



Université
de Toulouse

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par Université Toulouse III – Paul Sabatier

Discipline **Mathématiques appliquées**

Présentée par **Mathieu LEROY-LERÊTRE**

Soutenue le 27 octobre 2014

Titre de la thèse

Étude de la croissance tumorale via la modélisation agent-centré du comportement collectif des cellules au sein d'une population cellulaire

École doctorale **Mathématiques, informatique et télécommunications de Toulouse (MITT)**

Unité de recherche **Institut de mathématiques de Toulouse (IMT) – UMR 5219**

Thèse dirigée par Pierre DEGOND (Imperial College London) directeur
Valérie LOBJOIS (Université Toulouse III) codirecteur (biologie cellulaire)
Giacomo DIMARCO (University of Ferrara) coencadrant

Composition du jury

<i>Rapporteurs</i>	Mostafa ADIMY	directeur de recherche à l'INRIA (Lyon)
	Florence HUBERT	maitre de conférences HDR à l'université d'Aix-Marseille
<i>Examineurs</i>	Patrick LABORDE	professeur à l'université Toulouse III – <i>Président du jury</i>
	Olivier SAUT	chargé de recherche HDR au CNRS (Bordeaux)
<i>Directeurs de thèse</i>	Pierre DEGOND	chair professor
	Valérie LOBJOIS	maitre de conférences HDR

Ce document a été rédigé avec l'excellente et moderne classe L^AT_EX yathesis, développée par Denis Bitouzé pour faciliter la rédaction des mémoires de thèses préparées en France.

Mots-clés: modélisation agent-centré, minimisation sous contraintes, confrontation entre modèle et expériences, croissance tumorale, orientation de la division cellulaire, lignages cellulaires

Keywords: agent-based modeling, constrained minimization, comparison between modeling and experiments, tumor growth, orientation of the cell division axis, cell lineages

Cette thèse a été préparée dans les laboratoires suivants.

**Institut de mathématiques de Toulouse (IMT) –
UMR 5219**

Université Toulouse III
118, route de Narbonne
31062 TOULOUSE Cedex 9

☎ 05 61 55 67 90

📠 05 61 55 75 99

✉ direction.imt@math.univ-toulouse.fr

Site <http://www.math.univ-toulouse.fr/>



**Institut des technologies avancées en sciences du
vivant (ITAV) – USR 3505**

Centre Pierre Potier
1, place Pierre Potier Oncopole entrée B
BP 50624
31106 Toulouse Cedex 1

☎ 05 82 99 10 10

✉ direction@itav.fr

Site <http://www.itav-recherche.fr/>



*Au Génie d'Aladin,
"O Captain! My Captain!"*

If you're going through hell, keep
going.

Attribuée à Winston Churchill

Oh, to hell with dignity. I'll leave
when the job's done.

M (Dame Judi Dench), *Skyfall*

ÉTUDE DE LA CROISSANCE TUMORALE VIA LA MODÉLISATION AGENT-CENTRÉ DU COMPORTEMENT COLLECTIF DES CELLULES AU SEIN D'UNE POPULATION CELLULAIRE**Résumé**

L'objectif de cette thèse de mathématiques appliquées à la biologie est de comprendre des mécanismes fondamentaux impliqués dans la croissance de tumeurs cancéreuses non vascularisées.

Un nouveau modèle discret 2D de type agent-centré et son implémentation numérique sont établis pour étudier les liens entre, d'une part, la croissance et la division des cellules et, d'autre part, l'organisation de la population cellulaire ; ce modèle explore en particulier l'influence de l'orientation de la division cellulaire sur la population et ses lignages. Les cellules sont des disques se regroupant de façon isotrope, sans chevauchement : à chaque instant, leur position est obtenue par la résolution d'un problème de minimisation sous contraintes. La méthode utilisée pour l'implémentation numérique de cette minimisation repose sur l'algorithme d'Uzawa.

Des indicateurs statistiques ainsi qu'un procédé de visualisation des populations cellulaires sont mis en place, pour étudier à la fois les résultats mathématiques *in silico* et les résultats biologiques *in vitro* : ils mesurent des caractéristiques sur le comportement de la population et des différentes lignées qui la composent. Il en ressort une forte disparité entre les lignées situées au centre de la population et celles situées en périphérie. L'interaction mathématiques-biologie, inscrite au cœur de notre démarche, permet de réaliser une confrontation du modèle aux expériences : elle en apporte une première validation mais met en valeur des différences. Suivant une hypothèse formulée pour se rapprocher des expériences, un modèle enrichi est proposé : ses résultats sont en meilleur accord avec les expériences.

Mots-clés : modélisation agent-centré, minimisation sous contraintes, confrontation entre modèle et expériences, croissance tumorale, orientation de la division cellulaire, lignages cellulaires

AGENT-BASED MODELING OF THE COLLECTIVE BEHAVIOR OF CELLS IN A CELL POPULATION, TO STUDY TUMOR GROWTH**Abstract**

This thesis in Mathematics applied to Biology aims at understanding key behaviors involved in avascular tumor growth.

A new 2D discrete model, agent-based, and its numerical implementation are introduced to study the relation between, on the one hand, cell growth and division and, on the other hand, the organization of the cell population. This model is specially designed to investigate the impact of the orientation during division on the population and on its lineages. Cells are represented as disks grouping together in an isotropic way, without overlapping themselves: at each time, their positions are given by the solution of a constrained minimization problem. This numerical resolution uses an algorithm based on the Uzawa method.

Some statistical indicators and a visualisation process of cell populations are set up, to study both the mathematical results (*in silico*) and the biological ones (*in vitro*): they measure distinctive information on the behavior of the population and its lineages. A large disparity between central and peripheral lineages is highlighted. The interaction between Mathematics and Biology, placed at the heart of our approach, enables us to confront the modeling with the experiments: this brings a first step toward validation but also shows some differences. We then state a hypothesis, the exploration of which leads to an improved model: its results are a better match to the experiments.

Keywords: agent-based modeling, constrained minimization, comparison between modeling and experiments, tumor growth, orientation of the cell division axis, cell lineages

Institut de mathématiques de Toulouse (IMT) – UMR 5219

Université Toulouse III – 118, route de Narbonne – 31062 TOULOUSE Cedex 9

Remerciements

Cette expérience professionnelle dans le monde de la recherche m'a permis de réaliser, certes de manière fort détournée, un rêve de petit garçon : celui d'aller *where no one has gone before*¹. Est venu le temps des remerciements !

Je tiens en premier lieu à remercier le CNRS, la région Midi-Pyrénées et l'INSA Toulouse, auprès desquels j'ai obtenu les financements permettant de réaliser mes travaux de recherche. Sans mon directeur et mon coencadrant de thèse, cela n'aurait pas été possible : je ne sais comment les remercier ; au premier, je dois le choix d'un sujet de recherche qui m'a passionné dès le premier jour ; au second, je suis reconnaissant de l'accompagnement pendant les premières années.

À Valérie, ma codirectrice, j'apporte mes remerciements les plus chaleureux, pour s'être engagée dans ce voyage du premier au dernier jour, sans faire chavirer le navire et sans le quitter aux premiers signes de turbulences. Pour avoir apporté un cadre de travail humain, respectueux et productif ; pour des conseils avisés sur le présent et le futur ; pour un investissement jusqu'au jour de la soutenance : je suis heureux d'avoir pu travailler, apprendre et échanger avec toi.

Je remercie Raphaël, parrain de thèse toujours présent dans les moments où il le fallait. J'adresse à Jean-Michel Roquejoffre et Patrick Cattiaux une reconnaissance de leur professionnalisme en tant, respectivement, que directeurs d'école doctorale et d'unité de recherche, pour m'avoir donné l'opportunité de remettre en état de marche le navire endommagé : vous avez assurément joué un rôle nécessaire au bon aboutissement de ce voyage.

Je remercie mes rapporteurs, Florence Hubert et Mostafa Adimy, d'avoir accepté une grosse lecture estivale pour évaluer mon manuscrit de thèse. L'accueil par Florence au CIRM a toujours été agréable : j'en garde un très bon souvenir. Je remercie Olivier Saut et Patrick Laborde d'avoir accepté d'être membres de mon jury. J'ai rencontré Patrick lors de mes premières semaines à l'IMT : nous avons depuis eu l'occasion de discuter fort agréablement aux niveaux recherche, enseignement et vie de l'université : c'est un plaisir qu'il soit membre de mon jury.

Special thanks to special friends ! À ces amis avec qui j'ai gravi des montagnes en Écosse, joué à l'auto-temponeuse (grandeur nature) en Irlande, passé des nuits comprimées à

1. Tout Trekkie aura reconnu cette phrase... mais c'est quoi un Trekkie ? (question pour ceux qui ont décroché pendant la soutenance).

Cork et sur l'île d'Achill², vidé des Guinness en plein visite guidée à Dublin et trainé au Couleurs Café. Mention spéciale au propriétaire de la Saxo rouge, qui a un jour eu besoin d'un chauffeur pour rentrer rue Lavoisier, ce qui fut le début de grandes amitiés. À ces amis avec qui j'ai des souvenirs inoubliables de NYC, tous les plus originaux les uns que les autres : je nierai tout ce que vous pourrez raconter un jour.

