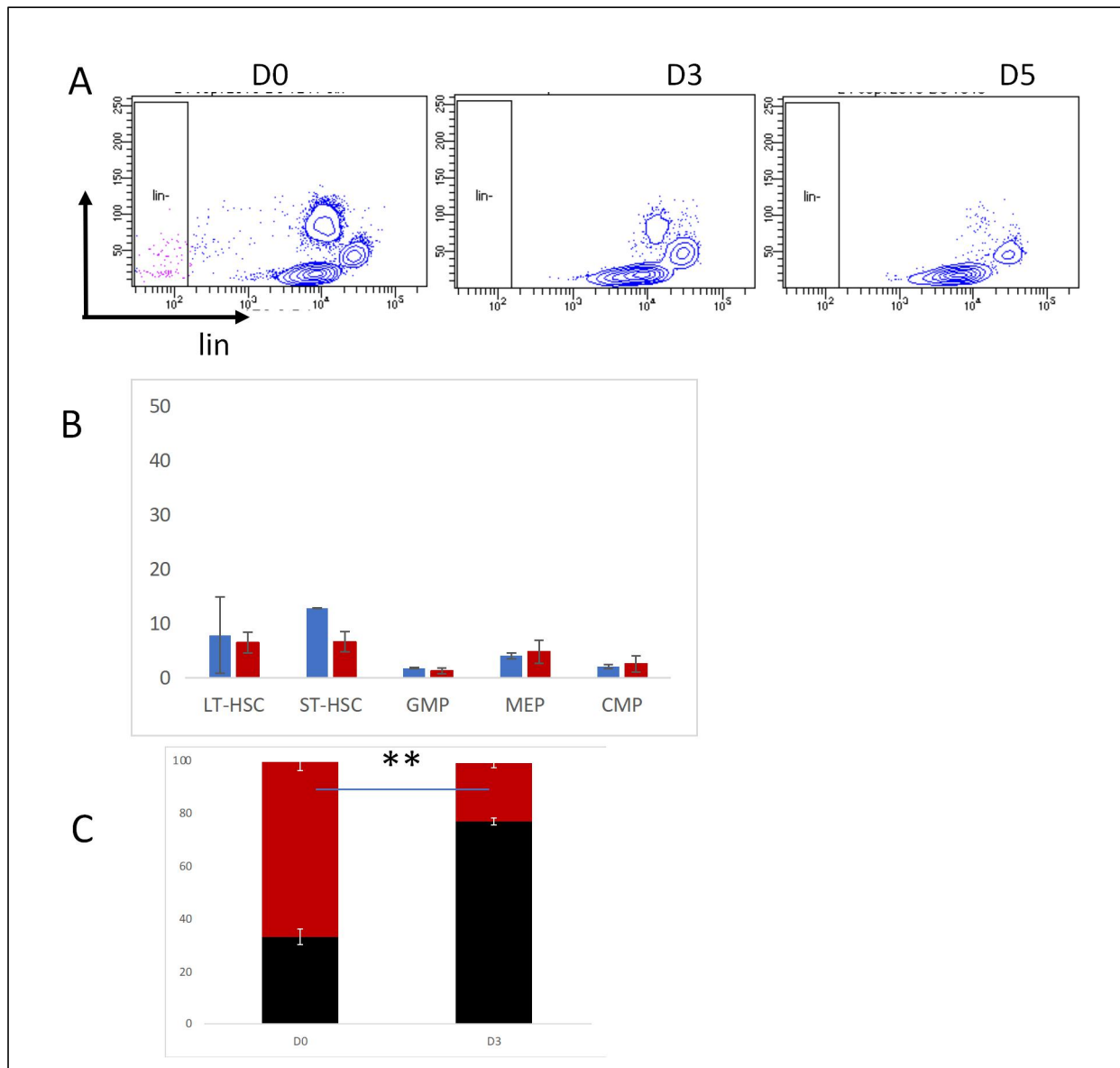


Figure S3: Analysis of peripheral blood, bone marrow and spleen after PHZ administration. (A) Cytometric analysis to detect the presence of lin⁻ cells in the peripheral blood at d 0, 1, 3 and 5 after PHZ administration. (B) Cytometric analysis of apoptosis (Annexin positive cells) in the different progenitor compartments at day 3 after PHZ treatment. (C) Three days after PHZ administration, lethally irradiated C57BL/6-Ly5.1 mice received 10^6 BM cells from PHZ-treated C57BL/6 (Ly5.2) mice, in competition with the same amount of BM cells from Ly5.1 mice. Presence of CD45.2⁺ PB cells of engrafted mice was assessed 3 months later by flow cytometry. Data are expressed as mean $\pm$ SEM. *P < .05; **P < .01; ***P < .001 (Student t test).

SUPPLEMENTAL DATA

1- Explanation of Figure 1F.

Table 1 contains data on the proportion at steady-state (day 0) of cells of each type that made at least one division during 20 hours. Here, we will relate these informations to the steady-state division rates of each cell type: τ_i^* (Figure 1F).

Before BrdU injection, the system is at equilibrium and all cells are unlabeled by BrdU. The numbers of BrdU labeled and unlabeled i-type cells at time t of the experiment are denoted by $N_i^m(t)$ and $N_i(t)$. Then $N_i^m(0)=0$ and $N_i(0)=x_i^*$.

During the whole duration of the experiment, when an unlabeled cell of type i divides, it becomes BrdU-labeled. Thus the dynamics of the unlabeled i-type cells number is expressed as an exponential decrease at rate τ_i^* . Throughout the experiment, the system is considered to be at steady state. Therefore, for any $i = 1, \dots, k-1$, and for any $t \in [0, 24h]$, $N_i^m(t) + N_i(t) = x_i^*$.

Finally, we can deduce the formula in Figure 1F.

2- Calibration of the regulated-erythropoiesis model

We calibrated the parameters of the regulated-model from the data of the 6 population sizes dynamics after PHZ induced hemolysis. In order to model the return to the equilibrium (given in Figure 4B), we considered day 3 as the initial time. Hence initial conditions of the model are given by :

$x_1(\text{day } 3) = 0.006 \cdot 10^6 \text{ cells/mouse}$, $x_2(\text{day } 3) = 0.0152 \cdot 10^6 \text{ cells/mouse}$,

$x_3(\text{day } 3) = 0.02 \cdot 10^6 \text{ cells/mouse}$, $x_4(\text{day } 3) = 0.093 \cdot 10^6 \text{ cells/mouse}$,

$x_5(\text{day } 3) = 0.395 \cdot 10^6 \text{ cells/mouse}$, $x_6(\text{day } 3) = 0.94 \cdot 10^{10} \text{ cells/mouse}$.

Using a stochastic optimization algorithm based on a CMA-ES method (Covariance Matrix Adaptation Matrix in Evolution Strategy), developed on Python by the INRIA team RandOpt [Hansen 2001] and least-squares cost function, we minimized the discrepancy between *in silico* simulated and experimental data. Results were summarized in Table 4.

Chapitre 2

Large fluctuations in multi-scale modeling for rest erythropoiesis

This chapter corresponds to paper [12] "Large fluctuations in multi-scale modeling for rest erythropoiesis" in collaboration with Sylvie Méléard and submitted to *Journal of Mathematical Biology*.

Abstract Erythropoiesis is a mechanism for the production of red blood cells by cellular differentiation. It is based on amplification steps due to an interplay between renewal and differentiation in the successive cell compartments from stem cells to red blood cells. We will study this mechanism with a stochastic point of view to explain unexpected fluctuations on the red blood cell numbers, as surprisingly observed by biologists and medical doctors in a rest erythropoiesis. We consider three compartments : stem cells, progenitors and red blood cells. The dynamics of each cell type is characterized by its division rate and by the renewal and differentiation probabilities at each division event. We model the global population dynamics by a three-dimensional stochastic decomposable branching process. We show that the amplification mechanism is given by the inverse of the small difference between the differentiation and renewal probabilities. Introducing a parameter K which scales simultaneously the size of the first component, the differentiation and renewal probabilities and the red blood cell death rate, we describe the asymptotic behavior of the process for large K . We show that each compartment has its own size scale and its own time scale. Focussing on the third component, we prove that the red blood cell population size, conveniently renormalized (in time and size), is expanded in an usual way inducing large fluctuations. The proofs are based on a fine study of the different scales involved in the model and on the use of different convergence and average techniques in the proofs.

Keywords Decomposable branching process . Multi-scale approximation . Stochastic slow-fast dynamical system . Large fluctuations . Rest Erythropoiesis . Amplification mechanism

Mathematics Subject Classification (2010) MSC 60F05 . MSC 92D25

2.1 Introduction

The model and the stochastic behavior we are studying in this paper are based on the biological mechanisms of rest (without stress) erythropoiesis. Erythropoiesis is a mechanism

for the production of red blood cells by cellular differentiation of stem cells. Stem cells, although in large numbers, produce even more red blood cells per day using a specific amplification mechanism.

We will study this amplification mechanism using a decomposable branching process (see [44]-Sect. 12, [4]-Sect. 6.9.1, [93]). Such process allows in particular to capture the genealogy of the cells, including the history of their types.

Let us firstly describe more precisely the biological dynamics, then we will introduce the mathematical model. The dynamics of erythropoietic cells, at rest, results in two distinct events, renewal and differentiation. Indeed, each cell of each type (except the last one) divides into two cells at a constant rate, depending on its type. These two new cells are either of the same type as the mother cell (renewal) or of the "next" cell type (differentiation). The final stage of differentiation corresponds to red blood cells which don't divide and can only die at a constant rate. The stem cells are those with the highest capacity for renewal, but not so high to prevent the cell population to explode. Further, the amplification from one compartment (characterized by one type) to the next one is proportional to the inverse of the small difference between the differentiation and renewal probabilities. Note also that the death rate in the last compartment plays a main role.

We are interested in describing the stochastic fluctuations of the compartment sizes for the rest erythropoiesis. In this case, the regulation doesn't play any role but nevertheless one observes unusual large oscillations at the red blood cell level. Indeed, the red blood cells number, in a human rest erythropoiesis, varies by 10% around its average value (cf. [92]). The order of magnitude of these variations is greater than the one of the classical variations for multi-type branching processes, which should be of the order 0.001%.

In this paper, we will model the differentiation steps by considering 3 types. These types correspond to stem cells (type 1), progenitors (type 2) with the ability in amplifying the cells number, and red blood cells (type 3). The number of stem cells in the initial state will be characterized by a (large) scaling parameter $K \in \mathbb{N}^*$.

Let us now introduce more precisely the parameters of the dynamics.

Cells of type 1 evolve according a critical linear birth and death process. Birth events correspond to renewal division events, occurring at rate $\frac{\tau_1}{2} > 0$, while death events correspond to differentiation events occurring at the same rate (a cell of type 1 divides in two cells of type 2). Cells of type 2 divide at rate $\tau_2 > 0$ in two cells of the same type (renewal event) with probability p_2^R and in two cells of type 3 (differentiation event) with probability $p_2^D = 1 - p_2^R \in]1/2, 1[$. The cells of type 3 are mature cells which die at rate $d_3 > 0$. We can summarize the dynamics as follows. If (N_1, N_2, N_3) denotes the vector of sub-population sizes, the transitions of the hematopoietic process are given by

$$\begin{array}{lll}
N_1 & \longrightarrow & N_1 + 1 & \text{at rate } (\tau_1/2) N_1 \\
(N_1, N_2) & \longrightarrow & (N_1 - 1, N_2 + 2) & \text{at rate } (\tau_1/2) N_1 \\
N_2 & \longrightarrow & N_2 + 1 & \text{at rate } \tau_2 p_2^R N_2 \\
(N_2, N_3) & \longrightarrow & (N_2 - 1, N_3 + 2) & \text{at rate } \tau_2 p_2^D N_2 \\
N_3 & \longrightarrow & N_3 - 1 & \text{at rate } d_3 N_3.
\end{array}$$

Here, we have assumed that each division is symmetric, so that

$$p_2^D + p_2^R = 1. \tag{2.1}$$

We could have included asymmetric division without changing the results of our study. Indeed it doesn't change the main characteristics of the dynamics.

As explained above, the number of cells of each type is large, but moreover, there is an amplification mechanism between the compartments, based on the small difference $p_2^D - p_2^R$ between the differentiation and renewal probabilities in compartment 2, and on the small death rate d_3 , in a way which is now defined.

We assume that

- the size of the type 1-cells population is of order K ,
- there exists a couple of positive parameters $(\gamma_2, \gamma_3) \in]0, 1[$ such that

$$p_2^D - p_2^R = K^{-\gamma_2} \quad \text{and} \quad d_3 = \tau_3 K^{-\gamma_3} \quad \text{with } \tau_3 > 0. \quad (2.2)$$

Let us note that (2.1) and (2.2) make the probabilities p_2^R and p_2^D depend on K ,

$$p_2^D = 1 - p_2^R = 1/2 + K^{-\gamma_2}/2.$$

Therefore the dynamics in this compartment is close to a critical process.

Assumptions (2.2) introduce the different time and size scales playing a role for the multi-scale population process describing the dynamics of each compartment size. Hence we will denote by N^K , the population process N previously defined.

We assume in the following that

$$\gamma_2 < \gamma_3 < 1. \quad (2.3)$$

This case is the most interesting mathematically and closest to the biological observations. Indeed, in a more realistic model with a larger number of compartments based on biological observations (see [12]), we observe that the red blood cell death rate drives the slowest time scale.

Our aim in this paper is to finely describe this dynamics, when K goes to infinity, using appropriate renormalizations.

We will see that a size renormalization is not enough to describe the dynamics of the last two components of the process. A time renormalization is also necessary. More precisely each compartment has its own size scale, of order K for Compartment 1, $K^{1+\gamma_2}$ for Compartment 2 (resp. $K^{1+\gamma_2+\gamma_3}$ for Compartment 3) and its own time scale, of order 1 for Compartment 1, K^{γ_2} for Compartment 2 (resp. K^{γ_3} for Compartment 3).

The next simulations show the dynamics of the process in the typical time scale of each compartment, namely K , K^{γ_2} and K^{γ_3} . We take as initial condition

$$N^K(0) = (K, 0, 0)$$

and choose $K = 2000$, $\gamma_2 = 0.55$, $\gamma_3 = 0.8$. Hence $K^{\gamma_2} \sim 60$ and $K^{\gamma_3} \sim 400$.

The others parameters are equal to 1.

Figure 2.1 shows the simulation of a trajectory of the process $(N^K(t), t \in [0, T])$ for $T \sim 1$, decomposed on the three compartments. Figure 2.2 shows the simulation of a trajectory of the process $(N^K(t), t \in [0, T])$ for $T \sim K^{\gamma_2}$ and Figure 2.3 shows the simulation of a trajectory of the process $(N^K(t), t \in [0, T])$ for $T \sim K^{\gamma_3}$. The horizontal

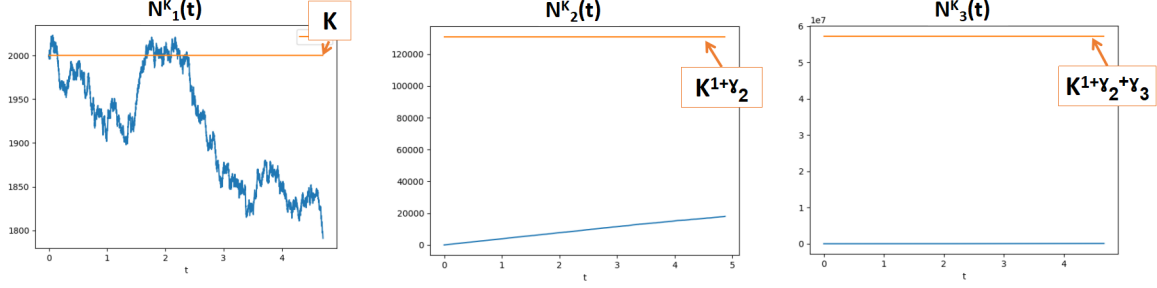


FIGURE 2.1 – A trajectory of the N^K process for $t \in [0, T]$ with $T = O(1)$

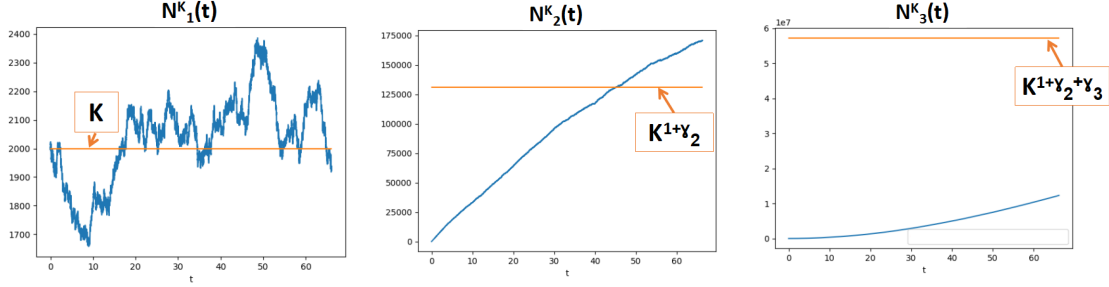


FIGURE 2.2 – A trajectory of the N^K process for $t \in [0, T]$ with $T = O(K^{\gamma_2})$

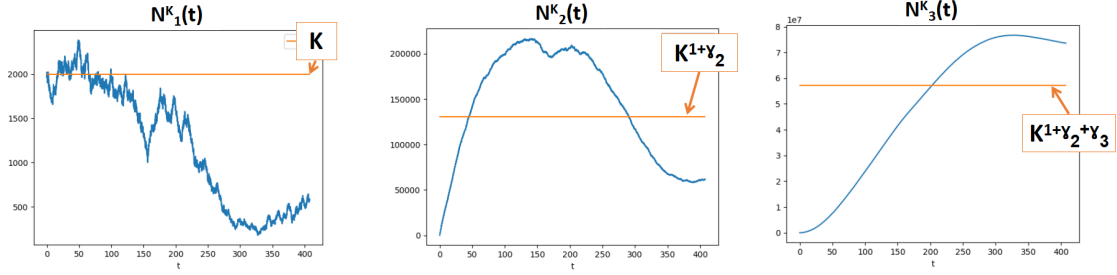


FIGURE 2.3 – A trajectory of the N^K process for $t \in [0, T]$ with $T = O(K^{\gamma_3})$

orange line gives the order of magnitude for each compartment size (K , resp. $K^{1+\gamma_2}$, $K^{1+\gamma_2+\gamma_3}$).

We observe that at a time scale of order 1, the two last components of the process N^K are far from their equilibrium size. We observe in Figure 2 that the two first components of $(N^K(t), t \in [0, T])$ for $T \sim K^{\gamma_2}$ evolve around their equilibrium size, which is not the case of the third one. In Figure 3, the process is considered on a longer period of time, $T \sim K^{\gamma_3}$ and one sees that the third component hits a neighborhood of its equilibrium. Furthermore, we observe the oscillations of the components of N^K around their equilibrium. We note that they are smoother and smoother from Compartment 1 to Compartment 3 and that the amplitude of the waves is longer and longer.

Compartment 3 illustrates the particular behavior of the red blood cells in a rest human erythropoiesis, highlighted above. Indeed, the expected variations should be around $K^{-(1+\gamma_2+\gamma_3)/2} \sim 0.01\%$, and we observe variations which seem to be of order $\frac{10^7}{K^{(1+\gamma_2+\gamma_3)/2}} \sim 17\%$.

Our aim is to prove and quantify these different behaviors and to explain these large fluctuations.

In Sect. 2.2, basic martingale properties are stated, some estimates are given for the moments of compartments sizes and a first study on the convergence of the process at a time-scale of order 1, when K tends to infinity, is given. We show that the two last components do not reach their equilibrium at this time scale. In Sect. 2.3, we study the process on an appropriate time-scale to capture the asymptotic behavior of the second and third types. We show that the limiting behavior of the two first components process at the time scale of order K^{γ_2} is given by an explicit deterministic function y . We also show that this time scale is not long enough to observe the dynamics of the third component. Hence, we study the limiting behavior of the process at the time scale of order K^{γ_3} . At this time scale, the second component process goes too fast and doesn't converge anymore. We consider the associated occupation measure, as already done in [62]. We prove its convergence in a weak sense to the Dirac measure at the unique equilibrium of the second component of the deterministic function y . Then we deduce the convergence of the third component to a limiting deterministic system involving this equilibrium. In Sect. 2.4, we study and quantify the large fluctuations observed in the simulations. We show that the first component behaves at the different time scales as a Brownian motion. This is not the case for the other two components. Theorem 3 describes the second and third order asymptotics of the second component on its typical time and size scale. The fluctuations around its deterministic limit are not Gaussian. They are described by a finite variation process integrating the randomness of the first component. An independent Brownian motion appears in the third order term. Theorem 4, which is the main theorem of the paper, describes the fluctuations associated with the third component dynamics. We show that the randomness induced by the dynamics of the two last components is negligible. To capture the effect of the randomness of the first component imposes a size-scale which allows to observe the large oscillations of the third component. We identify these oscillations as a finite variation process integrating as above the fluctuations of the first component.

The mathematical modeling of hematopoiesis has been firstly introduced in the seminal paper of Till, McCulloch and Siminovitch [93]. In this paper, the authors study a binary branching process and show, comparing with biological results, that the probabilistic framework is relevant. Since this pioneering work, many mathematical approaches have been proposed to describe more precisely the cell differentiation kinetics, based either on deterministic or stochastic models (a survey concerning many models can be found in [97]). A deterministic approach consists in introducing a dynamical system describing the behavior of the different compartments and in studying different properties of this system, in particular the equilibrium states (see for example [19], [71], [75], [3] and the references therein). One can also add a noise to model some random perturbation of these systems, with an eventual delay (see for example [65], [81]). In [29], a continuous description of the different cells types is also proposed, using a partial differential equation. Moreover, in all these papers, the authors are interested in modeling the regulation which happens when the system is perturbed by some stress and this nonlinearity involves many mathematical difficulties. Let us note that other stochastic models for hematopoiesis have been introduced ([1], [86], [57]) but they concentrate on a specific level (either stem cells or red blood cells). In all this literature, the questions studied by the authors don't concern the impact of the parameters on the amplification mechanism. We have found only two papers, [27] and [75], in which the question is mentioned. To our knowledge, the fluctuations generated by this amplification mechanism, have never been rigorously studied with such a space-time multi-scale point of view.

Most of the slow fast dynamical systems model interaction between species with different behaviors driving the time scales (see for example [56]). In such cases, slow and fast components appear naturally, contrary to our case, for which a fine study is needed to find the specific time scale of each compartment. In the other way, the authors in [54] have developed a general theorem for convergence and fluctuations of multi-scale processes. Their result can't explain our asymptotics. Indeed, in their result, the fluctuations around the deterministic behavior of the slow component are Gaussian, which is not the case of the red blood cells dynamics previously described. In their work, the martingale part of the limit is due to two sources of randomness : the slow component dynamics and the averaged effects of the fast components on the slow component dynamics. As previously explained, in our case the randomness of the slow component is only due to the fast ones and its intrinsic randomness is negligible.

Notation. $\mathcal{P}(E)$ and $\mathcal{L}(X)$ will denote respectively the space of probability measures on E and the law of a process X . As in [62], we will denote by $l_m(\mathbb{R}_+)$ the space of measures on $[0, \infty) \times \mathbb{R}_+$ such that $\mu([0, t] \times \mathbb{R}_+) = t$, for each $t \geq 0$.

2.2 Amplification mechanism : size-scale dynamics

2.2.1 The amplification mechanism

As explained in the introduction, we will reduce the model by simplicity to three compartments. The first one will describe the stem cells compartment, the second one will describe the compartment of progenitors and the third one will refer to red blood cells. Also by simplicity, we will describe the type of cells in each compartment by type 1, type 2 and type 3.

Let us now introduce the vector $N^K(t) = (N_1^K(t), N_2^K(t), N_3^K(t))$ of population sizes at time t . The process N^K is a decomposable multi-type branching process, that is a Markov jump process whose dynamics is given by the following equations.

We assume that for any fixed K , $N_1^K(0), N_2^K(0), N_3^K(0)$ are integrable.

Let us denote by $(\mathcal{N}_i^j)_{\substack{1 \leq i \leq 3 \\ j \in \{-, +\}}}$ independent Poisson point measures with intensity $ds du$ on $\mathbb{R}_+^2$ and introduce the filtration $(\mathcal{F}_t)_{t \geq 0}$ given by

$$\mathcal{F}_t = \sigma(\mathcal{N}_i^j([0, s] \times A); i \in \{1, \dots, 3\}, j \in \{-, +\}, s \leq t, A \in \mathcal{B}(\mathbb{R})).$$

Then we have

$$\begin{aligned} N_1^K(t) &= N_1^K(0) + \int_0^t \int_{\mathbb{R}_+} \mathbf{1}_{\{u \leq \tau_1 N_1^K(s^-)\}} \mathcal{N}_1^+(ds, du) - \int_0^t \int_{\mathbb{R}_+} \mathbf{1}_{\{u \leq \tau_1 N_1^K(s^-)\}} \mathcal{N}_1^-(ds, du) \\ N_2^K(t) &= N_2^K(0) + 2 \int_0^t \int_{\mathbb{R}_+} \mathbf{1}_{\{u \leq \tau_2 N_1^K(s^-)\}} \mathcal{N}_1^-(ds, du) + \int_0^t \int_{\mathbb{R}_+} \mathbf{1}_{\{u \leq \tau_2 p_2^R N_2^K(s^-)\}} \mathcal{N}_2^+(ds, du) \\ &\quad - \int_0^t \int_{\mathbb{R}_+} \mathbf{1}_{\{u \leq \tau_2 p_2^D N_2^K(s^-)\}} \mathcal{N}_2^-(ds, du) \\ N_3^K(t) &= N_3^K(0) + 2 \int_0^t \int_{\mathbb{R}_+} \mathbf{1}_{\{u \leq \tau_2 p_2^D N_2^K(s^-)\}} \mathcal{N}_2^-(ds, du) - \int_0^t \int_{\mathbb{R}_+} \mathbf{1}_{\{u \leq \tau_3 K^{-\gamma_3} N_3^K(s^-)\}} \mathcal{N}_3^-(ds, du) \end{aligned}$$

It can be written as

$$\begin{aligned}
\forall t \geq 0, \quad N_1^K(t) &= N_1^K(0) + M_1^K(t) \\
N_2^K(t) &= N_2^K(0) + \tau_1 \int_0^t N_1^K(s) ds - \tau_2 K^{-\gamma_2} \int_0^t N_2^K(s) ds + M_2^K(t) \\
N_3^K(t) &= N_3^K(0) + 2\tau_2 p_2^D \int_0^t N_2^K(s) ds - \tau_3 K^{-\gamma_3} \int_0^t N_3^K(s) ds + M_3^K(t)
\end{aligned} \tag{2.4}$$

where $M^K = (M_1^K, M_2^K, M_3^K)$ is a square-integrable martingale such that for all $t \geq 0$,

$$\begin{aligned}
\langle M_1^K \rangle_t &= \tau_1 \int_0^t N_1^K(s) ds \\
\langle M_2^K \rangle_t &= 2\tau_1 \int_0^t N_1^K(s) ds + \tau_2 \int_0^t N_2^K(s) ds \\
\langle M_3^K \rangle_t &= 4p_2^D \tau_2 \int_0^t N_2^K(s) ds + \tau_3 K^{-\gamma_3} \int_0^t N_3^K(s) ds \\
\langle M_1^K, M_2^K \rangle_t &= -\tau_1 \int_0^t N_1^K(s) ds \\
\langle M_2^K, M_3^K \rangle_t &= -2p_2^D \tau_2 \int_0^t N_2^K(s) ds.
\end{aligned} \tag{2.5}$$

Indeed, by standard localization and Gronwall's arguments applied to $(N_1^K(t))_t$, we can easily prove that for any $T > 0$ and $K \in \mathbb{N}^*$,

$$\mathbb{E} \left[\sup_{t \leq T} N_1^K(t) \right] \leq (2 + \mathbb{E}[N_1^K(0)]) e^{2\tau_1 T},$$

and then that

$$\mathbb{E} \left[\sup_{t \leq T} N_2^K(t) \right] < +\infty; \quad \mathbb{E} \left[\sup_{t \leq T} N_3^K(t) \right] < +\infty. \tag{2.6}$$

We obtain from (2.4) that the function $t \mapsto n(t) = \mathbb{E}[N^K(t)] = (n_1(t), n_2(t), n_3(t))$ satisfies the system of equations

$$\forall t \leq T, \begin{cases} n_1(t) &= \mathbb{E}[N_1^K(0)] \\ \frac{d}{dt} n_2(t) &= \tau_1 n_1(t) - \tau_2 K^{-\gamma_2} n_2(t) \\ \frac{d}{dt} n_3(t) &= 2\tau_2 p_2^D n_2(t) - \tau_3 K^{-\gamma_3} n_3(t). \end{cases} \tag{2.7}$$

As explained in the introduction,

$$\mathbb{E}[N_1^K(0)] \sim K.$$

Therefore there is a unique equilibrium given by

$$\begin{aligned}
\forall t \geq 0, \quad n_1^* &= \mathbb{E}[N_1^K(0)] \sim K \\
n_2^* &= \frac{\tau_1 n_1^*}{\tau_2} K^{\gamma_2} \sim K^{1+\gamma_2} \\
n_3^* &= \frac{2p_2^D \tau_2 n_2^*}{\tau_3} K^{\gamma_3} \sim K^{1+\gamma_2+\gamma_3}.
\end{aligned} \tag{2.8}$$

Remark 1. *In the above computation, we obtain the order of magnitude of each sub-population size at equilibrium, but we cannot deduce the order of magnitude of the time taken by the process to reach this equilibrium. We will keep this remark in mind in all the paper.*

Let us first begin by a lemma showing that for any K and t , the expectations of the sub-population sizes behave as expected from (2.8).

Lemma 1. *Let us now assume that*

$$\sup_K \mathbb{E}\left[\frac{N_1^K(0)}{K}\right] < +\infty, \quad \sup_K \mathbb{E}\left[\frac{N_2^K(0)}{K^{1+\gamma_2}}\right] < +\infty, \quad \sup_K \mathbb{E}\left[\frac{N_3^K(0)}{K^{1+\gamma_2+\gamma_3}}\right] < +\infty.$$

then

$$\sup_{K, t \in \mathbb{R}_+} \mathbb{E}\left[\frac{N_1^K(t)}{K}\right] < +\infty, \quad \sup_{K, t \in \mathbb{R}_+} \mathbb{E}\left[\frac{N_2^K(t)}{K^{1+\gamma_2}}\right] < +\infty, \quad \sup_{K, t \in \mathbb{R}_+} \mathbb{E}\left[\frac{N_3^K(t)}{K^{1+\gamma_2+\gamma_3}}\right] < +\infty.$$

Démonstration. The first assertion follows immediately.

From (2.7), we obtain that for all $t \geq 0$,

$$\mathbb{E}\left[\frac{N_2^K(t)}{K^{1+\gamma_2}}\right] = \frac{\tau_1}{\tau_2} \mathbb{E}\left[\frac{N_1^K(0)}{K}\right] + \left(\mathbb{E}\left[\frac{N_2^K(0)}{K^{1+\gamma_2}}\right] - \frac{\tau_1}{\tau_2} \mathbb{E}\left[\frac{N_1^K(0)}{K}\right]\right) e^{-\tau_2 K^{-\gamma_2} t}$$

and the proof of the second assertion follows.

Similarly, straightforward computation yields

$$\begin{aligned} \mathbb{E}\left[\frac{N_3^K(t)}{K^{1+\gamma_2+\gamma_3}}\right] &= \frac{2p_2^D \tau_1}{\tau_3} \mathbb{E}\left[\frac{N_1^K(0)}{K}\right] - \beta_K e^{-\tau_2 K^{-\gamma_2} t} \\ &\quad + \left(\mathbb{E}\left[\frac{N_3^K(0)}{K^{1+\gamma_2+\gamma_3}}\right] - \frac{2p_2^D \tau_1}{\tau_3} \mathbb{E}\left[\frac{N_1^K(0)}{K}\right] + \beta_K\right) e^{-\tau_3 K^{-\gamma_3} t}, \end{aligned}$$

with

$$\beta_K = 2p_2^D \tau_2 \frac{1}{\tau_2 K^{\gamma_3-\gamma_2} - \tau_3} \left(\mathbb{E}\left[\frac{N_2^K(0)}{K^{1+\gamma_2}}\right] - \frac{\tau_1}{\tau_2} \mathbb{E}\left[\frac{N_1^K(0)}{K}\right] \right).$$

Hence the third assertion is proved. □

2.2.2 Asymptotic behavior on a finite time interval

The parameter K is defined as the order of magnitude of the martingale N_1^K at time 0. The first result describes the dynamics of the process on a finite time interval. The proof of this proposition is left to the reader. It is classical and more difficult proofs in a similar spirit will be given later.

Proposition 2. *Let us introduce the jump process X^K defined for all $t \geq 0$ by*

$$X^K(t) = \left(\frac{N_1^K(t)}{K}, \frac{N_2^K(t)}{K^{1+\gamma_2}}, \frac{N_3^K(t)}{K^{1+\gamma_2+\gamma_3}} \right).$$

(i) Let us assume that there exists a vector $(x_1, 0, 0) \in \mathbb{R}_+^3$ such that the sequence $\left(\frac{N_1^K(0)}{K}, \frac{N_2^K(0)}{K^{1+\gamma_2}}, \frac{N_3^K(0)}{K^{1+\gamma_2+\gamma_3}}\right)_{K \in \mathbb{N}^*}$ converges in law to $(x_1, 0, 0)$ when K tends to infinity and such that

$$\sup_K \mathbb{E}\left[\frac{N_1^K(0)}{K}\right] < +\infty, \quad \sup_K \mathbb{E}\left[\frac{N_2^K(0)}{K^{1+\gamma_2}}\right] < +\infty \quad \text{and} \quad \sup_K \mathbb{E}\left[\frac{N_3^K(0)}{K^{1+\gamma_2+\gamma_3}}\right] < +\infty.$$

Then for all $T > 0$, the sequence $\left(\frac{N_1^K(t)}{K}, \frac{N_2^K(t)}{K^{1+\gamma_2}}, \frac{N_3^K(t)}{K^{1+\gamma_2+\gamma_3}}\right)_{K \in \mathbb{N}^*}$ converges in law in $\mathbb{D}([0, T], \mathbb{R}_+^3)$ to $(x_1, 0, 0)$.

(ii) Let us assume that there exists a vector $(x_1, 0, 0) \in \mathbb{R}_+^3$ such that the sequence $\left(\frac{N_1^K(0)}{K}, \frac{N_2^K(0)}{K}, \frac{N_3^K(0)}{K}\right)_{K \in \mathbb{N}^*}$ converges in law to $(x_1, 0, 0)$ when K tends to infinity and such that

$$\sup_K \mathbb{E}\left[\frac{N_1^K(0)}{K}\right] < +\infty, \quad \sup_K \mathbb{E}\left[\frac{N_2^K(0)}{K}\right] < +\infty \quad \text{and} \quad \sup_K \mathbb{E}\left[\frac{N_3^K(0)}{K}\right] < +\infty.$$

Then for all $T > 0$, the sequence $\left(\frac{N_1^K(t)}{K}, \frac{N_2^K(t)}{K}, \frac{N_3^K(t)}{K}\right)_{K \in \mathbb{N}^*}$ converges in law in $\mathbb{D}([0, T], \mathbb{R}_+^3)$ to $x_1(1, \tau_1 t, \frac{\tau_2}{2} t^2)$.

Let us underline that at this time scale, assertion (i) shows that the two last components do not reach their equilibrium order, as observed in the simulations. Assertion (ii) proves that the three compartments are only of order K during the time interval $[0, T]$.

2.3 Size-time multi-scale dynamics and asymptotic behavior

A size renormalization of the stochastic process is not enough to understand the dynamics of the model. We need to change the time scale as observed in the simulations. This part is devoted to identify and study two significant asymptotics for the process, corresponding to the two time scales K^{γ_2} and K^{γ_3} , when K is large.

2.3.1 Asymptotic behavior at a time-scale of order K^{γ_2}

In this section, we study the system composed of the two first components at the time scale K^{γ_2} . To this end, let us introduce the jump process Y^K defined for all $t \geq 0$ by

$$Y^K(t) = \left(\frac{N_1^K(t K^{\gamma_2})}{K}, \frac{N_2^K(t K^{\gamma_2})}{K^{1+\gamma_2}}\right).$$

Let us note that only the time scale differs between processes X^K and Y^K . Hence, at time $t = 0$, we have

$$(Y_1^K(0), Y_2^K(0)) = (X_1^K(0), X_2^K(0)).$$

The next theorem describes the approximating behavior of Y^K when K tends to infinity.

Theorem 1. Assume that there exists a vector $(x_1, x_2) \in \mathbb{R}_+^2$ such that the sequence $(Y^K(0))_{K \in \mathbb{N}^*}$ converges in law to (x_1, x_2) when K tends to infinity and such that

$$\sup_K \mathbb{E}[Y_1^K(0)^2 + Y_2^K(0)^2] < \infty.$$

Then for each $T > 0$, the sequence $(Y^K)_{K \in \mathbb{N}^*}$ converges in law (and hence in probability) in $\mathbb{D}([0, T], \mathbb{R}_+^2)$ to the continuous function $y = (y_1, y_2)$ such that for all $t \geq 0$,

$$\begin{cases} y_1(t) &= x_1 \\ y_2(t) &= \frac{\tau_1 x_1}{\tau_2} + \left(x_2 - \frac{\tau_1 x_1}{\tau_2}\right) e^{-\tau_2 t}. \end{cases} \quad (2.9)$$

Démonstration. By standard localization argument, use of Gronwall's Lemma and Doob's inequality, we easily prove (successively for the first and then for the second component) that for any $T > 0$,

$$\sup_K \mathbb{E}\left[\sup_{t \in [0, T]} (Y_1^K(t)^2 + Y_2^K(t)^2)\right] < \infty. \quad (2.10)$$

From (2.4) and (2.5), we can write

$$\begin{aligned} Y_1^K(t) &= Y_1^K(0) + \widehat{M}_1^K(t) \\ Y_2^K(t) &= Y_2^K(0) + \tau_1 \int_0^t Y_1^K(u) du - \tau_2 \int_0^t Y_2^K(u) du + \widehat{M}_2^K(t), \end{aligned} \quad (2.11)$$

where $\widehat{M}_1^K$ and $\widehat{M}_2^K$ are two square-integrable martingales satisfying

$$\begin{aligned} \langle \widehat{M}_1^K \rangle_t &= \frac{\tau_1}{K^{1-\gamma_2}} \int_0^t Y_1^K(u) du, \\ \langle \widehat{M}_2^K \rangle_t &= \frac{2\tau_1}{K^{1+\gamma_2}} \int_0^t Y_1^K(u) du + \frac{\tau_2}{K} \int_0^t Y_2^K(u) du, \\ \langle \widehat{M}_1^K, \widehat{M}_2^K \rangle_t &= -\frac{\tau_1}{K} \int_0^t Y_1^K(u) du. \end{aligned} \quad (2.12)$$

It is very standard to prove that the sequence of laws of (Y^K) is tight (using the moment estimates (2.10)) and that the martingale parts go to 0. The result follows using the method summarized for example in [5]. Each limiting value is proved to only charge the subset of continuous functions. Then introducing

$$\phi_t(y) = \begin{pmatrix} y_1(t) - y_1(0) \\ y_2(t) - y_2(0) - \int_0^t (\tau_1 y_1(s) - \tau_2 y_2(s)) ds \end{pmatrix}$$

and using the uniform integrability of the sequence $(\phi_t(Y^K))_K$, deduced from (2.10), we identify the limit as the unique continuous solution y of the deterministic system defined by $y(0) = (x_1, x_2)$ and

$$\forall t \geq 0, \quad \phi_t(y) = 0.$$

That concludes the proof. □

Remark 2. Since $\gamma_2 < \gamma_3$, the time scale K^{γ_2} is not large enough to observe the dynamics of the third component. The next proposition shows that at such a time scale, the third component converges to a trivial value.

Proposition 3. Under the same hypotheses as in Theorem 1, we assume furthermore that there exists $x_3 \in \mathbb{R}_+$ such that the sequence $(\frac{N_3^K(0)}{K^{1+\gamma_2+\gamma_3}})_{K \in \mathbb{N}^*}$ converges in law to x_3 when K tends to infinity and such that

$$\sup_K \mathbb{E}[(\frac{N_3^K(0)}{K^{1+\gamma_2+\gamma_3}})^2] < \infty.$$

Then for each $T > 0$, the sequence $(\frac{N_3^K(K^{\gamma_2})}{K^{1+\gamma_2+\gamma_3}})_{K \in \mathbb{N}^*}$ converges in probability in $\mathbb{D}([0, T], \mathbb{R}_+)$ to x_3 .

Démonstration. Following (2.4) and (2.5), let us write the semimartingale decomposition of the process $Y_3^K = \frac{N_3^K(K^{\gamma_2})}{K^{1+\gamma_2+\gamma_3}}$. We have for any $t \leq T$,

$$Y_3^K(t) = Y_3^K(0) + 2\tau_2 p_2^D K^{\gamma_2-\gamma_3} \int_0^t Y_2^K(s) ds - \tau_3 K^{\gamma_2-\gamma_3} \int_0^t Y_3^K(s) ds + \widehat{M}_3^K(t),$$

where $\widehat{M}_3^K$ is a square-integrable martingale such that

$$\langle \widehat{M}_3^K \rangle_t = 2\tau_2 p_2^D K^{\gamma_2-\gamma_3} \int_0^t Y_2^K(s) ds + \tau_3 K^{\gamma_2-\gamma_3} \int_0^t Y_3^K(s) ds.$$

Let us recall that $\gamma_2 < \gamma_3$, which makes $K^{\gamma_2-\gamma_3}$ tends to 0 when K tends to infinity. Using Theorem 1, we know that Y_2^K converges to the continuous function y_2 . By standard tightness argument, one can easily deduce that the process Y_3^K converges in probability to x_3 , on any finite time interval. $\square$

2.3.2 Asymptotic behavior at a time-scale of order K^{γ_3}

In order to catch the long time dynamics of the third component we will study the process N^K on the time scale K^{γ_3} . To this end, let us introduce the jump process Z^K defined for all $t \geq 0$ by

$$Z^K(t) = (\frac{N_1^K(t K^{\gamma_3})}{K}, \frac{N_2^K(t K^{\gamma_3})}{K^{1+\gamma_2}}, \frac{N_3^K(t K^{\gamma_3})}{K^{1+\gamma_2+\gamma_3}}).$$

Note that we still have

$$Z^K(0) = Y^K(0) = X^K(0).$$

At this time scale, the second component has time to reach the equilibrium of its deterministic approximation by an average procedure. By an adaptation of the proof in [54] to this specific framework, we are able to prove the following theorem.

Theorem 2. Assume that there exists $(x_1, x_2, x_3) \in \mathbb{R}_+^3$ such that the sequence $(Z^K(0))_{K \in \mathbb{N}^*}$ converges in law to (x_1, x_2, x_3) when K tends to infinity and such that

$$\sup_K \mathbb{E}[Z_1^K(0)] < +\infty, \quad \sup_K \mathbb{E}[Z_2^K(0)] < +\infty \quad \text{and} \quad \sup_K \mathbb{E}[Z_3^K(0)] < +\infty. \quad (2.13)$$

Let Γ_2^K be the $l_m(\mathbb{R}_+)$ -valued random variable given by

$$\Gamma_2^K([0, t] \times B) = \int_0^t \mathbf{1}_{\{B\}}(\mathbf{Z}_2^K(s)) ds.$$

Then for all $T > 0$, the sequence $(Z_1^K, \Gamma_2^K, Z_3^K)_{K \in \mathbb{N}^*}$ converges in law in $\mathbb{D}([0, T], \mathbb{R}_+) \times l_m(\mathbb{R}_+) \times \mathbb{D}([0, T], \mathbb{R}_+)$ to $(z_1, \delta_{z_2^*}(dz_2), z_3)$. The functions z_1 and z_3 are defined for all $t \leq T$ by

$$\begin{cases} z_1(t) &= x_1 \\ z_3(t) &= \frac{\tau_2}{\tau_3} z_2^* + \left(x_3 - \frac{\tau_2}{\tau_3} z_2^*\right) e^{-\tau_3 t} \end{cases} \quad (2.14)$$

and z_2^* is the value of y_2 at infinity :

$$z_2^* = \frac{\tau_1 x_1}{\tau_2}.$$

Let us first state a lemma in which all moment estimates are gathered.

Lemma 2. *Under Assumption (2.13), we obtain*

$$\sup_K \mathbb{E} \left[\sup_{t \in [0, T]} Z_1^K(t) \right] < +\infty ; \quad \forall t > 0 \quad \sup_K \mathbb{E} \left[\int_0^t Z_2^K(s) ds \right] < +\infty ;$$

and

$$\sup_K \mathbb{E} \left[\sup_{t \in [0, T]} Z_3^K(t) \right] < +\infty.$$

Proof of Lemma 2. The first and third estimates are obtained by usual arguments (localization, Doob's inequality and Gronwall's Lemma). Let us focus on the second one.

By positivity and definition of the process Z_2^K , we have for any $t > 0$

$$Z_2^K(t) = Z_2^K(0) + \tau_1 K^{\gamma_3 - \gamma_2} \int_0^t Z_1^K(s) ds - \tau_2 K^{\gamma_3 - \gamma_2} \int_0^t Z_2^K(s) ds + \widetilde{M}_2^K(t), \quad (2.15)$$

the latter term being a square-integrable martingale and

$$\langle \widetilde{M}_2^K \rangle_t = \frac{1}{K^{1+\gamma_2}} \left(2\tau_1 K^{\gamma_3 - \gamma_2} \int_0^t Z_1^K(s) ds + \tau_2 K^{\gamma_3} \int_0^t Z_2^K(s) ds \right). \quad (2.16)$$

In particular,

$$\mathbb{E} \left[\int_0^t Z_2^K(s) ds \right] = \frac{1}{\tau_2} K^{\gamma_2 - \gamma_3} \left(\mathbb{E}[Z_2^K(0)] - \mathbb{E}[Z_2^K(t)] \right) + \frac{\tau_1}{\tau_2} \int_0^t \mathbb{E}[Z_1^K(s)] ds.$$

Assumptions (2.13) and Lemma 1 ensure that the first term goes to 0 as K tends to infinity and the third term is bounded uniformly in K . That allows to conclude. $\square$

Proof of Theorem 2. Let Γ^K be the occupation measure of Z^K , a random measure belonging to the space $l_m(\mathbb{R}_+)$ of positive measures on $[0, \infty) \times \mathbb{R}_+$ with mass t on $[0, t] \times \mathbb{R}_+$ and defined for all Borelian set B and for $t > 0$ by

$$\Gamma^K([0, t] \times B) = \int_0^t \mathbf{1}_{\{B\}}(\mathbf{Z}^K(s)) ds.$$

Using Lemma 2.9 of [54] (cf. Appendix), we obtain that $(\Gamma^K)_K$ is relatively compact in $l_m(\mathbb{R}_+)$ endowed with a weak topology generated by the class of test functions defined in (2.39).

Let us denote by $\Gamma \in l_m(\mathbb{R}_+^3)$ a limiting value. Using [62] Lemma 1.4, one can show that there exists a $\mathcal{P}(\mathbb{R}_+^3)$ -valued process γ_s such that

$$\Gamma(dz \times ds) = \gamma_s(dz) ds.$$

Let us now introduce the function F^K

$$\forall z \in \mathbb{R}_+^3, \quad F^K(z) = (1 + \frac{1}{K^{\gamma_2}}) \tau_2 z_2 - \tau_3 z_3.$$

Then for all $t \geq 0$,

$$Z_3^K(t) = Z_3^K(0) + (1 + \frac{1}{K^{\gamma_2}}) \tau_2 \int_0^t Z_2^K(s) ds - \tau_3 \int_0^t Z_3^K(s) ds + \widetilde{M}_3^K(t) \quad (2.17)$$

$$= Z_3^K(0) + \int_0^t F^K(Z^K(s)) ds + \widetilde{M}_3^K(t)$$

$$Z_1^K(t) = Z_1^K(0) + \widetilde{M}_1^K(t) \quad (2.18)$$

with $(\widetilde{M}_1^K, \widetilde{M}_3^K)$ independent martingales and such that for all $t \geq 0$,

$$\langle \widetilde{M}_1^K \rangle_{t=K^{\gamma_3-1}} = \int_0^t 2\tau_1 Z_1^K(s) ds \quad (2.19)$$

$$\langle \widetilde{M}_3^K \rangle_{t=K^{-(1+\gamma_2+\gamma_3)}} = \int_0^t (2(1 + \frac{1}{K^{\gamma_2}}) \tau_2 Z_2^K(s) + \tau_3 Z_3^K(s)) ds. \quad (2.20)$$

By usual arguments involving Lemma 2 one can prove that the sequences of processes $(Z_3^K)_K$ and $(Z_1^K)_K$ are uniformly tight in $\mathbb{D}([0, T], \mathbb{R}_+)$. Let us also note that the distributions of any limiting value only charge processes with a.s. continuous trajectories.

Furthermore by Doob's inequality

$$\mathbb{E} \left[\sup_{t \leq T} |\widetilde{M}_3^K(t)|^2 \right] \leq 4 \mathbb{E} \left[\langle \widetilde{M}_3^K \rangle_T \right].$$

Using (2.20) and Lemma 2, we obtain that $\lim_{K \rightarrow \infty} \mathbb{E} \left[\sup_{t \in [0, T]} |\widetilde{M}_3^K(t)|^2 \right] = 0$ and a similar property for $\widetilde{M}_1^K$ since $\gamma_3 < 1$. Then, we deduce from Markov's inequality, that the processes $(\widetilde{M}_3^K)_K$ and $(\widetilde{M}_1^K)_K$ converge in probability for the uniform norm to 0. Hence they converge in law in $\mathbb{D}([0, T], \mathbb{R}_+)$ to 0.

Adding all these asymptotic behaviors, we deduce that there exists a subsequence of (Z_1^K, Z_3^K) converging in law in $\mathbb{D}([0, T], \mathbb{R}_+^2)$ to the deterministic limit (z_1, Z_3^∞) defined for all $t \geq 0$ by

$$\begin{aligned} z_1(t) &= x_1 \\ Z_3^\infty(t) &= x_3 + \int_{\mathbb{R}_+^3 \times [0, t]} (\tau_2 z_2 - \tau_3 z_3) \gamma_s(dz) ds. \end{aligned}$$

Then by convergence of $(Z_1^K)_K$, $(Z_3^K)_K$ and Γ^K , one can easily deduce that

$$\gamma_s(dz) = \delta_{\bar{z}_1}(dz_1)\delta_{Z_3^\infty(s)}(dz_3)\tilde{\gamma}_s(z_1, z_3, dz_2).$$

We have now to identify these measures $\tilde{\gamma}_s$, $s \in [0, T]$.

Let us write the generator of the process Z_2^K . For $h \in C_c^\infty$, it is given for $z \in \mathbb{R}_+^3$ by

$$\begin{aligned}\mathcal{L}_2^K(h)(z) = & \left(h(z_2 + 2K^{-(1+\gamma_2)}) - h(z_2)\right) \frac{\tau_1}{2} z_1 K^{1+\gamma_3} \\ & + \left(h(z_2 + K^{-(1+\gamma_2)}) - h(z_2)\right) p_2^R \tau_2 z_2 K^{1+\gamma_2+\gamma_3} \\ & + \left(h(z_2 - K^{-(1+\gamma_2)}) - h(z_2)\right) p_2^D \tau_2 z_2 K^{1+\gamma_2+\gamma_3}.\end{aligned}$$

Let us introduce the function g defined for $z \in \mathbb{R}_+^3$ by

$$g(z) = z_1 \tau_1 - \tau_2 z_2.$$

Then, by a Taylor expansion, we obtain that

$$\forall h \in C_c^\infty, \quad \lim_{K \rightarrow \infty} \sup_{z \in \mathbb{R}_+^3} |K^{\gamma_2-\gamma_3} \mathcal{L}_2^K(h)(z) - g(z)h'(z_2)| = 0. \quad (2.21)$$

Using (2.16), Lemma 3 and the same arguments as above, we obtain that the sequence of processes $\left(K^{\gamma_2-\gamma_3} \left(h(Z_2^K(t)) - h(Z_2^K(0)) - \int_0^t \mathcal{L}_2^K(h)(Z^K(s)) ds\right), \quad t \in [0, T]\right)$ converges in law in $\mathbb{D}([0, T], \mathbb{R})$ to 0. In the other hand, using that h is bounded and (2.21) and $gh' \in C_b$ (since h has compact support), we easily see that this sequence also converges to

$$- \int_0^t (\tau_1 x_1 - \tau_2 z_2) h'(z_2) \gamma_s(z_1, z_3, dz_2) ds.$$

We deduce that for any $t \in [0, T]$, for any $h \in C_c^\infty$

$$\int_0^t (\tau_1 x_1 - \tau_2 z_2) h'(z_2) \tilde{\gamma}_s(z_1, z_3, dz_2) ds = 0.$$

It implies that

$$\tilde{\gamma}_s(z_1, z_3, dz_2) = \frac{\delta \tau_1 x_1}{\tau_2}(dz_2).$$

To end the proof, we solve the equation satisfied by $Z_3^\infty(t)$ and obtain that

$$\forall t \geq 0, \quad Z_3^\infty(t) = \frac{\tau_1 x_1}{\tau_3} + \left(x_3 - \frac{\tau_1 x_1}{\tau_3}\right) e^{-\tau_3 t}.$$

Hence we have uniquely identified the limit of any converging subsequence. That ends the proof. □

2.4 Amplified fluctuations

In this section, we will quantify the large fluctuations observed on the simulations. As pointed out above, each component has its own typical size and time scale. Hence, we will study separately the fluctuations of the second and third types. Classical results easily imply that the first component behaves as a Brownian motion : for large K , for all $t \in [0, T]$,

$$N_1^K(t) \sim x_1 K + \sqrt{K} \sqrt{\tau_1 x_1} B_t.$$

The originality of our results concerns the large fluctuations of the two last types due to the amplification of these first type fluctuations.

2.4.1 The large fluctuations of the second type

As seen in Subsection 2.3.1, the typical size scale (respectively time scale) of the second type is $K^{1+\gamma_2}$ (respectively K^{γ_2}) and the first order asymptotics relative to this time scale is given by the function y defined by (2.9). We are also able to give an expansion of the process at the second and third orders on such time scale.

Theorem 3. *Let define the sequence $(U^K)_K$ by*

$$\forall t \geq 0, \quad U^K(t) = K^{(1-\gamma_2)/2} (Y^K(t) - y(t)).$$

(i) *Assume that there exists $U_0 = (U_0^{(1)}, U_0^{(2)}) \in \mathbb{R}^2$ such that $(U^K(0))_{K \in \mathbb{N}^*}$ converges in law to U_0 and that*

$$\sup_K \mathbb{E}[U_1^K(0)^2 + U_2^K(0)^2] < +\infty. \quad (2.22)$$

Then for each $T > 0$, the sequence $(U^K)_{K \in \mathbb{N}^}$ converges in law in $\mathbb{D}([0, T], \mathbb{R}^2)$ to the process $U = (U_1, U_2)$ defined for all $t \geq 0$ by*

$$\begin{aligned} U_1(t) &= U_0^{(1)} + \sqrt{\tau_1 x_1} B_1(t), \\ U_2(t) &= U_0^{(2)} + \tau_1 \int_0^t U_1(s) ds - \tau_2 \int_0^t U_2(s) ds, \end{aligned}$$

where B_1 is a standard Brownian motion.

(ii) *Furthermore, the sequence $(W_2^K)_{K \in \mathbb{N}^*}$ defined by*

$$\forall t \in [0, T], \quad W_2^K(t) = K^{\gamma_2/2} [U_2^K(t) - U_0^{(2)} - \tau_1 \int_0^t U_1^K(s) ds + \tau_2 \int_0^t U_2^K(s) ds],$$

converges in law in $\mathbb{D}([0, T], \mathbb{R})$, for each $T > 0$, to the process $(\sqrt{\tau_2 y_2(t)} B_2(t), \quad t \in [0, T])$ where B_2 is a standard Brownian motion independent of the process B_1 .

From this theorem, we can deduce the following expansion, which quantifies the large waves of fluctuations. Assuming that U_0 is equal to zero, we obtain for all t and large K ,

$$N_2^K(t) \sim K^{1+\gamma_2} y_2(t K^{-\gamma_2}) + K^{(1+3\gamma_2)/2} U_2(t K^{-\gamma_2}) + K^{(1+2\gamma_2)/2} \sqrt{\tau_2 y_2(t K^{-\gamma_2})} B_2(t K^{-\gamma_2})$$

where

$$\forall t, \quad U_2(t) = \tau_1 \sqrt{\tau_1 x_1} \int_0^t B_1(s) ds - \tau_2 \int_0^t U_2(s) ds$$

and B_1, B_2 are independent Brownian motions.

of Theorem 3. (i) First we deduce from (2.22) with similar arguments as above that

$$\sup_K \mathbb{E} \left[\sup_{t \in [0, T]} (U_1^K(t)^2 + U_2^K(t)^2) \right] < +\infty. \quad (2.23)$$

The tightness of the families $(\sup_{t \leq T} |U_1^K(t)|)_K$ and $(\sup_{t \leq T} |U_2^K(t)|)_K$ immediately follows.

We consider the semi-martingale decomposition of (U^K) and write

$$U_i^K(t) = U_i^K(0) + A_i^K(t) + K^{\frac{1-\gamma_2}{2}} \widehat{M}_i^K(t),$$

where $\widehat{M}^K$ has been defined in (2.11), $A_1^K = 0$ and $A_2^K(t) = \tau_1 \int_0^t U_1^K(s) ds - \tau_2 \int_0^t U_2^K(s) ds$.

Thanks to the above moment estimates, it is almost immediate to prove that the finite variation processes $\langle K^{\frac{1-\gamma_2}{2}} \widehat{M}_i^K \rangle$ and A_2^K satisfy the Aldous condition. Thanks to Aldous and Rebolledo criteria (see [53] and [5]), the uniform tightness of $\mathcal{L}(U^K)$ in $\mathcal{P}(\mathbb{D}([0, T], \mathbb{R}^2))$ follows.

We denote by simplicity by the same notation $(U^K)_K$ a subsequence converging in law in $\mathbb{D}([0, T], \mathbb{R}^2)$. Let Q be the limiting value of $(\mathcal{L}(U^K))_K$. It is easy to observe that

$$\sup_{t \in [0, T]} \|\Delta U^K(t)\| \leq 2K^{-(1+\gamma_2)/2}.$$

Therefore, by continuity of the mapping $x \rightarrow \sup_{t \in [0, T]} \|\Delta x(t)\|$ from $\mathbb{D}([0, T], \mathbb{R}^2)$ into $\mathbb{R}_+$, the probability measure Q only charges the processes with continuous paths.

The extended generator of U^K is defined for functions $f \in C_b^2(\mathbb{R}^2, \mathbb{R})$ as : $\forall u \in \mathbb{R}^2$,

$$\begin{aligned} \mathcal{L}^K(f, t)(u) &= (f(u_1 + K^{-(1+\gamma_2)/2}, u_2) - f(u)) \frac{\tau_1}{2} K^{1+\gamma_2} (K^{-(1-\gamma_2)/2} u_1 + x_1) \\ &\quad + (f(u_1 - K^{-(1+\gamma_2)/2}, u_2 + 2K^{-(1+3\gamma_2)/2}) - f(u)) \frac{\tau_1}{2} K^{1+\gamma_2} (K^{-(1-\gamma_2)/2} u_1 + x_1) \\ &\quad + (f(u_1, u_2 + K^{-(1+3\gamma_2)/2}) - f(u)) p_2^R \tau_2 K^{1+2\gamma_2} (K^{-(1-\gamma_2)/2} u_2 + y_2(t)) \\ &\quad + (f(u_1, u_2 - K^{-(1+3\gamma_2)/2}) - f(u)) p_2^D \tau_2 K^{1+2\gamma_2} (K^{-(1-\gamma_2)/2} u_2 + y_2(t)) \\ &\quad - K^{(1-\gamma_2)/2} \partial_2 f(u) (\tau_1 x_1 - \tau_2 y_2(t)). \end{aligned} \quad (2.24)$$

By a Taylor's expansion, we easily obtain that $\forall f \in C_b^2(\mathbb{R}^2, \mathbb{R})$,

$$\lim_{K \rightarrow \infty} \sup_{(u, t) \in \mathbb{R} \times \mathbb{R}_+} |\mathcal{L}^K(f, t)(u) - (\frac{\tau_1}{2} x_1 \partial_1^2 f(u) + (\tau_1 u_1 - \tau_2 u_2) \partial_2 f(u))| = 0.$$

In the other hand, let us define, for $f \in C_b^2(\mathbb{R}^2, \mathbb{R})$, $u \in \mathbb{D}([0, T], \mathbb{R}^2)$ and $t \in [0, T]$, the function $\xi_t^{K, f}$ by

$$\xi_t^{K, f}(u) = f(u_t) - f(u_0) - \int_0^t \mathcal{L}^K(f, s)(u_s) ds.$$

Then, by (2.24), Dynkin's formula and (2.23), we can easily prove that the processes $(\xi_t^{K,f}(U^K))_K$ are uniformly integrable martingales.

Therefore by standard arguments (cf. [33], [5]), the limiting process under Q is continuous and satisfies the following martingale problem :

$$\forall f \in C_b^2(\mathbb{R}^2, \mathbb{R}), \quad f(U(t)) - f(U_0) - \int_0^t \left(\frac{\tau_1}{2} x_1 \partial_1^2 f(U(s)) + (\tau_1 U_1(s) - \tau_2 U_2(s)) \partial_2 f(U(s)) \right) ds$$

is a martingale.

We conclude using a representation theorem (cf. [51] p.84) that for each $T > 0$, the sequence $(U^K)_{K \in \mathbb{N}^*}$ converges in law in $\mathbb{D}([0, T], \mathbb{R}^2)$ to the process $U = (U_1, U_2)$, unique solution of the following stochastic differential system : for all $t \in [0; T]$,

$$\begin{aligned} U_1(t) &= U_0^{(1)} + \sqrt{\tau_1 x_1} B_1(t), \\ U_2(t) &= U_0^{(2)} + \tau_1 \int_0^t U_1(s) ds - \tau_2 \int_0^t U_2(s) ds, \end{aligned}$$

with B_1 a Brownian motion.

(ii) Let us now expand the second component to the next order. We deduce from (2.11) that

$$\begin{aligned} U_1^K(.) - U^K(0) &= K^{(1-\gamma_2)/2} \widehat{M}_1^K(.), \\ W_2^K(.) &= \sqrt{K} \widehat{M}_2^K(.). \end{aligned}$$

Using (2.12), (2.10) and applying Theorem 7.1.4 of [33], we conclude the proof. $\square$

2.4.2 The large fluctuations of the third type

Let us now study the fluctuation process associated with the largest fluctuation scale of the third component. We have seen in Theorem 2 that at the time scale K^{γ_3} , the size of the population process in the third compartment is of order of magnitude $K^{1+\gamma_2+\gamma_3}$. In an usual setting, the Central Limit Theorem would lead to fluctuations of order $K^{(1+\gamma_2+\gamma_3)/2}$. We will see in the next theorem that they are of order $K^{(1+2\gamma_2+3\gamma_3)/2} \gg K^{(1+\gamma_2+\gamma_3)/2}$, since amplified by the fluctuations of the first compartment.

Using (2.14), (2.15) and (2.17), we know that for all $t \geq 0$,

$$\begin{aligned} (Z_3^K(t) - z_3(t)) &= (Z_3^K(0) - z_3(0)) + \tau_2 \int_0^t (Z_2^K(s) - z_2^*) ds - \tau_3 \int_0^t (Z_3^K(s) - z_3(s)) ds \\ &\quad + \tau_2 \frac{1}{K^{\gamma_2}} \int_0^t (Z_2^K(s) - z_2^*) ds + \widetilde{M}_3^K(t) \end{aligned} \tag{2.25}$$

where

$$\begin{aligned} (Z_2^K(t) - z_2^*) &= (Z_2^K(0) - z_2^*) + K^{\gamma_3-\gamma_2} \tau_1 \int_0^t (Z_1^K(s) - x_1) ds \\ &\quad - K^{\gamma_3-\gamma_2} \tau_2 \int_0^t (Z_2^K(s) - z_2^*) ds + \widetilde{M}_2^K(t). \end{aligned} \tag{2.26}$$

Our goal is to quantify the effect of the first component fluctuations on the dynamics of the third component. Considering the expressions of the martingale quadratic variation (2.19) imposes the choice of the rescaling parameter $K^{(1-\gamma_3)/2}$ in front of $(Z_1^K(t) - x_1)$. We will see that to keep the effect of the first component on the third component, we need to rescale $(Z_2^K(t) - Z_2^K(0))$ by $\frac{K^{(1-\gamma_3)/2}}{K^{\gamma_3-\gamma_2}}$ and $(Z_3^K(t) - z_3(t))$ by $K^{(1-\gamma_3)/2}$.

Let us now state the main theorem of this section.

Theorem 4. *Let us define the three processes*

$$\forall t \geq 0, \quad \begin{cases} V_1^K(t) &= K^{(1-\gamma_3)/2} (Z_1^K(t) - x_1) \\ V_2^K(t) &= \frac{K^{(1-\gamma_3)/2}}{K^{\gamma_3-\gamma_2}} (Z_2^K(t) - Z_2^K(0)) \\ V_3^K(t) &= K^{(1-\gamma_3)/2} (Z_3^K(t) - z_3(t)) \end{cases}.$$

Let us assume that there exists $V_0 = (V_0^{(1)}, V_0^{(3)})$ a $\mathbb{R}^2$ -valued random vector such that the sequence $(V_1^K(0), V_3^K(0))_{K \in \mathbb{N}^*}$ converges in law to V_0 and such that

$$\sup_K \mathbb{E}[V_1^K(0)^4] < +\infty \quad ; \quad \sup_K \mathbb{E}[Z_2^K(0)^2] < +\infty. \quad (2.27)$$

$$\sup_K \mathbb{E}[|V_3^K(0)|] < +\infty. \quad (2.28)$$

Then for all $T > 0$, the sequence $(V_1^K, V_3^K)_{K \in \mathbb{N}^*}$ converges in law in $\mathbb{D}([0, T], \mathbb{R}^2)$ to (V_1, V_3) such that for all t ,

$$\begin{aligned} V_1(t) &= V_0^{(1)} + \sqrt{\tau_1 x_1} W_1(t) \\ V_3(t) &= V_0^{(3)} + \tau_1 \int_0^t V_1(s) ds - \tau_3 \int_0^t V_3(s) ds, \end{aligned}$$

where W_1 is a standard Brownian motion.

Let us interpret this result in terms of fluctuations. Assuming that the initial vector V_0 is equal to zero, we obtain that for any t and large K ,

$$N_3^K(t) \sim K^{1+\gamma_2+\gamma_3} z_3(t K^{-\gamma_3}) + K^{(1+2\gamma_2+3\gamma_3)/2} V_3(t K^{-\gamma_3}) \quad (2.29)$$

where

$$\forall t, \quad V_3(t) = \tau_1 \sqrt{\tau_1 x_1} \int_0^t W_1(s) ds - \tau_3 \int_0^t V_3(s) ds$$

and W_1 is a standard Brownian motion.

The order of magnitude appearing in the fluctuation second order term (2.29) summarizes the cumulative effects of the third dynamics driven by the fluctuations of the first level. That can explain the exceptionally large fluctuations observed for the size of terminal cells populations, in hematopoietic systems.

As a first step in the proof of Theorem 4, we will prove that the sequence of processes $(V_2^K)_K$ converges to 0 uniformly in $\mathbb{L}^2$ on any finite time interval.