Very special thanks to a very special friend! À toi qui m'a soutenu et relevé dans les moments les plus difficiles, qui m'a permis de retrouver mon chemin lorsque je m'égarais et qui m'a accueilli sous tes toits successifs (j'en compte quatre !) quand le besoin s'en faisait sentir. La petite twingo bleu, des jours (sauf le dernier ;-)) alimentés d'une mémorable mousse au chocolat blanc, des situations ambiguës en soirées ker lannaises et des fous rires intempestifs sont des souvenirs inoubliables.

L'histoire de cette expérience dans la recherche commence bien évidemment en terres bretonnes, à Ker Lann, au sein d'une petite structure récemment devenue grande : l'ENS Rennes. Que de personnes à remercier ! Tout d'abord, les membres de l'équipe du département math auprès desquels, même plusieurs années après avoir quitté KL, on trouve toujours quelqu'un pour donner des nouvelles et quérir des conseils. De tous ces membres, il y en a un qui mérite mes plus sincères remerciements !

Cher Greg, tour à tour président de l'association des magistériens de Rennes, membre du BdE inventif (... [bip]...), chargé de TP et de cours en analyse numérique, et préparateur hors pair à l'agrégation, je marche du mieux que je peux dans tes pas³. J'ai plaisir et fierté à te rendre un vibrant hommage, car c'est grâce à toi que j'ai donné mes premiers enseignements, que j'ai « canalisé » mon volontarisme à enseigner \LaTeX et que j'ai eu envie d'explorer plus en détail la « modélisation mathématique ». Pour beaucoup tu es, et bien plus que tu ne t'en doutes, dans mon affectation actuelle : mille mercis ! Maintenant que le titre de « dieu Ker Lannais » est libre (le dernier tenant ayant atteint le stade de « dieu émérite »), je suis sûr que l'unanimité se ferait pour te l'attribuer⁴ !

Je ne me risquerais pas à remercier les Ker Lannais qui comptent pour moi, tant il y en a ! Alors merci à ceux qui m'ont permis de faire des « Inter-ENS culturelles 2007 », les premières Inter Q ker lannaises, une réussite gravée à jamais en moi⁵. Merci à tous ceux qui ont fait du BdE une aventure inoubliable, en s'en occupant avant, pendant ou après. Merci aux math.2008, qui ont cobayé mes compétences d'enseignant, dignement fêtées rue de la Soif le soir même.

Le cœur de cette histoire se déroule dans la ville rôasse, au sein de l'institut de mathématiques de Toulouse.

Des remerciements personnels à Marie-Laure, pour des sorties nocturnes mémorables ; à Tamara, pour des discussions interminables sous les douches ; à Delphine, pour un soutien permanent ; à « Céline de l'IMT » pour un texto et des conseils toujours avisés, y compris sur l'avenir ; à mes cobureaux, avec qui nous avons tant partagé, parfois

2. Néanmoins le jeu d'auto-temponeuse a rendu ces nuits moins comprimées qu'elles n'auraient pu l'être...

3. Je n'ai cependant jamais été préparateur à l'agrégation !

4. En fait, après un rapide sondage, l'unanimité pointait déjà son nez aux 10 ans du département.

5. Il reste encore des tee-shirts...

tard : Laetitia qui m'accueillit, Fabien qui prit soin de son cactus et Anne-Charline qui donna tout son sens à l'issue de secours (je fais snif, snif à l'idée de la fin du bureau 302). Aux anciens docteurs et parfois actuels (et j'espère futur ;-) collègues, que j'ai plaisir à recroiser, ce fût bien sympathique de travailler à vos cotés : Alice, Amic, Benjamin, Jacques, Laetitia, Laurent, Léo, Marion, Max Fév, Max Dele, Roubly, Sandrine ; et bien sûr Erwan et Julien, que je suis ou qui me suivent, selon les destinations. À Julien pour le club piscine, très assidu. À Miclo, parce qu'un rang A qui s'assoie tout seul au milieu d'une table de doctorants pour déjeuner, c'est rare et tellement bien. Aux permanents qui se sont mouillés pour me soutenir, même par quelques mots, et qui se reconnaîtront.

Des remerciements professionnels, sur l'enseignement : à JBHU, que j'ai connu de loin via un très bon livre utilisé en prépa agreg, que j'ai mieux connu lors d'un excellent cours de M2 et que j'ai encore mieux connu en tant que collègue de TD pour un enseignement des plus agréables ; à Fanny, pour ces bons moments passés en salle info à (essayer de) faire comprendre l'algorithmique, une fois que les ordi étaient opérationnels ; à Philippe, qui m'a donné l'opportunité d'enseigner de la modélisation en M1 (même si la charge fût lourde !) et m'a permis d'implanter la formation \LaTeX à Toulouse ; à Violaine, toujours prête le dimanche soir pour filtrer des signaux ; à Pascal, toujours prêt à fournir un corrigé des exos de TD (alors même que l'énoncé vient d'être envoyé !) ; à Emmanuel et Patrick, avec qui j'ai commencé et j'ai fini ces années d'enseignement.

Des remerciements professionnels, sur la recherche : à Fanny, qui m'a accompagné tout au long de la quatrième année de thèse, qui a permis au manuscrit d'être ce qu'il est, et qui s'est déplacée pour la répétition de soutenance, avec mon plus jeune auditeur. Merci à Sébastien et Thomas de m'avoir accueilli au CSCAMM, qui fût un environnement de travail serein pour rebondir. Merci aux membres de l'ITAV, toujours prêts à accueillir un math(i)eu dans leur bureau (et j'en ai testé, des bureaux !). L'ITAV fut mon deuxième lieu de travail au cours de cette thèse, où la bonne humeur était à chaque fois présente, en chacun, y compris le vendredi soir... ; où je pouvais venir me ressourcer en travaillant. Remerciements particuliers à Marie-Laure et Martine, qui ont effectué un travail énorme pour réaliser et post-traiter les expériences biologiques ayant alimenté ce travail de thèse : le chapitre 4 de ce manuscrit vous doit beaucoup ! et moi aussi ! Le travail de fourni avec les gommettes a été pleinement exploité et les films diffusés ! Remerciements à Bernard, aux idées riches et aux points de vue toujours pertinents, et qui a retenu l'histoire du saladier avec les clémentines (plutôt que la minimisation d'un potentiel sous contraintes de non-recouvrement). Odile m'a permis de garder contact avec la (grande) famille Dallaporta : quelle heureuse surprise ce fût de le découvrir !

Des remerciements professionnels, sur... mais que reste-t-il en milieu universitaire ? sur la vie collective à l'université, ceux qui font tourner la machine, sans qui pas grand chose ne fonctionnerait. Merci au staff administratif de l'IMT que je n'ai pas encore cité, et en particulier à Marie-Line et Monique, à l'accent toulousain qui égaye les matings. Un grand merci aux informatologues, et en particulier à Jean, qui a dépensé beaucoup de temps pour faire fonctionner la visioconférence le jour de la soutenance, et à Lionel, qui a voulu « réparer » mon ordi au plein milieu de la rédaction du manuscrit. J'ai pu voir à l'œuvre, à certaines réunions (du pôle) ou face au traitement RH dont des

doctorants furent la cible, l'action d'un directeur de l'IMT, Patrick Cattiaux, qui s'est battu comme un ours⁶ pour défendre son personnel : que de travail accompli, au service du *collectif* ; avec par ailleurs de l'énergie et la volonté pour résoudre des situations individuelles : un grand merci, Patrick, pour la « mission exfiltration » à DC. Merci à Agnès, au dynamisme débordant, et à Martine, aux nombreuses qualités, pour avoir suivi le dossier à l'école doctorale. Je suis enfin très heureux d'avoir fréquenté l'association UdA et j'ai appris beaucoup : des gens volontaires, aux idées parfois tranchées mais toujours prêts à débattre dans le but améliorer l'université Toulouse III ; merci d'avoir accueilli un doctorant parmi vous !

Il reste... mais qui reste-t-il ? Il reste les doctorants actuels de l'IMT : je ne vous oublie pas ! Impossible ! Pour des déjeuners reposants ou plein de débats, pour des rando, pour de fréquents soutiens, je vous remercie : Claire, Claire et Claire (dans cet ordre⁷), les trois petits math.2008 (Laurent, Nil, Pierre) devenus grands depuis leur WEI, Benoit, Magali, Malika, ainsi que Guillem, pour une rando plein de neige et Mélanie, organisatrice de talent.

À Denis Bitouzé, qui ne rentre dans aucune des cases précédentes mais qui mérite sa case à lui tout seul : pour avoir coécrit l'Œuvre [BC10], pour m'avoir embarqué dans le beta-testage (que dis-je ? le gamma !) de la classe de thèse yathesis⁸ et pour m'avoir fait faire du L^AT_EX au bout du monde (non, pardon, c'était à Dunkerque).

Aux membres de ma famille, et en particulier ceux qui habitent près du CIRM, de NYC et de DC, je vous remercie de votre présence en dehors du boulot. Tout comme aux Leroy de Pins-Justaret : Richard et Fanny (et Valentin (et Raphaël (et Amandine (et... ?⁹)))) ; il y en a deux qui n'ont toujours pas eu leur milkshake !