Proposition 4. *Under the assumption (2.27), we obtain*

$$\forall T \geq 0, \quad \lim_{K \rightarrow \infty} \mathbb{E} \left[\sup_{t \in [0, T]} V_2^K(t)^2 \right] = 0. \quad (2.30)$$

Démonstration. Using (2.15), we obtain that

$$\begin{aligned} \forall t \geq 0, \quad V_2^K(t) &= K^{\frac{1-3\gamma_3}{2} + \gamma_2} (Z_2^K(t) - Z_2^K(0)) \\ &= \tau_1 \int_0^t V_1^K(s) ds - \tau_2 K^{\gamma_3 - \gamma_2} \int_0^t V_2^K(s) ds + \mathcal{R}^K(t), \end{aligned}$$

where $\mathcal{R}^K$ is the square-integrable martingale defined by

$$\forall t \geq 0, \quad \mathcal{R}_t^K = K^{\frac{1-3\gamma_3}{2} + \gamma_2} \widetilde{M}_2^K(t) \quad (2.31)$$

and satisfying

$$\forall t, \quad \langle \mathcal{R}^K \rangle_t = K^{-2\gamma_3} \left(\tau_1 \int_0^t Z_1^K(s) ds + \tau_2 K^{\gamma_2} \int_0^t Z_2^K(s) ds \right). \quad (2.32)$$

Let us first show that

$$\forall t > 0, \quad \sup_K \mathbb{E} \left[\int_0^t V_2^K(s)^4 ds \right] < \infty.$$

Itô's formula immediately implies that $\forall t \geq 0$,

$$\begin{aligned} V_2^K(t)^4 &= V_2^K(0)^4 + 4 \int_0^t V_2^K(s)^3 d\mathcal{R}_s^K + 4 \int_0^t V_2^K(s)^3 (\tau_1 V_1^K(s) - K^{\gamma_3 - \gamma_2} \tau_2 V_2^K(s)) ds \\ &\quad + 6 \int_0^t V_2^K(s)^2 d\langle \mathcal{R}^K \rangle_s. \end{aligned}$$

By standard localization arguments, we prove using (2.27) that for any $t \geq 0$,

$$\forall t \geq 0, \quad \sup_K \mathbb{E} \left[\int_0^t V_1^K(s)^4 ds \right] < \infty. \quad (2.33)$$

Let us now introduce the stopping time

$$T_n = \inf\{t \geq 0, |V_2^K(t)| \geq n\}.$$

Then, applying the following inequality

$$\begin{aligned} 4(\tau_1 v_1 v_2 - K^{\gamma_3 - \gamma_2} \tau_2 v_2^2) &= 4 \tau_2 K^{\gamma_3 - \gamma_2} \left(\left[\frac{v_1 \tau_1}{2 \tau_2 K^{\gamma_3 - \gamma_2}} \right]^2 - \left[v_2 - \frac{v_1 \tau_1}{2 \tau_2 K^{\gamma_3 - \gamma_2}} \right]^2 \right) \\ &\leq \frac{v_1^2 \tau_1^2}{\tau_2 K^{\gamma_3 - \gamma_2}} \end{aligned} \quad (2.34)$$

to $v_1 = V_1^K(s)$ and $v_2 = V_2^K(s)$, we obtain the following upper-bound

$$\begin{aligned} \mathbb{E}[V_2^K(t \wedge T_n)^4] &\leq \mathbb{E}[V_2^K(0)^4] + \frac{\tau_1^2}{\tau_2 K^{\gamma_3 - \gamma_2}} \int_0^{t \wedge T_n} \mathbb{E}[V_2^K(s)^2 V_1^K(s)^2] ds \\ &\quad + 6 \mathbb{E} \left[\int_0^{t \wedge T_n} V_2^K(s)^2 d\langle \mathcal{R}^K \rangle_s \right]. \end{aligned} \quad (2.35)$$

Using (2.32), we obtain for all $t \in [0, T]$,

$$\mathbb{E}\left[\int_0^{t \wedge T_n} V_2^K(s)^2 d\langle \mathcal{R}^K \rangle_s\right] = K^{\gamma_2 - 2\gamma_3} \mathbb{E}\left[\int_0^{t \wedge T_n} V_2^K(s)^2 (K^{-\gamma_2} \tau_1 Z_1^K(s) + \tau_2 Z_2^K(s)) ds\right].$$

Writing Z_2^K in function of V_2^K , we find the following upper bound,

$$\begin{aligned} \mathbb{E}\left[\int_0^{t \wedge T_n} V_2^K(s)^2 d\langle \mathcal{R}^K \rangle_s\right] &\leq (\tau_1 K^{-2\gamma_3} + \tau_2 K^{-(1+\gamma_3)/2} + \tau_2 K^{\gamma_2 - 2\gamma_3}) \int_0^{t \wedge T_n} \mathbb{E}[V_2^K(s)^4] ds \\ &\quad + \tau_1 K^{-2\gamma_3} \int_0^{t \wedge T_n} \mathbb{E}[Z_1^K(s)^2] ds \\ &\quad + \tau_2 (K^{-(1+\gamma_3)/2} + K^{\gamma_2 - 2\gamma_3}) (\mathbb{E}[Z_2^K(0)^2] + 1) T. \end{aligned} \quad (2.36)$$

We deduce from (2.35) using Lemma 2, (2.27) and Gronwall's Lemma that

$$\forall t \in [0, T], \quad \sup_K \mathbb{E}\left[\int_0^{t \wedge T_n} V_2^K(s)^4 ds\right] < \infty.$$

Let us now come back to the proof of (2.30). Itô's formula yields

$$V_2^K(t)^2 = 2 \int_0^t V_2^K(s) d\mathcal{R}_s^K + 2 \int_0^t V_2^K(s) (\tau_1 V_1^K(s) - K^{\gamma_3 - \gamma_2} \tau_2 V_2^K(s)) ds + \langle \mathcal{R}^K \rangle_t.$$

Therefore, using again (2.34) and Doob's inequality, we obtain

$$\begin{aligned} \mathbb{E}\left[\sup_{t \in [0, T \wedge T_n]} V_2^K(t)^2\right] &\leq 8 \mathbb{E}\left[\int_0^{T \wedge T_n} V_2^K(s)^2 d\langle \mathcal{R}^K \rangle_s\right] + \frac{\tau_1^2}{\tau_2 K^{\gamma_3 - \gamma_2}} \mathbb{E}\left[\int_0^{T \wedge T_n} V_1^K(s)^2 ds\right] \\ &\quad + \mathbb{E}[\langle \mathcal{R}^K \rangle_{T \wedge T_n}]. \end{aligned}$$

Finally, we deduce from Lemma 2, (2.36), (2.27) and (2.33) that for any K , T_n tends almost surely to $+\infty$ and that

$$\forall T \geq 0, \quad \lim_{K \rightarrow \infty} \mathbb{E}\left[\sup_{t \in [0, T]} V_2^K(t)^2\right] = 0.$$

□

Let us now come back to the proof of Theorem 4. It has been inspired by the proof of the main result in [54].

of Theorem 4. Using similar convergence arguments as in Theorem 3 and (2.27), we firstly observe that the sequence $(V_1^K)_K$ converges in law in $\mathbb{D}([0, T], \mathbb{R})$ to a continuous process V_1 defined by

$$\forall t, \quad V_1(t) = V_0^{(1)} + \sqrt{\tau_1 x_1} W_1(t) \quad (2.37)$$

with W_1 a standard Brownian motion.

Let us recall that from (2.25), (2.26) and (2.31), that for all t ,

$$V_3^K(t) = V_3^K(0) + \tau_2 \left(1 + \frac{1}{K^{\gamma_2}}\right) \int_0^t K^{(1-\gamma_3)/2} (Z_2^K(s) - z_2^*) ds - \tau_3 \int_0^t V_3^K(s) ds + K^{(1-\gamma_3)/2} \widetilde{M}_3^K(t)$$

with

$$\tau_2 \int_0^t K^{(1-\gamma_3)/2} (Z_2^K(s) - z_2^*) ds = K^{(1-\gamma_3)/2} \frac{Z_2^K(0) - Z_2^K(t)}{K^{\gamma_3-\gamma_2}} + \int_0^t \tau_1 V_1^K(s) ds + \mathcal{R}_t^K.$$

Hence, for all t ,

$$V_3^K(t) = V_3^K(0) + (1 + \frac{1}{K^{\gamma_2}}) \int_0^t \tau_1 V_1^K(s) ds - \tau_3 \int_0^t V_3^K(s) ds - (1 + \frac{1}{K^{\gamma_2}}) V_2^K(t) + \mathcal{M}_t^K, \quad (2.38)$$

where $\mathcal{M}^K$ is the square-integrable martingale

$$\mathcal{M}^K(t) = (1 + \frac{1}{K^{\gamma_2}}) \mathcal{R}_t^K + K^{(1-\gamma_3)/2} \widetilde{M}_3^K(t).$$

We deduce from (2.20), (2.32), Lemma 2 and Doob's inequality, that

$$\lim_{K \rightarrow \infty} \mathbb{E} \left[\sup_{t \in [0, T]} |\mathcal{M}^K(t)|^2 \right] = 0.$$

Then it turns out from Markov's inequality that the sequence $(\mathcal{M}^K)_K$ converges in probability to 0 for the uniform norm and hence $(\mathcal{M}^K)_K$ converges in law in $\mathbb{D}([0, T], \mathbb{R})$ to 0.

Furthermore using (2.38), (2.33), (2.28) and Proposition 4 we obtain

$$\sup_K \mathbb{E} \left[\sup_{t \leq T} |V_3^K(t)| \right] < \infty.$$

We are now able to prove the tightness of the family $(\sup_{t \leq T} |V_3^K(t)|)_K$. Indeed, let us introduce stopping times S, S' satisfying $S \leq S' \leq (S + \delta) \wedge T$, with $\delta > 0$. Using (2.38), we have

$$\begin{aligned} \mathbb{P}(|V_3^K(S') - V_3^K(S)| > \epsilon) &\leq \frac{1}{\epsilon} \mathbb{E}[|V_3^K(S') - V_3^K(S)|] \\ &\leq \frac{1}{\epsilon} \mathbb{E}[|V_3^K(S') - V_3^K(S)|^2]^{1/2} \\ &\leq \frac{1}{\epsilon} \left[\delta T \left(\tau_1 (1 + \frac{1}{K^{\gamma_2}}) (\mathbb{E}[\sup_{t \leq T} V_1^K(t)^2] + 1) + \tau_3 \mathbb{E}[\sup_{t \leq T} |V_3^K(t)|] \right) \right. \\ &\quad \left. + 4 \mathbb{E}[\sup_{t \leq T} \mathcal{M}^K(t)^2] + 4 p_2^D \mathbb{E}[\sup_{t \leq T} V_2^K(t)^2] \right]^{1/2}. \end{aligned}$$

Then from (2.38), (2.33), (2.28) and Proposition 4, we deduce that Aldous conditions (see [53] and [5]) are satisfied and obtain the tightness of $(V_3^K)_K$.

Finally, using Proposition 4, the convergence in law in $\mathbb{D}([0, T], \mathbb{R})$ of the processes $\mathcal{M}^K$ and V_1^K respectively to zero and V_1 and the convergence in law of $V_3^K(0)$ to $V_0^{(3)}$, we obtain that the sequence $(V_3^K)_K$ converges in law in $\mathbb{D}([0, T], \mathbb{R})$ to the process V_3 , unique solution of the following SDE

$$\forall t \in [0, T], \quad V_3(t) = V_0^{(3)} + \tau_1 (1 + \frac{1}{K^{\gamma_2}}) \int_0^t V_1(s) ds - \tau_3 \int_0^t V_3(s) ds,$$

where V_1 has been defined in (2.37). That ends the proof. $\square$

2.5 Appendix

Lemma 3 (Lemma 2.9 of [54]). *Let V^N be a sequence of $\mathbb{R}_+^3$ -valued processes. We consider its occupation measure defined for D a Borelian set by*

$$\Gamma_N(D \times [0, t]) = \int_0^t \mathbf{1}_{\{D\}}(\mathbf{V}^N(s)) \, ds.$$

Let us assume that there exists a function $\psi : \mathbb{R}_+^3 \rightarrow [1, \infty)$ locally bounded such that $\lim_{v \rightarrow +\infty} \psi(v) = +\infty$ and such that for each $t > 0$,

$$\sup_N \mathbb{E} \left[\int_0^t \psi(V^N(s)) ds \right] < +\infty.$$

Then Γ_N is relatively compact, and if Γ_N converges in law to Γ , then for $f_1, \dots, f_m \in D_\psi$,

$$\begin{aligned} & \left(\int_0^\cdot f_1(V_N(s)) \, ds, \dots, \int_0^\cdot f_m(V^N(s)) \, ds \right) \xrightarrow{\mathcal{L}} \\ & \left(\int_{\mathbb{R}_+^3} f_1(v) \Gamma(dv \times [0, \cdot]), \dots, \int_{\mathbb{R}_+^3} f_m(v) \Gamma(dv \times [0, \cdot]) \right) \end{aligned}$$

where D_ψ denote the collection of continuous functions f satisfying

$$\sup_{v \in \mathbb{R}_+^3} \frac{|f(v)|}{\psi(v)} < \infty \quad \text{and} \quad \lim_{k \rightarrow \infty} \sup_{v \in \mathbb{R}_+^3, \|v\| > k} \frac{|f(v)|}{\psi(v)} = 0. \quad (2.39)$$

Acknowledgements We warmly thank Vincent Bansaye, the hematologist Stéphane Giraudier and the biologist Evelyne Lauret for exciting and fruitful discussions which have motivated this work. We also thank Vincent Bansaye for his precious comments on our paper. This work was supported by a grant from Région Île-de-France.

Chapitre 3

Influence de la quiescence dans le développement d'un mutant : une approche déterministe par morceaux

Résumé Ce chapitre est constitué d'un travail en cours dans lequel nous étudions l'influence de longues phases de quiescence des cellules souches mutantes sur l'évolution au cours du temps du nombre de globules rouges mutants produits par celles-ci. Nous nous intéressons à la dynamique d'un processus stochastique de dimension 4 décrivant l'évolution au cours du temps du nombre de cellules souches et du nombre de globules rouges sains et mutants. Le modèle prend en compte une dynamique d'amplification entre les cellules souches et les globules rouges ainsi qu'une régulation de cette amplification en fonction du nombre total de globules rouges (sains et mutants). Une seule cellule souche mutante basculant aléatoirement d'un état actif à un état quiescent est considérée tandis que les autres populations sont en grand nombre. Nous montrons la convergence en loi de ce modèle vers un processus de Markov déterministe par morceaux (PDMP). Nous étudions ensuite le comportement en temps long de ce processus limite. Nous montrons l'existence et l'unicité d'une mesure de probabilité invariante absolument continue par rapport à la mesure de Lebesgue pour le PDMP limite obtenu. Nous décrirons le support de la probabilité invariante et montrerons que le processus converge en variation totale vers celle-ci. Des simulations numériques de trajectoires de ce processus limite ainsi que des perspectives d'évolution de ce travail seront finalement données en fin de chapitre.

3.1 Introduction

Dans ce chapitre nous nous intéressons à la dynamique des cellules souches hématopoïétiques et à celle des globules rouges qu'elles produisent. La cellule souche est une cellule dont la spécialité est de fournir un apport infini de cellules fraîchement différenciées, en particulier des globules rouges [69]. Deux types de cellules souches sont souvent distingués, les cellules souches quiescentes et les cellules souches actives [32]. Lorsqu'une cellule souche active se divise, chacune des deux cellules filles peut soit rester une cellule souche active comme la cellule parentale, soit entrer irréversiblement dans une succession d'étapes qui mènent à la différenciation terminale [69]. Les cellules souches quiescentes quant à elle ne peuvent pas se diviser. Le changement d'état des cellules souches, activation ou passage en quiescence, se produit en fonction de l'état de stress du système [99].

Nous nous intéressons plus particulièrement à la dynamique d'une seule cellule souche mutante et aux effets de ses phases de quiescence sur la dynamique des globules rouges mutants mais également sur la dynamique des cellules résidentes. Les cellules résidentes sont constituées des cellules souches saines et des globules rouges sains et supposées en grand nombre. Nous verrons qu'elles sont affectées par la dynamique de la cellule souche mutante à travers différentes régulations.

La modélisation mathématique du comportement dynamique des cellules souches mutantes est un enjeu important. De nombreuses mutations au niveau des cellules souches conduisent à des cancers dont les symptômes et le nom diffèrent en fonction de la mutation et de la lignée cellulaire affectée par le dysfonctionnement engendré.

De nombreux modèles mathématiques ont été développés pour comprendre la dynamique de ces différentes cellules souches mutantes ([38, 77]). Ces modèles sont stochastiques ([14, 25, 26, 60]) ou déterministes ([90, 10, 24]) et ne prennent pas en compte des dynamiques à la fois déterministes et stochastiques comme dans notre modèle. Notre modèle suppose l'existence d'une seule cellule souche mutante tandis que les cellules dites résidentes sont en grand nombre. Nous verrons que ces différentes échelles de tailles impliquent une différence d'échelles de temps dans lesquelles les dynamiques des cellules se produisent.

D'autres auteurs allient des méthodes stochastiques et déterministes pour modéliser la dynamique d'un même type cellulaire ([50, 85, 47, 43]). Leur objectif est d'approcher une certaine réalité biologique.

Introduisons un paramètre d'échelle $K \in \mathbb{N}^*$. Nous supposons que le nombre total de cellules souches saines, actives et quiescentes, est constant au cours du temps et égal à K . Ainsi une fois le nombre de cellules souches actives déterminé, on en déduit immédiatement le nombre de cellules souches quiescentes.

Notre modèle décrit la dynamique de quatre types cellulaires au cours du temps. Leurs effectifs respectifs sont notés,

N_1^K : le nombre de cellules souches actives ;

N_2^K : le nombre de globules rouges sains ;

N_3^K : le nombre de globules rouges mutants.

Nous verrons que chacune de ces trois populations sont en grand nombre à la différence de l'unique cellule souche mutante dont l'état sera désigné par le processus I^K ,

$I^K = 1$: lorsque la cellule souche mutante est active ;

$I^K = 0$: lorsque celle-ci est quiescente.

La dynamique de notre modèle est la suivante. Les cellules souches saines (respectivement mutantes) basculent de l'état actif à l'état quiescent à taux constant $a > 0$ (respectivement a_M). Elles basculent de l'état quiescent à l'état actif à un taux q^K (respectivement q_M^K). Les taux q^K et q_M^K sont des fonctions continues croissantes dépendant du nombre de globules rouges sains et du nombre de globules rouges mutants. Une cellule souche active engendre un globule rouge en restant en vie à chaque division asymétrique qu'elle effectue à taux τ pour les saines et à taux τ_M pour la mutante. Ces taux de division sont régulés par les nombres de globules rouges sains et mutants. Ils dépendent de K de la façon suivante,

$$\tau = K^\alpha r^K, \quad \tau_M = K^\beta r_M^K$$

avec $\alpha > 0$ et $\beta > 0$. Les fonctions r^K et r_M^K sont continues décroissantes et bornées. Elles modélisent la régulation de la production des globules rouges sains, respectivement mutant en fonction du nombre à la fois de globules rouges sains et de globules rouges

mutants dans le système. Les puissances α et β résument l'amplification entre le nombre de cellules souches et celui des globules rouges, dû à une succession d'étapes de différenciation intermédiaires ([69]).

Les formes explicites des différents taux seront données en (3.3).

Les globules rouges ont quant à eux un taux de mort individuel constant noté $d > 0$ pour les sains et $d_M > 0$ pour les mutants.

Résumons la dynamique du modèle à l'aide d'un schéma où N^K désigne le vecteur (N_1^K, N_2^K, N_3^K) .

$$\begin{array}{lll}
N_1^K & \longrightarrow & N_1^K + 1 & \text{à taux } a(K - N_1^K) \\
N^K & \longrightarrow & (N_1^K - 1, N_2^K, N_3^K) & \text{à taux } q^K(N_2^K, N_3^K) N_1^K \\
N^K & \longrightarrow & (N_1^K, N_2^K + 1, N_3^K) & \text{à taux } K^\alpha r^K(N_2^K, N_3^K) N_1^K \\
N_2^K & \longrightarrow & N_2^K - 1 & \text{à taux } d N_2^K \\
(N^K, I^K) & \longrightarrow & (N_2^K, N_3^K + 1, I^K) & \text{à taux } K^\beta r_M^K(N_2^K, N_3^K) I^K \\
N_3^K & \longrightarrow & N_3^K - 1 & \text{à taux } d_M N_3^K \\
I^K & \longrightarrow & I^K + 1 & \text{à taux } a_M(1 - I^K) \\
(N_2^K, N_3^K, I^K) & \longrightarrow & (N_2^K, N_3^K, I^K - 1) & \text{à taux } q_M^K(N_2^K, N_3^K) I^K.
\end{array}$$

Rappelons que K représente le nombre total de cellules souches saines au cours du temps. Ainsi $K - N_1^K$ représente le nombre de cellules souches quiescentes saines.

Nos objectifs, dans ce chapitre, sont de décrire une approximation au premier ordre de ce modèle lorsque K tend vers l'infini puis d'étudier le comportement en temps long du processus limite pour comprendre l'influence des phases aléatoires de quiescence de la cellule souche mutante sur la dynamique des trois autres types cellulaires considérés.

Ainsi en Section 3.2, nous présentons le modèle à l'aide de mesures ponctuelles de Poisson. Par des contrôles de moments du processus préalablement renormalisé, nous montrons qu'il admet une décomposition en semi-martingales et qu'il converge en loi, lorsque K tend vers l'infini, vers un processus de Markov déterministe par morceaux (PDMP). Une telle convergence, pour laquelle la limite admet une composante à valeurs discrètes, a déjà été étudiée par Crudu et al (Théorème 3.1 [20]). Nous avons choisi de démontrer ce résultat, par une preuve plus directe (cf Théorème 13).

Les PDMPs sont des processus stochastiques impliquant un mouvement déterministe ponctué de sauts aléatoires. Cette classe de modèles a été introduite dans la littérature par Davis [21, 22] (voir aussi [52]). Ces processus interviennent en biologie ([16]), en épidémiologie [67], en écologie [8], sur le déplacement des bactéries [39] ou sur l'expression génique ([72, 103]). Chacun de ces modèles admet des particularités et des difficultés différentes concernant l'étude de leurs comportements en temps long [73].

Nous étudions ensuite le comportement en temps long du PDMP obtenu en Section 3.3. Dans cette section nous nous sommes inspirés des résultats de [6] (Théorème 2.6) et [67] (Théorème 9) pour développer un ensemble de résultats intermédiaires permettant ensuite de démontrer l'existence et l'unicité d'une probabilité invariante, absolument continue par rapport à la mesure de Lebesgue et vers laquelle le processus converge en variation totale (cf Théorème 14). La démonstration du Théorème 14 repose principalement sur la vérification d'une condition de crochet faible (ou de Hörmander) par les champs de vecteurs

associés au processus permettant d'établir l'absolue continuité de la mesure invariante par rapport à la mesure de Lebesgue. Nous identifions ensuite la densité associée comme solution d'un système d'équations aux dérivées partielles.

Finalement en Section 3.4 nous montrons des simulations numériques de différentes trajectoires du PDMP limite et donnons quelques perspectives.

Notations

Soit $E = [0, 1] \times \mathbb{R}_+^2$ muni de la norme suivante

$$\forall x \in E, \|x\|^2 = \sum_i x_i^2.$$

Soit $\mathcal{E}$ le sous-ensemble de $\mathbb{R}_+^4$

$$\mathcal{E} = E \times \{0, 1\}$$

et l'ensemble

$$C_b^1(\mathcal{E}) = \{f : \mathcal{E} \rightarrow \mathbb{R} \text{ bornée dont la restriction à } E \text{ est } C^1\}. \quad (3.1)$$

3.2 Convergence vers un PDMP

Comme vu en introduction, nous considérons 4 types cellulaires dont les effectifs respectifs au temps t des cellules souches saines actives, des globules rouges sains et des globules mutants sont notés $N^K(t) = (N_1^K(t), N_2^K(t), N_3^K(t))$ tandis que l'état de l'unique cellule souche mutante est noté $I^K(t)$. L'amplification entre les compartiments de cellules souches et de globules rouges est modélisée par un facteur multiplicatif respectivement K^α pour les saines et K^β pour les mutantes. Nous verrons au Lemme 1 que ces facteurs traduisent bien une différence d'ordre de grandeur entre les processus

$$N_1^K \sim K, \quad N_2^K \sim K^{1+\alpha}, \quad N_3^K \sim K^\beta.$$

Afin de modéliser les différentes régulations du système, nous supposons que les fonctions q^K, q_M^K, r^K et r_M^K s'écrivent pour tout $(x_2, x_3) \in \mathbb{R}_+^2$,

$$\begin{cases} q^K(x_2, x_3) = q\left(\frac{x_2}{K^{1+\alpha}}, \frac{x_3}{K^\beta}\right) \\ q_M^K(x_2, x_3) = q_M\left(\frac{x_2}{K^{1+\alpha}}, \frac{x_3}{K^\beta}\right) \\ r^K(x_2, x_3) = r\left(\frac{x_2}{K^{1+\alpha}}, \frac{x_3}{K^\beta}\right) \\ r_M^K(x_2, x_3) = r_M\left(\frac{x_2}{K^{1+\alpha}}, \frac{x_3}{K^\beta}\right) \end{cases} \quad (3.2)$$

où pour tout $(y_2, y_3) \in \mathbb{R}_+^2$,

$$\begin{cases} q(y_2, y_3) = q_1 + q_2 y_2 + q_3 y_3, \\ q_M(y_2, y_3) = q_{1,M} + q_{2,M} y_2 + q_{3,M} y_3 \\ r(y_2, y_3) = \frac{c_1}{1 + c_2 y_2 + c_3 y_3}, \\ r_M(y_2, y_3) = \frac{c_{1,M}}{1 + c_{2,M} y_2 + c_{3,M} y_3} \end{cases} . \quad (3.3)$$

Nous supposons que

- (H1) Les constantes $a, q_1, q_{1,M}, c_1, c_{1,M}, d$ et d_M sont strictement positives ;
- (H2) Les constantes $q_2, q_3, c_2, c_3, q_{2,M}, q_{3,M}, c_{2,M}, c_{3,M}$ sont positives ou nulles.

Notons que les fonctions q et q_M ne sont pas bornées mais vérifient

$$\forall x \in \mathbb{R}^4, \quad 0 \leq |q(x)| + |q_M(x)| \leq q_1 + q_{1,M} + \max(q_2 + q_{2,M}, q_3 + q_{3,M}) \|x\|. \quad (3.4)$$

Rappelons les hypothèses précédemment introduites

$$p.s \quad \forall t \geq 0, \quad N_1^K(t) \in [0, K] \quad \text{et} \quad I^K(t) \in \{0, 1\}. \quad (3.5)$$

Supposons de plus que pour tout $K \in \mathbb{N}^*$, le couple de variables aléatoires réelles positives $(N_2^K(0), N_3^K(0))$ soit intégrable.

Alors N^K définit un processus de saut markovien dont la dynamique est décrite par les équations suivantes.

Soient $(\mathcal{N}_i^j)_{i \in \{1,2,3\}, j \in \{+, -\}}$ huit mesures ponctuelles de Poisson sur $(\mathbb{R}_+ \times \mathbb{R}_+, \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{B}(\mathbb{R}_+))$, indépendantes et d'intensité $dsdu$.

Notons $(\mathcal{F}_t)_{t \geq 0}$ la filtration associée,

$$\mathcal{F}_t = \sigma(\mathcal{N}_i^j([0, s], A); i \in \{1, 2, 3, 4\}, j \in \{+, -\} s \leq t, A \in \mathcal{B}(\mathbb{R}_+)).$$

Alors pour tout $t \geq 0$,

$$\begin{aligned} N_1^K(t) &= N_1^K(0) + \int_0^t \int_{\mathbb{R}_+} \mathbf{1}_{\{u \leq a(\mathbf{K} - N_1^K(s^-))\}} \mathcal{N}_1^+(ds, du) \\ &\quad - \int_0^t \int_{\mathbb{R}_+} \mathbf{1}_{\{u \leq q^K(N_2^K(s^-), N_3^K(s^-)) N_1^K(s^-)\}} \mathcal{N}_1^-(ds, du) \\ N_2^K(t) &= N_2^K(0) + \int_0^t \int_{\mathbb{R}_+} \mathbf{1}_{\{u \leq K^{1+\alpha} r_K(N_2^K(s^-), N_3^K(s^-)) N_1^K(s^-)\}} \mathcal{N}_2^+(ds, du) \\ &\quad - \int_0^t \int_{\mathbb{R}_+} \mathbf{1}_{\{u \leq d N_2^K(s^-)\}} \mathcal{N}_2^-(ds, du) \end{aligned} \quad (3.6)$$

$$\begin{aligned} N_3^K(t) &= N_3^K(0) + \int_0^t \int_{\mathbb{R}_+} \mathbf{1}_{\{u \leq K^\beta r_{K,M}(N_2^K(s^-), N_3^K(s^-)) I^K(s^-)\}} \mathcal{N}_4^+(ds, du) \\ &\quad - \int_0^t \int_{\mathbb{R}_+} \mathbf{1}_{\{u \leq d_M N_3^K(s^-)\}} \mathcal{N}_4^-(ds, du) \\ I^K(t) &= I^K(0) + \int_0^t \int_{\mathbb{R}_+} \mathbf{1}_{\{u \leq a_M(1 - I^K(s^-))\}} \mathcal{N}_3^+(ds, du) \\ &\quad - \int_0^t \int_{\mathbb{R}_+} \mathbf{1}_{\{u \leq q_M^K(N_2^K(s^-), N_3^K(s^-)) I^K(s^-)\}} \mathcal{N}_3^-(ds, du). \end{aligned} \quad (3.7)$$

où les fonctions q^K, q_M^K, r^K et r_M^K sont définies en (3.2)-(3.3).

Définissons le processus stochastique X^K par

$$X^K = \left(\frac{N_1}{K}, \frac{N_2^K}{K^{1+\alpha}}, \frac{N_3^K}{K^\beta} \right).$$

Présentons un résultat de contrôle uniforme des moments d'ordre 2 de la suite $(X^K)_K$.

Lemme 1. *Supposons que le vecteur aléatoire $(X^K(0), I^K(0))$ à valeurs dans $\mathcal{E}$ vérifie*

$$\sup_K \mathbb{E}[\|X^K(0)\|^2] < \infty. \quad (3.8)$$

Alors pour tout $T > 0$,

$$\sup_K \mathbb{E}[\sup_{t \leq T} \|X^K(t)\|^2] < \infty.$$

Démonstration. En utilisant la formule d'Itô, on montre d'après les hypothèses (3.5), les définitions (3.3), un argument de localisation et le lemme de Gronwall (voir par exemple [5]) que

pour tout $T > 0$ et $K \in \mathbb{N}^*$,

$$\mathbb{E}[\sup_{t \in [0, T]} (X_2^K(t))^2] \leq (\mathbb{E}[(X_2^K(0))^2] + 3T) e^{3c_1 T} < \infty$$

et

$$\mathbb{E}[\sup_{t \in [0, T]} (X_3^K(t))^2] \leq (\mathbb{E}[(X_3^K(0))^2] + 3T) e^{3c_{1,M} T} < \infty.$$

On en déduit à l'aide de (3.8) que pour tout $T > 0$,

$$\sup_K \mathbb{E}[\sup_{t \leq T} \|X^K(t)\|^2] < \infty.$$

□

Énonçons finalement le résultat principal de cette section décrivant le comportement asymptotique au premier ordre du processus (X^K, I^K) sur un intervalle de temps fini.

Théorème 13. *Soient $T > 0$ et X^K le processus stochastique à valeurs dans $\mathbb{D}([0, T], E)$ défini par*

$$\forall t \in [0, T], \quad X^K(t) = \left(\frac{N_1^K(t)}{K}, \frac{N_2^K(t)}{K^{1+\alpha}}, \frac{N_3^K(t)}{K^\beta} \right).$$

Supposons que le vecteur aléatoire $(X^K(0), I^K(0))$ converge en loi vers $(x_0, i_0) \in \mathcal{E}$ et vérifie

$$\sup_K \mathbb{E}[\|X^K(0)\|^2] < \infty. \quad (3.9)$$

Alors la suite $((X^K(t), I^K(t)), t \in [0, T])_K$ converge en loi dans $\mathbb{D}([0, T], \mathcal{E})$, lorsque K tend vers l'infini, vers l'unique processus stochastique (X, I) à valeurs dans $\mathcal{E}$ de condition initiale (x_0, i_0) et de générateur infinitésimal $\mathcal{L}$ défini pour $f \in C_b^1(\mathcal{E})$ par

$\forall x \in E, i \in \{0, 1\},$

$$\mathcal{L}f(x, i) = \sum_{j=1}^3 \left(\frac{\partial f}{\partial x_j}(x, i) g_j(x, i) \right) + a_M (1-i)(f(x, i+1) - f(x, i)) + q_M(x_2, x_3) i(f(x, i-1) - f(x, i))$$

où

$$g(x, i) = (a - (a + q(x_2, x_3))x_1, r(x_2, x_3)x_1 - dx_2, r_M(x_2, x_3)i - d_M x_3)^T. \quad (3.10)$$

Les fonctions q , q_M , r et r_M sont définies en (3.3) et l'ensemble $C_b^1(\mathcal{E})$ est défini en (3.1).

Démonstration. On déduit de l'hypothèse (3.9), du Lemme 1 et de l'expression (3.6), la décomposition en semi-martingales du processus (X^K, I^K) suivante

$\forall t \geq 0$,

$$\begin{aligned} X_1^K(t) &= X_1^K(0) + \int_0^t \left(a(1 - X_1^K(s)) - q(X_2^K(s), X_3^K(s))X_1^K(s) \right) ds + M_1^K(t) \\ X_2^K(t) &= X_2^K(0) + \int_0^t \left(r(X_2^K(s), X_3^K(s))X_1^K(s) - dX_2^K(s) \right) ds + M_2^K(t) \\ X_3^K(t) &= X_3^K(0) + \int_0^t \left(r_M(X_2^K(s), X_3^K(s))I^K(s) - d_M X_3^K(s) \right) ds + M_3^K(t) \\ I^K(t) &= N_3^K(0) + \int_0^t \left(a_M(1 - I^K(s)) - q_M(X_2^K(s), X_3^K(s))I^K(s) \right) ds + \widehat{M^K}(t). \end{aligned} \quad (3.11)$$

où les martingales de carrés intégrables $M_i^K, \widehat{M^K}$ admettent pour variations quadratiques

$$\begin{aligned} \langle M_1^K \rangle_t &= K^{-1} \int_0^t \left(a(1 - X_1^K(s)) + q(X_2^K(s), X_3^K(s))X_1^K(s) \right) ds \\ \langle M_2^K \rangle_t &= K^{-(1+\alpha)} \int_0^t \left(r(X_2^K(s), X_3^K(s))X_1^K(s) + dX_2^K(s) \right) ds \\ \langle M_3^K \rangle_t &= K^{-\beta} \int_0^t \left(r_M(X_2^K(s), X_3^K(s))I^K(s) + d_M X_3^K(s) \right) ds \\ \langle \widehat{M^K} \rangle_t &= \int_0^t \left(a_M(1 - I^K(s)) + q_M(X_2^K(s), X_3^K(s))I^K(s) \right) ds \\ \langle M_i^K, M_j^K \rangle_t &= \langle M_i^K, \widehat{M^K} \rangle_t = 0 \quad \text{pour } i \neq j. \end{aligned} \quad (3.12)$$

On vérifie alors aisément les critères de tension d'Aldous et de Rebolledo (voir [53] et [5]). On en déduit la tension uniforme de la suite de loi de $(X^K, I^K)_K$ dans l'ensemble des mesures de probabilités sur $\mathbb{D}([0, T], \mathbb{R}_+^4)$. D'après le théorème de Prohorov (voir Billingsley), il existe alors une probabilité μ limite d'une sous-suite extraite de la suite de loi de $(X^K, I^K)_K$. Dans la suite, la notation $(X^K, I^K)_K$ désignera la sous-suite extraite.

De plus, nous savons que

$$\sup_{t \in [0, T]} \|\Delta X^K(t)\| \leq K^{-1 \wedge \beta}, \quad \beta > 0.$$

Ainsi par continuité de la fonction $(x, i) \in \mathbb{D}([0, T], \mathbb{R}^4) \rightarrow \sup_{t \in [0, T]} \|\Delta x(t)\| \in \mathbb{R}_+$, nous en déduisons que la mesure limite μ charge uniquement l'ensemble des processus à valeurs dans $\mathbb{R}_+^4$ dont les trois premières composantes sont presque sûrement à trajectoires continues.

De plus, d'après l'inégalité de Doob, on sait que pour tout $T > 0$,

$$\mathbb{E} \left[\sup_{t \leq T} |M_i^K(t)|^2 \right] \leq 4 \mathbb{E} \left[\langle M_i^K \rangle_T \right].$$

On obtient alors par (3.12) et Lemme 1, que $\lim_{K \rightarrow \infty} \mathbb{E} \left[\sup_{t \in [0, T]} |M_i^K(t)|^2 \right] = 0$. On en déduit à l'aide de l'inégalité de Markov que les martingales $(M_i^K)_{i \in \{1, 2, 3\}}$ convergent en probabilité pour la norme uniforme vers 0. En particulier elles convergent vers 0 en loi dans $\mathbb{D}([0, T], \mathbb{R}_+)$.