Maintenant, il est l'heure pour moi de dormir¹⁰ : ces dernières semaines, entre les cours à préparer, les cours à préparer et les cours à préparer (et une soutenance de thèse, accessoirement), ces remerciements se finissent en ce 26 octobre.¹¹

And here today, I remember this, I believe, from Tennyson:
 “We are not now that strength which in old days
 Moved earth and heaven; that which we are, we are;
 One equal temper of heroic hearts,
 Made weak by time and fate, but strong in will
 To strive, to seek, to find, and *not to yield*.”

M (Dame Judi Dench), *Skyfall*

6. Certes, parfois piqué par des abeilles (c'était plutôt des frelons).

7. Mais ne t'inquiète pas Claire : c'est bien toi qui est en premier !

8. À vous qui lisez ces lignes et qui prévoyez de rédiger un jour un manuscrit de thèse ou d'HDR : allez voir du côté de cette classe, elle vous simplifiera la vie.

9. Cela devient difficile typographiquement parlant.

10. En fait, non. Il est l'heure d'aller voir Harry Potter à la télé, c'est le dernier épisode : celui où Voldemort perd ; est-ce prémonitoire ?

11. Pour info, il n'y a pas à ma connaissance de limite au nombre de footnotes que contient un document : bien heureusement pour moi !

Table des matières

Résumé	xiii
Remerciements	xv
Table des matières	xix
Liste des figures	xxiii
Liste des tableaux	xxvii
Liste des algorithmes	xxix
Introduction générale	1
Généralités sur la croissance tumorale	2
Les « caractéristiques fondamentales » d'une cellule cancéreuse	2
Les principales étapes du développement tumoral	4
Cycle cellulaire et mitose	5
Positionnement biologique et problématiques	7
Quelques exemples de modèles mathématiques de croissance tumorale	8
Modèles discrets sur grille (<i>on lattice models</i>)	9
Modèles agent-centré (<i>agent-based models</i>)	10
Modèles continus	10
Modèles hybrides	11
Un nouveau modèle agent-centré de croissance tumorale	12
Choix d'une modélisation agent-centré	12
Forme des agents : disques incompressibles	13
Règle de comportement : interaction entre les agents par minimisation d'un potentiel global sous contraintes de non-interpénétration	14
Règles de comportement : croissance jusqu'à doublement, puis division par déformation « en 8 »	15
Choix de l'orientation de division	16
Échanges continus et confrontations mutuelles mathématiques-biologie	17
Organisation du manuscrit	18
Corps	18
Pages liminaires, annexe et pages finales	19

1	Principes de modélisation	21
1.1	Mise en place du modèle	22
1.1.1	Choix d'un modèle discret, agent-centré	22
1.1.2	Choix des agents et des règles de comportement	23
1.2	Loi de positionnement	26
1.2.1	Principe général de la loi de positionnement	26
1.2.2	Valeurs des paramètres et des fonctions de modélisation	27
1.3	Loi de croissance	28
1.3.1	Croissance linéaire en volume	28
1.3.2	Croissance linéaire en volume avec bruit	29
1.4	Loi de division	30
1.4.1	Principe général de la loi de division	30
1.4.2	Processus de division d'une cellule : la transformation	32
1.4.3	Processus de division d'une cellule : l'interaction	35
1.4.4	Choix d'orientation	37
2	Implémentation numérique du modèle	39
2.1	Structure générale de l'algorithme	40
2.2	Étape de positionnement	40
2.2.1	Algorithme de positionnement	42
2.2.2	Discussion sur les valeurs attribuées aux pas de l'algorithme	47
2.2.3	Améliorations autour de l'algorithme de positionnement	50
2.2.4	Bilan de l'étape de positionnement	51
2.3	Étape de croissance	52
2.4	Étape de division d'une cellule	52
2.4.1	Processus de division à l'étape τ_k pour k fixé	55
2.4.2	Initialisation du processus de division et synthèse de l'étape	58
2.5	Configurations de l'algorithme	59
2.5.1	Étape d'initialisation	59
2.5.2	Valeurs numériques des paramètres	60
3	Mise en place des indicateurs statistiques et étude des simulations	63
3.1	Représentation par diagramme de Voronoï adapté	64
3.1.1	Motivation	64
3.1.2	Rappels sur les diagrammes de Voronoï	65
3.1.3	Modification et utilisation des diagrammes de Voronoï	66
3.2	Définition des diagnostics	70
3.2.1	Motivation des diagnostics	71
3.2.2	Représentation d'une population cellulaire en terme de graphe	72
3.2.3	Définition des diagnostics portant sur la population cellulaire	75
3.2.4	Définition des diagnostics portant sur les lignages	76
3.3	Implémentation des diagnostics	78
3.3.1	Calculs préliminaires	78
3.3.2	Calculs des diagnostics	79

3.4	Protocole des simulations et résultats des diagnostics	81
3.4.1	Protocole d'étude	81
3.4.2	Résultats de chacun des diagnostics	82
3.4.3	Conclusions communes des différents résultats	95
4	Expériences biologiques et confrontation avec le modèle	97
4.1	Utilisation des données expérimentales	98
4.1.1	Post-traitement des données expérimentales	98
4.1.2	Préparation des données expérimentales post-traitées	99
4.2	Application des diagnostics aux données expérimentales post-traitées .	104
4.2.1	Protocole d'analyse	104
4.2.2	Utilisation des diagnostics sur les données expérimentales	105
4.2.3	Observations et conclusions des diagnostics	109
4.3	Confrontation avec les résultats des simulations	112
4.3.1	Comparaison entre les résultats des simulations et les résultats des expériences	112
4.3.2	Bilan de la modélisation grâce à la confrontation	115
5	Retour sur le modèle	117
5.1	Améliorations des principes de modélisation	118
5.1.1	Ajout d'une permutation entre cellules	118
5.1.2	Choix par zone du potentiel global W	119
5.2	Adaptation de l'algorithme	120
5.2.1	Étape de positionnement	122
5.2.2	Étape de division	125
5.2.3	Étape de permutation	128
5.2.4	Valeurs numériques des nouveaux paramètres de modélisation .	128
5.3	Étude du modèle enrichi avec le filtre des diagnostics	129
5.3.1	Impacts de chacune des deux améliorations	129
5.3.2	Impacts des deux améliorations combinées et conclusions	132
	Conclusion et perspectives	135
	Conclusion	135
	Perspectives	136
	Bibliographie	137
	A Protocole expérimental	145
	Glossaire	147

Liste des figures

1	Les trois grandes étapes d'évolution d'une tumeur	4
2	Cycle cellulaire et stade G0 de quiescence	6
1.1	Transformation au cours de la division	32
1.2	Forme d'une cellule au cours de la division	33
1.3	Dôme sphérique	34
2.1	Vision schématique de la totalité de l'algorithme	41
2.2	Illustration de la succession de la croissance et du positionnement	53
2.3	Vision schématique de l'étape de division	54
2.4	Illustration de la succession des deux actions du processus de division	56
3.1	Représentation par des disques de configurations issues de l'algorithme	64
3.2	Diagrammes de Voronoï pour 4, 5, 6, 9, 15 et 20 points de \mathbb{R}^2	67
3.3	Représentation par diagramme de Voronoï, avec différents choix de sites fictifs	69
3.4	Représentation par diagramme de Voronoï adapté de configurations issues de l'algorithme	70
3.5	Évolution de lignages au cours du temps	73
3.6	Représentation du polygone de bord pour une population cellulaire	80
3.7	Représentation de la direction principale de trois lignages	80
3.8	Diagnostic n° 1 sur le rapport aire versus périmètre de la population (R_1), pour les simulations	83
3.9	Diagnostic n° 2 sur le rapport aire versus aire de l'enveloppe convexe de la population (R_2), pour les simulations	84
3.10	Diagnostic n° 3 sur le rapport aire versus périmètre par composante connexe de lignage (R_3), pour les simulations	86
3.11	Diagnostic n° 4 sur le nombre de composantes connexes par lignage (R_4), pour les simulations	86
3.12	Diagnostic n° 3 sur le rapport aire versus périmètre par composante connexe de lignage (R_3), pour chacune des six situations des simulations	87
3.13	Diagnostic n° 4 sur le nombre de composantes connexes par lignage (R_4), pour chacune des six situations des simulations	88

3.14	Diagnostic n° 5 sur le nombre de cellules par composante connexe de lignage (R_5), pour les simulations	91
3.15	Diagnostic n° 6 sur l'écart de l'orientation des lignages à la direction tangentielle ($ R_6 $), pour les simulations	91
3.16	Diagnostic n° 5 sur le nombre de cellules par composante connexe de lignage (R_5), pour chacune des six situations des simulations	92
3.17	Diagnostic n° 6 sur l'écart de l'orientation des lignages à la direction tangentielle ($ R_6 $), pour chacune des six situations des simulations	93
3.18	Diagnostic n° 6 sur l'excentricité des lignages (\tilde{R}_6), pour chacune des six situations des simulations	94
4.1	Expérience : suivi de certains lignages	99
4.2	Expérience : délimitation des régions détectées par la segmentation	100
4.3	Expérience : représentation après la segmentation	100
4.4	Expérience : représentation par disques des noyaux	102
4.5	Expérience : représentation par points des (centres des) noyaux	103
4.6	Expérience : représentation par diagramme de Voronoï adapté de la population cellulaire	103
4.7	Expérience : représentation du polygone de bord de la population	104
4.8	Diagnostic n° 1 sur le rapport aire versus périmètre de la population (R_1), pour les expériences	106
4.9	Diagnostic n° 2 sur le rapport aire versus aire de l'enveloppe convexe de la population (R_2), pour les expériences	107
4.10	Diagnostic n° 3 sur le rapport aire versus périmètre par composante connexe de lignage (R_3), pour les expériences	107
4.11	Diagnostic n° 4 sur le nombre de composantes connexes par lignage (R_4), pour les expériences	108
4.12	Diagnostic n° 5 sur le nombre de cellules par composante connexe de lignage (R_5), pour les expériences	109
4.13	Diagnostic n° 6 sur l'écart de l'orientation des lignages à la direction tangentielle ($ R_6 $), pour les expériences	110
4.14	Diagnostic n° 6 sur l'excentricité des lignages (\tilde{R}_6), pour les expériences	110
4.15	Diagnostic n° 3 : comparaison entre les simulations et les expériences	113
4.16	Diagnostic n° 4 : comparaison entre les simulations et les expériences	113
4.17	Diagnostic n° 5 : comparaison entre les simulations et les expériences	114
4.18	Diagnostic n° 6 : comparaison entre les simulations et les expériences	114
5.1	Vision schématique de la totalité de l'algorithme pour le modèle enrichi	121
5.2	Vision schématique de l'étape de division pour le modèle enrichi	127
5.3	Diagnostic n° 4 sur le morcellement d'un lignage, pour le modèle enrichi par l'ajout de permutation : en zone centrale avec différentes valeurs de la probabilité p	130

5.4	Diagnostic n° 5 sur la taille des fragments d'un lignage, pour le modèle enrichi par l'ajout de permutation : en zone centrale avec différentes valeurs de la probabilité p	130
5.5	Diagnostic n° 6 sur l'orientation d'un lignage, pour le modèle enrichi par l'ajout de permutation : zone centrale pour différentes valeurs de la probabilité p	130
5.6	Diagnostic n° 4 sur le morcellement d'un lignage, pour le modèle enrichi par changement de potentiel : évolution en zone périphérique pour différentes valeurs de la probabilité N_s	131
5.7	Diagnostic n° 5 sur la taille des fragments d'un lignage, pour le modèle enrichi par changement de potentiel : évolution en zone périphérique pour différentes valeurs de la probabilité N_s	131
5.8	Diagnostic n° 6 sur l'orientation d'un lignage, pour le modèle enrichi par changement de potentiel : évolution en zone périphérique pour différentes valeurs de la probabilité N_s	131
5.9	Diagnostic n° 4 sur le morcellement d'un lignage : comparaison entre le modèle enrichi, le modèle de départ et les données expérimentales, en zone centrale	132
5.10	Diagnostic n° 4 sur le morcellement d'un lignage : comparaison entre le modèle enrichi, le modèle de départ et les données expérimentales, en zone périphérique	133
5.11	Diagnostic n° 5 sur la taille des fragments d'un lignage : comparaison entre le modèle enrichi, le modèle de départ et les données expérimentales, en zone centrale	133
5.12	Diagnostic n° 5 sur la taille des fragments d'un lignage : comparaison entre le modèle enrichi, le modèle de départ et les données expérimentales, en zone périphérique	133

Liste des tableaux

2.1	Valeurs numériques des paramètres de modélisation, données par défaut dans les simulations	60
2.2	Valeurs numériques des paramètres de modélisation, à choisir pour chaque simulation	61
2.3	Valeurs numériques des paramètres de l’algorithme, données par défaut dans les simulations	61
5.1	Valeurs numériques des paramètres de modélisation (du modèle enrichi), données par défaut dans les simulations	128

Liste des algorithmes

- 2.1 Positionnement 46
- 2.2 Retour 51
- 2.3 Croissance 52
- 2.4 Positionnement en cours de division 57
- 2.5 Initialisation du processus de division 58

- 5.1 Positionnement modifié 123
- 5.2 Positionnement modifié en cours de division 125
- 5.3 Prise de décision sur la permutation 126
- 5.4 Permutation 128

Introduction générale

L'objectif de ce chapitre est de déterminer les problématiques motivant notre démarche de modélisation et de situer notre approche par rapport aux autres modèles existants.

Nous commençons par apporter quelques connaissances sur la croissance tumorale, à l'attention du lecteur non biologiste, qui aboutissent à la formulation des problématiques guidant notre travail. Nous nous recentrons ensuite sur les modèles mathématiques : dans un premier temps en fournissant un panorama de différents types de modèles utilisés en croissance tumorale ; dans un second temps en nous recentrant davantage sur les modèles agent-centré. Enfin, nous terminons en donnant la structure du manuscrit.

Sommaire du chapitre

Généralités sur la croissance tumorale	2
Les « caractéristiques fondamentales » d'une cellule cancéreuse	2
Les principales étapes du développement tumoral	4
Cycle cellulaire et mitose	5
Positionnement biologique et problématiques	7
Quelques exemples de modèles mathématiques de croissance tumorale	8
Modèles discrets sur grille (<i>on lattice models</i>)	9
Modèles agent-centré (<i>agent-based models</i>)	10
Modèles continus	10
Modèles hybrides	11
Un nouveau modèle agent-centré de croissance tumorale	12
Choix d'une modélisation agent-centré	12
Forme des agents : disques incompressibles	13
Règle de comportement : interaction entre les agents par minimisation d'un potentiel global sous contraintes de non-interpénétration	14
Règles de comportement : croissance jusqu'à doublement, puis division par déformation « en 8 »	15
Choix de l'orientation de division	16
Échanges continus et confrontations mutuelles mathématiques-biologie	17
Organisation du manuscrit	18
Corps	18
Pages liminaires, annexe et pages finales	19

Généralités sur la croissance tumorale

Cette section vise à expliquer quelques notions de biologie liées à la croissance tumorale, qui nous permettront de formuler les problématiques auxquelles nous nous intéressons. Il ne s'agit pas de donner des informations exhaustives sur « le cancer » mais de fournir des informations ciblées, utiles à la compréhension et à la motivation des sections et chapitres suivants.

La suite est présentée ainsi : nous décrivons de manière générale ce qu'est une cellule cancéreuse, puis les grandes étapes d'évolution d'une tumeur et enfin le cycle cellulaire. Nous terminons en posant les questionnements biologiques qui motivent notre travail.

Note. Les termes composés en caractères *inclinés* correspondent à des mots dont la définition figure dans le glossaire en page 147.

Les « caractéristiques fondamentales » d'une cellule cancéreuse

En biologie, la cellule joue un rôle particulier et central. Il s'agit de l'élément de base de tout être vivant : on dit qu'elle est l'unité structurale, fonctionnelle et reproductrice du vivant. En particulier, elle contient dans son noyau les chromosomes constitués de l'*acide désoxyribonucléique (ADN)* associé à des protéines. Ce rôle fondamental nous amène à placer la cellule au cœur de notre discussion sur la croissance tumorale.

Une cellule, au cours de son existence, dispose de divers mécanismes qui lui permettent d'assurer son bon fonctionnement. Une cellule cancéreuse est, en quelque sorte, une cellule au comportement anormal échappant à ces mécanismes de contrôle, suite à des dérégulations qui sont souvent associées à des modifications de l'information génétique (appelées mutations).

Qu'est-ce qu'une cellule cancéreuse ? Après des années de recherche au cours du siècle dernier, l'article « The hallmarks of cancer » [HW00] publié en 2000 tend à définir le cancer, ou plus précisément à en extraire des « caractéristiques fondamentales ». Ces caractéristiques sont définies comme des capacités fonctionnelles qui permettent à une cellule cancéreuse de survivre, de proliférer et de se disséminer ; elles sont chacune la conséquence d'une brèche dans les mécanismes de défense anticancéreuse de la cellule et de l'organisme. Ces propriétés correspondent à des altérations de la physiologie cellulaire, acquises au cours des différentes étapes du développement tumoral, par le biais de mécanismes variés. Les six *caractéristiques fondamentales*, distinctes et complémentaires, sont supposées être présentes dans la plupart des cancers humains, et peut-être même tous. Elles sont rapidement décrites ci-dessous.

Autonomie de croissance. La prolifération des cellules est normalement la conséquence de signaux stimulant la croissance : une cellule cancéreuse a la faculté de croître indépendamment de ces signaux.

Insensibilité à la non-prolifération. L'organisme dispose de mécanismes pour arrêter la croissance cellulaire : une cellule cancéreuse a la faculté d'ignorer ces signaux antiprolifératifs.

Résistance à l'apoptose. L'organisme peut demander à une cellule déficiente de déclencher l'*apoptose*, c'est-à-dire la mort cellulaire, ce qui permet d'éliminer cette cellule anormale : une cellule cancéreuse a la faculté d'ignorer les signaux d'autodestruction.

Réplication à l'infini. Un mécanisme intracellulaire limite le nombre de fois où une cellule peut se reproduire : une cellule cancéreuse n'est plus soumise à cette limite¹.

Angiogenèse. L'*angiogenèse* est un processus consistant à créer de nouveaux vaisseaux sanguins : cela permet d'irriguer en nutriments et en oxygène des populations cellulaires, et ainsi de contribuer à leur prolifération. Les cellules cancéreuses ont la capacité de réaliser l'angiogenèse.