Le g n rateur associ  au processus (X^K, I^K) s' crit pour toute fonction $f \in C_b^1(\mathcal{E})$ et $(x, i) \in \mathcal{E}$,

$$\begin{aligned}\mathcal{L}^K(f)(x, i) = & (f(x + e_1 K^{-1}, i) - f(x, i)) a (1 - x_1) K + (f(x - e_1 K^{-1}, i) - f(x, i)) q(x_2, x_3) x_1 K \\ & + (f(x + e_2 K^{-(1+\alpha)}, i) - f(x, i)) r(x_2, x_3) x_1 K^{1+\alpha} \\ & + (f(x + e_3 K^{-\beta}, i) - f(x, i)) r_M(x_2, x_3) i K^\beta \\ & + (f(x, i + 1) - f(x, i)) a_M (1 - i) + (f(x, i - 1) - f(x, i)) q_M(x_2, x_3) i.\end{aligned}$$

Notons pour $f \in C_b^1(\mathcal{E})$ et $(x, i) \in \mathcal{E}$

$$\mathcal{L}f(x, i) = \sum_{j=1}^3 \left(\frac{\partial f}{\partial x_j}(x, i) g_j(x, i) \right) + a_M (1-i)(f(x, i+1) - f(x, i)) + q_M(x_2, x_3) i (f(x, i-1) - f(x, i)) \quad (3.13)$$

avec

$$g(x, i) = (a - (a + q(x_2, x_3))x_1, r(x_2, x_3)x_1 - dx_2, r_M(x_2, x_3)i - d_M x_3)^T.$$

Alors par un d veloppement de Taylor on obtient

$$\forall f \in C_b^1, \quad \lim_{K \rightarrow \infty} \sup_{(x, i) \in \mathcal{E}} |\mathcal{L}^K(f)(x, i) - \mathcal{L}(f)(x, i)| = 0. \quad (3.14)$$

Soient (X, I) un processus de loi μ et $\xi_t^{K, f}$ la fonction d finie pour $f \in C_b^1(\mathcal{E})$ sur $\mathbb{D}([0, T], \mathcal{E})$ par

$$\xi_t^{K, f}(x, i) = f(x_t, i_t) - f(x_0, i_0) - \int_0^t \mathcal{L}^K(f)(x_s, i_s) ds.$$

D'apr s la formule de Dynkin, de (3.9) et du Lemme 1, la suite $(\xi_t^{K, f}(X^K, N_3^K))_K$ d finit une suite de martingales uniform ment int grables. On en d duit   l'aide de (3.13) que la limite X est une fonction continue solution   I fix  du syst me d' quations

$$\frac{dX(t)}{dt} = g(X(t), I(t))$$

de condition initiale x_0 . D'apr s le Th or me de Cauchy-Lipschitz   I fix  une telle solution est unique. De plus, la loi du processus I est solution du probl me de martingale suivant, pour toute fonction f born e,

$$f(I_t) - f(i_0) - \int_0^t (f(I_s + 1) - f(I_s)) a_M (1 - I_s) + (f(I_s - 1) - f(I_s)) q_M(X_2(s), X_3(s)) I_s ds$$

est une martingale. En effet, le g n rateur associ    I^K s' crit

$$\forall f, \forall t, \quad \mathcal{L}_I^K f(i, t) = a_M (1-i)(f(i+1, t) - f(i, t)) + q_M(X_2^K(t), X_3^K(t)) i (f(i-1, t) - f(i, t)).$$

On en d duit l'existence d'une martingale M de sauts purs telle que

$$\forall t \in [0, T], \quad I(t) = I(0) + \int_0^t a_M - \left(a_M + q_M(X_2(s), X_3(s)) \right) I(s) ds + M(t) \in \{0, 1\}. \quad (3.15)$$

On peut alors appliquer la formule d'Itô à cette expression pour la fonction carrée. Sachant que $I^2 = I$, on obtient, par unicité de la décomposition en semi-martingales de Doob-Meyer, que M est une martingale intégrable de variation quadratique donnée par

$$\langle M \rangle_t = \int_0^t a_M(1 - I(s)) + q_M(X_2(s), X_3(s))I(s) ds.$$

La martingale M de carré intégrale ainsi définie est unique (Théorème 22 p.66 de [66]). On en déduit l'existence d'une mesure ponctuelle de Poisson N sur $(\mathbb{R}_+ \times \mathbb{R}_+, \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{B}(\mathbb{R}_+))$ d'intensité $dsdu$ telle que

$$I(t) = I(0) + \int_0^t \int_{\mathbb{R}_+} \mathbf{1}_{\{u \leq a_M(1 - I(s^-))\}} - \mathbf{1}_{\{a_M(1 - I(s^-)) < u \leq a_M + (q_M(X_2(s), X_3(s)) - a_M)I(s^-)\}} \mathbf{N}(du, ds).$$

L'unicité trajectorielle de (X, I) implique l'unicité de la loi limite μ . On en déduit alors la convergence en loi du processus (X^K, I^K) dans $\mathbb{D}([0, T], \mathcal{E})$ vers le PDMP de générateur infinitésimal $\mathcal{L}$ défini par (3.13). $\square$

Etudions maintenant le comportement en temps du processus limite (X, I) obtenu.

3.3 Comportement en temps long du processus limite.

L'objectif de cette section est de montrer que le processus limite (X, I) , défini au Théorème 13 à l'aide de son générateur infinitésimal $\mathcal{L}$, admet une unique probabilité invariante, absolument continue par rapport à la mesure de Lebesgue, vers laquelle il converge en variation totale.

Pour cela nous allons énoncer différents résultats intermédiaires en suivant le squelette de l'article [7]. Nous allons dans un premier temps définir puis montrer l'existence d'un compact positivement invariant par rapport aux flots associés à la dynamique du processus X . Nous définirons et décrirons ensuite l'ensemble des points accessibles du processus (X_t, I_t) en montrant l'existence d'un unique point stationnaire globalement asymptotiquement stable. L'ensemble de ces résultats intermédiaires nous permettra de démontrer le résultat principal de cette section (Théorème 14). Nous donnerons finalement une description de la probabilité invariante à l'aide d'un système d'équations aux dérivées partielles.

Précisons tout d'abord les notations utilisées.

$$\forall (x, i) \in E \times \{0, 1\} = \mathcal{E}, \forall f \in C_b^1(\mathcal{E}),$$

$$\mathcal{L}f(x, i) = \sum_{j=1}^3 \left(\frac{\partial f}{\partial x_j}(x, i) g_j(x, i) \right) + a_M(1 - i)(f(x, i+1) - f(x, i)) + q_M(x_2, x_3)i(f(x, i-1) - f(x, i))$$

où

$$g(x, i) = (a - (a + q(x_2, x_3))x_1, r(x_2, x_3)x_1 - dx_2, r_M(x_2, x_3)i - d_Mx_3)^T. \quad (3.16)$$

Rappelons les expressions des fonctions de régulations, pour tout $(x_2, x_3) \in \mathbb{R}_+^2$,

$$\begin{cases} q(x_2, x_3) = q_1 + q_2 x_2 + q_3 x_3, \\ q_M(x_2, x_3) = q_{1,M} + q_{2,M} x_2 + q_{3,M} x_3 \end{cases} \quad \text{et} \quad \begin{cases} r(x_2, x_3) = \frac{c_1}{1 + c_2 x_2 + c_3 x_3}, \\ r_M(x_2, x_3) = \frac{c_{1,M}}{1 + c_{2,M} x_2 + c_{3,M} x_3} \end{cases}. \quad (3.17)$$

Notons pour $i \in \{0, 1\}$, ϕ^i le flot associé au champ de vecteurs $g(., i)$, autrement dit la fonction $t \mapsto \phi_t^i(x_0) = \phi^i(x_0, t)$ qui est la solution du problème de Cauchy

$$\frac{dx}{dt}(t) = g(x(t), i) = (g_1(x(t), i), \dots, g_3(x(t), i))^T$$

avec $x(t=0) = x_0$.

Montrons dans un premier temps l'existence d'un compact B positivement invariant par les flots $(\phi^i)_i$.

Proposition 2. *Soit l'ensemble compact $B \subset E$ tel que*

$$B = [b_1, B_1] \times [b_2, \frac{c_1}{d}] \times [0, \frac{c_{1,M}}{d_M}]$$

avec

$$b_1 = \frac{a}{a + q_1 + q_2 \frac{c_1}{d} + q_3 \frac{c_{1,M}}{d_M}} \quad , \quad b_2 = \frac{c_1 b_1}{d (c_2 \frac{c_1}{d} + c_3 \frac{c_{1,M}}{d_M})}$$

et

$$B_1 = \frac{a}{a + q_1},$$

où les paramètres sont définis en (3.17) et vérifient les hypothèses (H2) et (H3).

Alors B est positivement invariant par rapport aux flots ϕ^i , $i \in \{0, 1\}$, autrement dit

$$\forall i \in \{0, 1\}, \forall t \geq 0, \quad \phi_t^i(B) \subset B.$$

De plus pour toute condition initiale appartenant à $E \times \{0, 1\}$ du processus (X, I) , il existe $t \geq 0$ tel que

$$(X(t), I(t)) \in B \times \{0, 1\}.$$

Démonstration. Remarquons dans un premier temps que les deux premières composantes du champ de vecteur $g(., i)$ sont indépendantes de i . Ainsi

$$p.s. \quad \forall t \geq 0, \quad X_1(t) \leq 1$$

et

$$p.s. \quad \forall t \geq 0, \quad X_2(t) \leq \frac{c_1}{d}(1 - e^{-dt}) + X_2(0)e^{-dt}.$$

On en déduit que

$$\forall (x_0, i_0) \in E \times \{0, 1\}, \quad \exists t_1 \geq 0 \text{ tel que } \forall t \geq t_1, \quad X_2(t) \leq \frac{c_1}{d}.$$

De la même manière, on sait que

$$p.s. \quad \forall t \geq 0, \quad I(t) \leq 1 \quad \text{et} \quad X_i(t) \geq 0 \quad \text{pour } i \in \{2, 3\}$$

On en déduit que presque sûrement,

$$\forall t \geq 0, \quad X_3(t) \leq \frac{c_{1,M}}{d_M}(1 - e^{-d_M t}) + X_2(0)e^{-d_M t} \quad \text{et} \quad X_1(t) \leq \frac{a}{a + q_1}(1 - e^{-(a+q_1)t}) + X_1(0)e^{-(a+q_1)t}$$

puis que

$$\forall (x_0, i_0) \in E \times \{0, 1\}, \quad \exists t_2 \geq 0 \text{ tel que } \forall t \geq t_2, \quad X_3(t) \leq \frac{c_{1,M}}{d_M} \quad \text{et} \quad X_1(t) \leq B_1.$$

Par des arguments similaires, on obtient la minoration suivante

$$\forall (x_0, i_0) \in E \times \{0, 1\}, \quad \exists t_3 \geq 0 \text{ tel que } \forall t \geq t_3, \quad X_2(t) \geq b_2 \quad \text{et} \quad X_1(t) \geq b_1.$$

En effet, les fonctions

$$t \rightarrow b_1 \left(1 - e^{-\left(a + q_1 + q_2 \frac{c_1}{d} + q_3 \frac{c_{1,M}}{d_M}\right) t} \right) + x_1(0) e^{-\left(a + q_1 + q_2 \frac{c_1}{d} + q_3 \frac{c_{1,M}}{d_M}\right) t}$$

et

$$t \rightarrow b_2 (1 - e^{-dt}) + x_2(0) e^{-dt}$$

sont respectivement solutions des équations

$$\frac{dy(t)}{dt} = a - \left(a + q_1 + q_2 \frac{c_1}{d} + q_3 \frac{c_{1,M}}{d_M}\right) y(t) \quad \text{et} \quad \frac{dz(t)}{dt} = \frac{c_1 b_1}{1 + \frac{c_2 c_1}{d} + \frac{c_3 c_{1,M}}{d_M}} - dz(t).$$

On a ainsi prouvé et décrit l'existence d'un compact positivement invariant par rapport aux flots ϕ^i , pour $i \in \{0, 1\}$. $\square$

Sans perte de généralité, nous supposons pour la suite que

$$(X(0), I(0)) = (x_0, i_0) \in B \times \{0, 1\}.$$

Définissons la notion de point accessible pour le processus (X, I) .

Posons pour tout $n \in \mathbb{N}^*$,

$$\mathbb{T}_n = \{(\bar{i}, \bar{u}) = (i_1, \dots, i_n), (u_1, \dots, u_n) \in \{0, 1\}^n \times \mathbb{R}_+^n\}.$$

Alors les trajectoires de (X, I) peuvent s'écrire à l'aide des flots $(\phi^i)_i$ de la façon suivante, pour $x \in B$ et $(\bar{i}, \bar{u}) \in \mathbb{T}_n$,

$$\Phi_{\bar{u}}^{\bar{i}}(x) = \phi_{u_n}^{i_n} \circ \dots \circ \phi_{u_1}^{i_1}(x).$$

Soit $x \in B$, alors l'orbite positive de x est donnée par

$$\gamma^+(x) = \{\Phi_{\bar{u}}^{\bar{i}}(x) : (\bar{i}, \bar{u}) \in \bigcup_{n \in \mathbb{N}^*} \mathbb{T}_n\}.$$

On dit alors qu'un point x est accessible d'un singleton $\{y\}$ si $x \in \overline{\gamma^+(y)}$. De façon plus générale, on définit l'ensemble des points accessibles.

Définition 1. *L'ensemble des points accessibles est défini par*

$$\Gamma = \bigcap_{x \in B} \overline{\gamma^+(x)}.$$

La proposition suivante nous permet de décrire plus précisément l'ensemble Γ .

Proposition 3. *L'ensemble des points accessibles du processus (X, I) est donné par $\Gamma = \overline{\gamma^+(p)}$ avec $p = (p_1, p_2, 0) \in B$ tel que*

- *Si $q_2 = 0$, alors*

$$p_1 = \frac{a}{a + q_1} \quad \text{et} \quad p_2 = \begin{cases} \frac{c_1}{d} p_1 & \text{si } c_2 = 0 \\ \frac{-d + \sqrt{d(d + 4c_1c_2 p_1)}}{2d c_2} & \text{sinon} \end{cases}.$$

- *Si $q_2 \neq 0$ et $c_2 = 0$, alors*

$$\begin{cases} p_1 = \frac{d}{2q_2c_1} \left(\sqrt{(a + q_1)^2 + \frac{4aq_2c_1}{d}} - (a + q_1) \right) \\ p_2 = \frac{c_1}{d} p_1 \end{cases}.$$

- *Si $q_2 \neq 0$ et $c_2 \neq 0$ alors p_1 est l'unique solution de l'équation*

$$p_1 = \frac{2dc_2a}{2dc_2(a + q_1) + q_2 (\sqrt{d(d + 4c_1c_2 p_1)} - d)}$$

et

$$p_2 = \frac{-d + \sqrt{d(d + 4c_1c_2 p_1)}}{2d c_2}.$$

Démonstration. Par définition de Γ , on sait que pour tout $p \in B$,

$$\Gamma \subset \overline{\gamma^+(p)}.$$

Montrons que

$$\overline{\gamma^+(p)} \subset \Gamma$$

pour p solution de l'équation

$$g(p, 0) = 0.$$

Commençons par montrer l'unicité d'un tel équilibre. Nous détaillerons ici uniquement le cas où $c_2 \neq 0$ et $q_2 \neq 0$, les autres cas se prouvent par des arguments similaires.

On déduit de la stricte positivité de la constante d et de l'expression de la fonction $g_3(p, 0)$ que $p_3 = 0$. Alors le couple (p_1, p_2) est solution du système

$$\begin{cases} 0 = a(1 - p_1) - (q_1 + q_2 p_2) p_1 \\ 0 = \frac{c_1 p_1}{1 + c_2 p_2} - d p_2 \end{cases}.$$

Le réel p_2 est donc racine du polynôme $P(x) = dc_2 x^2 + dx - c_1 p_1$ dont le discriminant $\Delta = d^2 + 4c_1dc_2 p_1$ est strictement positif. Par conséquent p_2 est défini de manière unique en fonction de p_1 comme racine positive de P ,

$$p_2 = p_2(p_1) = \frac{-d + \sqrt{d(d + 4c_1c_2 p_1)}}{2d c_2}.$$

De plus, p_1 est l'unique solution positive de l'équation

$$p_1 = \frac{a}{a + q_1 + q_2 p_2(p_1)}.$$

En effet, la fonction $p_1 \in [0, 1] \mapsto \frac{2dc_2a}{2dc_2(a+q_1)+q_2(\sqrt{d(d+4c_1c_2p_1)-d})} \in]0, \frac{a}{a+q_1}]$ est strictement décroissante. Elle intersecte donc bien la fonction identité, strictement croissante, en un unique point compris entre 0 et 1.

Pour conclure il suffit donc de montrer que pour tout $\mathcal{U}$ voisinage de p ,

$$\forall x \in B, \exists u \in \mathbb{R}_+ \text{ tel que } \phi_u^{i=0}(x) \in \mathcal{U}.$$

La décroissance vers 0 de la troisième composante du flot $t \mapsto \phi_t^0(x)$ pour tout $x \in B$ est claire puisque la dynamique de celle-ci est donnée par $\forall t \geq 0, \frac{dx_3(t)}{dt} = -d_M x_3(t)$.

Ainsi il nous suffit d'étudier le comportement dynamique du flot associé au champ de vecteurs de $\mathbb{R}_+^2 : G = (g_1(.,., 0, 0), g_2(.,., 0, 0))$. Nous avons montré à la Proposition 2 que ce flot est confiné dans un compact

$$\widehat{B} = [b_1, B_1] \times [b_2, \frac{c_1}{d}].$$

Alors d'après le théorème de Poincaré-Bendixson [9], soit le champ de vecteurs G admet une orbite périodique, soit pour toute la condition initiale appartenant à $\widehat{B}$, le flot associé à G converge vers l'unique point stationnaire appartenant à $\widehat{B}$, c'est à dire (p_1, p_2) .

La divergence du champ de vecteurs G ne s'annule pas sur le compact $\widehat{B}$:

$$\operatorname{div} G(x) = -a - q_1 x_2 - \frac{c_1 c_2 x_1}{(1 + c_2 x_2)^2} - d < 0.$$

Par conséquent G n'admet pas d'orbite périodique (voir Proposition 4 en Annexe). Nous avons donc montré que $p \in \Gamma$ et donc que

$$\overline{\gamma^+(p)} = \Gamma.$$

□

Nous pouvons désormais prouver le théorème principal de cette section en utilisant la théorie développée dans [6] et [67]. Autrement dit nous allons montrer que le processus (X, I) admet une unique probabilité invariante, qu'elle est absolument continue par rapport à la mesure de Lebesgue, que son support est donné par $\Gamma \times \{0, 1\}$ et que le processus converge vers elle en variation totale.

Pour cela, introduisons la condition dite de crochet faible ainsi que les conditions (i) et (ii) que nous allons vérifier pour un point $p \in \Gamma$ bien choisi.

Définition 2. Notons $\mathcal{F}_0 = \{g(., i)\}_{i \in \{0, 1\}}$ puis $\mathcal{F}_{j+1} = g(., j) \cup \{[g(., i), V], V \in \mathcal{F}_k, i \in \{0, 1\}\}$ où

$$[V, W](x) = DW(x)V(x) - DV(x)W(x), \quad x \in \mathbb{R}^d.$$

Alors on dit que la condition de crochet faible est vérifiée en x si

$$\operatorname{Vect}(V(x) : V \in \cup_{j \geq 0} \mathcal{F}_j) = \mathbb{R}^3.$$

Nous dirons de plus que la condition (i) est vérifiée s'il existe $s \in \mathbb{R}$ tel que

$$s g(p, 0) + (1 - s) g(p, 1) = 0$$

et que la condition (ii) est vérifiée s'il existe $x \in B$ accessible de $\{p\}$ en lequel la condition de crochet faible est vérifiée.

Le théorème de convergence en temps long du processus (X, I) s'énonce alors.

Théorème 14. *Supposons que l'ensemble des constantes du modèle vérifient les hypothèses (H2) et (H3) formulées en Section 3.2 et*

$$c_3 + q_3 > 0. \quad (3.18)$$

Alors le processus (X, I) admet une unique probabilité invariante π absolument continue par rapport à la mesure de Lebesgue et de support $\Gamma \times \{0, 1\}$. De plus il existe des constantes strictement positives C et γ telles que pour tout $t \geq 0$ et pour tout $(x, i) \in B \times \{0, 1\}$,

$$\| P_t((x, i), \cdot) - \pi \|_{VT} \leq C e^{-\gamma t}.$$

Démonstration. La preuve repose sur la vérification des conditions (i) et (ii).

La condition (i) est vérifiée par le point stationnaire accessible p défini au Théorème 3. En effet, $g(p, 0) = 0$ par conséquent $s = 1$ vérifie la condition.

Montrons maintenant que la condition (ii) est également vérifiée.

Pour simplifier les expressions, nous noterons

$$T(x) = 1 + c_2 x_2 + c_3 x_3 > 0, \quad T_M(x) = 1 + c_{2,M} x_2 + c_{3,M} x_3 > 0.$$

Le champ de vecteurs $g(\cdot, i)$ s'écrit alors pour $(x, i) \in \mathcal{E}$,

$$g(x, i) = \left(a(1 - x_1) - q(x_2, x_3)x_1, \frac{c_1 x_1}{T(x)} - dx_2, \frac{c_{1,M}}{T_M(x)} i - d_M x_3 \right)^T \quad (3.19)$$

et sa différentielle s'écrit

$$Dg(\cdot, i)(x) = \begin{pmatrix} -a - q(x_2, x_3) & -q_2 x_1 & -q_3 x_1 \\ \frac{c_1}{T(x)} & \frac{-c_1 c_2 x_1}{T(x)^2} - d & \frac{-c_1 c_3 x_1}{T(x)^2} \\ 0 & \frac{-c_{1,M} c_{2,M} i}{T_M(x)^2} & \frac{-c_{1,M} c_{3,M} i}{T_M(x)^2} - d_M \end{pmatrix}.$$

Par (2), on sait que

$$\forall x \in B, \quad [g(\cdot, 0), g(\cdot, 1)](x) = Dg(\cdot, 0)(x) g(x, 1) - Dg(\cdot, 1)(x) g(x, 0).$$

Ainsi par le calcul,

$$[g(\cdot, 0), g(\cdot, 1)](x) = \left(-q_3 x_1 \frac{c_{1,M}}{T_M(x)}, -\frac{c_1 c_3 x_1 c_{1,M}}{T(x)^2 T_M(x)}, V_3(x) \right)$$

$$\text{avec } V_3(x) = -d_M \frac{c_{1,M}}{T_M(x)} + \frac{c_{1,M} c_{2,M} i}{T_M(x)^2} \left(\frac{c_1 x_1}{T(x)} - dx_2 \right) - \frac{c_{1,M} c_{3,M} i}{T_M(x)^2} d_M x_3.$$

Vérifier la condition (ii) signifie prouver l'existence de $x \in \Gamma$ tel que

$$\forall(\alpha_1, \alpha_2) \in \mathbb{R}^2, \quad [g(., 0), g(., 1)](x) \neq \alpha_1 g(x, 0) + \alpha_2 g(x, 1).$$

Ainsi nous pouvons vérifier la condition (ii) à l'aide du raisonnement par l'absurde suivant :

Supposons que pour tout $x \in \Gamma$, il existe $\alpha \in \mathbb{R}$ tel que

$$\begin{cases} -q_3 x_1 \frac{c_{1,M}}{T_M(x)} = \alpha \left(a(1 - x_1) - q(x_2, x_3) x_1 \right) \\ -\frac{c_1 c_3 x_1 c_{1,M}}{T(x)^2 T_M(x)} = \alpha \left(\frac{c_1 x_1}{T(x)} - d x_2 \right) \end{cases}. \quad (3.20)$$

Alors pour tout $x \in \Gamma$, tel que

$$a(1 - x_1) \neq q(x_2, x_3) x_1, \quad (3.21)$$

on obtient l'égalité suivante

$$\frac{c_1 c_3}{T(x)^2} = \frac{q_3}{a(1 - x_1) - q(x_2, x_3) x_1} \left(\frac{c_1 x_1}{T(x)} - d x_2 \right).$$

D'après l'hypothèse (3.18), on en déduit que, pour tout $x \in \Gamma$ vérifiant (3.21),

$$x_1 = \frac{c_1 c_3 a + d q_3 T(x)^2 x_2}{q_3 c_1 T(x) + c_1 c_3 (a + q(x_2, x_3))}. \quad (3.22)$$

Il existe alors deux cas de figures :

- $q_2 \neq 0$ ou $q_3 \neq 0$. Notons $x = \phi_t^{i=1}(p)$ avec $t > 0$ et $y = \phi_u^{i=0}(x)$ avec $u > 0$, deux points appartenant à Γ . Pour t suffisamment grand et u suffisamment petit, les points x et y vérifient (3.21). De plus supposons que le point x vérifie également (3.22). Dans le cas contraire on obtiendrait une contradiction avec l'hypothèse (3.20).

Remarquons que le champ de vecteurs $g(x, i)$ ne dépend de i qu'au travers de sa troisième composante. Ainsi pour u suffisamment petit, on obtient $x_1 = y_1$, $x_2 = y_2$ et $x_3 \neq y_3$. Nous allons utiliser cette remarque pour montrer l'existence de u suffisamment petit pour lequel y ne vérifie pas (3.22). Nous aurons alors une contradiction avec l'hypothèse (3.20) et vérifié la condition (ii).

Afin d'étudier les variations de la fonction

$$x_3 \mapsto \frac{c_1 c_3 a + d q_3 T(x)^2 x_2}{q_3 c_1 T(x) + c_1 c_3 (a + q(x_2, x_3))} \quad (3.23)$$

réécrivons la sous la forme simplifiée suivante

$$x \mapsto \frac{S_1 + S_2(1 + S_3 x)^2}{S_4 + S_5 x}$$

où $S_i \geq 0$ et $S_4 = q_3 c_1 (1 + c_2 x_2) + c_1 c_3 (a + q_1 + q_2 x_2) > 0$.

La dérivée d'une telle fonction admet pour numérateur le polynôme de degré maximum 2 suivant

$$P(x) = 2S_2S_3(1 + S_3x)(S_4 + S_5x) - (S_1 + S_2(1 + S_3x)^2)S_5.$$

Un tel polynôme ne peut être nul sur l'intervalle $[x_3, y_3]$, sans admettre une infinité de racines et être nul sur $\mathbb{R}_+$. Or si $c_3 \neq 0$, la fonction (3.23) n'est pas constante. Ainsi si $c_3 \neq 0$, il existe u suffisamment petit tel que $y \in \Gamma$ ne puisse vérifier à la fois (3.21) et (3.22).

Finalement si $c_3 = 0$, par l'hypothèse (3.18), $q_3 \neq 0$. Dans ce cas il existe $t > 0$ tel que

$$g_1(x, 0) \neq 0 \quad \text{et} \quad g_2(x, 1) \neq 0.$$

Par conséquent, la condition (3.20) n'est pas vérifiée en x car $x_1 > 0$ et $c_{1,M} > 0$.

- $q_2 = q_3 = 0$. Alors, il n'existe pas de point $x \in \Gamma$, vérifiant la condition (3.21). En effet, pour tout $x \in \Gamma$, $x_1 = \frac{a}{a+q_1}$ car la dynamique de X_1 est alors indépendante du processus stochastique I .

Dans ce cas pour vérifier la condition (ii), il suffit de prouver l'existence de $x \in \Gamma$ tel que

$$\forall \alpha \in \mathbb{R}, \quad (g_2(x, 0), g_3(x, 0)) \neq \alpha (g_2(x, 1), g_3(x, 1)).$$

Or il existe $t > 0$ tel que $x = \phi_t^{i=1}(p) \in \Gamma$ vérifie cette condition. En effet sachant que p est l'unique point stationnaire du champ $g(\cdot, 0)$, pour t suffisamment grand, on a d'une part que $g_2(x, 0) = g_2(x, 1) \neq 0$ et d'autre part que $g_3(x, 0) \neq g_3(x, 1)$.

On peut alors conclure la preuve en utilisant le Théorème 16 en Annexe. $\square$

Nous terminons cette section en décrivant plus précisément, à l'aide du générateur de (X, I) , la probabilité invariante π obtenue au Théorème 14.

Théorème 15. *Supposons les hypothèses du Théorème 14 vérifiées. Alors il existe un couple de fonctions intégrables positives (h_0, h_1) solution du système d'équations aux dérivées partielles suivant sur $\overset{\circ}{B} =]b_1, B_1[\times]b_2, \frac{c_1}{d}[\times]0, \frac{c_{1,M}}{d_M}[$,*

$$\begin{cases} g(x, 0)^T \nabla h_0(x) + h_0(x) \left[\sum_{j=1}^3 \partial_j g_j(x, 0) - a_M \right] & = -q_M(x_2, x_3) h_1(x) \\ g(x, 1)^T \nabla h_1(x) + h_1(x) \left[\sum_{j=1}^3 \partial_j g_j(x, 1) - q_M(x_2, x_3) \right] & = -a_M h_0(x) \end{cases}$$

telles que pour tout $i \in \{0, 1\}$,

$$\pi(dx, \{i\}) = \sum_{j=0}^1 \delta_{ij} h_j(x) dx \quad \text{et} \quad \int_B (h_0(x) + h_1(x)) dx = 1.$$

Démonstration. Notons $C_c^1(\overset{\circ}{B})$ (respectivement $C^1(B)$) l'ensemble des fonctions $f : \mathbb{R}_+^3 \times \{0, 1\}$, telles que pour tout $i \in \{0, 1\}$, la fonction $f(\cdot, i)$ est C^1 à support dans l'intérieur de B (respectivement dans B).

On souhaite trouver $\pi \in \mathcal{P}(B \times \{0, 1\})$ telle que

$$\forall f \in C^1(B), \quad \int_{B \times \{0,1\}} \mathcal{L}f(z) \pi(dz) = 0 \quad (3.24)$$

où

$$\mathcal{L}f(x, i) = g(x, i)^T \nabla_x f(x, i) + lf(x, i)$$

et

$$lf(x, i) = a_M(1 - i)(f(x, i + 1) - f(x, i)) + q_M(x_2, x_3) i (f(x, i - 1) - f(x, i)).$$

On sait d'après le Théorème 14 qu'il existe deux fonctions h_0 et h_1 intégrables telles que pour $i \in \{0, 1\}$,

$$\pi(dx, \{i\}) = \sum_{j=0}^1 \delta_{ij} h_j(x) dx.$$

Ainsi

$$\int_{B \times \{0,1\}} \mathcal{L}f(z) \mu(dz) = \int_B \sum_{i=0}^1 h_i(x) [g(x, i)^T \nabla_x f(x, i) + lf(x, i)] dx. \quad (3.25)$$

Or $\forall x \in B$,

$$\sum_{i=0}^1 h_i(x) lf(x, i) = (f(x, 1) - f(x, 0))(a_M h_0(x) - q_M(x_2, x_3) h_1(x)).$$

De plus, en intégrant par parties (3.25), on obtient pour tout $i \in \{0, 1\}$, $j \in \{1, 2, 3\}$ et pour toute fonction $f \in C_c^1(\overset{\circ}{B})$,

$$\int_B h_i(x) g_j(x, i) \partial_j f(x, i) dx = 0 - \int_B \partial_j (h_i(x) g_j(x, i)) f(x, i) dx.$$

Finalement, pour toute fonction $f \in C_c^1(\overset{\circ}{B})$,

$$0 = \sum_{i=0}^1 \int_B \left[\sum_{j=1}^3 \partial_j (h_i(x) g_j(x, i)) + (1 - 2i) (q_M(x_2, x_3) h_1(x) - a_M h_0(x)) \right] f(x, i) dx. \quad (3.26)$$

Pour $\overset{\circ}{B} =]b_1, B_1[\times]b_2, \frac{c_1}{d}[\times]0, \frac{c_{1,M}}{d_M}[$,

$$\forall f, g \in L^2(\overset{\circ}{B} \times \{0, 1\}), \quad \langle f, g \rangle = \int_{\Omega \times \{0,1\}} f(x, i) g(x, i) dx (\delta_0 + \delta_1)(i)$$

définit un produit scalaire dans $L^2(\overset{\circ}{B} \times \{0, 1\})$.

On en déduit que le couple de fonctions (h_0, h_1) est solution du système d'équations aux dérivées partielles suivant sur $\overset{\circ}{B}$,

$$\begin{cases} g(x, 0)^T \nabla h_0(x) + h_0(x) \left[\sum_{j=1}^3 \partial_j g_j(x, 0) - a_M \right] & = -q_M(x_2, x_3) h_1(x) \\ g(x, 1)^T \nabla h_1(x) + h_1(x) \left[\sum_{j=1}^3 \partial_j g_j(x, 1) - q_M(x_2, x_3) \right] & = -a_M h_0(x) \end{cases} \quad (3.27)$$

Ceci conclut la preuve. $\square$

Remarque 5. D'après le Théorème 14, le support de la probabilité invariante π est donné par $\Gamma \times \{0, 1\} \subset B \times \{0, 1\}$. On en déduit qu'en dehors de B , plus précisément en dehors de Γ , les fonctions $(h_i)_{i \in \{0, 1\}}$ sont nulles.

Remarquons de plus que si $q_2 = q_3 = 0$ alors $]b_1, B_1[= \emptyset$. Dans ce cas,

$$h_0(x) = \delta_{\{x_1 = \frac{a}{a+q_1}\}} H_0(x_2, x_3), \quad h_1(x) = \delta_{\{x_1 = \frac{a}{a+q_1}\}} H_1(x_2, x_3)$$

avec pour tout $(y, z) \in]b_2, \frac{c_1}{d}[\times]0, \frac{c_{1,M}}{d_M}[$,

$$\begin{cases} f_1(y, z) \partial_1 H_0(y, z) - d_M z \partial_2 H_0(y, z) + (f_2(y, z) - a_M) H_0(y, z) & = -q_M(y, z) H_1(y, z) \\ f_1(y, z) \partial_1 H_1(y, z) + f_3(y, z) \partial_2 H_1(y, z) + (f_4(y, z) - q_M(y, z)) H_1(y, z) & = -a_M H_0(y, z) \end{cases} \quad (3.28)$$

où

$$f_1(y, z) = \frac{c_1 a}{(a + q_1)(1 + c_2 y + c_3 z)} - d y, \quad f_2(y, z) = \frac{c_1 c_2 a}{(a + q_1)(1 + c_2 y + c_3 z)^2} - d - d_M$$

et

$$f_3(y, z) = \frac{c_{1,M}}{1 + c_{2,M} y + c_{3,M} z} - d_M y, \quad f_4(y, z) = f_2(y, z) + \frac{c_{1,M} c_{3,M}}{(1 + c_{2,M} y + c_{3,M} z)^2}.$$

3.4 Simulations et perspectives

Dans cette section nous présentons dans un premier temps une trajectoire du processus (X, I) puis nous donnons quelques perspectives d'évolution du travail présenté en nous appuyant sur des simulations numériques.

La figure 3.1 représente une trajectoire du processus $(KX_1(t), K^{1+\alpha}X_2(t), K^\beta X_3(t), I)_t$ de condition initiale donnée par $(Kp_1, K^{1+\alpha}p_2, K^\beta p_3, 1)$ où p est l'unique point stationnaire du champ de vecteur $g(\cdot, 0)$ défini en Proposition 3. Les constantes du processus sont toutes supposées égales à 1, $K = 10^4$, $\alpha = 1.5$ et $\beta = 1.7$.

Nous avons vu dans ce chapitre que les nombres de cellules souches actives saines, de globules rouges sains et de globules rouges mutants sont approchés lorsque K est grand par

$$N_1^K \sim KX_1, \quad N_2^K \sim K^{1+\alpha}X_2 \quad \text{et} \quad N_3^K \sim K^\beta X_3.$$

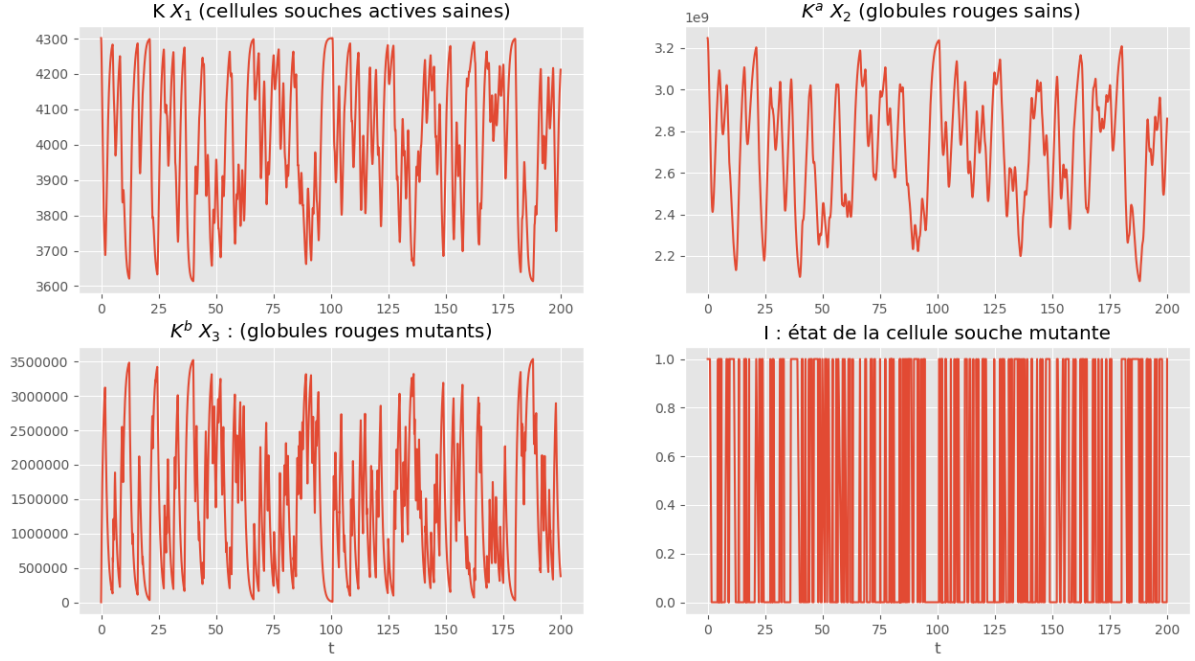


FIGURE 3.1 – Une trajectoire du processus $(KX_1(t), K^{1+\alpha}X_2(t), K^\beta X_3(t), I)_t$ pour $t \in [0, 200]$ (unité de temps). Le terme $a = 1 + \alpha$ et le terme $b = \beta$.

On observe que le nombre de globules rouges mutants réagit très rapidement et fortement au changement d'état de l'unique cellule souche mutante. Il serait donc intéressant de considérer un nombre toujours fini et faible mais plus élevé de cellules souches mutantes.

De plus l'amplification entre les cellules souches et les globules rouges dans le modèle considéré ici est représentée par un taux de division asymétrique des cellules souches proportionnel à une puissance de K . Une telle modélisation génère la réaction presque instantanée et très forte observée en Figure 3.1 sur la dynamique du nombre de globules rouges mutants. Or nous avons étudié dans le Chapitre 2, une dynamique d'amplification différente et plus réaliste d'un point de vue biologique. En effet dans le modèle du Chapitre 2, un compartiment intermédiaire de cellules ayant la capacité d'effectuer des divisions symétriques joue le rôle d'amplificateur. Plus précisément ces cellules intermédiaires se divisent à taux constant en deux cellules de type intermédiaire avec une probabilité p^R ou en deux globules rouges avec probabilité $p^D = 1 - p^R$. L'amplification en K^α se faisait alors à l'aide d'une hypothèse sur la différence entre ces deux probabilités

$$p^D - p^R \sim K^{-\alpha}$$

et/ou d'une hypothèse sur les taux de mort

$$d \sim K^{-\alpha}.$$

Nous avons vu au Chapitre 2 que de telles amplifications ne s'observaient qu'au bout d'un temps d'ordre K^α . Ainsi allier en un seul modèle la régulation considérée dans le Chapitre 3 avec la dynamique d'amplification lente du modèle étudié au Chapitre 2 semble tout à fait pertinent pour étudier l'influence de longues phases de quiescence des cellules souches mutantes sur la dynamique des cellules hématopoïétiques mutantes et saines. En effet, reprenons la simulation précédente pour laquelle nous supposons que les taux de mort des

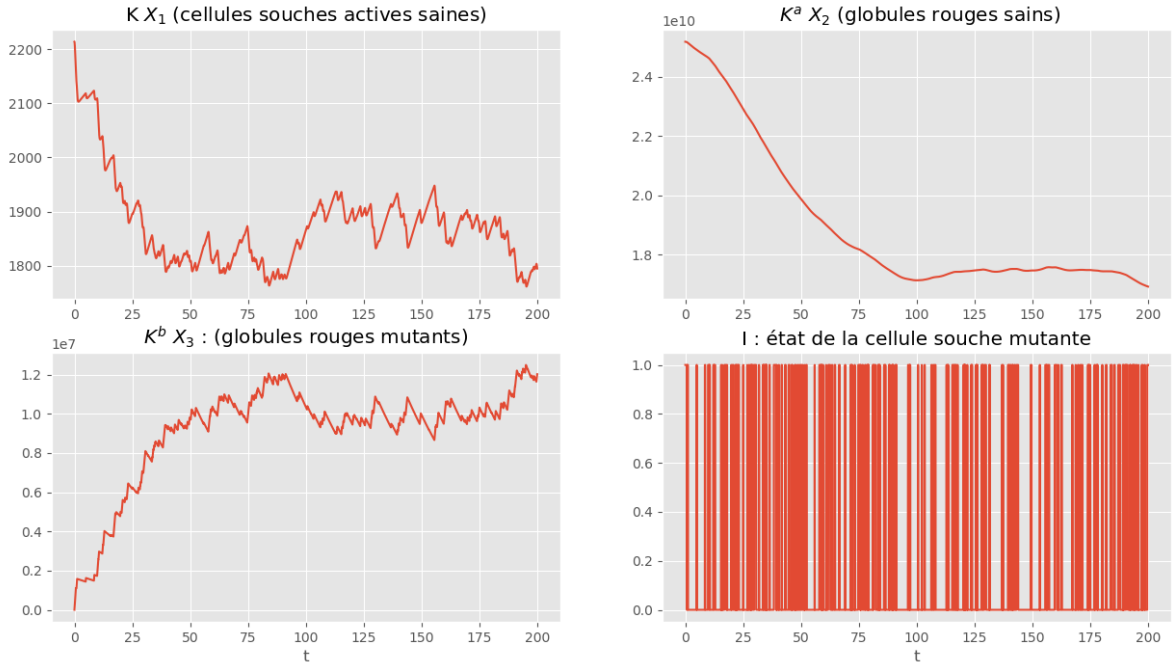


FIGURE 3.2 – Une trajectoire du processus $(KX_1(t), K^{1+\alpha}X_2(t), K^\beta X_3(t), I)_t$ pour $t \in [0, 200]$ (jours), et $d = d_M = 1/40$. Le terme $a = 1 + \alpha$ et le terme $b = \beta$.

globules rouges sains et mutants sont plus faibles $d = d_M = 1/40$. On obtient la trajectoire présentée en Figure 3.2

On remarque donc bien que lorsque le taux de mort est supposé faible, la dynamique des globules rouges mutants est très différente.

Toutefois l'ajout de compartiments intermédiaires de cellules saines et mutants impliquerait l'étude d'un PDMP de dimension 6, ce qui semble constituer un intéressant challenge.

3.5 Annexe

Théorème 16. (Théorème 2.6 et Corollaire 2.7 de [6], Théorème 9 de [67])

Notons (X, I) un processus de Markov déterministe par morceaux à valeurs dans $B \times \{1, \dots, N\}$, où B est un compact positivement invariant par les flots associés à la dynamique de X et $N \in \mathbb{N}^*$. Notons Γ l'ensemble des points accessibles du processus X et $(g(\cdot, i))_{i \in \{1, \dots, N\}}$ les champs de vecteurs associés à la dynamique de X .

On suppose que

(i) Il existe $\alpha_1, \dots, \alpha_N \in \mathbb{R}$ tels que $\sum_i \alpha_i = 1$ et $p \in B$ un point accessible tel que

$$\sum_i \alpha_i g(p, i) = 0,$$

(ii) Il existe $x \in B$ un point accessible de p pour lequel la condition de crochet faible est vérifiée.

Alors le processus (X, I) admet une unique probabilité invariante π absolument continue par rapport à la mesure de Lebesgue et de support $\Gamma \times \{1, \dots, N\}$. De plus il existe des constantes C et γ telles que pour tout $t \geq 0$ et $(x, i) \in B \times \{1, \dots, N\}$,

$$\|P_t((x, i), \cdot) - \pi\|_{VT} \leq C e^{-\gamma t}$$

où P_t est le semi groupe associé à (X, I) .

Proposition 4. (Proposition 8.25 de [23]). Soit $X : U \rightarrow \mathbb{R}^2$ un champ de vecteurs de classe C^1

$$X(x, y) = (f(x, y), g(x, y)).$$

On suppose que l'ouvert $U \subset \mathbb{R}^2$ est simplement connexe et que la divergence

$$\operatorname{div} X := \frac{\partial f}{\partial x} + \frac{\partial g}{\partial y}$$

ne s'annule pas. Alors X n'a pas d'orbite périodique non stationnaire.

Démonstration. On procède par l'absurde. Soit γ une orbite périodique non triviale de X . Puisque l'ouvert U est simplement connexe, γ délimite un domaine $\Omega \subset U$. Ainsi à l'aide de la formule de Green-Riemann, on en déduit (en fonction du sens de parcours de γ) que

$$\int_{\gamma} f dy - g dx = \pm \int_{\Omega} \left(\frac{\partial f}{\partial x} + \frac{\partial g}{\partial y} \right) dx dy = \pm \int_{\Omega} \operatorname{div} X dx dy$$

est non nul par hypothèse. On montre alors une contradiction, en utilisant un changement de variable et le fait que γ est une orbite périodique,

$$\int_{\gamma} [f dy - g dx] = \int_{\gamma} [f \circ g(\gamma(t)) - g \circ f(\gamma(t))] dt = 0.$$

□

Chapitre 4

Multiscale population dynamics in reproductive biology.

This chapter corresponds to paper [11] "Multiscale population dynamics in reproductive biology : singular perturbation reduction in deterministic and stochastic models", in collaboration with K. Chahour, F. Clément, M. Postel and R. Yvinec and accepted in *ESAIM : Proceedings and surveys*.

Abstract. In this study, we describe different modeling approaches for ovarian follicle population dynamics, based on either ordinary (ODE), partial (PDE) or stochastic (SDE) differential equations, and accounting for interactions between follicles. We put a special focus on representing the population-level feedback exerted by growing ovarian follicles onto the activation of quiescent follicles. We take advantage of the timescale difference existing between the growth and activation processes to apply model reduction techniques in the framework of singular perturbations. We first study the linear versions of the models to derive theoretical results on the convergence to the limit models. In the nonlinear cases, we provide detailed numerical evidence of convergence to the limit behavior. We reproduce the main semi-quantitative features characterizing the ovarian follicle pool, namely a bimodal distribution of the whole population, and a slope break in the decay of the quiescent pool with aging.

Résumé. Dans cette étude, nous décrivons différentes approches de modélisation de la dynamique des populations de follicules ovariens, basées sur des équations différentielles ordinaires (EDO), aux dérivées partielles (EDP) ou stochastiques (SDE), et tenant compte des interactions entre follicules. Nous avons mis un accent particulier sur la représentation des rétro-actions exercées par les follicules en croissance sur l'activation des follicules quiescents. Nous tirons parti de la différence d'échelle de temps entre les processus de croissance et d'activation pour appliquer des techniques de réduction de modèle dans le cadre des perturbations singulières. Nous étudions d'abord les versions linéaires du modèle afin d'en déduire des résultats théoriques sur la convergence vers le modèle limite. Dans le cas non linéaire, nous fournissons des arguments numériques détaillés sur la convergence vers le comportement limite. Nous reproduisons les principales caractéristiques semi-quantitatives caractérisant le pool de follicules ovariens, à savoir une distribution bimodale de la population totale et une rupture de pente dans la décroissance du pool de follicules quiescents avec le vieillissement.

Introduction

In mammals, the pool of oocytes (egg cells) available for a female throughout her reproductive life is fixed very early, either during the fetal life or in the perinatal period. All along their maturation, oocytes are sheltered within spheroidal somatic structures called ovarian follicles. Folliculogenesis is the process of growth and maturation undergone by ovarian follicles from the time they leave the pool of quiescent, primordial follicles until ovulation, when they release a fertilizable oocyte.

Follicle growth is first due to the enlargement of the oocyte, then to the proliferation of somatic cells organized into successive concentric cell layers, and finally to the inflation of a fluid-filled cavity (antrum) that forms above a critical size. The activation of primordial follicles can occur at any time once they are formed [78], even if they can remain quiescent for up to tens of years [84]. Growing follicles can progress along the first developmental stages (known as “basal development”) before puberty. The final developmental stages (known as “terminal development”) can only occur after puberty; they are related to the dynamics of ovarian cycles, involving endocrine feedback loops between the ovaries on one side, and the hypothalamus and pituitary gland on the other side. The whole sequence of development spans several months, as assessed by cell kinetics studies [45] or grafts of ovarian cortex [28]. The terminal stages are the shortest; they cover a few weeks at most. Since follicle activation is asynchronous, all developmental stages can be observed in the ovaries at any time during reproductive life. The follicle distribution (mostly studied using the size as a maturity marker) has a characteristic bimodal pattern, which is remarkably preserved between species. This pattern remains similar with ovarian aging, yet with a decreased amplitude [30], as a result of the progressive exhaustion of the quiescent pool. Such a distribution is shaped not only by the differences in the follicle activation times, but also by the hormonal interactions between follicles [80]. In particular, the activation and growth rates in the earliest stages are moderated by the Anti-Müllerian Hormone (AMH) secreted locally by the subpopulation of “intermediary” follicles (rigorously speaking : from the fully activated one-layer stage to the pre-antral and small antral stages) [95]. At the other end, the selection of ovulatory follicles results from a competition-like process operating amongst terminally developing follicles [15], which is mediated by endocrine controls and associated with a species-specific number of ovulations. Namely, inhibin (a peptid hormone) and estradiol, produced by the mature follicles, feedback onto the pituitary gland, leading to a drop in a pituitary hormone (the Follicle-Stimulating Hormone, FSH) supporting follicle survival.

Less than one in a thousand of the follicles manage to reach the ovulatory stage. All others disappear through a degeneration process (atresia) associated with the death of the somatic cells (during mid and terminal folliculogenesis) or oocyte (in the quiescent pool and during early folliculogenesis). For instance, in humans, the quiescent pool size is of the order of 1 million follicles, amongst which only some hundreds will reach ovulation [79].

Experimentalist investigators have proposed classifying follicle development into different stages, according to morphological and functional criteria such as follicle and oocyte diameters, number of cell layers, number of somatic cells, antrum formation [45, 82]. Hence, a natural formalism to consider when modeling follicle population dynamics is that of compartmental modeling, using either deterministic or stochastic rates for transfer (λ_i) and exit (μ_i) rates (see Figure 4.1). Pioneering studies (see e.g. [36]) have focused on fitting the parameters entering these rates according to follicle numbers available in each deve-

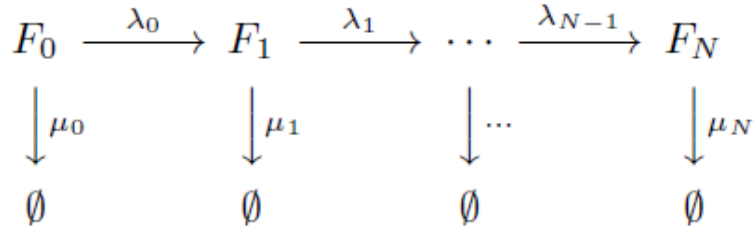


FIGURE 4.1 – Scheme of the dynamics

lopmental stage. However, these studies remained rather descriptive and considered at best possibly time-varying (piecewise constant) rates [37], yet with no follicle interaction. Most of the classification criteria change in a continuous manner. In addition, the most common variable available to monitor follicle development on the ovarian scale, follicle size, is an intrinsically continuous variable. Hence, another suitable modeling formalism is that of PDE models for structured population dynamics. Although the interest of such a formalism has been pointed out quite early [76], it has yet not been implemented.

Finally, in some situations, a discrete stochastic formalism can be useful both to handle finite-size effects and follow individual follicle trajectories. This is especially true for the relatively small cohort of terminally developing follicles, and for transient physiological regimes when follicle pools are either still replenishing, or, on the contrary getting progressively exhausted. In any case, such a formalism gives insight into the fluctuations around the average deterministic behavior.

In this study, we describe different modeling approaches for follicle population dynamics, based on either ODE, PDE or SDE, and accounting for interactions between follicles. We put a special focus on representing the population-level modulation exerted by growing ovarian follicles on the activation of quiescent, primordial follicles. We take advantage of the timescale difference between the growth and activation processes to apply model reduction techniques in the framework of singular perturbations (slow/fast systems).

The paper is organized as follows. We successively introduce the ODE, PDE and SDE formulation of the model for follicle population dynamics. We describe the initial (non-rescaled) model in the ODE case. In each case, we introduce (i) the model in rescaled timescale exhibiting a slow/fast structure with a small perturbation parameter ($\varepsilon > 0$) and (ii) the model in the limit $\varepsilon \rightarrow 0$. We discuss the well-posedness of the limit models in two situations : the linear formulation and a weakly nonlinear formulation in which only the quiescent follicle population is subject to a feedback from the remaining of the population. In the linear case, we prove the convergence of the rescaled to the limit models. In the nonlinear case, we provide detailed numerical evidence of convergence. The numerical illustrations are settled within a biologically-realistic framework, allowing us to reproduce the main semi-quantitative features characterizing the dynamics of the ovarian follicle pool, namely a bimodal distribution of the whole population and a slope break in the decay of the quiescent pool with aging.

4.1 Compartmental, ODE-based model

4.1.1 Initial model

Starting from the schematic model (see Eq. (4.1)), we formalize a system of nonlinear ordinary differential equations (ODE) as follows. Let $d \in \mathbb{N}^*$ and $y = (y_0, \dots, y_d)$ be a function such that, for all $i \in \{0, \dots, d\}$, $y_i : t \in \mathbb{R}_+ \mapsto y_i(t)$ represents the time evolution of the number of follicles of maturity i . Follicles in the first compartment ($i = 0$) are named quiescent follicles, and their maturation and death rates are denoted by $\bar{\lambda}_0$ and $\bar{\mu}_0$, respectively. Follicles in the intermediate compartments ($1 \leq i \leq d-1$) are the growing follicles, and may either mature and go to the next maturation stage $i+1$, at rate λ_i , or die at rate μ_i . Follicles in the last compartment ($i = d$) are named the mature follicles and can only die at rate μ_d , *i.e.* $\lambda_d = 0$ (death in this compartment corresponds to either degeneration or ovulation). All the rates (μ_i, λ_i) may depend on the growing and mature follicles population (non-local interactions), which leads to the following nonlinear ODE system

$$\begin{cases} \frac{dy_0(t)}{dt} &= -(\bar{\lambda}_0(y(t)) + \bar{\mu}_0(y(t)))y_0(t), \\ \frac{dy_1(t)}{dt} &= \bar{\lambda}_0(y(t))y_0(t) - (\lambda_1(y(t)) + \mu_1(y(t)))y_1(t), \\ \frac{dy_i(t)}{dt} &= \lambda_{i-1}(y(t))y_{i-1}(t) - (\lambda_i(y(t)) + \mu_i(y(t)))y_i(t), \quad i \in \{2, \dots, d\}, \end{cases} \quad (4.1)$$

where, for $i = 0$,

$$\bar{\lambda}_0(y) = \frac{\bar{f}_0}{1 + K_{1,0} \sum_{j=1}^d a_j y_j}, \quad \bar{\mu}_0(y) = \bar{g}_0 \left(1 + K_{2,0} \sum_{j=1}^d b_j y_j \right), \quad (4.2)$$

with non-negative parameter constants $\bar{f}_0, \bar{g}_0, K_{1,0}, K_{2,0}$, and $a_j \in [0, 1], b_j \in [0, 1]$. For $i \in \{1, \dots, d\}$,

$$\lambda_i(y) = \frac{f_i}{1 + K_{1,i} \sum_{j=1}^d \omega_{1,j} y_j}, \quad \mu_i(y) = g_i \left(1 + K_{2,i} \sum_{j=1}^d \omega_{2,j} y_j \right), \quad (4.3)$$

with non-negative parameter constants, f_i ($f_d = 0$), $K_{1,i}, g_i, K_{2,i}$ and $\omega_{1,j} \in [0, 1], \omega_{2,j} \in [0, 1]$.

The formulation of the growth and death rates is based on the following, biologically-grounded principles, and motivated by biological knowledge (see Introduction). Parameters f_i (resp. $\bar{f}_0$) and g_i (resp. $\bar{g}_0$) set the “basal levels” of growth or death rates, as they would be ideally observed in standardized situations where follicle dynamics would be uncoupled from one another. Such a situation is rather well approximated by innovative devices of *in vitro* culture of isolated ovarian follicles. From this basal situation, switching the coupling on would result in lowering the growth rate and increasing the death rate. Parameters $K_{1,i}$ (resp. $K_{2,i}$) tunes the sensitivity of the growth rate (resp. death rate)

of a given developmental stage i to the other follicles' feedback. Parameters a_j and $\omega_{1,j}$ (resp. b_j and $\omega_{2,j}$) are used to weight, or even filter (when their values are set to zero) the contribution of follicles from any developmental stage $j \geq 1$ to the control of the maturation rates (resp. death rates) of quiescent and growing follicles. For instance, the maturation rates of the earliest stages are slowed down by AMH, which is secreted by the small growing follicles. On the other end, the survival of the terminally developing follicles is highly sensitive to the sufficient supply of FSH, whose levels are in turn down regulated by hormones secreted by the cohort of terminally developing follicles itself. With a non-negative vector $y^{in} \in \mathbb{R}_+^{d+1}$ as initial data, one can see that Eq. (4.1) generates a unique non-negative solution for all times (the right-hand side is globally Lipschitz, with positive off-diagonal entries). Moreover, one can obtain immediately the following conservation law,

$$\frac{d}{dt} \sum_{i=0}^d y_i(t) = -\bar{\mu}_0(y(t))y_0(t) - \sum_{i=1}^d \mu_i(y(t))y_i(t) \leq 0, \quad (4.4)$$

which shows that any follicle sub-population y_i is globally bounded.

4.1.2 Rescaled model

As outlined in the Introduction, before reproductive senescence, quiescent follicles are very numerous compared to the growing and mature follicles, follicle activation dynamics are much slower than growth dynamics, yet the flow of follicles between each compartment is of the same order. In consistency with this timescale contrast, we introduce a small positive parameter $\varepsilon \ll 1$, such that

$$\bar{f}_0 = \varepsilon f_0, \quad \bar{g}_0 = \varepsilon g_0, \quad y_0^{in} = \frac{x_0^{in}}{\varepsilon}, \quad (4.5)$$

with non-negative constants f_0, g_0 and positive initial data x_0^{in} , independent of ε . Note that the initial flow $\bar{f}_0 y_0^{in} = f_0 x_0^{in}$ is preserved.

We then define the rescaled solution $x = (x_0, \dots, x_d)$ by, for all $t \geq 0$,

$$x_0(t) = \varepsilon y_0(t/\varepsilon), \quad \text{and for all } i \geq 1, \quad x_i(t) = y_i(t/\varepsilon). \quad (4.6)$$

Then x is solution of the following system

$$\begin{cases} \frac{dx_0(t)}{dt} = -(\lambda_0(x(t)) + \mu_0(x(t)))x_0(t), \\ \varepsilon \frac{dx_i(t)}{dt} = \lambda_{i-1}(x(t))x_{i-1}(t) - (\lambda_i(x(t)) + \mu_i(x(t)))x_i(t), \quad i \in \{1, \dots, d\}, \end{cases} \quad (4.7)$$

with initial condition given by $x_0(t=0) = x_0^{in}$, and $x_i(t=0) = x_i^{in} := y_i^{in}$, for $i \in \{1, \dots, d\}$, and where, for $i=0$,

$$\lambda_0(x) = \frac{f_0}{1 + K_{1,0} \sum_{i=1}^d a_i x_i}, \quad \mu_0(x) = g_0 \left(1 + K_{2,0} \sum_{j=1}^d b_j x_j \right), \quad (4.8)$$

and λ_i and μ_i , for $i \in \{1, \dots, d\}$, are defined in Eq. (4.3). We note that the conservation law (4.4) becomes

$$\frac{d}{dt}x_0(t) + \varepsilon \frac{d}{dt} \sum_{i=1}^d x_i(t) = - \sum_{i=0}^d \mu_i(x(t))x_i(t) \leq 0. \quad (4.9)$$

We now consider the limit for which the small parameter ε tends to 0 and the associated sequence (x^ε) solution of system (4.7). In such a case, system (4.7) is called a “slow-fast” system (x_0^ε is the slow variable, $(x_1^\varepsilon, \dots, x_d^\varepsilon)$ are the fast variables) and the study of the limit behavior when $\varepsilon \rightarrow 0$ is a singular perturbation problem (see for instance [96]).

4.1.3 Limit model

Formally, setting $\varepsilon = 0$ in system (4.7) leads to the following system :

$$\begin{cases} \frac{d\bar{x}_0(t)}{dt} &= -(\lambda_0(\bar{x}(t)) + \mu_0(\bar{x}(t)))\bar{x}_0(t), \\ 0 &= \lambda_{i-1}(\bar{x}(t))\bar{x}_{i-1}(t) - (\lambda_i(\bar{x}(t)) + \mu_i(\bar{x}(t)))\bar{x}_i(t), \quad i \in \{1, \dots, d\}, \end{cases} \quad (4.10)$$

with initial condition given by $\bar{x}_0(t=0) = x_0^{in}$, and $\bar{x}_i(t=0) = \bar{x}_i^{in} \geq 0$, that satisfies the second line of Eq. (4.10) at $t=0$. Note that system (4.10) is a differential-algebraic system, in which the variable $(\bar{x}_1, \dots, \bar{x}_d)$ can be seen as reaching instantaneously (at any time t) a quasi-steady state, “driven” by the time-dependent variable $\bar{x}_0(t)$.

System (4.10) is not necessarily well-posed, as there may be several solutions to the second line of Eq. (4.10). In the next two specific examples, we can prove that system (4.10) does admit a single positive solution, which is a natural limit candidate for the sequence x^ε .

Example 1 (Linear case). *Let us suppose that $K_{1,i} = K_{2,i} = 0$ for all $i \in \{0, \dots, d\}$, and $f_i + g_i > 0$ for all $i \in \{1, \dots, d\}$. Then, system (4.10) becomes linear, and has a unique solution given by $\bar{x} = (\bar{x}_0, \dots, \bar{x}_d)$ such that for all $t \geq 0$,*

$$\begin{cases} \bar{x}_0(t) &= x_0^{in} \exp(-(f_0 + g_0)t), \\ \bar{x}_i(t) &= \left(\prod_{j=0}^{i-1} \frac{f_j}{f_{j+1} + g_{j+1}} \right) \bar{x}_0(t), \quad i \in \{1, \dots, d\}. \end{cases} \quad (4.11)$$

Example 2 (Feedback onto quiescent follicle activation and death rates). *Let us suppose that $K_{1,i} = 0$, $K_{2,i} = 0$ and $f_i + g_i > 0$ for $i \in \{1, \dots, d\}$, yet $K_{1,0} > 0$ and $K_{2,0} \geq 0$. Then, system (4.10) with positivity requirement ($x_i \geq 0$ for $i \in \{0, \dots, d\}$) can be rewritten as :*

$$\begin{cases} \frac{d\bar{x}_0(t)}{dt} &= -(\lambda_0(\bar{x}(t)) + \mu_0(\bar{x}(t)))\bar{x}_0(t), \\ \bar{x}_1(t) &= \frac{-(f_1 + g_1) + \sqrt{(f_1 + g_1)^2 + 4f_0\bar{x}_0(t)(f_1 + g_1)K_{1,0} \sum_{i=1}^d a_i \prod_{j=1}^{i-1} \frac{f_j}{f_{j+1} + g_{j+1}}}}{2(f_1 + g_1)K_{1,0} \sum_{i=1}^d a_i \prod_{j=1}^{i-1} \frac{f_j}{f_{j+1} + g_{j+1}}}, \\ \bar{x}_i(t) &= \left(\prod_{j=1}^{i-1} \frac{f_j}{f_{j+1} + g_{j+1}} \right) \bar{x}_1(t), \quad i \in \{2, \dots, d\}, \end{cases} \quad (4.12)$$

which admits a unique solution. Indeed, one can verify that $\bar{x}_1$ is the only positive root of a polynomial of degree 2, namely

$$f_0\bar{x}_0 = (f_1 + g_1) \left(1 + K_{1,0} \sum_{i=1}^d a_i \prod_{j=1}^{i-1} \frac{f_j}{f_{j+1} + g_{j+1}} \bar{x}_1 \right) \bar{x}_1.$$

4.1.4 Convergence in the linear case

In this paragraph, we assume that $K_{1,i} = K_{2,i} = 0$ for all $i \in \{0, \dots, d\}$, and $f_i + g_i > 0$ for all $i \in \{1, \dots, d\}$ as in Example 1. In such a case, one can solve system (4.7) explicitly for each $\varepsilon > 0$. The solution is given by, using vectorial notations, for all $t \geq 0$,

$$\begin{cases} x_0^\varepsilon(t) &= x_0^{in} \exp(-(f_0 + g_0)t), \\ (x_1^\varepsilon, \dots, x_d^\varepsilon)^T(t) &= e^{-\frac{Bt}{\varepsilon}} (x_1^{in}, \dots, x_d^{in})^T + \int_0^t \frac{1}{\varepsilon} e^{-\frac{B(t-s)}{\varepsilon}} e_1 f_0 x_0^\varepsilon(s) ds, \end{cases} \quad (4.13)$$

where $e_1 = (1, 0, \dots, 0)^T$ and

$$B = \begin{pmatrix} (f_1 + g_1) & 0 & \dots & 0 \\ -f_1 & \ddots & & \vdots \\ 0 & \ddots & \ddots & \vdots \\ \vdots & & & 0 \\ 0 & 0 & -f_{d-1} & (f_d + g_d) \end{pmatrix}.$$

It is thus clear that $x_0^\varepsilon = \bar{x}_0$ is a constant sequence in ε (as both x_0^ε and $\bar{x}_0$ have same initial conditions, and same evolution equation). For the fast variables, we prove the following

Proposition 1. *Assume that $K_{1,i} = K_{2,i} = 0$ for all $i \in \{0, \dots, d\}$, and $f_i + g_i > 0$ for $i \in \{1, \dots, d\}$. Then, for all $\eta > 0$, we have*

$$\limsup_{\varepsilon \rightarrow 0} \max_{t > \eta} \max_{i \in \{1, \dots, d\}} |x_i^\varepsilon(t) - \bar{x}_i(t)| = 0, \quad (4.14)$$

where $(\bar{x}_1, \dots, \bar{x}_d)$ is given in Eq. (4.11).

proof. From Eq. (4.13) and initial condition, and using integration by parts, we obtain

$$\int_0^t \frac{1}{\varepsilon} e^{-\frac{B(t-s)}{\varepsilon}} e_1 f_0 x_0^\varepsilon(s) ds = \int_0^t B^{-1} e^{-\frac{B(t-s)}{\varepsilon}} e_1 f_0 (f_0 + g_0) x_0^\varepsilon(s) ds + B^{-1} e_1 f_0 x_0^\varepsilon(t) - B^{-1} e^{-\frac{Bt}{\varepsilon}} e_1 f_0 x_0^{in}.$$

As x_0^ε is uniformly bounded (both in ε and time) by x_0^{in} , we obtain, taking the 1-norm,

$$\begin{aligned} \left\| \int_0^t B^{-1} e^{-\frac{B(t-s)}{\varepsilon}} e_1 f_0 (f_0 + g_0) x_0^\varepsilon(s) ds \right\| &\leq f_0 (f_0 + g_0) x_0^{in} \left\| \int_0^t B^{-1} e^{-\frac{B(t-s)}{\varepsilon}} e_1 ds \right\| \\ &\leq f_0 (f_0 + g_0) x_0^{in} \varepsilon \left\| B^{-2} (Id - e^{-\frac{Bt}{\varepsilon}}) e_1 \right\| \leq f_0 (f_0 + g_0) x_0^{in} \varepsilon \left\| B^{-2} \right\| (1 + e^{-\min(f_i + g_i) \frac{t}{\varepsilon}}), \end{aligned}$$

with $\left\| B^{-2} \right\| = \sup_{x \in \mathbb{R}^d, x \neq 0} \frac{\|B^{-2}x\|}{\|x\|}$. The third inequality above was deduced from

$$\forall t \geq 0, \quad \left\| e^{-\frac{Bt}{\varepsilon}} e_1 \right\| \leq \|e_1\| e^{-\min(f_i + g_i) \frac{t}{\varepsilon}}.$$

We verify that for all $t \geq 0$, $B^{-1} e_1 f_0 x_0^\varepsilon(t) = (\bar{x}_1, \dots, \bar{x}_d)(t)$. Then, we obtain,

$$\begin{aligned} \sup_{t > \eta} \|(x_1^\varepsilon, \dots, x_d^\varepsilon)(t) - (\bar{x}_1, \dots, \bar{x}_d)(t)\| &\leq (\|(x_1^{in}, \dots, x_d^{in})\| + \left\| B^{-1} \right\| f_0 x_0^{in}) e^{-\min(f_i + g_i) \frac{\eta}{\varepsilon}} \\ &\quad + 2f_0 (f_0 + g_0) x_0^{in} \left\| B^{-2} \right\| \varepsilon, \end{aligned} \quad (4.15)$$

which converges to 0 as ε converges to 0. $\square$

Remark 1. The proof of Proposition 1 can also be obtained as a direct application of Tikhonov theorem [96].

Remark 2. It is apparent in formula (4.13) that one cannot hope to obtain a convergence on a time interval starting from 0 (unless the initial data is “well-prepared”), and that standard Ascoli-Arzelà theorem would not apply in such a case, as the time derivative of $(x_1^\varepsilon, \dots, x_d^\varepsilon)(t)$ is not uniformly bounded as $\varepsilon \rightarrow 0$. However, from estimate (4.15), one can see that convergence holds uniformly in time on any interval $[\eta_\varepsilon, \infty)$ for $\frac{\eta_\varepsilon}{\varepsilon} \rightarrow \infty$ as $\varepsilon \rightarrow 0$. For instance, one can take $\eta_\varepsilon = \varepsilon \ln(1/\varepsilon)$.

4.1.5 Numerical convergence

In this paragraph, we illustrate the convergence of $(x_0^\varepsilon, x_1^\varepsilon, \dots, x_d^\varepsilon)$ towards $(\bar{x}_0, \bar{x}_1, \dots, \bar{x}_d)(t)$ in a nonlinear scenario. The chosen scenario and the parameter values are detailed in the Appendix (section 4.4).