Invasion. Les cellules cancéreuses ont la capacité de se propager dans les tissus environnants et même dans tout l'organisme, où elles forment de nouvelles colonies, appelées *métastases*.

Ces six caractéristiques fondamentales ont été complétées, en 2011, par l'article « Hallmarks of cancer : the next generation » [HW11] : deux nouvelles caractéristiques, dites « caractéristiques fondamentales émergentes » et qui restent à être fermement établies, sont proposées en sus des six précédentes. Il s'agit des deux suivantes.

Résistance au système immunitaire. L'organisme surveille constamment les cellules, par l'intermédiaire du système immunitaire ; il est apte à détecter et à détruire celles qui sont en dysfonctionnement. Une cellule cancéreuse a la faculté de ne pas être détectée ou de résister à sa destruction par le système immunitaire : cela empêche l'organisme d'éradiquer les cellules cancéreuses.

Dérégulation du métabolisme. Le métabolisme énergétique regroupe l'ensemble des réactions liées à la production d'énergie au sein d'une cellule ; cette énergie, qui provient des nutriments et de l'oxygène, est en particulier utilisée lors de la croissance et de la division. Une cellule cancéreuse, pour s'adapter aux ressources disponibles éventuellement limitées, a la capacité de modifier son métabolisme énergétique : cela lui permet de reprogrammer ses besoins en nutriments et oxygène, dans le but d'intensifier sa reproduction.

Comment le cancer démarre-t-il ? Si de nombreux aspects de la réponse à cette question restent encore inconnus, on sait que différentes mutations génétiques sont nécessaires et que de nombreux facteurs environnementaux contribuent à son apparition. [HW11] met en valeur les deux « facteurs déclenchants » suivants, qui rendraient possible l'acquisition des caractéristiques fondamentales précédentes.

Instabilité du génome et mutation. Les lésions courantes de l'*ADN* sont normalement détectées et corrigées par l'organisme, afin de limiter l'impact des mutations spontanées. Les cellules cancéreuses sont le résultat de nombreuses altérations

1. La maintenance de ce mécanisme de réplication est liée à une enzyme, la télomérase, dont la découverte fut récompensée par un prix Nobel en 2009. Cette enzyme est exprimée dans les cellules cancéreuses alors qu'elle ne le devrait pas.

de leur *génome*, irrémédiablement endommagé, dont les gènes mutants leur confèrent des avantages sur les cellules saines.

Inflammation protumorale. L'inflammation est une réaction de défense du système immunitaire, censée combattre ou réparer les dommages subis par l'organisme. Les cellules immunitaires produisent paradoxalement l'effet inverse : par le déploiement de leurs molécules chimiques, qui se trouvent être très mutagènes sur des cellules présentant certaines des *caractéristiques fondamentales*, elles favorisent l'émergence de toutes ces caractéristiques.

Bilan La cellule est la brique de base du vivant, ce qui en fait un objet d'étude principal du biologiste : cela justifie en particulier que la cellule sera l'élément central de notre modèle mathématique. Les comportements des cellules cancéreuses, tant individuels que collectifs, sont encore peu connus et leur compréhension reste un enjeu majeur de la recherche sur le cancer. Nous savons cependant qu'elles s'organisent entre elles et forment des tumeurs : cela est abordé dans la section suivante sur les trois grandes étapes du développement d'une tumeur.

Les principales étapes du développement tumoral

Une tumeur est une structure tridimensionnelle composée d'une population hétérogène de cellules : on peut trouver des cellules en prolifération, des cellules ayant stoppé leur croissance et des cellules mortes. Parmi les tumeurs, on distingue les tumeurs bénignes et les tumeurs malignes, appelées tumeurs cancéreuses. La terminologie « un cancer » peut aussi bien désigner « une tumeur cancéreuse » qu'une maladie caractérisée par l'apparition de tumeurs cancéreuses.

Lors du développement d'une tumeur cancéreuse, trois grandes étapes bien distinctes peuvent être exhibées. Il s'agit, dans l'ordre chronologique de développement, des trois suivantes, qui sont schématiquement illustrées en figure 1.

1. La phase d'initiation, où la tumeur commence à croître en utilisant les nutriments et l'oxygène qu'elle capte dans le microenvironnement.
2. La phase vasculaire où la tumeur, trop grosse pour n'être alimentée que par le système sanguin pré-existant, développe ses propres vaisseaux sanguins qui vont l'alimenter.
3. La phase métastatique, où les cellules cancéreuses se disséminent à travers l'organisme afin de fonder des tumeurs, appelées *métastases*, dans d'autres organes.

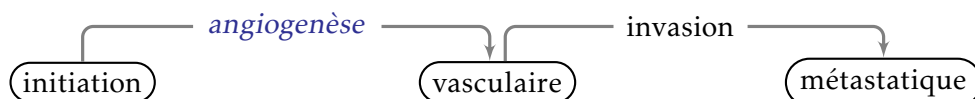


FIGURE 1 – Les trois grandes étapes d'évolution d'une tumeur

Dans le cadre de ce travail de thèse, nous nous intéressons à la première de ces trois grandes étapes : la croissance de tumeurs non vascularisées. Le choix d'étudier cette phase d'initiation est motivé par les arguments biologiques suivants :

- comprendre la situation non vascularisée devrait permettre de comprendre l'évolution des zones peu vascularisées d'une tumeur en phase vasculaire ;
- ce stade de développement s'apparente aux très nombreuses microtumeurs qui se développent chez l'homme et qui sont le plus souvent spontanément éliminées [BH11] ;
- les expériences biologiques sont faciles à réaliser et à reproduire, ce qui fournit une richesse qualitative et quantitative de données expérimentales ; cela n'est en effet pas le cas de la troisième étape, pour laquelle réaliser une expérience nécessite d'avoir un modèle vivo métastasé. Les résultats mathématiques peuvent donc être confrontés aux résultats biologiques, démarche essentielle lors d'une modélisation mathématique.

De plus, la situation biologique est moins complexe dans cette étape, ce qui rend accessible une modélisation mathématique centrée sur les phénomènes importants. Les phases vasculaire et métastatique n'en sont néanmoins pas moins importantes : ce sont elles qui provoquent la mort du patient donc leur compréhension, d'un point de vue clinique, est cruciale.

Par ailleurs, nous allons nous intéresser aux premiers jours de vie d'une population cellulaire tumorale, qui est initialement composée de plusieurs dizaines de cellules pour en atteindre des centaines au bout de quelques jours. Dans cette situation, le modèle biologique expérimental sur lequel nous travaillons est uniquement constitué de cellules cancéreuses, situées dans un milieu contenant tous les nutriments nécessaires à son développement. En particulier, il n'existe pas de système immunitaire et nous n'étudions pas la transformation de cellules saines en cellules cancéreuses. Toutes ces raisons nous amènent à nous concentrer sur les quatre premières *caractéristiques fondamentales* du cancer précédemment énoncées, dans lesquelles la prolifération tient un rôle majeur.

Bilan Cette description schématique du cancer nous a permis de remplir deux objectifs. D'une part donner une vision d'ensemble, très simplifiée, du phénomène biologique de croissance tumorale. D'autre part, centrer notre étude sur la phase d'initiation et plus précisément sur des populations cellulaires assez petites, composées uniquement de cellules en prolifération, où les quatre premières *caractéristiques fondamentales* du cancer jouent un rôle primordial. Le cycle cellulaire, et en particulier la mitose, étant un phénomène au cœur de la prolifération, nous l'abordons ci-après.

Cycle cellulaire et mitose

La prolifération cellulaire correspond au développement et à la multiplication de cellules : une cellule, dite « cellule mère », grossit jusqu'à doubler son volume pour ensuite donner naissance à deux cellules identiques, dites « cellules filles ». L'ensemble de ce processus est appelé le *cycle cellulaire* ; il est essentiellement constitué de deux

temps : l'*interphase* et la *mitose*, décrits ci-après.

L'interphase. Elle correspond à la croissance de la cellule et à la préparation de sa division. Elle est usuellement découpée en trois stades :

- le stade G₁, où la cellule croît et prépare le stade S ;
- le stade S, où la réplication de l'*ADN* a lieu, via la duplication des chromosomes ;
- le stade G₂, où la cellule prépare le stade M.

La mitose. Elle correspond au stade M du cycle cellulaire. Il s'agit de la division de la cellule mère en deux cellules filles, au cours de laquelle les chromosomes dupliqués au stade S se répartissent équitablement entre les deux cellules filles.

Le *cycle cellulaire* est illustré en figure 2. Il mentionne en particulier le stade G₀, qui correspond à des cellules *quiescentes* ayant temporairement quitté le cycle : il s'agit par exemple de cellules au cours du processus de *différenciation* ou bien de cellules en manque de nutriments ou d'oxygène pour effectuer leur croissance.