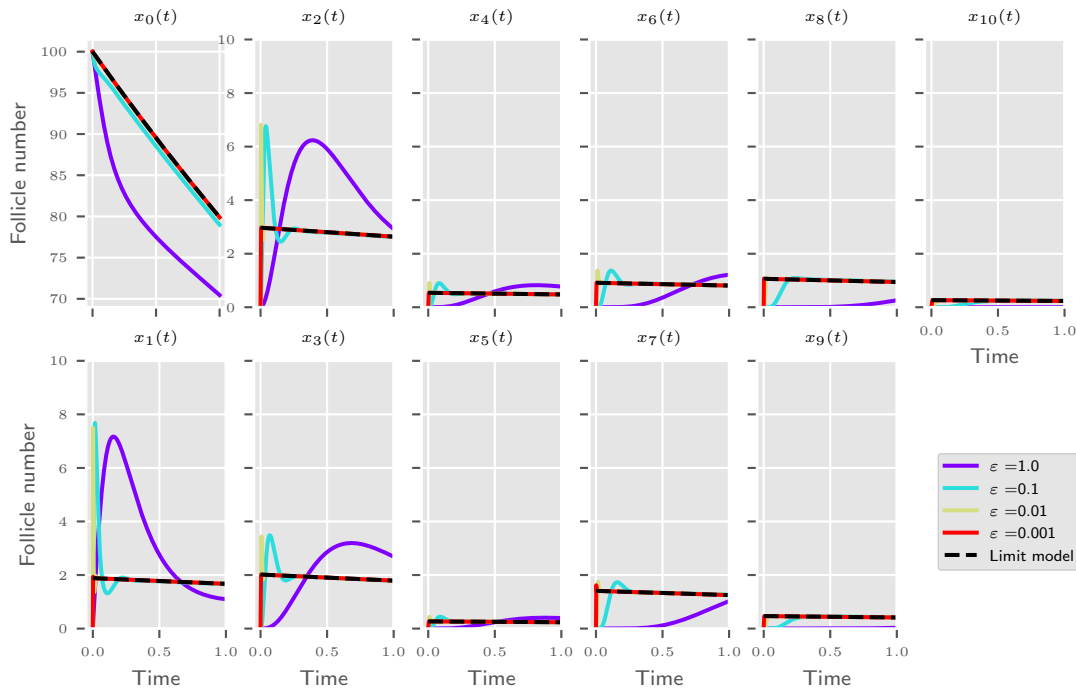


FIGURE 4.2 – Trajectories in each maturity compartment ($d = 10$), for the rescaled variables x_i^ε , for different ε (solid colored lines, see legend insert) and the reduced limit variables $\bar{x}_i$ (black dashed lines). See the Appendix (section 4.4) for details on the parameter values used in the numerical simulations.

In Figure 4.2, we plot the trajectories $x_i(t)$ in each maturity compartment ($d = 10$) for the rescaled and limit models on a time horizon $t \in (0, 1)$. In each compartment, the trajectories of the rescaled model get closer and closer to the limit model as $\varepsilon \rightarrow 0$. For $\varepsilon = 0.001$, they are almost indistinguishable. Note however that, for the growing follicles, the initial conditions of the rescaled and limit models are different, and the convergence holds only for positive times.

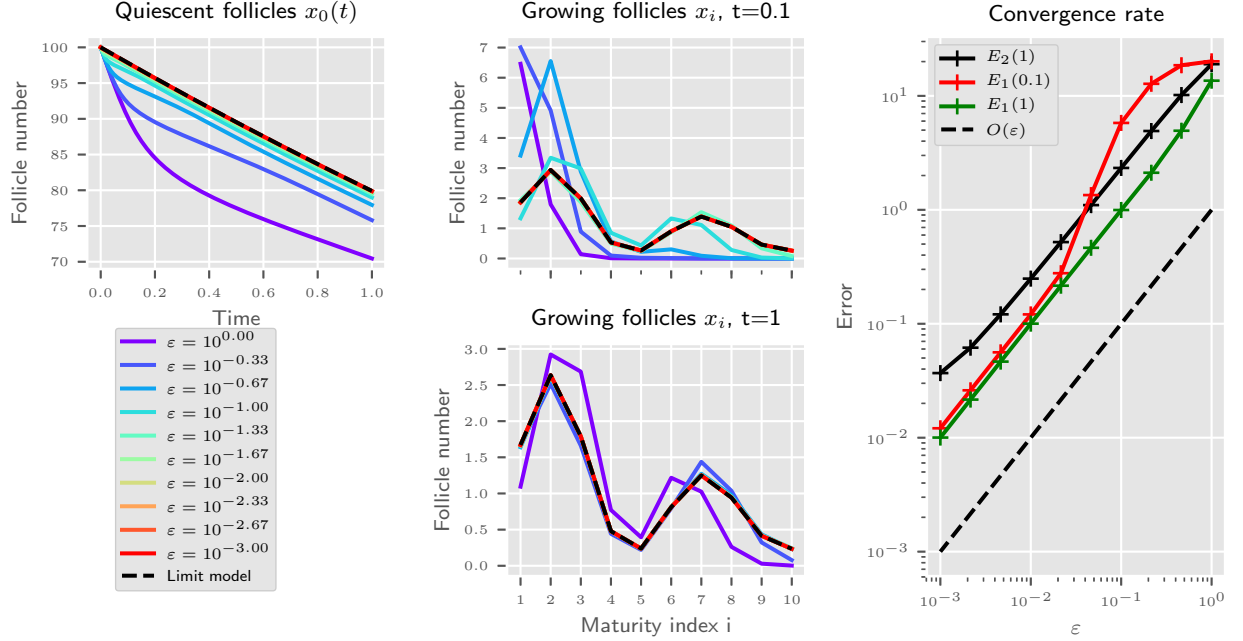


FIGURE 4.3 – Trajectories in the quiescent follicle compartment (top left panel) and distribution of the growing follicle population according to the maturity index i at time $t = 0.1$ (center top panel) and time $t = 1$ (center bottom panel), for the rescaled variables x_i^ε , for different ε (solid colored lines, see legend insert) and the reduced limit variables $\bar{x}_i$ (black dashed lines). On the right panel, we plot the discrete l^1 norm error $E_1(t)$ at a fixed time $t = 0.1$ and $t = 1$ (solid red and green lines, resp.) and the l^1 -cumulative error $E_2(1)$ on $t \in (0, 1)$ (black solid line) as a function of ε (see details in the text). The black dashed line is the straight line of slope 1 according to ε . See the Appendix (section 4.4) for details on the parameter values used in the numerical simulations.

In Figure 4.3, using the same parameters as in Figure 4.2, we display the maturity distribution in the growing follicle population, for various ε . We can expect from Figure 4.2 that the convergence gets better for larger times. We confirm this fact in Figure 4.3 where we compare the maturity distribution in the growing follicle population at two times, $t = 0.1$ and $t = 1$. We can quantify the error between the rescaled model and the limit model by computing the l^1 error at time t ,

$$E_1(t) = \sum_{i=0}^d |X_i^\varepsilon(t) - \bar{X}_i(t)|, \quad (4.16)$$

and the cumulative error on time interval $[0, T]$,

$$E_2(T) = \int_0^T \sum_{i=0}^d |X_i^\varepsilon(t) - \bar{X}_i(t)| dt, \quad (4.17)$$

which can be assessed numerically as

$$\tilde{E}_2(T) = \sum_{k=0}^{N_T} \delta_t \sum_{i=0}^d |X_i^\varepsilon(t_k) - \bar{X}_i(t_k)|, \quad (4.18)$$

where $t_k = k\delta_t$, for $k = 0 \dots N_T$. From the right panel of Figure 4.3, we can see that, at least for a small enough ε , the error is inversely proportional to ε , $E_i \approx C^{te}\varepsilon^{-1}$, where the constant pre-factor may depend on the chosen norm or the particular time t .

4.2 PDE model

When considering a continuous maturity variable, the PDE formalism is more suited for representing the follicle population dynamics. In this section, we skip the rescaling procedure, which follows an analogous reasoning as that detailed in section 4.1, and present directly the rescaled model.

4.2.1 Rescaled model

Denoting by $\rho_0(t)$ the number of quiescent follicles and by $\rho(t, x)$ the population density of follicles of maturity x , we consider the following coupled ODE-PDE system, for all $t \geq 0$,

$$\begin{cases} \frac{d\rho_0(t)}{dt} &= -(\lambda_0(\rho(t, \cdot)) + \mu_0(\rho(t, \cdot)))\rho_0(t), \\ \varepsilon \partial_t \rho(t, x) &= -\partial_x(\lambda(\rho(t, \cdot), x)\rho(t, x)) - \mu(\rho(t, \cdot), x)\rho(t, x), \quad \text{for } x \in (0, 1), \\ \lim_{x \rightarrow 0} \lambda(\rho(t, \cdot), x)\rho(t, x) &= \lambda_0(\rho(t, \cdot))\rho_0(t), \end{cases} \quad (4.19)$$

where

$$\lambda_0(\rho(t, \cdot)) = \frac{f_0}{1 + K_{1,0} \int_0^1 a(y)\rho(t, y)dy}, \quad \mu_0(\rho(t, \cdot)) = g_0 \left(1 + K_{2,0} \int_0^1 b(y)\rho(t, y)dy \right), \quad (4.20)$$

and, for all $x \in (0, 1)$,

$$\lambda(\rho(t, \cdot), x) = \frac{f(x)}{1 + K_1(x) \int_0^1 \omega_1(y)\rho(t, y)dy}, \quad \mu(\rho(t, \cdot), x) = g(x) \left(1 + K_2(x) \int_0^1 \omega_2(y)\rho(t, y)dy \right), \quad (4.21)$$

with initial condition

$$\rho_0(t=0) = \rho_0^{in}, \quad \rho(t=0, x) = \rho^{in}(x), \quad x \in (0, 1). \quad (4.22)$$

We assume that $f, g, K_1, K_2, a, b, w_1, w_2, \rho^{in}$ are regular enough functions, and will admit the existence and uniqueness of solutions of system (4.19). A standard fixed point argument, based on the mild formulation, could be used (see for instance [17, 35, 34] and references therein), yet this is beyond the scope of this work. We can write the following conservation law, that gives (at least formally)

$$\frac{d}{dt}\rho_0(t) + \varepsilon \frac{d}{dt} \int_0^1 \rho(t, x)dx = -\mu_0(\rho(t, \cdot))\rho_0(t) - \int_0^1 \mu(\rho(t, \cdot), x)\rho(t, x)dx - \lim_{x \rightarrow 1} \lambda(\rho(t, \cdot), x)\rho(t, x) \leq 0. \quad (4.23)$$

In the following, we consider a sequence $(\rho_0^\varepsilon, \rho^\varepsilon)$ of solutions of system (4.19) in the limit $\varepsilon \rightarrow 0$.

4.2.2 Limit model

Formally, setting $\varepsilon = 0$ in system (4.19) leads to the following system : for all $t \geq 0$,

$$\begin{cases} \partial_t \bar{\rho}_0(t) &= -(\lambda_0(\bar{\rho}(t, \cdot)) + \mu_0(\bar{\rho}(t, \cdot))) \bar{\rho}_0(t), \\ \partial_x (\lambda(\bar{\rho}(t, \cdot), x) \bar{\rho}(t, x)) &= -\mu(\bar{\rho}(t, \cdot)) \bar{\rho}(t, x), \quad \text{for } x \in (0, 1), \\ \lim_{x \rightarrow 0} \lambda(\bar{\rho}(t, \cdot), x) \bar{\rho}(t, x) &= \lambda_0(\bar{\rho}(t, \cdot)) \bar{\rho}_0(t), \end{cases} \quad (4.24)$$

with an initial condition given by $\bar{\rho}_0(t = 0) = \rho_0^{in}$, and $\bar{\rho}(t = 0, \cdot) = \bar{\rho}^{in}$, that satisfies the second and third lines of Eq. (4.24) at $t = 0$. System (4.24) is not necessarily well-posed, as there may be several solutions $\bar{\rho}$ for a given $\bar{\rho}_0$. In the next two specific examples, we can prove that system (4.24) does admit a single positive solution, which is a natural limit candidate for the sequence $(\rho_0^\varepsilon, \rho^\varepsilon)$.

Example 3 (Linear case). *Let us suppose that $K_{1,0} = K_{2,0} = 0$ and $K_1 \equiv 0$, $K_2 \equiv 0$. Assume furthermore that $f(0) > 0$. Then, system (4.24) becomes linear, and has a unique solution given by*

$$\begin{cases} \bar{\rho}_0(t) &= \rho_0^{in} \exp(-(f_0 + g_0)t), \\ \bar{\rho}(t, x) &= \frac{f_0}{f(0)} \bar{\rho}_0(t) e^{-\int_0^x \frac{g(y) + f'(y)}{f(y)} dy}, \quad \text{for } x \in (0, 1). \end{cases} \quad (4.25)$$

Example 4 (Feedback onto quiescent follicle activation and death rates). *Let us suppose that $K_1 \equiv 0$, $K_2 \equiv 0$, and $K_{1,0} > 0$ and $K_{2,0} \geq 0$. Assume furthermore that $f(0) > 0$. Then, system (4.24) with positivity requirement $\bar{\rho}(t, x) \geq 0$ can be rewritten as*

$$\begin{cases} \frac{d\bar{\rho}_0(t)}{dt} &= -(\lambda_0(\bar{\rho}(t, \cdot)) + \mu_0(\bar{\rho}(t, \cdot))) \bar{\rho}_0(t), \\ \bar{\rho}(t, 0) &= \frac{-f(0) + \sqrt{(f(0))^2 + 4f_0\bar{\rho}_0(t)f(0)K_{1,0} \int_0^1 a(x) e^{\int_0^x \frac{g(y) + f'(y)}{f(y)} dy} dx}}{2f(0)K_{1,0} \int_0^1 a(x) e^{\int_0^x \frac{g(y) + f'(y)}{f(y)} dy} dx}, \\ \bar{\rho}(t, x) &= \bar{\rho}(t, 0) e^{\int_0^x \frac{g(y) + f'(y)}{f(y)} dy}, \quad \text{for } x \in (0, 1), \end{cases} \quad (4.26)$$

which admits a unique solution. Indeed, the functional expression of $\bar{\rho}$ comes directly from solving the second line of Eq. (4.24). Using the boundary condition in the third line of Eq. (4.24), one can verify that $\bar{\rho}(t, 0)$ is the positive root of a polynomial of degree 2, namely

$$f_0 \bar{\rho}_0(t) = f(0) \bar{\rho}(t, 0) \left(1 + K_{1,0} \bar{\rho}(t, 0) \int_0^1 a(x) e^{\int_0^x \frac{g(y) + f'(y)}{f(y)} dy} dx \right).$$

The limit system (4.26) is a nonlinear ODE. To simplify notations, we introduce

$$H(x) = e^{-\int_0^x \frac{g(y) + f'(y)}{f(y)} dy}, \quad H_a = \int_0^1 a(x) H(x) dx, \quad H_b = \int_0^1 b(x) H(x) dx,$$

from which we rewrite system (4.26) as

$$\begin{cases} \frac{d\bar{\rho}_0(t)}{dt} &= -(\lambda_0(\bar{\rho}(t, \cdot)) + \mu_0(\bar{\rho}(t, \cdot)))\bar{\rho}_0(t), \\ \bar{\rho}(t, 0) &= \frac{-1 + \sqrt{1 + 4\bar{\rho}_0(t)K_{1,0}H_a \frac{f_0}{f(0)}}}{2K_{1,0}H_a}, \\ \bar{\rho}(t, x) &= \bar{\rho}(t, 0)H(x), \quad \text{for } x \in (0, 1), \\ \lambda_0(\bar{\rho}(t, \cdot)) &= \frac{f_0}{1 + K_{1,0}\bar{\rho}(t, 0)H_a}, \\ \mu_0(\bar{\rho}(t, \cdot)) &= g_0(1 + K_{2,0}\bar{\rho}(t, 0)H_b). \end{cases}$$

We can thus solve $\bar{\rho}_0(t)$ as the solution of an autonomous nonlinear ODE

$$\frac{d}{dt}\bar{\rho}_0(t) = G(\bar{\rho}_0(t)), \quad \bar{\rho}_0(0) = \rho_0^{ini}, \quad (4.27)$$

with

$$G(\rho) = - \left[\frac{2f_0}{1 + \sqrt{1 + 4\frac{f_0}{f(0)}K_{1,0}H_a\rho}} + g_0 \left(1 + K_{2,0} \frac{-1 + \sqrt{1 + 4\frac{f_0}{f(0)}K_{1,0}H_a\rho}}{2K_{1,0}H_a} H_b \right) \right] \rho,$$

from which we then compute $\bar{\rho}(t, 0)$ and eventually $\bar{\rho}(t, x)$.

4.2.3 Convergence in the linear case

In this paragraph, we assume that $K_{1,0} = K_{2,0} = 0$, $K_1 \equiv 0$, $K_2 \equiv 0$, and $f(0) > 0$ as in Example 3. In such a case, one can solve explicitly system (4.19) for each $\varepsilon > 0$ using the characteristics method. We obtain

$$\begin{cases} \rho_0^\varepsilon(t) &= \rho_0^{in} \exp(-(f_0 + g_0)t), \\ \rho^\varepsilon(t, x) &= \begin{cases} e^{-\int_{X(0;t,x)}^x \frac{g(y)+f'(y)}{f(y)} dy} \rho^{ini}(X(0;t,x)), & \text{if } t \leq \int_0^x \frac{\varepsilon}{f(y)} dy, \\ \frac{f_0}{f(0)} \rho_0^{in} e^{-(f_0+g_0)(t-\int_0^x \frac{\varepsilon}{f(y)} dy)} e^{-\int_0^x \frac{g(y)+f'(y)}{f(y)} dy}, & \text{if } t > \int_0^x \frac{\varepsilon}{f(y)} dy. \end{cases} \end{cases} \quad (4.28)$$

where $X(0;t,x)$ is the location of the characteristic at time 0, given that it goes through the point x at time t , namely :

$$\frac{d}{ds}X(s;t,x) = \frac{1}{\varepsilon}f(X(s;t,x)), \quad X(t;t,x) = x. \quad (4.29)$$

It is thus clear that $\rho_0^\varepsilon = \bar{\rho}_0$ is a constant sequence in ε . For the population density ρ^ε , which is here the fast unknown, we prove the following

Proposition 2. Assume that $K_{1,0} = K_{2,0} = 0$, $K_1 \equiv 0$, $K_2 \equiv 0$, and $f(0) > 0$ with $\int_0^1 \frac{1}{f(x)} dx < \infty$ and $\int_0^1 \frac{g(x)+f'(x)}{f(x)} dx < \infty$. Then, for all $\eta > 0$, we have

$$\limsup_{\varepsilon \rightarrow 0} \sup_{t > \eta} \sup_{x \in (0,1)} |\rho^\varepsilon(t, x) - \bar{\rho}(t, x)| = 0, \quad (4.30)$$

where $\bar{\rho}$ is given in Eq. (4.25).

Démonstration. It is clear that for any $\eta > 0$, there exists ε' such that for all $\varepsilon < \varepsilon'$, and all $t > \eta$ we have

$$t > \int_0^x \frac{\varepsilon}{f(y)} dy.$$

Then, comparing the solutions of Eq. (4.28) and Eq. (4.25) allows us to conclude that, for all $t > \eta$,

$$\sup_{t > \eta} \sup_{x \in (0,1)} |\rho^\varepsilon(t, x) - \bar{\rho}(t, x)| \leq \frac{f_0}{f(0)} \rho_0^{in} \left(\sup_{x \in (0,1)} e^{-\int_0^x \frac{g(y) + f'(y)}{f(y)} dy} \right) \left(e^{(f_0 + g_0)\varepsilon \int_0^1 \frac{1}{f(y)} dy} - 1 \right) \rightarrow 0, \quad \text{as } \varepsilon \rightarrow 0$$

□

Remark 3. As in Remark 1, we can see that during a time of order εt , we cannot get the convergence of the rescaled model towards the reduced one, which precludes uniform convergence in time starting from $t=0$. Also, as in Remark 2, from the representation formula (4.28), one can see that convergence holds uniformly in time on any interval $[\eta_\varepsilon, \infty)$ for $\frac{\eta_\varepsilon}{\varepsilon} \rightarrow \infty$ as $\varepsilon \rightarrow 0$. For instance, one can take $\eta_\varepsilon = \varepsilon \ln(1/\varepsilon)$.

4.2.4 Numerical study

In this paragraph, we detail the numerical schemes that we have designed to solve both systems (4.19) and (4.24), and illustrate the consistency and convergence of these algorithms using the exact solutions.

Numerical scheme for the limit model

We design a finite difference scheme to compute a numerical solution to the PDE limit system (4.24). This system is nonlinear due to the dependence of λ_0 , λ and μ upon the solution $\bar{\rho}(t, x)$, which itself depends on $\bar{\rho}_0$. We propose to treat this nonlinearity with a fixed point scheme. At each time step $t_n = n\Delta_t$, for $n = 0, \dots, N$ with $T = N\Delta_t$ we build a sequence $\rho^\ell(t_n, x)$ such that

$$\lim_{\ell \rightarrow \infty} \rho^\ell(t_n, x) = \bar{\rho}(t_n, x).$$

Let $x_k = k\Delta_x$, for $k = 0, \dots, M$, with $M\Delta_x = 1$. We introduce the discretized approximations

$$\rho_k^{n,\ell} \approx \rho^\ell(t_n, x_k), \quad \bar{\rho}_k^n \approx \lim_{\ell \rightarrow \infty} \rho^\ell(t_n, x_k) \quad \text{and} \quad \bar{\rho}_0^n \approx \bar{\rho}_0(t_n),$$

which we compute as follows. Let $\eta \ll 1$.

1. Initialization.

$$\rho_k^{0,0} = \bar{\rho}^{in}(x_k) \quad \text{and} \quad \bar{\rho}_0^0 = \rho_0^{in}.$$

2. For $n = 0 \nearrow N$ compute $\bar{\rho}_k^n = \lim_{\ell \rightarrow \infty} \rho_k^{\ell,n}$ iteratively then update $\bar{\rho}_0^{n+1}$ as follows.
 - (a) Initialize residual $R_\ell^n = 1$ and set $\ell = 0$
 - (b) While $R_\ell^n > \eta$ do

- Compute the PDE parameters $\bar{\lambda}_0^{n,\ell}$, $\lambda_k^{n,\ell}$ and $\mu_k^{n,\ell}$ (for $k = 0, \dots, M$) by standard trapezoidal rules

$$\begin{aligned}\bar{\lambda}_0^{n,\ell} &= \frac{f_0}{1 + K_{1,0}\Delta_x \left(\frac{1}{2}a(0)\rho_0^{n,\ell} + \sum_{j=1}^{M-1} a(x_j)\rho_j^{n,\ell} + \frac{1}{2}a(x_M)\rho_M^{n,\ell} \right)}, \\ \lambda_k^{n,\ell} &= \frac{f(x_k)}{1 + K_1(x_k)\Delta_x \left(\frac{1}{2}\omega_1(0)\rho_0^{n,\ell} + \sum_{j=1}^{M-1} \omega_1(x_j)\rho_j^{n,\ell} + \frac{1}{2}\omega_1(x_M)\rho_M^{n,\ell} \right)}, \\ \mu_k^{n,\ell} &= g(x_k) \left(1 + K_2(x_k)\Delta_x \left(\frac{1}{2}\omega_2(0)\rho_0^{n,\ell} + \sum_{j=1}^{M-1} \omega_2(x_j)\rho_j^{n,\ell} + \frac{1}{2}\omega_2(x_M)\rho_M^{n,\ell} \right) \right).\end{aligned}$$

- Enforce boundary condition at $x = 0$

$$\lambda_0^{n,\ell} \rho_0^{n,\ell+1} = \bar{\lambda}_0^{n,\ell} \bar{\rho}_0^n.$$

- Integrate numerically the PDE in x

$$\lambda_k^{n,\ell} \left(\rho_{k+1}^{n,\ell+1} - \rho_k^{n,\ell+1} \right) = -\Delta_x \mu_k^{n,\ell} \rho_k^{n,\ell+1}, \quad k = 0, \dots, M-1.$$

- Compute residual between ℓ and $\ell + 1$ iterations

$$R_\ell^n = \max_{k=0,\dots,M} |\rho_k^{n,\ell+1} - \rho_k^{n,\ell}|.$$

- (c) Set the new population density equal to the value obtained at the end of the fixed point procedure $\bar{\rho}_k^n = \rho_k^{n,\ell+1}$
- (d) Compute the ODE parameters from the fixed point value, using standard trapezoidal rules

$$\begin{aligned}\bar{\bar{\lambda}}_0^n &= \frac{f_0}{1 + K_{1,0}\Delta_x \left(\frac{1}{2}a(0)\bar{\rho}_0^n + \sum_{j=1}^{M-1} a(x_j)\bar{\rho}_j^n + \frac{1}{2}a(x_M)\bar{\rho}_M^n \right)}, \\ \bar{\bar{\mu}}_0^n &= g_0 \left(1 + K_{2,0}\Delta_x \left(\frac{1}{2}b(0)\bar{\rho}_0^n + \sum_{j=1}^{M-1} b(x_j)\bar{\rho}_j^n + \frac{1}{2}b(x_M)\bar{\rho}_M^n \right) \right).\end{aligned}$$

- (e) Integrate numerically the ODE $\bar{\rho}_0$ between t_n and t_{n+1} with a classic explicit Euler scheme

$$\bar{\rho}_0^{n+1} = \bar{\rho}_0^n - \Delta_t \left(\bar{\bar{\lambda}}_0^n + \bar{\bar{\mu}}_0^n \right) \bar{\rho}_0^n.$$

Convergence of the numerical scheme for the limit model.

For the nonlinear scenario described in Example 4 we obtain a “pseudo exact” solution using scipy library ODE solver odeint to solve Eq. (4.27) and compute $\bar{\rho}_0(t)$, and we use Eq. (4.26) to compute $\bar{\rho}(t, \cdot)$. We use this reference solution to assess the performances of the numerical scheme described in paragraph 4.2.4. We display on Figure 4.4 the simulations performed with the parameter values detailed in the Appendix (section 4.4), on a time horizon $t \in (0, 1)$, and an increasingly fine discretization

$$\Delta_x \in \{0.05, 0.025, 0.0125, 0.00625, 0.003125\}.$$

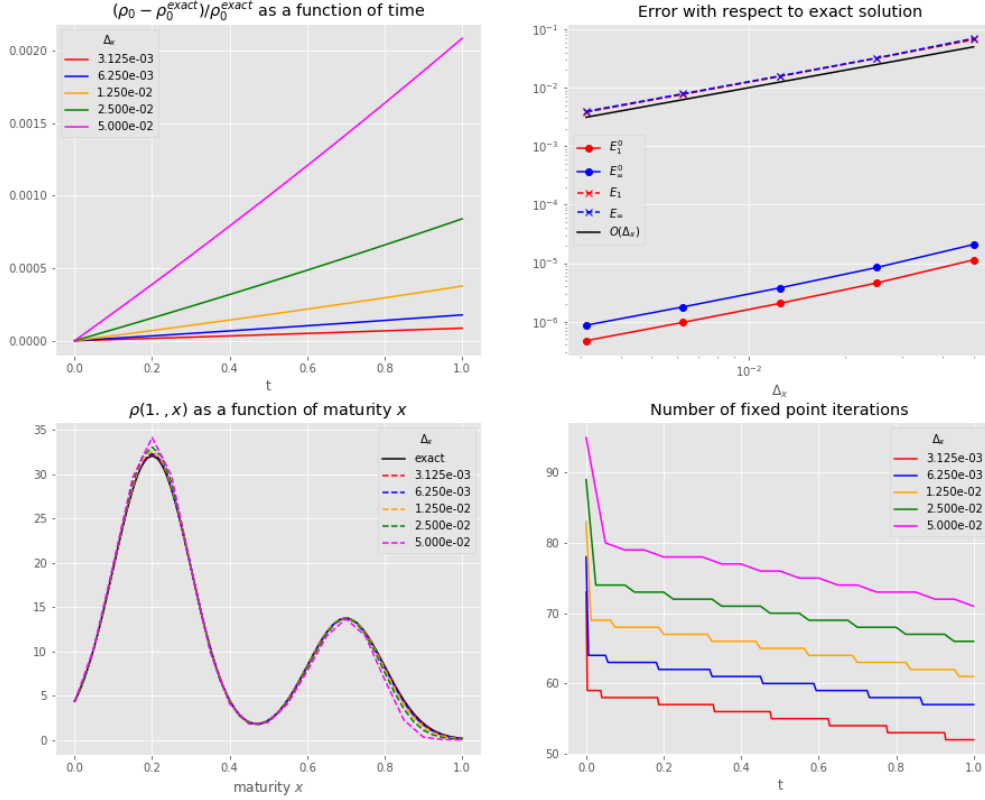


FIGURE 4.4 – Convergence of the numerical scheme for the limit model as a function of $\Delta_x = \Delta_t$. Top left panel : relative error of $\bar{\rho}_0^n \approx \rho_0(t_n)$, computed by means of the fixed point algorithm, compared the numerical solution of (4.27), computed with the odeint python ODE solver. Bottom left panel : comparison of $\bar{\rho}_k^N = \bar{\rho}(t = 1, x_k)$ and the numerical solution of (4.26). Top right panel : relative errors with respect to the pseudo exact solution of (4.26). Bottom right panel : number of iterations in the fixed point algorithm.

The top left panel shows the difference between the pseudo exact solution $\bar{\rho}_0(t_n)$ of Eq. (4.27) and the numerical solutions $\bar{\rho}^n$ computed with the fixed point algorithm, as a function of $t \in [0, T]$. The bottom left panel shows the pseudo exact solution $\bar{\rho}(T, x_k)$ and the numerical solutions $\bar{\rho}_k^N$ at final time $T = 1$, as a function of $x \in [0, 1]$. In the top right panel, we display the relative errors in L^1 and L^∞ norms between the pseudo exact and the numerical solution $\bar{\rho}_0(t)$, as a function of $\Delta_t = \Delta_x = 1/N$:

$$E_1^0(\Delta_x) = \frac{\sum_{n=0}^N |\bar{\rho}_0(t_n) - \bar{\rho}^n|}{\sum_{n=0}^N |\bar{\rho}_0(t_n)|},$$

$$E_\infty^0(\Delta_x) = \frac{\max_{n=0, \dots, N} |\bar{\rho}_0(t_n) - \bar{\rho}^n|}{\max_{n=0, \dots, N} |\bar{\rho}_0(t_n)|},$$

and the relative errors between the pseudo exact and the numerical solutions $\bar{\rho}(T, x)$

$$E_1(\Delta_x) = \frac{\sum_{k=0}^M |\bar{\rho}(T, x_k) - \bar{\rho}_k^N|}{\sum_{k=0}^M |\bar{\rho}(T, x_k)|},$$

$$E_\infty(\Delta_x) = \frac{\max_{k=0, \dots, M} |\bar{\rho}(T, x_k) - \bar{\rho}_k^N|}{\max_{k=0, \dots, M} |\bar{\rho}(T, x_k)|}.$$

As expected, the errors are linear in Δ_x , which means that the order of numerical convergence is one. In the bottom right panel, we display the number of iterations performed in the inner loop of the fixed point algorithm, as a function of time. The number of iterations needed to converge decreases with time and the number of time points. This tendency is not really surprising, since, at each time step, we start with the solution obtained at the previous time step as initial condition for the fixed point loop. Since the solution decreases with time, the distance between the fixed point initial condition and the solution decreases with both the time and time step, hence convergence requires less iterations.

Numerical scheme for the rescaled model

We design an explicit finite volume scheme to compute a numerical solution to the rescaled model (Eq. 4.19). The discretized unknowns are at each time step $t_n = n\Delta_t$, for $n = 0, \dots, N$ with $T = N\Delta_t$

$$\bar{\rho}_0^{\varepsilon, n} \approx \rho_0^{\varepsilon}(t_n),$$

and, for $x_k = k\Delta_x$, $k = 0, \dots, M$, with $M\Delta_x = 1$,

$$\rho_k^{\varepsilon, n} \approx \int_{x_k}^{x_{k+1}} \rho^{\varepsilon}(t_n, x) dx, \quad k = 0, \dots, M-1.$$

We integrate numerically the PDE between t_n and t_{n+1} and over $[x_k, x_{k+1}]$ by freezing the nonlinear coefficients $\lambda(\bar{\rho}(t, \cdot), x)$ and $\mu(\bar{\rho}(t, \cdot))$ at time t_n

$$\rho_k^{\varepsilon, n+1} - \rho_k^{\varepsilon, n} = -\frac{\Delta_t^n}{\varepsilon \Delta_x} (\lambda_k^{\varepsilon, n} \rho_k^{\varepsilon, n} - \lambda_{k-1}^{\varepsilon, n} \rho_{k-1}^{\varepsilon, n}) - \frac{\Delta_t^n}{\varepsilon} \mu_k^{\varepsilon, n} \rho_k^{\varepsilon, n}, \quad k = 1, \dots, M. \quad (4.31)$$

At each time step, we compute both the PDE and ODE coefficients using the midpoint rule and the numerical solution $(\rho_k^{\varepsilon, n})_{k=0, \dots, M}$ as a piecewise constant solution

$$\begin{cases} \lambda_k^{\varepsilon, n} = \frac{f(x_{k+1/2})}{1 + K_1(x_{k+1/2})\Delta_x \sum_{j=0}^{M-1} \omega_1(x_{j+1/2})\rho_j^{\varepsilon, n}}, \\ \bar{\lambda}_0^{\varepsilon, n} = \frac{f_0}{1 + K_{1,0}\Delta_x \sum_{j=0}^{M-1} a(x_{j+1/2})\rho_j^{\varepsilon, n}}, \\ \mu_k^{\varepsilon, n} = g(x_{k+1/2}) \left(1 + K_2(x_{k+1/2})\Delta_x \sum_{j=0}^{M-1} \omega_2(x_{j+1/2})\rho_j^{\varepsilon, n} \right), \\ \bar{\mu}_0^{\varepsilon, n} = g_0 \left(1 + K_{2,0}\Delta_x \sum_{j=0}^{M-1} b(x_{j+1/2})\rho_j^{\varepsilon, n} \right). \end{cases} \quad (4.32)$$

For the explicit scheme (4.31), two stability conditions must be satisfied

— CFL-like stability condition :

$$\frac{\Delta_t^n}{\varepsilon \Delta_x} \lambda_k^{\varepsilon, n} \leq C_{cfl} < 1, \quad k = 1, \dots, M,$$

which can be rewritten as

$$\frac{\Delta_t^n}{\varepsilon \Delta_x} \leq \frac{C_{cfl}}{\max_k \lambda_k^{\varepsilon, n}}. \quad (4.33)$$

- Positivity conservation condition : $1 + \frac{\Delta_t^n}{\varepsilon} \left(\frac{\lambda_k^{\varepsilon,n}}{\Delta_x} - \mu_k^{\varepsilon,n} \right) \geq 0$, $k = 1, \dots, M$ we impose that if $\rho_i^{\varepsilon,n} = \delta_{ik}$ then $\rho_i^{\varepsilon,n+1} \geq 0$ for all k and all i which leads to

$$1 - \frac{\Delta_t^n}{\varepsilon} \left(\frac{\lambda_k^{\varepsilon,n}}{\Delta_x} + \mu_k^{\varepsilon,n} \right) \geq 0, \quad k = 1, \dots, M,$$

which can be rewritten as

$$\frac{\Delta_t^n}{\varepsilon \Delta_x} \leq \frac{1}{\max_k (\lambda_k^{\varepsilon,n} + \Delta_x \mu_k^{\varepsilon,n})}. \quad (4.34)$$

The overall numerical scheme proceeds as follows :

1. Initialization :

$$\rho_k^{\varepsilon,0} = \rho^{in}(x_k) \quad \text{and} \quad \bar{\rho}_0^{\varepsilon,0} = \rho_0^{in}.$$

2. For $n = 0 \nearrow N$ compute $\rho_k^{\varepsilon,n+1}$ for $k = 1, \dots, M$ then update $\bar{\rho}_0^{\varepsilon,n+1}$ and finally compute $\rho_0^{\varepsilon,n+1}$:

- (a) Compute the PDE and ODE coefficients $\bar{\lambda}_0^{\varepsilon,n}$, $\bar{\mu}_0^{\varepsilon,n}$, $\lambda_k^{\varepsilon,n}$ and $\mu_k^{\varepsilon,n}$ (Eq. 4.32)
- (b) Compute Δ_t^n satisfying stability conditions (4.33) and (4.34)
- (c) Integrate numerically the PDE in x at time t_n using (4.31)
- (d) Integrate numerically the ODE $\bar{\rho}_0$ between t_n and t_{n+1}

$$\bar{\rho}_0^{\varepsilon,n+1} = \bar{\rho}_0^{\varepsilon,n} - \Delta_t^n (\bar{\lambda}_0^{\varepsilon,n} + \bar{\mu}_0^{\varepsilon,n}).$$

- (e) Enforce the boundary condition at $x = 0$

$$\lambda_0^{\varepsilon,n+1} \rho_0^{\varepsilon,n+1} = \bar{\lambda}_0^{\varepsilon,n+1} \bar{\rho}_0^{\varepsilon,n+1}.$$

Convergence of the numerical scheme for the rescaled model.

We start by checking the convergence of the numerical scheme in a case where we know an exact solution, that is the linear case $K_{10} = K_{20} = K_1 = K_2 = 0$. We test several discretizations

$$\Delta_x \in \{0.1, 0.05, 0.025, 0.0125, 0.00625, 0.003125\},$$

for the parameter values of the linear scenario detailed in the Appendix (section 4.4). The results are displayed in Figure 4.5. In the left panel, we see that $\rho_0(t)$ is computed exactly since all curves corresponding to different discretizations are superimposed (as expected in the linear case). In the center panel, we display the solution at a fixed time T as a function of x , which does depend on the discretization. The right panel shows the relative error curves, which exhibit a convergence rate better than linear.

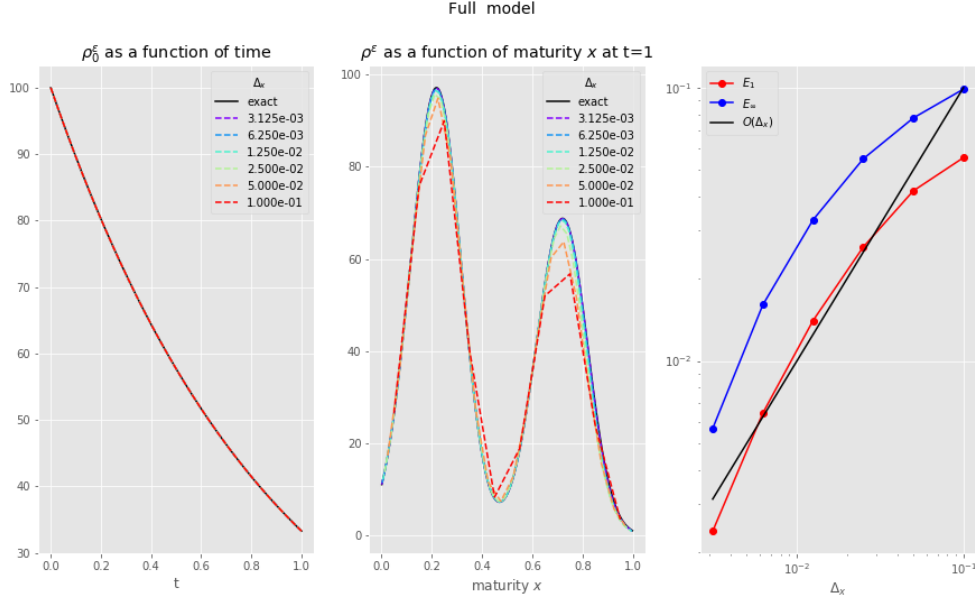


FIGURE 4.5 – Convergence of the numerical scheme for the rescaled model as a function of Δ_x , for $\varepsilon = 0.5$. Left panel : ρ_0^ε as a function of time (exact and finite volume scheme), center panel : ρ^ε as a function of x at final time $t = 1$, right panel : L_∞ and L_1 relative errors with respect to the exact solution.

ε -convergence towards the limit model

So far we have proved the convergence of the rescaled model towards the limit model when $\varepsilon \rightarrow 0$ in the linear case. We can only test it numerically in the general case. To minimize the numerical error arising from solving the limit model numerically, we illustrate the ε -convergence of $(\rho_0^\varepsilon, \rho^\varepsilon)$ towards $(\bar{\rho}_0, \bar{\rho})$ in the nonlinear scenario of example 4 (see details on the parameter values in the Appendix (section 4.4)), for which the pseudo-exact solution of the limit model is available. As in Figure 4.4, pseudo exact solutions $(\bar{\rho}_0, \bar{\rho})$ are simulated using Eq. (4.27) and Eq. (4.26). The results are displayed in Figure 4.6 for a set of ε values and two discretizations $M = N = 100$ and $M = N = 200$. The agreement of $\rho_0(t)$ with the limit model solution (top left panel), and that of $\rho^\varepsilon(t, x)$ (bottom left panel), are qualitatively good as soon as $\varepsilon \leq 10^{-2}$. The relative error for $\rho_0(t)$ (top right panel) exhibits a linear behavior in ε . The error curves for the convergence of the solution $\rho^\varepsilon(t, x)$ in the domain (bottom right panel) are not linear, and remain beyond a threshold when ε goes to 0. However, both the value of ε for which the error approaches this threshold, and the value of the error itself decrease when we refine the discretization. This indicates that we should refine the discretization when we decrease ε . Since our current numerical scheme is explicit in time, any refinement must be simultaneous in Δ_x and Δ_t , and as a consequence the cost in CPU time depends quadratically in ε^{-1} .

In practice, realistic values for ε should remain tractable. However this behavior is a good incentive to study a more economical numerical scheme, namely an implicit one, which would provide accurate results with coarser discretizations.

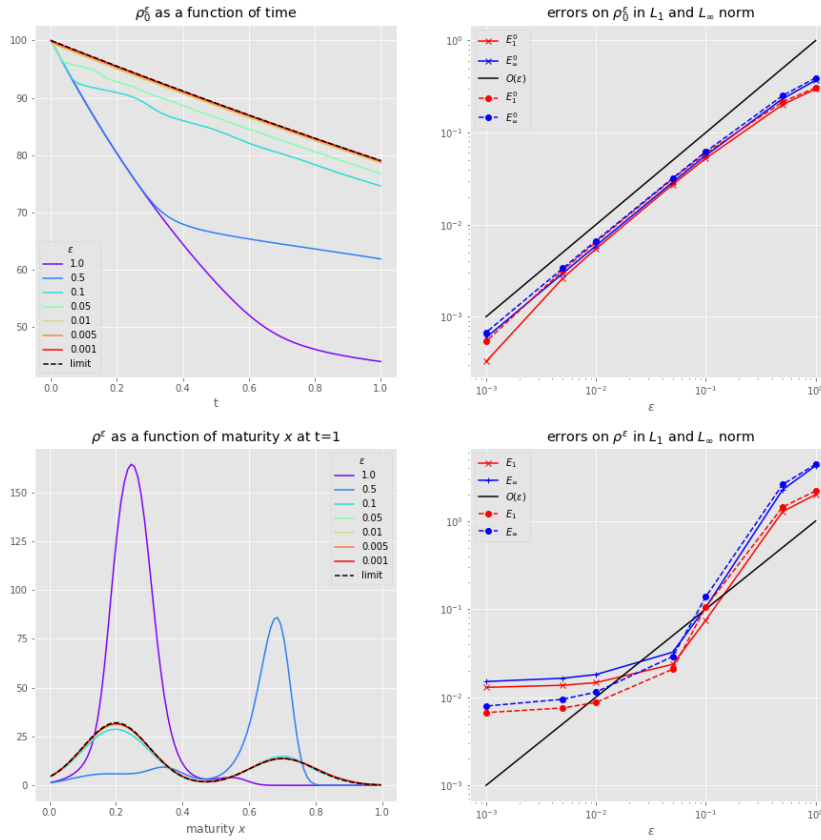


FIGURE 4.6 – Numerical assessment of the ε -convergence of the rescaled model towards the limit model. Top left panel : $\bar{\rho}_0^{\varepsilon,n}$ and $\bar{\rho}_0$. Bottom left panel : $\rho_k^{\varepsilon,N}$ and $\bar{\rho}(1, \cdot)$. Top right panel : relative error in L_1 and L_∞ norms between $\rho_0^\varepsilon(t)$ and $\bar{\rho}_0$. Bottom right panel : relative error in L_1 and L_∞ norms between $\rho^\varepsilon(1, \cdot)$ and $\bar{\rho}(1, \cdot)$. In both right panels, solid lines correspond to solutions computed with $M = N = 100$ and dashed lines to solutions computed with $M = N = 200$.

4.3 SDE model

We now turn to a stochastic model for the follicle population dynamics. We skip the rescaling procedure, which follows an analogous reasoning as that detailed in section 4.1, and present directly the rescaled model.

4.3.1 Rescaled model

We consider the following coupled Poisson-driven SDE system, given by, for all $t \geq 0$,

$$\left\{ \begin{array}{l} X_0^\varepsilon(t) = X_0^{\varepsilon, in} - \varepsilon P_0^+ \left(\int_0^t \frac{\lambda_0(X^\varepsilon(s))}{\varepsilon} X_0^\varepsilon(s) ds \right) \\ \quad - \varepsilon P_0^- \left(\int_0^t \frac{\mu_0(X^\varepsilon(s))}{\varepsilon} X_0^\varepsilon(s) ds \right) \\ X_i^\varepsilon(t) = X_i^{\varepsilon, in} + P_{i-1}^+ \left(\int_0^t \frac{\lambda_{i-1}(X^\varepsilon(s))}{\varepsilon} X_{i-1}^\varepsilon(s) ds \right) \\ \quad - P_i^+ \left(\int_0^t \frac{\lambda_i(X^\varepsilon(s))}{\varepsilon} X_i^\varepsilon(s) ds \right) \\ \quad - P_i^- \left(\int_0^t \frac{\mu_i(X^\varepsilon(s))}{\varepsilon} X_i^\varepsilon(s) ds \right), \quad i \in \{1, \dots, d\}, \end{array} \right. \quad (4.35)$$

where X^ε represents the vectorial process $(X_0^\varepsilon, \dots, X_d^\varepsilon)$ of the follicle number in each maturity stage, the functions λ_i, μ_i , for $i \in \{0, \dots, d\}$ are given by Eq. (4.3) and (4.8) and $(P_i^+, P_i^-)_{i \in \{0, \dots, d\}}$ are independent standard Poisson processes. In Eq. (4.35), $X_0^{\varepsilon, in}$ is an $\varepsilon\mathbb{N}$ -valued random variable, and each $X_i^{\varepsilon, in}$, $i \in \{1, \dots, d\}$, is a $\mathbb{N}$ -valued random variable.

SDE (4.35) defines uniquely (in law) a continuous time Markov chain in $\varepsilon\mathbb{N} \times \mathbb{N}^d$. We note the following conservation law,

$$X_0^\varepsilon(t) + \varepsilon \sum_{i=1}^d X_i^\varepsilon(t) = X_0^\varepsilon(0) + \varepsilon \sum_{i=1}^d X_i^\varepsilon(0) - \varepsilon \sum_{i=0}^d P_i^- \left(\int_0^t \frac{\mu_i(X^\varepsilon(s))}{\varepsilon} X_i^\varepsilon(s) ds \right). \quad (4.36)$$

In the following, we consider a sequence X^ε of solutions to system (4.35) in the limit ε tends to 0.

4.3.2 Limit model

Formally, setting $\varepsilon = 0$ in system (4.35) leads to the following system for $(\bar{x}_0, \bar{f}_t)$, coupling the dynamics of a deterministic continuous function on $\mathbb{R}_+$, $\bar{x}_0$, with those of a time-dependent measure on $\mathbb{N}^d$, $\bar{f}_t$, for all $t \geq 0$,

$$\left\{ \begin{array}{l} \frac{d}{dt} \bar{x}_0(t) = -(\bar{\lambda}_0 + \bar{\mu}_0) \bar{x}_0(t), \quad \bar{x}_0(0) = x_0^{in}, \\ \bar{\lambda}_0 = \sum_{x \in \mathbb{N}^d} \lambda_0(x) \bar{f}_t(x), \\ \bar{\mu}_0 = \sum_{x \in \mathbb{N}^d} \mu_0(x) \bar{f}_t(x), \\ 0 = \sum_{x \in \mathbb{N}^d} \bar{A}_{\bar{x}_0(t)} \psi(x) \bar{f}_t(x), \quad \forall \psi \in B(\mathbb{N}^d). \end{array} \right. \quad (4.37)$$

In system (4.37), x_0^{in} is a real positive constant, and, for any $x_0 > 0$, $\bar{A}_{x_0}$ is an operator, defined, for all bounded functions ψ on $\mathbb{N}^d$ and for all $x = (x_1, \dots, x_d) \in \mathbb{N}^d$, by

$$\begin{aligned} \bar{A}_{x_0} \psi(x) = & \lambda_0(x) x_0 [\psi(x + e_1) - \psi(x)] + \sum_{i=1}^d \lambda_i(x) x_i [\psi(x - e_i + e_{i+1}) - \psi(x)] \\ & + \sum_{i=1}^d \mu_i(x) x_i [\psi(x - e_i) - \psi(x)], \end{aligned} \quad (4.38)$$

where, for $i \in \{1, \dots, d\}$, e_i is a unit vector of $\mathbb{N}^d$, with coordinate 1 in the i^{th} position and zero elsewhere, and e_{d+1} is the null vector.

Remark 4. Although the limit system (4.37) may appear quite different from its deterministic counterpart (4.10), it has the same flavour : the fast variable is in a “quasi-equilibrium” at any time t . Its law f_t thus needs to solve the equilibrium of the Kolmogorov equations associated with SDE (4.35), which are written here with the help of the infinitesimal generator of the fast variable to ease the notations. More precisely, let A^ε be the infinitesimal generator associated with the process X^ε , solution of (4.35), then for all functions $\phi : \mathbb{R}_+ \times \mathbb{N}^d \rightarrow \mathbb{R}$ bounded and independent of the first variable ($\forall (x_0, x) \in \mathbb{R}_+ \times \mathbb{N}^d$, $\phi(x_0, x) = \psi(x)$), we have

$$\forall (x_0, x) \in \mathbb{R}_+ \times \mathbb{N}^d, \quad A^\varepsilon \phi(x_0, x) = \frac{1}{\varepsilon} \bar{A}_{x_0} \psi(x).$$

Thus $\bar{f}_t$ is the stationary solution corresponding to A^ε when the slow variable is “frozen”.

System (4.37) is not necessarily well-posed, as there may be several solutions $\bar{f}_t$ for a given $\bar{x}_0$. In the next two specific examples, we can prove that system (4.37) does admit a single solution, which is a natural limit candidate for the sequence X^ε .

Example 5 (Linear case). Let us suppose that $K_{1,i} = K_{2,i} = 0$ for all $i \in \{0, \dots, d\}$, and $(f_i + g_i) > 0$ for all $i \in \{1, \dots, d\}$. Then, system (4.37) becomes linear and has a unique solution. The invariant measure $\bar{f}_t$ has a product measure form

$$\forall x \in \mathbb{N}^d, \quad \bar{f}_t(x_1, \dots, x_d) = \prod_{i=1}^d \bar{f}_t^i(x_i), \quad (4.39)$$

with $\bar{f}_t^i$ a Poisson law on $\mathbb{N}$ of mean parameter $p_i x_0(t)$, with

$$p_i = \prod_{j=0}^{i-1} \frac{f_j}{f_{j+1} + g_{j+1}}. \quad (4.40)$$

System (4.37) then reduces to

$$\left\{ \begin{array}{ll} \bar{x}_0(t) &= x_0^{in} \exp(-(f_0 + g_0)t), \\ \forall x \in \mathbb{N}^d, \quad \bar{f}_t(x) &= \prod_{i=1}^d (p_i \bar{x}_0(t))^{x_i} \frac{e^{-p_i \bar{x}_0(t)}}{x_i!}. \end{array} \right. \quad (4.41)$$

It is classical that stationary distributions associated with the generator (4.38) are of product form for constant coefficients λ_i, μ_i (see for instance [42, 58, 55]). Taking the product form in Eq. (4.39) for granted, the following calculus shows that each marginal distribution has to be a Poisson law.

Indeed, for a function ψ which depends on the first variable only

$$\forall x_1, \dots, x_d, \quad \psi(x_1, \dots, x_d) = \psi_1(x_1),$$

and for $\bar{x}_0 > 0$, we obtain, using expression (4.38),

$$\begin{aligned} \sum_{x \in \mathbb{N}^d} \bar{A}_{\bar{x}_0} \psi(x) \bar{f}_t(x) &= \sum_{x \in \mathbb{N}^d} \{ [f_0 \bar{x}_0 (\psi_1(x_1 + 1) - \psi_1(x_1)) + (f_1 + g_1) x_1 (\psi_1(x_1 - 1) - \psi_1(x_1))] \\ &\quad \bar{f}_t^1(x_1) \prod_{i=2}^d \bar{f}_t^i(x_i) \} , \\ &= \left(\sum_{x \in \mathbb{N}} \{ f_0 \bar{x}_0 (\psi_1(x + 1) - \psi_1(x)) + (f_1 + g_1) x (\psi_1(x - 1) - \psi_1(x)) \} \bar{f}_t^1(x) \right) \\ &\quad \left(\sum_{(x_2, \dots, x_d) \in \mathbb{N}^{d-1}} \prod_{i=2}^d \bar{f}_t^i(x_i) \right) . \end{aligned}$$

Hence, the solution $\bar{f}_t$ is such that for any bounded function ψ_1

$$0 = \sum_{x \in \mathbb{N}} \psi_1(x) \{ f_0 \bar{x}_0 (\bar{f}_t^1(x - 1) \mathbf{1}_{x \geq 1} - \bar{f}_t^1(x)) + (f_1 + g_1) ((x + 1) \bar{f}_t^1(x + 1) - x \bar{f}_t^1(x)) \} .$$

This holds in particular for $\psi_1(x) = \mathbf{1}_{x=n}$ for any $n \in \mathbb{N}$, so that we obtain after calculus

$$\forall x \in \mathbb{N}, \quad \bar{f}_t^1(x) = \frac{1}{x!} \left(\frac{f_0 \bar{x}_0}{f_1 + g_1} \right)^x \bar{f}_t^1(0) ,$$

and then, by same arguments

$$\forall i \in \{1, \dots, d\}, \forall x \in \mathbb{N}, \quad \bar{f}_t^i(x) = \frac{1}{x!} (p_i \bar{x}_0)^x \bar{f}_t^i(0) ,$$

where p_i is defined in Eq. (4.40).

Example 6 (A single feedback onto the quiescent follicle death rate). *Let us suppose that $K_{1,i} = 0$ for all $i \in \{0, \dots, d\}$ and $K_{2,i} = 0$ for $i \in \{1, \dots, d\}$ but $K_{2,0} > 0$. Then, system (4.37) can be simplified as*

$$\begin{cases} \frac{d}{dt} \bar{x}_0(t) &= - (f_0 + \bar{\mu}_0(\bar{x}_0(t))) \bar{x}_0(t), \\ \bar{\mu}_0(\bar{x}_0) &= g_0 (1 + \bar{x}_0 K_{2,0} \sum_{j=1}^d b_j p_j), \\ \forall x \in \mathbb{N}^d, \quad \bar{f}_t(x) &= \prod_{i=1}^d (p_i \bar{x}_0(t))^{x_i} \frac{e^{-p_i \bar{x}_0(t)}}{x_i!}, \end{cases} \quad (4.42)$$

where p_i is defined in Eq. (4.40), and has a unique solution. The justification of system (4.42) follows that of Example (5) for the measure f_t , which is not directly modified by the feedback term onto the quiescent follicle death rates.

4.3.3 Convergence in the linear case

In this paragraph, we assume that $K_{1,i} = K_{2,i} = 0$ for all $i \in \{0, \dots, d\}$ as in Example 5, and we assume further that the initial condition $X^{\varepsilon, in} = (X_0^{\varepsilon, in}, X_1^{\varepsilon, in}, \dots, X_d^{\varepsilon, in})$ of system (4.35) is such that $\frac{1}{\varepsilon} X_0^{\varepsilon, in}, X_1^{\varepsilon, in}, \dots, X_d^{\varepsilon, in}$ are independent Poisson random variables

of mean respectively $x_0^{in}/\varepsilon, x_1^{in}, \dots, x_d^{in}$. Denoting by G^{in} the probability generating function of the integer-valued random vector $(\frac{1}{\varepsilon}X_0^{\varepsilon, in}, X_1^{\varepsilon, in}, \dots, X_d^{\varepsilon, in})$, we thus have

$$G^{in}(z) = \mathbb{E} \left\{ z_0^{X_0^{\varepsilon, in}/\varepsilon} \prod_{j=1}^d z_j^{X_j^{\varepsilon, in}} \right\} = \exp \left(\frac{x_0^{in}}{\varepsilon} (z_0 - 1) \right) \prod_{i=1}^d \exp (x_i^{in} (z_i - 1)) . \quad (4.43)$$

In such a case, one can solve explicitly system (4.35) for each $\varepsilon > 0$. We briefly sketch the formal arguments and computations (see [42] for details). We define the probability generating function of the vector $(X_0^\varepsilon(t)/\varepsilon, X_1^\varepsilon(t), \dots, X_d^\varepsilon(t))$, for $\mathbf{z} \in \mathbb{R}^{d+1}$, by

$$G^\varepsilon(\mathbf{z}, t) = \mathbb{E} \left\{ z_0^{X_0^\varepsilon(t)/\varepsilon} \prod_{j=1}^d z_j^{X_j^\varepsilon(t)} \right\} = \sum_{i=0}^d \sum_{n_i=0}^\infty \prod_{j=0}^d z_j^{n_j} \mathbb{P} \{ X_0^\varepsilon(t) = \varepsilon n_0, X_1^\varepsilon(t) = n_1, \dots, X_d^\varepsilon(t) = n_d \} . \quad (4.44)$$

The infinitesimal generator $\tilde{A}^\varepsilon$ of the process $\tilde{X}^\varepsilon(t) := (X_0^\varepsilon(t)/\varepsilon, X_1^\varepsilon(t), \dots, X_d^\varepsilon(t))$, is given, for all bounded functions ψ on $\mathbb{N}^{d+1}$ and for all $n = (n_0, n_1, \dots, n_d) \in \mathbb{N}^{d+1}$, by

$$\begin{aligned} \tilde{A}^\varepsilon \psi(n) &= f_0 n_0 [\psi(n - e_0 + e_1) - \psi(n)] + g_0 n_0 [\psi(n - e_0) - \psi(n)] \\ &\quad + \sum_{i=1}^{d-1} \frac{f_i}{\varepsilon} n_i [\psi(n - e_i + e_{i+1}) - \psi(n)] + \sum_{i=1}^d \frac{g_i}{\varepsilon} n_i [\psi(n - e_i) - \psi(n)] . \end{aligned} \quad (4.45)$$

Using the Kolmogorov backward equation,

$$\frac{d}{dt} \mathbb{E} [\psi(\tilde{X}^\varepsilon(t))] = \mathbb{E} [\tilde{A}^\varepsilon \psi(\tilde{X}^\varepsilon(t))] , \quad (4.46)$$

with $\psi(n) = \prod_{j=0}^d z_j^{n_j}$ (a truncation procedure is required to deal rigorously with such test functions), we obtain, using linearity of expectation, a first-order partial differential equation on G^ε , given by

$$\begin{aligned} \frac{\partial}{\partial t} G^\varepsilon(\mathbf{z}, t) &= -(f_0 + g_0)(z_0 - 1) \frac{\partial}{\partial z_0} G^\varepsilon(\mathbf{z}, t) + f_0(z_1 - 1) \frac{\partial}{\partial z_0} G^\varepsilon(\mathbf{z}, t) \\ &\quad + \sum_{i=1}^{d-1} (z_{i+1} - 1) \frac{f_i}{\varepsilon} \frac{\partial}{\partial z_i} G^\varepsilon(\mathbf{z}, t) - \sum_{i=1}^d (z_i - 1) \frac{1}{\varepsilon} (f_i + g_i) \frac{\partial}{\partial z_i} G^\varepsilon(\mathbf{z}, t) . \end{aligned} \quad (4.47)$$

It turns out that the unique solution to (4.47) is given by

$$G^\varepsilon(\mathbf{z}, t) = \exp \left(\frac{x_0^\varepsilon(t)}{\varepsilon} (z_0 - 1) \right) \prod_{i=1}^d \exp (x_i^\varepsilon(t) (z_i - 1)) . \quad (4.48)$$

where $(x_i^\varepsilon(t))_{i=0, \dots, d}$ is solution of the very same ODE as the linear version of Eq. (4.7) (whose solutions are given by Eq. (4.13)), with initial condition $x_i^\varepsilon(t=0) = x_i^{in}$, for all $i = 0 \dots d$. Thus, at any time t , $X_0^\varepsilon(t)/\varepsilon, X_1^\varepsilon(t), \dots, X_d^\varepsilon(t)$ are independent Poisson random variables of mean (respectively) $\frac{x_0^{in}(t)}{\varepsilon}, x_1^{in}(t), \dots, x_d^{in}(t)$.

Proposition 3. Assume that $K_{1,i} = K_{2,i} = 0$ for all $i \in \{0, \dots, d\}$, that $f_i + g_i$ is strictly positive for $i \in \{1, \dots, d\}$ and that the initial condition $X^{\varepsilon, in}$ is such that Eq. (4.43) holds. Then, for all $t > 0$, $X_0^\varepsilon(t)$ converges (in law) towards the deterministic value $\bar{x}_0(t) = x_0^{in} \exp(-(f_0 + g_0)t)$, and, for all $i = 1, \dots, d$, $X_i^\varepsilon(t)$ converges (in law) towards a Poisson random variable of mean $\bar{x}_i(t) = p_i \bar{x}_0(t)$, where p_i is given by Eq. (4.40).

Démonstration. The proof is a direct consequence of the explicit solution (4.48) for the probability generating function G^ε combined with Proposition 1. As $X_0^\varepsilon(t)/\varepsilon$ is a Poisson random variable of mean $\frac{\bar{x}_0(t)}{\varepsilon}$, it is clear that $\mathbb{E}[X_0^\varepsilon(t)] = \bar{x}_0(t)$ and $\text{var}(X_0^\varepsilon(t)) = \varepsilon \bar{x}_0^2(t) \leq \varepsilon x_0^{\text{in}} \rightarrow 0$ as $\varepsilon \rightarrow 0$, which implies that $X_0^\varepsilon(t)$ converges in law towards $\bar{x}_0(t)$. For all $i = 1, \dots, d$, and $t > 0$, as $x_i^\varepsilon(t)$ converges to $p_i \bar{x}_0(t)$ as $\varepsilon \rightarrow 0$, $X_i^\varepsilon(t)$ converges (in law) towards a Poisson random variable of mean $p_i \bar{x}_0(t)$. $\square$

Remark 5. The slow variable X_0^ε is independent of the fast variables and its dynamics are reduced to a “death and death” process with constant rate. Then convergence of this variable is given by Theorem 8.1 of [63], which enables us to obtain stronger results with fewer hypotheses on the initial condition. Suppose that

$$\lim_{\varepsilon \rightarrow 0} X_0^{\varepsilon, \text{in}} = x_0^{\text{in}}, \quad \text{a.s.},$$

then

$$\limsup_{\varepsilon \rightarrow 0} \sup_{s \leq t} |X_0^\varepsilon(s) - \bar{x}_0(s)| = 0 \quad \text{a.s. for all } t > 0.$$

4.3.4 Numerical convergence

In this paragraph, we illustrate the convergence of $(X_0^\varepsilon, X_1^\varepsilon, \dots, X_d^\varepsilon)$ as $\varepsilon \rightarrow 0$. The chosen scenario and the parameter values are detailed in the Appendix (section 4.4).

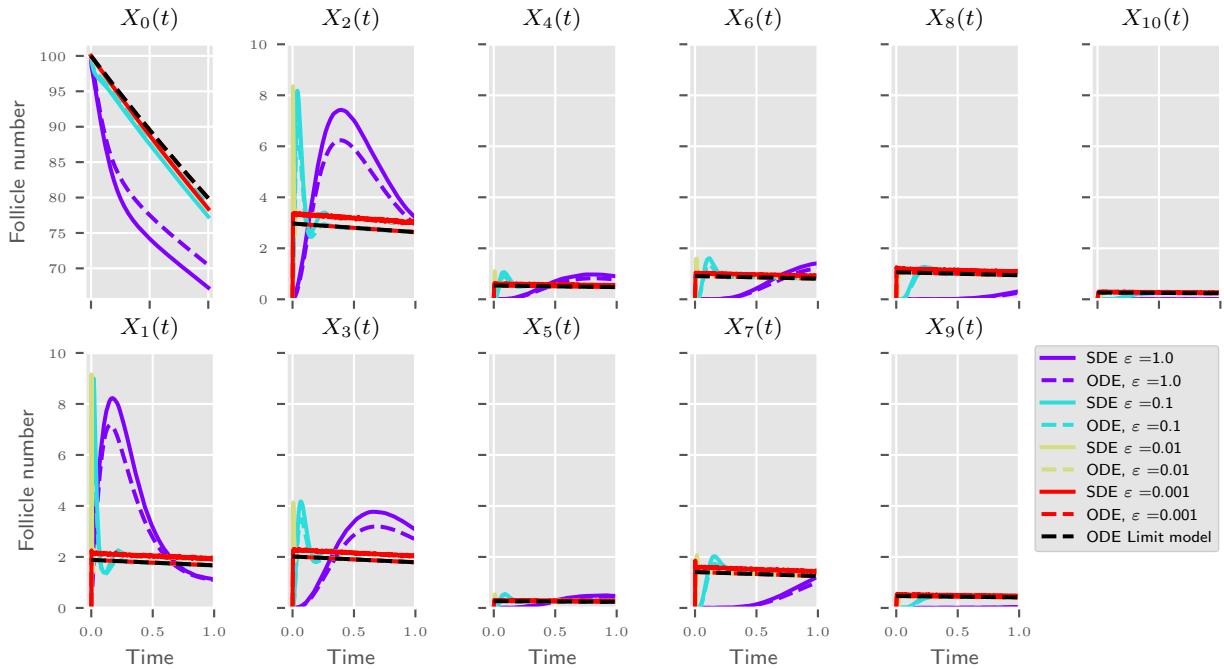


FIGURE 4.7 – In the same way as in Figure 4.2, we plot the trajectories in each maturity compartment ($d = 10$), for the rescaled variables X_i^ε of the SDE system (4.35) (solid lines) and the ODE system (4.7) (dashed lines), for different ε (see legend). For the SDE, we plot the empirical mean computed over 10000 trajectories. The limit variable $\bar{X}_i$ of the ODE system corresponds to the black dashed line.

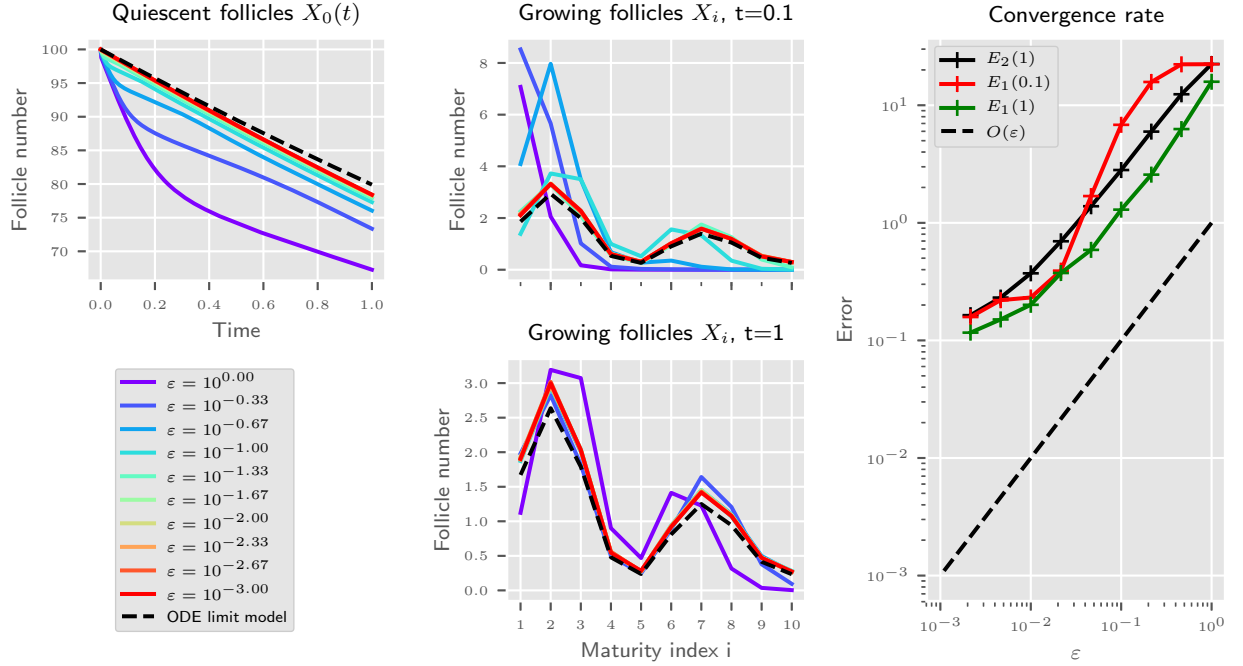


FIGURE 4.8 – Trajectories in the quiescent follicle compartment (top left panel) and distribution of the growing follicle population at time $t = 0.1$ (top center panel) and at time $t = 1$ (bottom center panel), for the rescaled variables X_i^ε , for different ε (solid colored lines, see legend insert) and the limit variable $\bar{X}_i$ (black dashed line). On the right panel, we plot the discrete l^1 norm error $E_1(t)$ at a fixed time $t = 0.1$ and $t = 1$ (solid red and green lines, resp.) and the l^1 -cumulative error $E_2(1)$ on $t \in (0, 1)$ (black solid line) as a function of ε (see details in body text). The black dashed line is the straight line of slope 1 according to ε .

In Figure 4.7, we plot the empirical mean trajectories (computed over 10^4 sampled trajectories) in each maturity compartment ($d = 10$) for the rescaled model on a time horizon $t \in (0, 1)$ in the nonlinear scenario ($K_{1,0} > 0$), together with the trajectories of the analogous ODE rescaled system (4.7) and its limit (4.10). We observe that, for each compartment, the empirical mean of the SDE seems to converge to a limit value, yet this limit does not superimpose with the ODE limit solution (which is expected in a nonlinear scenario, as the ODE and SDE limits are different).

In Figure 4.8, using the same parameters as in Figure 4.7, we display the maturity distribution in the growing follicle population, for various ε . We empirically quantify the convergence rate using the following error, at time t ,

$$E_1(t) = \sum_{i=0}^d | \mathbb{E}X_i^\varepsilon(t) - \mathbb{E}\bar{X}_i(t) |, \quad (4.49)$$

and the cumulative error on time interval $[0, T]$,

$$E_2(T) = \int_0^T \sum_{i=0}^d | \mathbb{E}X_i^\varepsilon(t) - \mathbb{E}\bar{X}_i(t) | dt, \quad (4.50)$$

which can be assessed numerically as

$$\tilde{E}_2(T) = \sum_{k=0}^{N_T} \delta_t \sum_{i=0}^d | \mathbb{E} X_i^\varepsilon(t_k) - \mathbb{E} \bar{X}_i(t_k) |, \quad (4.51)$$

where $t_k = k\delta_t$, for $k = 0 \cdots N_T$. In practice, we also replace the limit model $\bar{X}$ by the numerically evaluated limit model X^ε with $\varepsilon = 0.001$. We then observe that the error decreases roughly linearly with ε .