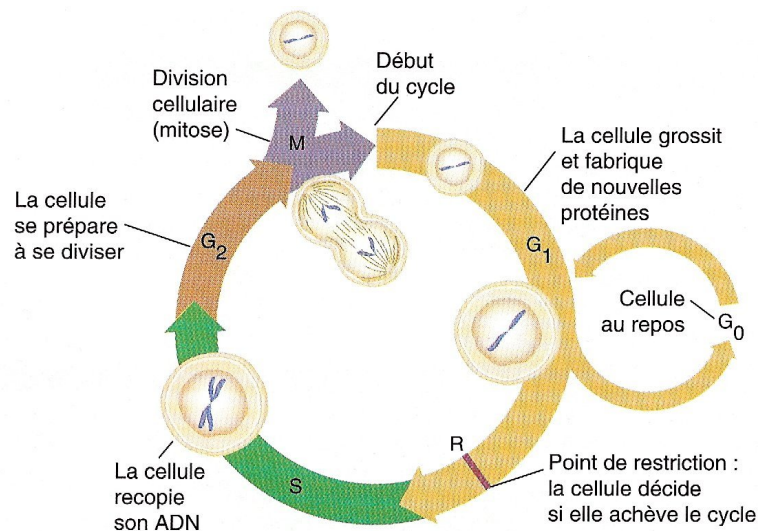


FIGURE 2 – Cycle cellulaire et stade G₀ de quiescence

Au cours de la division cellulaire, une succession de phénomènes particuliers auxquels nous nous intéressons se déroule :

- les paires de chromosomes s'alignent au centre de la cellule mère, le long d'un plan équatorial appelé le *plan de division (ou plan de clivage)* ;
- ensuite les paires de chromosomes se séparent équitablement et s'éloignent dans une direction perpendiculaire à ce plan de division, vers des pôles opposés de la cellule ;
- enfin, autour de chacun de ces deux pôles, émergent les deux cellules filles résultant de la division cellulaire.

Nous ne décrivons pas davantage les différentes étapes de la mitose, car des informations supplémentaires ne sont pas nécessaires à l'établissement de notre modèle.

Par ailleurs, nous n'aborderons pas les phénomènes biologiques de régulation du cycle cellulaire, dont la dérégulation est à l'origine de l'apparition et de la multiplication des cellules cancéreuses. De plus amples informations sur le cycle cellulaire, et ses liens avec le cancer, figurent dans [Mei03 ; VVB03].

Nous nous intéressons cependant plus en détail au *plan de division*, et plus précisément à son orientation ; en effet, la détermination ainsi que l'impact de l'orientation du plan de division sont très peu connus des biologistes.

- La détermination de l'orientation du plan de division a longtemps été considérée comme une conséquence de la forme de la cellule, et en particulier de sa direction d'élongation maximale ; des études plus récentes ont cependant mis en évidence des phénomènes plus complexes [TB06]. La géométrie de la cellule [MBC11], l'influence des cellules voisines (étudiée pour des cellules de la peau dans [Gib+11]) ou encore des forces mécaniques extérieures [Fin+11] peuvent être des facteurs influant la détermination de l'orientation du plan de division.
- L'impact de l'orientation du plan de division est lui aussi l'objet de travaux récents : ils mettent en valeur le rôle d'une division asymétrique dans la *différenciation* des cellules souches [LF05 ; WSS07]. Par ailleurs, de récentes découvertes tendent à établir un lien entre le cancer et une désorientation du plan de division [PT11] : des cellules en dysfonctionnement peuvent conduire à une désorientation et, à l'inverse, une désorientation peut favoriser le développement du cancer.

Cette discussion sur l'orientation du plan de division nous amène à poser la définition suivante, dont nous allons nous servir tout au long du manuscrit.

Définition (Orientation de division). Nous appelons *orientation de division* l'orientation dans laquelle s'éloignent, au cours de la mitose, les deux lots de chromosomes répliqués en interphase. Cette orientation est perpendiculaire au *plan de division* (formé par les paires de chromosomes avant leur séparation) ; elle va déterminer l'orientation dans laquelle se fera la déformation de la cellule mère en deux cellules filles, ainsi que le placement relatif de ces cellules filles.

Bilan Du *cycle cellulaire*, nous retenons l'alternance entre deux périodes : la croissance puis la division. Au cours de la division, les deux cellules filles identiques s'éloignent dans des sens opposés suivant une même direction, donnée par le *plan de division*. Sans rentrer dans les détails des mécanismes intracellulaires, notamment ceux concernant l'*ADN*, nous prenons en compte l'orientation de division.

Positionnement biologique et problématiques

Au cours des précédentes sections, nous sommes partis de la thématique générale et très vaste du « cancer », puis nous nous sommes recentrés sur la phase non vascularisée. Nous avons par ailleurs restreint notre champ d'étude aux petites populations cellulaires, composées uniquement de cellules en prolifération. Un des paramètres clés de la prolifération étant la mitose, nous nous sommes intéressés à la division cellulaire et plus particulièrement à un de ses nombreux aspects : le *plan de division*.

Cela nous amène à formuler deux problématiques, qui portent sur l'étude de l'organisation d'une population cellulaire tumorale. Pour cela, nous introduisons la notion suivante : étant donné une cellule, et après plusieurs divisions, il est possible de considérer l'ensemble de ses descendants, qui constituent son lignage.

Problématique 1. Quel est l'impact de la multiplication des cellules sur l'organisation de l'ensemble de la population cellulaire ?

Problématique 2. L'*orientation de division* a-t-elle une influence sur l'évolution de la population cellulaire et sur l'organisation de ses lignages ?

Quelques exemples de modèles mathématiques de croissance tumorale

Les premiers modèles mathématiques sur la croissance tumorale dont nous avons connaissance remontent aux années 1930 [Hil28 ; May32] ; c'est cependant essentiellement vers la fin du xx^e siècle qu'ils ont été nombreux à être développés [AM04 ; Byr10]. Parmi les différents modèles, on distingue usuellement les modèles discrets et les modèles continus. Dans un modèle discret, chaque élément est modélisé comme une entité à part, équipée de caractéristiques et suivie individuellement ; un certain nombre de règles de modélisation définissent ses actions. Le comportement de chaque entité est alors précisément défini, ce qui rend accessible l'ajout de nouvelles règles. Par ailleurs, le devenir de toutes les entités peut être connu, ce qui facilite les comparaisons entre le modèle mathématique et des données expérimentales. Ces avantages atteignent cependant leurs limites sur de grandes populations cellulaires : en effet, il devient alors nécessaire de suivre de nombreuses entités, ce qui se révèle couteux en terme de résolution numérique. L'utilisation de modèles continus permet de pallier cet inconvénient et de modéliser de grandes populations, sans toutefois conserver les avantages des modèles discrets. En effet, dans un modèle continu, les éléments sont décrits en terme de densité de population et leurs actions sont modélisées par des équations aux dérivées partielles : c'est un avantage pour étudier les propriétés mathématiques du système modélisé mais cela rend difficile l'établissement de liens directs entre les paramètres du modèle et les mesures physiques [Mac+10].

Notre objectif dans cette section est de proposer une vue d'ensemble non exhaustive de différents modèles mathématiques utilisés en croissance tumorale, à travers des exemples. Ce panorama est cependant volontairement axé sur les modèles discrets : c'est en effet un tel modèle qui est développé dans ce travail de thèse. Le choix a été retenu de les classer selon le *type* de modèle, plutôt que selon la situation biologique modélisée : cela permet d'illustrer les différentes possibilités mathématiques de modélisation et de situer les modèles agent-centré par rapport aux autres types de modèles.

La suite est présentée ainsi : les modèles discrets sont exposés en deux temps (modèles sur grille et modèles agent-centré), puis viennent les modèles continus et enfin les modèles hybrides.

Modèles discrets sur grille (*on lattice models*)

Il s'agit de modèles dans lesquels l'espace est découpé en plusieurs petites cases : ces différentes cases portent le nom de *sites*² et forment une discrétisation appelée *grille* (ou réseau). Cette grille détermine l'emplacement des entités, puisqu'elles sont nécessairement placées sur des sites. Elle fournit par ailleurs des relations intrinsèques de voisinage, qui permettent de définir les interactions entre les entités. Chaque site est caractérisé par son *état* qui, à chaque pas de temps, évolue en fonction d'un jeu de règles préalablement définies. Les avantages de ces modèles sont une implémentation numérique aisée et une facilité de modification des règles de modélisation. Concernant leurs inconvénients, nous pouvons dire que les déplacements (et en particulier leur direction), les positions ainsi que la forme des entités sont contraints par le réseau : ces modèles ne peuvent donc pas correctement capter des comportements d'entités qui ne se dérouleraient pas sur la grille. Parmi les modèles sur grille, on distingue :

les modèles d'automates cellulaires, constitués d'un réseau où chaque site modélise une entité ;

les modèles de Potts cellulaires, constitués d'un réseau où une entité est modélisée par plusieurs sites.

Dans les modèles d'automates cellulaires (*cellular automata*), chaque site du réseau correspond à une entité de la population modélisée. Chacun de ces sites est dans un état régi par les règles de l'automate et change d'état au cours du temps, selon ces mêmes règles. Ces modèles ont été fortement utilisés dans la modélisation mathématique appliquée à la biologie [DD05] ; on peut citer par exemple des modèles pour étudier la croissance tumorale non vasculaire dans le cerveau [Kan+00] ou pour étudier la migration de cellules tumorales [Aub+06]. Dans un grand nombre de modèles d'automates cellulaires, comme les deux précédents, les sites du réseau sont carrés ou cubiques ; il existe cependant des modèles où les sites possèdent d'autres formes géométriques : [Aub+08] étudie la migration de cellules associées aux neurones, au travers d'un réseau composé de sites hexagonaux. Certaines extensions existent, par exemple des modèles hybrides [GA07 ; Qua+08]. On pourra consulter [RCM07] pour d'autres modèles d'automates cellulaires appliqués à la croissance tumorale non vascularisée. Cependant, une des limitations majeures des automates cellulaires est que les changements de taille ou de forme des cellules ne peuvent pas être pris en compte.