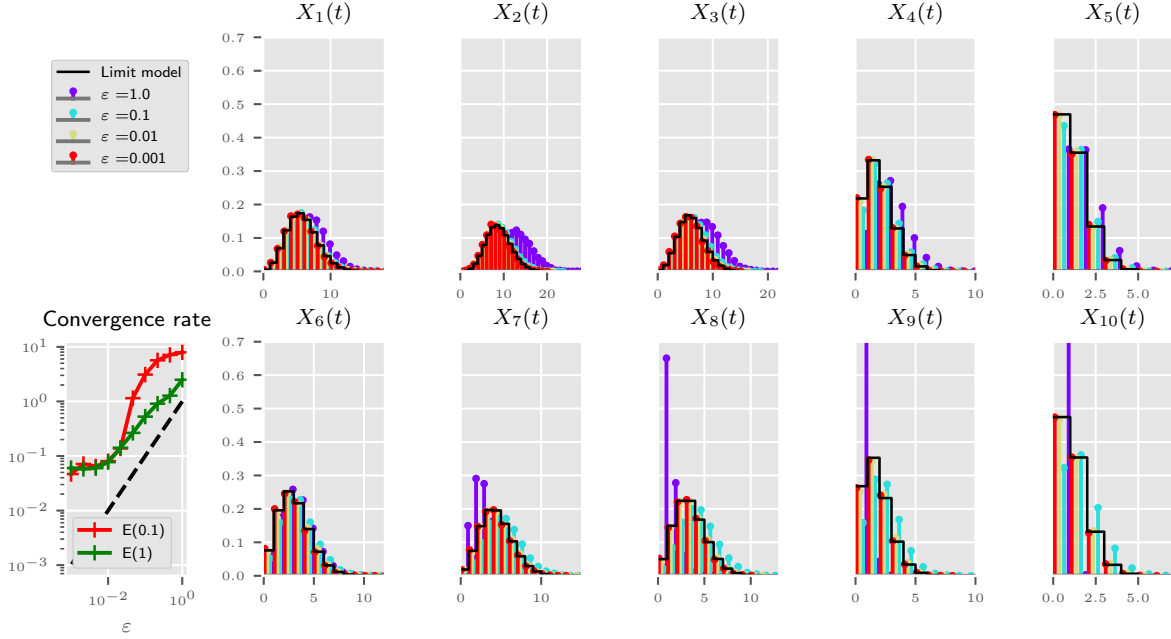


FIGURE 4.9 – Empirical law of X_i^ε in each maturity compartment at time $t = 1$ for different ε in colored bars (see legend insert) and the limit distribution $\bar{X}_i$ (black solid lines). On the bottom leftmost panel, we plot the total variation error $E(t)$ at a fixed time $t = 0.1$ and $t = 1$ (solid red and green lines, resp.). The black dashed line is the straight line of slope 1 according to ε .

In Figure 4.9, we use the linear scenario ($K_{1,0} = 0$) detailed in the Appendix (section 4.4) to visualize the convergence of the fast variable of the rescaled SDE to the “quasi-stationary” distribution of the limit model. The marginals of the rescaled model are evaluated over 10^4 sampled trajectories at time $t = 0.1$ and $t = 1$. The errors between the marginal laws of the rescaled and limit models are quantified by the total variation (restricted on the support of the numerically assessed limit model) :

$$E(t) = \sum_{i=1}^d d_{TV}(X_i^\varepsilon(t), \bar{X}_i(t)), \quad (4.52)$$

where

$$d_{TV}(X, Y) = \max\{ | \pi_X(i) - \pi_Y(i) |, i \in \mathbb{N} \}, \quad (4.53)$$

and π_X, π_Y are, respectively, the laws of X and Y . The error seems to decrease in a sub-linear manner with ε , with a plateau for $\varepsilon < 10^{-2}$, which is probably due to the limited finite sampling size (10^4).

4.4 Appendix - parameter values

In the numerical illustrations provided throughout the previous sections, we refer to either a linear or nonlinear scenario. As far as parameter values, the only difference is that parameter $K_{1,0}$ is set to 0 in the linear scenario. All other parameters are identical and chosen as explained below.

Numerical simulation of the PDE model

We begin by shaping the desired solution $H(x)$ and we choose functions f and g accordingly :

$$H(x) = e^{-\int_0^x \frac{g(y) + f'(y)}{f(y)} dy}.$$

$$\text{Choosing } g(x) = 1, \quad \text{we get} \quad f(x) = \frac{c - \int_0^x H(y) dy}{H(x)}.$$

Motivated by our biological application, we more specifically select a two-bump function

$$H(x) = \frac{p_1 e^{-\frac{(x-x_1)^2}{2s^2}} + p_2 e^{-\frac{(x-x_2)^2}{2s^2}}}{p_1 e^{-\frac{x_1^2}{2s^2}} + p_2 e^{-\frac{x_2^2}{2s^2}}},$$

with $s = 0.1$, $x_1 = 0.2$, $p_1 = 0.7$, $x_2 = 0.7$ and $p_2 = 0.3$.

Except $K_{1,0}$, all coefficients weighting the nonlinear terms, $K_{2,0}$, K_1 , K_2 , are set to zero. As a result, functions b , ω_1 and ω_2 , that were introduced for the sake of genericity, are not used in the numerical illustrations.

The basal activation rate f_0 is set to $f_0 = 1$, and the basal death rate in quiescent follicles g_0 is set to $g_0 = 0.1$.

Since the population feedback onto the activation rate is mainly exerted by follicles in an intermediate maturity stage, we choose $a(x) = 1_{[0.3, 0.7]}$, given that the state space lies in $x \in [0, 1]$. The feedback gain is set to $K_{1,0} = 2$.

Finally, the time horizon covers $t \in (0, 1)$, and the initial condition is given by $\rho_0^{ini} = 100$ and $\rho^{ini} \equiv 0$.

The parameter values are summed up in Table 4.1 and illustrated on Figure 4.10.

s	p_1	x_1	p_2	x_2	c		
0.1	0.7	0.2	0.3	0.7	2.61		
f_0	g_0	$K_{1,0}$	$K_{2,0}$	ρ_{ini}^0	a	$K_{1,d}, d \neq 0$	$K_{2,d}, d \neq 0$
1	0.1	2	0	100	$1_{[0.3, 0.7]}$	0	0

TABLE 4.1 – Parameter values for the numerical simulations

Numerical simulation of the ODE/SDE model

For the ODE- and SDE-based models, we first set the number of compartments, $d = 10$, and define $x_i = i/d$. Then, we adapt the functions selected in the continuous PDE case and set, for $i \in \{1, d\}$,

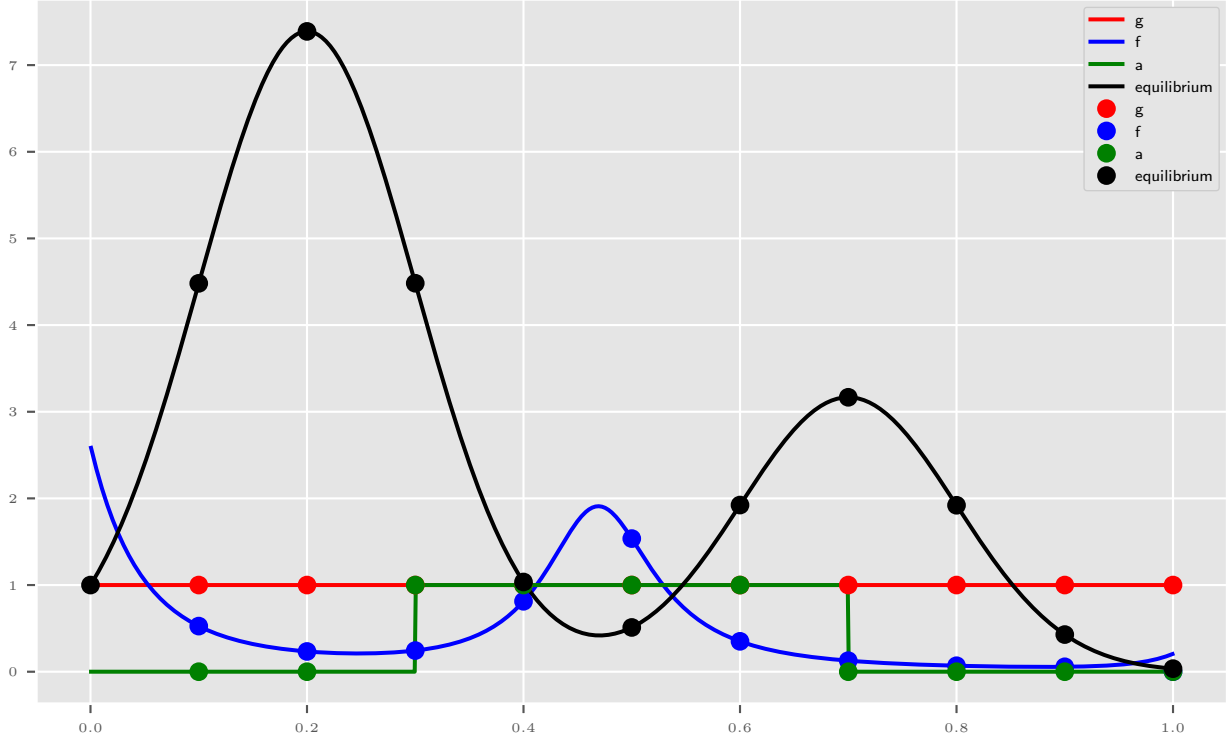


FIGURE 4.10 – Parametric functions used in the numerical simulations, for the PDE (plain lines), the ODE and the SDE (dot symbols).

- $a_i = a(x_i)$,
- $f_i = f(x_i)/d$,
- $g_i = g(x_i)$,
- $K_{1,i} = K_{2,i} = 0$.

All other parameters are kept as in Table 4.1, while the initial condition is chosen as $Y_0^{ini} = 100$ and $Y_i^{ini} = 0$, $i \in \{1, d\}$.

To simulate the ODE model, we use the standard python `scipy.odeint`, while, to simulate the SDE model, we use an exact stochastic simulation algorithm (Gillespie).

Bibliographie

- [1] J L Abkowitz, D Golinelli, D E Harrison, and P Gutter. In vivo kinetics of murine hemopoietic stem cells. *Blood*, 96(10) :3399–3405, 2000.
- [2] O. Angulo, F. Crauste, and J. C. López-Marcos. Numerical integration of an erythropoiesis model with explicit growth factor dynamics. *Journal of Computational and Applied Mathematics*, 330 :770–782, 2018.
- [3] O Arino and M Kimmel. Stability analysis of models of cell production systems. *Mathematical Modelling*, 7(9-12) :1269–1300, 1986.
- [4] DE Axelrod and M Kimmel. *Branching Processes in Biology*. Springer, New York, 2002.
- [5] V. Bansaye and S. Méléard. *Stochastic Models for Structured Populations : Scaling Limits and Long Time Behavior*. Mathematical Biosciences Institute Lecture Series. Springer International Publishing, 2015.
- [6] M Benaïm, T Hurth, and E Strickler. A user-friendly condition for exponential ergodicity in randomly switched environments. *Electronic Communications in Probability*, 23, 2018.
- [7] M Benaïm, S Le Borgne, F Malrieu, and P-A Zitt. Qualitative properties of certain piecewise deterministic markov processes. In *Annales de l’IHP Probabilités et statistiques*, volume 51, pages 1040–1075, 2015.
- [8] M Benaïm and C Lobry. Lotka–volterra with randomly fluctuating environments or “how switching between beneficial environments can make survival harder”. *The Annals of Applied Probability*, 26(6) :3754–3785, 2016.
- [9] I Bendixson. Sur les courbes définies par des équations différentielles. *Acta Mathematica*, 24 :1–88, 1901.
- [10] A Besse, G D Clapp, S Bernard, F E Nicolini, D Levy, and T Lepoutre. Stability analysis of a model of interaction between the immune system and cancer cells in chronic myelogenous leukemia. *Bulletin of mathematical biology*, 80(5) :1084–1110, 2018.
- [11] C. Bonnet, K. Chahour, F. Clément, M. Postel, and R. Yvinec. Multiscale population dynamics in reproductive biology : singular perturbation reduction in deterministic and stochastic models. *arXiv preprint arXiv :1903.08555*, 2019.
- [12] C Bonnet, P Gou, S Girel, V Bansaye, C Lacout, K Bailly, M-H Schlagetter, E Lauret, S Meleard, and S Giraudier. Modeling the behavior of hematopoietic compartments from stem to red cells in murine steady state and stress hematopoiesis. *Submitted to Cell Stem Cell*, 2020.
- [13] K. Busch, K. Klapproth, M. Barile, M. Flossdorf, T. Holland-Letz, S.M. Schlenner, M. Reth, T. Höfer, and H-R. Rodewald. Fundamental properties of unperturbed haematopoiesis from stem cells in vivo. *Nature*, 18(518) :542–546, 2015.

- [14] S N Catlin, P Gutterorp, and J L Abkowitz. The kinetics of clonal dominance in myeloproliferative disorders. *Blood*, 106(8) :2688–2692, 2005.
- [15] F. Clément and D. Monniaux. Multiscale modelling of ovarian follicular selection. *Prog. Biophys. Mol. Biol.*, 113(3) :398–408, 2013.
- [16] B Cloez, R Dessalles, A Genadot, F Malrieu, A Marguet, and R Yvinec. Probabilistic and piecewise deterministic models in biology. *ESAIM : Proceedings and Surveys*, 60 :225–245, 2017.
- [17] J.F Collet and T. Goudon. On solutions of the Lifshitz-Slyozov model. *Nonlinearity*, 13(4) :1239, 2000.
- [18] F. Crauste, I. Demin, O. Gandrillon, and V. Volpert. Mathematical study of feedback control roles and relevance in stress erythropoiesis. *Journal of theoretical biology*, 263(3) :303–316, 2010.
- [19] F. Crauste, L. Pujo-Menjouet, S. Génieys, C. Molina, and O. Gandrillon. Adding self-renewal in committed erythroid progenitors improves the biological relevance of a mathematical model of erythropoiesis. *Journal of theoretical biology*, 250(2) :322–338, 2008.
- [20] A Crudu, A Debussche, A Muller, and O Radulescu. Convergence of stochastic gene networks to hybrid piecewise deterministic processes. *The Annals of Applied Probability*, 22(5) :1822–1859, 2012.
- [21] M HA Davis. Piecewise-deterministic markov processes : A general class of non-diffusion stochastic models. *Journal of the Royal Statistical Society : Series B (Methodological)*, 46(3) :353–376, 1984.
- [22] MHA Davis. Markov models and optimization., 1993.
- [23] D.Hulin. Equations différentielles ordinaires. etudes qualitatives. *Cours L3 MFA Université Paris-Sud*.
- [24] D Dingli and F Michor. Successful therapy must eradicate cancer stem cells. *Stem cells*, 24(12) :2603–2610, 2006.
- [25] D Dingli and J M Pacheco. Stochastic dynamics and the evolution of mutations in stem cells. *BMC biology*, 9(1) :41, 2011.
- [26] D Dingli, A Traulsen, and F Michor. (a) symmetric stem cell replication and cancer. *PLoS computational biology*, 3(3), 2007.
- [27] D Dingli, A Traulsen, and JM Pacheco. Compartmental architecture and dynamics of hematopoiesis. *PlosOne*, 2007.
- [28] J. Donnez and M.-M. Dolmans. Transplantation of ovarian tissue. *Best Pract. Res. Clin. Obstet. Gynaecol.*, 28(8) :1188–1197, 2014.
- [29] M Doumic, A Marciniak-Czochra, B Perthame, and J Zubelli. Mathematical model of stem cell differentiation and tissue regeneration with stochastic noise. *SIAM Journal on Applied Mathematics*, 71(6) :1918–1940, 2011.
- [30] M.-A. Driancourt, A. Gougeon, D. Monniaux, D. Royère, and C. Thibault. Folliculogénèse et ovulation. In C. Thibault and M.C. Levasseur, editors, *La reproduction chez les mammifères et l’homme*, pages 317–347. INRA, 2001.
- [31] H Ema, K Sudo, J Seita, A Matsubara, Y Morita, M Osawa, K Takatsu, S Takaki, and H Nakauchi. Quantification of self-renewal capacity in single hematopoietic stem cells from normal and lmk-deficient mice. *Developmental cell*, 8(6) :907–914, 2005.

- [32] M AG Essers, S Offner, W E Blanco-Bose, Z Waibler, U Kalinke, M A Duchosal, and A Trumpp. Ifn α activates dormant haematopoietic stem cells in vivo. *Nature*, 458(7240) :904–908, 2009.
- [33] S. Ethier and T. Kurtz. *Markov processes – characterization and convergence*. Wiley Series in Probability and Mathematical Statistics : Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 1986.
- [34] J. Evers, S. Hille, and A. Muntean. Measure-Valued Mass Evolution Problems with Flux Boundary Conditions and Solution-Dependent Velocities. *SIAM J. Math. Anal.*, 48(3) :1929–1953, January 2016.
- [35] J.H.M. Evers, S.C. Hille, and A. Muntean. Mild solutions to a measure-valued mass evolution problem with flux boundary conditions. *J. Differential Equations*, 259(3) :1068–1097, 2015.
- [36] M. J. Faddy, E.C. Jones, and R.G. Edwards. An analytical model for ovarian follicle dynamics. *J. Exp. Zool.*, 197(2) :173–185, 1976.
- [37] M.J. Faddy and E.C. Jones. Fitting time-dependent multicompartment models : a case study. *Biometrics*, 44(2) :587–593, 1988.
- [38] A Fasano and A Sequeira. Blood and cancer. In *Hemomath*, pages 295–330. Springer, 2017.
- [39] J Fontbona, H Guérin, and F Malrieu. Quantitative estimates for the long-time behavior of an ergodic variant of the telegraph process. *Advances in Applied Probability*, 44(4) :977–994, 2012.
- [40] A. Foudi, K. Hochedlinger, D. Van Buren, J. W. Schindler, R. Jaenisch, V. Carey, and H. Hock. Defining hematopoietic stem and progenitor cell turnover by analysis of histone 2b-gfp dilution. *Nature biotechnology*, 27(1) :84, 2009.
- [41] N. Fournier and S. Méléard. A microscopic probabilistic description of a locally regulated population and macroscopic approximations. *The Annals of Applied Probability*, 14(4) :1880–1919, 2004.
- [42] C. Gadgil, C.H. Lee, and H.G. Othmer. A stochastic analysis of first-order reaction networks. *Bull. Math. Biol.*, 67 :901–946, 2005.
- [43] I Glauche, K Horn, M Horn, L Thielecke, M AG Essers, Andreas Trumpp, and I Roeder. Therapy of chronic myeloid leukaemia can benefit from the activation of stem cells : simulation studies of different treatment combinations. *British journal of cancer*, 106(11) :1742–1752, 2012.
- [44] M Gonzalez, I M Puerto, R Martinez, M Molina, M Mota, and A Ramos. *Workshop on branching processes and their applications*, volume 197. Springer Science and Business Media, 2010.
- [45] A. Gougeon. Regulation of ovarian follicular development in primates : facts and hypotheses. *Endocr. Rev.*, 11 :121–155, 1996.
- [46] S Haas, J Hansson, D Klimmeck, D Loeffler, L Velten, H Uckelmann, S Wurzer, Á M Prendergast, A Schnell, and K Hexel. Inflammation-induced emergency megakaryopoiesis driven by hematopoietic stem cell-like megakaryocyte progenitors. *Cell stem cell*, 17(4) :422–434, 2015.
- [47] H Haeno, R L Levine, D G Gilliland, and F Michor. A progenitor cell origin of myeloid malignancies. *Proceedings of the National Academy of Sciences*, 106(39) :16616–16621, 2009.

- [48] N. Hansen and A. Ostermeier. Completely derandomized self-adaptation in evolution strategies. *Evolutionary computation*, 9(2) :159–195, 2001.
- [49] M. Herbert, J. Wolstenholme, A.P. Murdoch, and T.J. Butler. Mitotic activity during preimplantation development of human embryos. *Reproduction*, 103(2) :209–214, 1995.
- [50] M Horn, M Loeffler, and I Roeder. Mathematical modeling of genesis and treatment of chronic myeloid leukemia. *Cells Tissues Organs*, 188(1-2) :236–247, 2008.
- [51] N. Ikeda and S. Watanabe. *Stochastic Differential Equations and Diffusion Processes*, 2nd ed. North-Holland, 1989.
- [52] M Jacobsen. *Point process theory and applications : marked point and piecewise deterministic processes*. Springer Science & Business Media, 2006.
- [53] A. Joffe and M. Métivier. Weak convergence of sequences of semimartingales with applications to multitype branching processes. *Advances in Applied Probability*, 18(1) :20–65, 1986.
- [54] H-W Kang, T G Kurtz, and L Popovic. Central limit theorems and diffusion approximations for multiscale markov chain models. *The Annals of Applied Probability*, 24(2) :721–759, 2014.
- [55] F. P. Kelly. *Reversibility and Stochastic Networks*. Wiley, Chichester, 1979.
- [56] M Khammash, B Munsky, and S Peleš. Reduction and solution of the chemical master equation using time scale separation and finite state projection. *The Journal of chemical physics*, 125(20), 2006.
- [57] M Kimmel and M Wazewska-Czyzewska. Stochastic approach to the process of red cell destruction. *Applicaciones Mathematicae*, 2(17) :217–225, 1982.
- [58] J.F.C. Kingman. Markov Population Processes. *J. Appl. Probab.*, 6(1) :1–18, 1969.
- [59] M. Klose, M. C. Florian, A. Gerbault, H. Geiger, and I. Glauche. Hematopoietic stem cell dynamics are regulated by progenitor demand : Lessons from a quantitative modeling approach. *Stem cells*, 37(7) :948–957, 2019.
- [60] N L Komarova and D Wodarz. Effect of cellular quiescence on the success of targeted cml therapy. *PLoS one*, 2(10), 2007.
- [61] M. J. Koury and M. C. Bondurant. Erythropoietin retards dna breakdown and prevents programmed death in erythroid progenitor cells. *Science*, 248(4953) :378–381, 1990.
- [62] T G Kurtz. Averaging for martingale problems and stochastic approximation. In *Applied Stochastic Analysis*, pages 186–209. Springer, 1992.
- [63] T.G. Kurtz. *Approximation of population processes*. SIAM, 1981.
- [64] A Y Lai and M Kondo. Asymmetrical lymphoid and myeloid lineage commitment in multipotent hematopoietic progenitors. *The Journal of experimental medicine*, 203(8) :1867–1873, 2006.
- [65] J Lei and M C Mackey. Stochastic differential delay equation, moment stability, and application to hematopoietic stem cell regulation system. *SIAM journal on applied mathematics*, 67(2) :387–407, 2007.
- [66] J-P Lepeltier and B Marchal. Problème des martingales et équations différentielles stochastiques associées à un opérateur intégro-différentiel. In *Annales de l’IHP Probabilités et statistiques*, volume 12, pages 43–103, 1976.

- [67] D Li and S Liu. Threshold dynamics and ergodicity of an sirs epidemic model with markovian switching. *Journal of Differential Equations*, 263(12) :8873–8915, 2017.
- [68] H. Li, A. Natarajan, J. Ezike, M. I. Barrasa, Y. Le, Z. A. Feder, H. Yang, C. Ma, S. Markoulaki, and H. F. Lodish. Rate of progression through a continuum of transit-amplifying progenitor cell states regulates blood cell production. *Developmental cell*, 49(1) :118–129, 2019.
- [69] H. Lodish, A. Berk, S. L. Zipursky, P. Matsudaira, D. Baltimore, and J. Darnell. Molecular cell biology 4th edition. *National Center for Biotechnology Information, Bookshelf*, 2000.
- [70] M Loeffler, K Pantel, H Wulff, and HE Wichmann. A mathematical model of erythropoiesis in mice and rats part 1 : Structure of the model. *Cell Proliferation*, 22(1) :13–30, 1989.
- [71] M Loeffler and HE Wichmann. A comprehensive mathematical model of stem cell proliferation which reproduces most of the published experimental results. *Cell Proliferation*, 13(5) :543–561, 1980.
- [72] M C Mackey, M Tyran-Kaminska, and R Yvinec. Dynamic behavior of stochastic gene expression models in the presence of bursting. *SIAM Journal on Applied Mathematics*, 73(5) :1830–1852, 2013.
- [73] F Malrieu. Some simple but challenging markov processes. In *Annales de la Faculté des sciences de Toulouse : Mathématiques*, volume 24, pages 857–883, 2015.
- [74] E Manesso, J Teles, D Bryder, and C Peterson. Dynamical modelling of haemato-poiesis : an integrated view over the system in homeostasis and under perturbation. *Journal of the Royal Society Interface*, 10(80) :20120817, 2013.
- [75] A. Marciniak-Czochra, T. Stiehl, A. D. Ho, W. Jäger, and W. Wagner. Modeling of asymmetric cell division in hematopoietic stem cells—regulation of self-renewal is essential for efficient repopulation. *Stem cells and development*, 18(3) :377–386, 2009.
- [76] J.C. Mariana and C. Millier. Application de quelques modèles de dynamique des populations à l’étude de la folliculogenèse ovarienne. *Ann. Biol. Anim. Biochim. Biophys.*, 17 :193–206, 1977.
- [77] F Michor. Mathematical models of cancer stem cells. *Journal of Clinical Oncology*, 26(17) :2854–2861, 2008.
- [78] D. Monniaux. Factors influencing establishment of the ovarian reserve and their effects on fertility. *Anim. Reprod.*, 15(Suppl.1) :635–647, 2018.
- [79] D. Monniaux, V. Cadoret, F. Clément, R. Dalbès-Tran, S. Elis, S. Fabre, V. Maillard, P. Monget, and S. Uzbekova. Folliculogenesis. In I. Huhtaniemi and L. Martini, editors, *Encyclopedia of Endocrine Diseases*, pages 377–398. Elsevier, (second edition) edition, 2019.
- [80] D. Monniaux, F. Clément, R. Dalbès-Tran, A. Estienne, S. Fabre, C. Mansanet, and P. Monget. The ovarian reserve of primordial follicles and the dynamic reserve of antral growing follicles : what is the link ? *Biol. Reprod.*, 90(4) :85, 2014.
- [81] P R Pazdziorek. Mathematical model of stem cell differentiation and tissue regeneration with stochastic noise. *Bulletin of mathematical biology*, 76(7) :1642–1669, 2014.

- [82] T. Pedersen and H. Peters. Proposal for a classification of oocytes and follicles in the mouse ovary. *J. Reprod. Fert.*, 17 :555–557, 1968.
- [83] M. Ramalho-Santos and H. Willenbring. On the origin of the term “stem cell”. *Cell stem cell*, 1(1) :35–38, 2007.
- [84] P. Reddy, W. Zheng, and K. Liu. Mechanisms maintaining the dormancy and survival of mammalian primordial follicles. *Trends Endocrinol. Metab.*, 21(2) :96–103, 2010.
- [85] I Roeder, M Horn, I Glauche, A Hochhaus, M C Mueller, and M Loeffler. Dynamic modeling of imatinib-treated chronic myeloid leukemia : functional insights and clinical implications. *Nature medicine*, 12(10) :1181–1184, 2006.
- [86] I Roeder and M Loeffler. A novel dynamic model of hematopoietic stem cell organization based on the concept of within-tissue plasticity. *Experimental hematology*, 30(8) :853–861, 2002.
- [87] K B Schoedel, M NF Morcos, T Zerjatke, I Roeder, T Grinenko, D Voehringer, J R Göthert, C Waskow, A Roers, and A Gerbaulet. The bulk of the hematopoietic stem cell population is dispensable for murine steady-state and stress hematopoiesis. *Blood*, 128(19) :2285–2296, 2016.
- [88] J-W Shin, A Buxboim, K R Spinler, J Swift, D A Christian, C A Hunter, C Léon, C Gachet, PC D P Dingal, and I L Ivanovska. Contractile forces sustain and polarize hematopoiesis from stem and progenitor cells. *Cell stem cell*, 14(1) :81–93, 2014.
- [89] T. Stiehl, A. D. Ho, and A. Marciniak-Czochra. Mathematical modeling of the impact of cytokine response of acute myeloid leukemia cells on patient prognosis. *Scientific reports*, 8(1) :1–11, 2018.
- [90] T Stiehl and A Marciniak-Czochra. Mathematical modeling of leukemogenesis and cancer stem cell dynamics. *Mathematical Modelling of Natural Phenomena*, 7(1) :166–202, 2012.
- [91] T. Tak, G. Prevedello, G. Simon, N. Paillon, K. R. Duffy, and L. Perie. Simultaneous tracking of division and differentiation from individual hematopoietic stem and progenitor cells reveals within-family homogeneity despite population heterogeneity. *bioRxiv*, page 586354, 2019.
- [92] P Thirup. Haematocrit. *Sports Medicine*, 33(3) :231–243, 2003.
- [93] J E Till, E A McCulloch, and L Siminovitch. A stochastic model of stem cell proliferation, based on the growth of spleen colony-forming cells. *Proceedings of the National Academy of Sciences of the United States of America*, 51(1) :29, 1964.
- [94] Vladimir Alekseevich Vatutin and Elena Evgen’evna Dyakonova. Decomposable branching processes with a fixed extinction moment. *Proceedings of the Steklov Institute of Mathematics*, 290(1) :103–124, 2015.
- [95] J.A. Visser and A.P.N. Themmen. Anti-müllerian hormone and folliculogenesis. *Mol. Cell. Endocrinol.*, 234(1) :81–86, 2005.
- [96] W Wasow. *Asymptotic expansions for ordinary differential equations*. Dover Publications Inc, 1987.
- [97] Z L Whichard, C A Sarkar, M Kimmel, and S J Corey. Hematopoiesis and its disorders : a systems biology approach. *Blood, The Journal of the American Society of Hematology*, 115(12) :2339–2347, 2010.

- [98] HE Wichmann, M Loeffler, K Pantel, and H Wulff. A mathematical model of erythropoiesis in mice and rats part 2 : Stimulated erythropoiesis. *Cell Proliferation*, 22(1) :31–49, 1989.
- [99] A Wilson, E Laurenti, G Oser, R C van der Wath, W Blanco-Bose, M Jaworski, S Offner, C F Dunant, L Eshkind, and E Bockamp. Hematopoietic stem cells reversibly switch from dormancy to self-renewal during homeostasis and repair. *Cell*, 135(6) :1118–1129, 2008.
- [100] M Wu, H Y Kwon, F Rattis, J Blum, C Zhao, R Ashkenazi, T L Jackson, N Gaiano, T Oliver, and T Reya. Imaging hematopoietic precursor division in real time. *Cell stem cell*, 1(5) :541–554, 2007.
- [101] H Wulff, HE Wichmann, K Pantel, and M Loeffler. A mathematical model of erythropoiesis in mice and rats part 3 : Suppressed erythropoiesis. *Cell Proliferation*, 22(1) :51–61, 1989.
- [102] R Yamamoto, Y Morita, J Ooehara, S Hamanaka, M Onodera, K L Rudolph, H Ema, and H Nakauchi. Clonal analysis unveils self-renewing lineage-restricted progenitors generated directly from hematopoietic stem cells. *Cell*, 154(5) :1112–1126, 2013.
- [103] R Yvinec, C Zhuge, J Lei, and M C Mackey. Adiabatic reduction of a model of stochastic gene expression with jump markov process. *Journal of mathematical biology*, 68(5) :1051–1070, 2014.
- [104] F. Ziebell, A. Martin-Villalba, and A. Marciniak-Czochra. Mathematical modelling of adult hippocampal neurogenesis : effects of altered stem cell dynamics on cell counts and bromodeoxyuridine-labelled cells. *Journal of The Royal Society Interface*, 11(94) :20140144, 2014.

Titre : Différenciation cellulaire, régulation des cellules souches et impact des mutations : une approche probabiliste.

Mots clés : Erythropoïèse, Modélisation mathématique, Processus Markovien, Traitement de données *in vivo*, Système lent/rapide, Processus de Markov Déterministe par Morceaux.

Résumé : Cette thèse porte sur la compréhension des mécanismes de différenciation cellulaire des cellules souches permettant la production des globules rouges (mécanisme appelé érythropoïèse). Nous avons élaboré différents modèles mathématiques permettant une compréhension à différents niveaux.

Dans un premier temps, nous avons construit et calibré un modèle à 8 équations différentielles ordinaires pour décrire la dynamique de 6 populations de cellules en érythropoïèse de repos et de stress. L'étude de données expérimentales *in vivo*, recueillies par nos collaborateurs Stéphane Giraudier (hématologue) et Evelyne Lauret (INSERM), a montré la nécessité d'ajouter deux équations pour modéliser les régulations érythropoïétiques. La calibration du modèle a été effectuée à l'aide des données biologiques et d'un algorithme d'optimisation stochastique appelé CMA-ES. Ce modèle nous a permis de mettre en lumière l'importance de la capacité d'auto-renouvellement des cellules érythropoïétiques dans la production des globules rouges.

L'élaboration d'un modèle probabiliste de dimension 3 nous a ensuite permis de comprendre les

conséquences dynamiques de cette capacité sur la production des globules rouges. L'étude de ce modèle a nécessité des changements d'échelles en taille et en temps révélant un système dit lent/rapide. A l'aide de méthodes dites de moyennisation nous avons décrit l'approximation en grande population du nombre de cellules de chaque type. Nous avons également quantifié mathématiquement les grandes fluctuations biologiquement observées au niveau du nombre de globules rouges.

Nous avons finalement construit un modèle pour comprendre l'influence des longues phases d'inactivité, connues, des cellules souches mutantes dans la production des globules rouges. Les cellules souches mutantes, en faible nombre dans l'organisme comparé aux cellules saines, basculent aléatoirement d'un état actif à un état inactif. Les différences d'échelle de taille entre les populations de cellules, nous a conduit à étudier la dynamique d'un processus de Markov déterministe par morceaux de dimension 4. Nous avons montré l'existence d'une unique probabilité invariante vers laquelle le processus converge en variation totale, et identifié cette limite.

Title : Cell differentiation, stem cell regulation and impact of the mutations: a stochastic approach.

Keywords : Erythropoiesis, Mathematical modeling, Markovian process, *In vivo* data exploitation, Slow/fast system, Piecewise Deterministic Markovian Process.

Abstract : This thesis focuses on understanding the mechanisms of stem cell differentiation leading to the production of red blood cells (a mechanism called erythropoiesis). To this end, we have developed different mathematical modelings leading to an understanding at different levels.

Firstly, we have built and calibrated a model with 8 ordinary differential equations to describe the dynamics of 6 populations of cells in steady-state and stress erythropoiesis. The study of *in vivo* experimental data, realized by our collaborators Stéphane Giraudier (hematologist) and Evelyne Lauret (INSERM), showed the need of two equations to model erythropoiesis regulations. Modeling calibration was performed using biological data and a stochastic optimization algorithm called CMA-ES. This model highlighted the importance of the self-renewal capacity of the erythropoietic cells in the production of red blood cells.

The development of a 3-dimensional probabilistic model then allowed us to understand the dynamic

consequences of this capacity on the production of red blood cells. The study of this model required changes of scale in size and time revealing a so-called slow/fast system. Using averaging methods, we described the large population approximation of the number of each cell type. We have also mathematically quantified the large fluctuations in the number of red blood cells, biologically observed.

Finally, we constructed a model to understand the influence of long periods of inactivity of mutant stem cells in the production of red blood cells. Mutant stem cells, which are in low numbers in the organism compared to healthy cells, randomly switch between an active and an inactive state. The different size scale between the cell populations led us to study the dynamics of a 4-dimensional piecewise deterministic Markov process. We showed the existence of a unique invariant probability measure towards which the process converges in total variation, and we identified this limits.