Les modèles de Potts cellulaires (*cellular Potts models*), aussi appelés modèles de Glazier-Graner-Hogewed, sont une extension des automates cellulaires dans lesquels *une* entité occupe *plusieurs* sites du réseau : ils permettent ainsi de rendre compte de la déformation et du changement de taille d'une entité. Les premiers travaux datent des années 1990 [GG92 ; GG93] ; un exemple décrivant les bases des modèles de Potts cellulaires est brièvement présenté dans [SP12] et l'on peut se référer à [ACR07, Chap. II] pour une étude exhaustive. Ces modèles peuvent être utilisés pour étudier la migration de cellules [SP13], l'arrangement des cellules par chimiotaxie [KHM06] ou l'angiogenèse

2. Une autre terminologie, qui donne son nom aux automates cellulaires, est *cellule*. Cependant, pour éviter toute confusion avec les cellules vivantes du monde biologique, nous ne l'utilisons pas.

[BJJ09 ; Shi+09]. Enfin, certaines extensions de ces modèles existent : dans [SP13] figure un modèle dont l'originalité est que chaque cellule est représentée par son noyau et le milieu dans lequel il baigne ; un modèle hybride reposant sur un modèle de Potts cellulaire est proposé dans [SP12].

Une description plus détaillée de ces différents modèles, dont sont inspirés les paragraphes précédents, est disponible dans [KRA12 ; RM10].

Modèles agent-centré (*agent-based models*)

Il s'agit de modèles discrets pour lesquels, à la différence des modèles sur grille précédents, il n'existe pas de réseau préalablement fixé qui contraigne les entités : leur placement, leur forme ainsi que leur mouvement sont alors totalement libres. Les entités sont représentées dans l'espace par des objets individuels, appelés « agents », équipés de certaines caractéristiques (forme, position, taille, état interne, etc.) à choisir. Un avantage de cette liberté de mouvement est de fournir une modélisation plus réaliste ; l'inconvénient est que l'implémentation numérique est plus ardue que dans les modèles sur grille.

Un des premiers choix à adopter pour utiliser ces modèles est la forme des entités. Les plus répandus adoptent une représentation sphérique : ils étudient la croissance d'une tumeur [DF00 ; DH05 ; GLD05], l'invasion de cellules saines par des cellules cancéreuses [GHA09 ; DHB07] ou encore des cancers spécifiques, comme des cancers du sein [Mac+10 ; Nor+10]. D'autres représentent les entités par des ellipses, ce qui donne une notion intrinsèque d'orientation : par exemple, pour étudier exclusivement le déplacement de cellules [PO00 ; Pal08]. Il est aussi possible d'utiliser des diagrammes de Voronoï³ pour représenter sous forme polygonale les entités : ont ainsi été étudiées les croissances de sphéroïdes [SM05] ou de cryptes intestinales [MPL01 ; Van+09]. Il existe enfin des modèles de frontières immergées (*immersed boundary models*), dans lesquels chaque entité peut avoir une forme particulière et est modélisée comme un corps élastique immergé dans un fluide visqueux incompressible ; par exemple, pour étudier la croissance de sphéroïdes [Rej07] ou l'invasion de cellules [Qua+08]. Ce modèle permet de rendre compte de la morphologie des cellules puisqu'elles sont pleinement déformables.

Une fois la forme des entités choisie, il reste à définir leurs règles d'évolution. Nous détaillons cela en section « [Un nouveau modèle agent-centré de croissance tumorale](#) » page 12.

Modèles continus

Dans les modèles continus, les entités ne sont pas directement modélisées : ce sont des populations d'entités qui sont décrites, par des fonctions de densité. De tels modèles permettent d'étudier une population, non plus en regardant les comportements individuels de chacune de ses entités mais en la regardant de manière globale. Le modèle

3. Leur signification est rappelée en section 3.1.2 page 65.

proposé dans [SC01], basé sur des équations de réaction-diffusion, vise à reproduire la croissance d'un sphéroïde ; il considère trois types de cellules : proliférante, *quiescente* et nécrotique. Une amélioration de ce modèle peut consister à mieux décrire le *cycle cellulaire*, notamment la division cellulaire : le modèle étudié dans [Bre+10] différencie, parmi les cellules en phase de prolifération, celles qui sont en *interphase* de celles qui sont en *mitose* ; ce modèle est utilisé pour étudier l'invasion de cellules tumorales sur des cellules saines. Les modèles continus peuvent aussi servir à étudier des interactions entre la tumeur et les tissus environnants [AM06] ; différents modèles continus pour décrire la croissance tumorale non vascularisée, ce qui correspond au cadre biologique auquel nous nous intéressons, sont abordés dans [RCM07].

Par ailleurs, certains modèles continus sont créés en lien avec des modèles discrets : [Osb+10] propose deux modèles discrets et un modèle continu pour étudier l'invasion de cellules saines par des cellules cancéreuses, dans le cadre d'une crypte colorectale. Dans d'autres situations, un modèle discret a d'abord été développé puis, dans un second temps, un modèle continu et une comparaison au modèle discret ont été réalisés [SM05 ; SM06], [BD09 ; DH05].

Modèles hybrides

La complexité de la réalité biologique nécessite parfois de prendre en compte des phénomènes qui se déroulent à différentes échelles spatiales ; par exemple, un comportement observé au niveau cellulaire peut provenir de mécanismes moléculaires se déroulant à l'intérieur des cellules. Il est aussi possible, là encore pour rendre compte de la complexité biologique, de mettre en place des modèles mélangeant le discret et le continu. Tous ces modèles sont souvent qualifiés de « modèles hybrides », même si cette terminologie regroupe des situations très différentes. Les distinctions suivantes sont issues de [PT09] :

- les modèles emboîtés (*nested models*)**, dans lesquels sont inclus des phénomènes se déroulant à plusieurs échelles spatiales. Par exemple aux niveaux cellulaire et intracellulaire ;
- les modèles hybrides (*hybrid models*)**, qui utilisent une approche discrète pour modéliser certains constituants et une approche continue pour d'autres. Par exemple, les cellules sont modélisées par des variables discrètes et le milieu extracellulaire par des variables continues ;
- les modèles d'interface hybrides (*interfacing hybrid models*)**, où des constituants sont décrits de manière discrète en certains endroits du domaine spatial et par des variables continues ailleurs. Cela permet d'adapter localement la modélisation, aux temps et aux endroits appropriés.

Parmi les modèles emboîtés, des phénomènes intracellulaires sont pris en compte dans [Ram+08 ; Ram+09] pour étudier l'influence de la β -caténine et de la E-cadhérine sur l'adhésion entre les cellules. Dans [AMD05 ; ZAD07], des comportements intracellulaires sont couplés à un modèle sur grille et visent à explorer l'impact de l'observation

suivante sur la croissance tumorale : des cellules cancéreuses du cerveau humain sont soit en prolifération, soit en migration (mais pas les deux en même temps).

Parmi les modèles hybrides, nombreux sont ceux qui utilisent une approche discrète pour les cellules et une approche continue pour le microenvironnement. Par exemple, dans [And05], la dynamique de chimiotaxie et le milieu extracellulaire sont contrôlés par un modèle continu, et les cellules par un automate cellulaire. On retrouve le même principe dans [Pat+01], pour étudier l'invasion et la densité des vaisseaux. Dans [ACR07 ; And+09 ; Qua+08], un modèle d'invasion tumorale, dit « *hybrid discrete-continuous* », est proposé : il couple une modélisation du microenvironnement avec une modélisation du déplacement et de l'interaction des cellules. La première est faite par un système de réaction-diffusion ; la seconde par un automate cellulaire. Ce modèle permet d'explorer les interactions entre les cellules et le microenvironnement.

Parmi les modèles d'interface hybrides, [KSO07] utilise une modélisation discrète pour les cellules proliférantes (qui correspondent aux endroits où les phénomènes de croissance et de division sont fréquents), alors que les cellules *quiescentes* et le milieu extracellulaire sont régis par un modèle continu.

Un nouveau modèle agent-centré de croissance tumorale

L'objectif de cette section est de fournir un état des lieux, autour des modèles agent-centré utilisés en croissance tumorale, axé en totalité sur les différents points de notre modèle. Cet état des lieux, en proposant des modélisations similaires ou différentes de la notre, permet de présenter et de comparer les grandes lignes de notre modèle.

La suite est présentée ainsi : nous commençons par justifier la volonté d'utiliser un modèle agent-centré ; puis nous discutons la construction de ce modèle, via les choix de la forme des agents et des différentes règles de comportement de ces agents. Enfin, nous abordons deux points précis : le choix de l'orientation au cours de la division, afin d'explorer notre problématique 2 ; puis les forts liens entre mathématiques et biologie, car ils sont inscrits au cœur de notre démarche.

Choix d'une modélisation agent-centré

La situation biologique à laquelle nous nous intéressons est une petite population cellulaire, composée de 20 à 500 cellules, évoluant sur un plan 2D. Dans ce cadre, nous cherchons à étudier l'impact de la croissance des cellules, ainsi que l'influence de l'*orientation de division*, sur l'organisation de la population. Par ailleurs, nous inscrivons au cœur de notre démarche un lien fort avec les biologistes, que ce soit pour l'établissement du modèle, pour le choix des problématiques à étudier ou pour les comparaisons entre les résultats du modèle mathématique et des expériences biologiques.

Dans cet objectif, nous devons placer la cellule au centre de notre modélisation mathématique et être en mesure de suivre individuellement chacune des entités du modèle : nous nous orientons donc vers un modèle discret. En effet, un modèle continu ne donne pas accès à un suivi individualisé des entités modélisées.

Ensuite, nous devons disposer d'une complète liberté dans la direction de déplacement et dans la direction de division des cellules : nous écartons donc les modèles sur grille. En effet, ils imposent des contraintes sur la position et le déplacement des cellules qui peuvent introduire des biais artificiels sur l'organisation des cellules ; de plus, ils ne peuvent pas explorer en profondeur les questions d'orientation, puisque le nombre d'orientations possibles pour une cellule est limité par la géométrie du réseau (elle est par exemple de quatre dans un réseau carré) [Mac+10]. Dans [RCM07], cette influence artificielle du réseau sur le comportement global est comparée aux conséquences, parfois observées, du choix d'un schéma numérique lors de résolutions d'équations aux dérivées partielles (autant dans ce dernier cas les conséquences sont quantifiables après une analyse mathématique, autant dans le cas des automates cellulaires une telle quantification est extrêmement difficile).

Au final, nous retenons donc le choix d'un modèle discret, de type agent-centré. Un tel choix simplifie l'exploration d'hypothèses, puisque les règles de modélisation sont facilement modifiables : cela permet d'étudier quels mécanismes microscopiques (au niveau des cellules) peuvent avoir un impact sur les comportements macroscopiques (au niveau de la population cellulaire) [DH05]. Par ailleurs, ce choix facilite l'accès aux résultats sur le suivi individualisé des agents : il est idéal pour des comparaisons entre le modèle mathématique et les expériences biologiques [Mac+10] ; en particulier, ce suivi nous sera d'une grande utilité pour explorer la problématique 2. Afin de réaliser cette comparaison, nous allons obtenir au niveau biologique un même repérage individualisé des cellules vivantes, grâce au suivi et à la segmentation réalisés lors des analyses d'images : en quelque sorte, nous convertirons les données biologiques sous une forme agent-centré.

Remarque (Modélisation agent-centré en dehors de la croissance tumorale). La création et l'utilisation des modèles agent-centré dans les domaines de la biologie et du médical dépassent le cadre de la croissance tumorale : plusieurs modèles sur le développement des structures tissulaires sont décrits, et classés par système biologique ou maladie, dans [TBP07]. Les exemples s'étendent à bien d'autres disciplines, par exemple : en écologie, pour étudier la réaction de populations d'oiseaux de rivage en réponse à la perte de leur habitat [Gos+06] ou pour étudier le développement résidentiel dans des systèmes urbains [BR06] ; en criminalité, pour étudier la formation de gangs de rue [Heg+11] ; en économie, pour étudier les phénomènes de régulation sur les marchés financiers [Bon02]. D'autres exemples figurent dans [GR13] et dans [MN09], ce dernier proposant une introduction à la modélisation agent-centré.

Forme des agents : disques incompressibles

Parmi les différentes formes des agents, il est très fréquent de trouver des disques ou des sphères [DF00 ; DL01]. De nombreux modèles proposent de déformer ces sphères en cas de forte compressibilité [DH05 ; DH12], par exemple en aplatissant la zone de contact [GLD05] ou même en les transformant en des polyèdres convexes [SM05 ; BM07]. D'autres modèles utilisent des ellipses, notamment pour donner une notion intrinsèque

d'orientation ou de polarisation : elles peuvent être déformables, à volume constant, en ellipses de paramètres différents pour rendre compte de compressions [PO00 ; Pal08] ; dans ces deux précédents articles, il n'y a pas de croissance ni de division mais il est possible de les ajouter [KSO07]. Certains modèles représentent les cellules sous forme polygonale : par exemple en suivant les centres des cellules avant d'en déduire les polygones par diagramme de Voronoï ou en suivant directement tous les sommets de ces polygones (ces deux approches sont abordées dans [Osby+10]).

Dans notre modèle, bien que nous souhaitions explorer une problématique d'orientation, nous utilisons une représentation par des disques (la notion d'orientation ne se trouve donc pas intrinsèquement modélisée par la forme des agents mais en sera un paramètre interne). Par ailleurs, ces disques sont incompressibles : cela est en accord avec la règle de comportement sur l'interaction entre les agents décrite ci-après (en particulier relative à la contrainte de non-interpénétration qui est fortement liée à la forme des agents).

Règle de comportement : interaction entre les agents par minimisation d'un potentiel global sous contraintes de non-interpénétration

Concernant les différentes règles d'interaction entre les agents, de nombreux modèles sont basés sur des forces d'interaction deux à deux, en considérant des forces répulsives et des forces attractives : l'état du système est alors donné par l'application du principe fondamental de la dynamique [KSO07 ; Mac+10 ; PO00 ; SM05]. Les forces attractives peuvent consister en des adhésions entre les agents [Mac+10 ; SM05]. Les forces de répulsion sont en général des forces élastiques [SM05] mais peuvent être plus génériques [Mac+10] ; elles traduisent une résistance à la déformation [KSO07 ; PO00 ; SM05] ou à la compression [Mac+10]. Par ailleurs, il est aussi possible de considérer des forces de friction entre les agents [BM07 ; KSO07 ; PO00].

D'autres modèles raisonnent en terme d'énergie d'interaction avant d'en déduire des forces de déplacement : par exemple, en modélisant de l'attraction et de la répulsion [DH05] ou de l'attraction et de la friction [GHA09]. Enfin, certains modèles ajoutent de la chimiotaxie : pour guider le déplacement [PO00] ou pour modéliser des interactions longue distance selon un gradient de concentration [BM07].

Nous choisissons un tout autre axe de modélisation : l'interaction entre les agents est représentée par la minimisation d'un potentiel *global* auquel sont soumis les agents. Par rapport aux modèles précédemment cités, le potentiel permet de modéliser de manière générique des interactions biologiques très différentes. Dans notre cas, nous modélisons un comportement de regroupement autour d'une position particulière : le potentiel peut par exemple représenter une compression, une force extérieure ou un point d'attraction des cellules. À notre connaissance, il s'agit d'une approche novatrice dans la modélisation agent-centré de populations cellulaires tumorales. Une utilité possible de cette manière de modéliser les interactions entre les agents est de pouvoir considérer aisément deux types d'agents, chacun soumis à son potentiel : cela est utilisé dans la thèse en cours [Peu], qui modélise des populations contenant des cellules (représentées par des disques)

et des fibres (représentées par des segments), pour étudier l'auto-organisation cellulaire au sein des tissus adipeux.

Notre potentiel modélise un phénomène extérieur qui agit sur les cellules. Pour rendre compte des interactions, nous considérons la minimisation de ce potentiel sous contraintes de non-interpénétration des agents ; en effet, les cellules vivantes ne peuvent pas s'interpénétrer. Cela permet d'inscrire au cœur de notre modèle le fait que le mouvement d'une cellule est déclenché par la poussée des cellules voisines, de part leur croissance ou leur division ; autrement dit, la contrainte de congestion est la source de déplacement. Cette notion de déplacement provoqué par les voisins se retrouve dans [DL01] mais l'étude y est faite pour un environnement 1D, c'est-à-dire une chaîne de cellules : notre travail se place en 2D et anticipe une future généralisation à la 3D ; par ailleurs, il ne s'agit pas du cœur de leur modèle. Des mouvements de type poussée sont aussi considérés dans [DH12] mais le modèle présenté intègre une composante de mouvement pour les cellules : dans notre modèle, le mouvement est *uniquement* engendré par la congestion. Une interprétation possible de notre minimisation sous contraintes est que les cellules veulent toutes atteindre une certaine *position souhaitée* mais qu'elles y sont empêchées par le non-recouvrement : cette vision est très proche de la notion de *vitesse souhaitée* présente dans un modèle de mouvement de foule orienté vers la gestion de configurations très denses [MV07 ; MV11], dont sont inspirés nos travaux.

Pour la résolution numérique du problème mathématique de minimisation sous contraintes, nous nous servons de la correspondance entre un problème de minimisation sous contraintes et un problème de recherche de point selle [All12]. Notre résolution numérique s'inspire de l'algorithme d'Uzawa, qui est une méthode numérique de recherche de point selle, et qui peut être vu comme la transformation d'un problème d'optimisation sous contraintes en une suite de problèmes d'optimisation sans contraintes.

Règles de comportement : croissance jusqu'à doublement, puis division par déformation « en 8 »

Parmi les différents modèles étudiant des populations cellulaires, quelques modèles, uniquement destinés à étudier le déplacement, ne considèrent aucune croissance des agents [PO00 ; Pal08]. Tous ceux qui étudient la prolifération cellulaire contiennent naturellement une modélisation du *cycle cellulaire* : ils tiennent alors compte des phases de croissance (*interphase*) et de division (*mitose*).

Certains modèles mélangent croissance et division [DHB07 ; Mac+12], c'est-à-dire qu'au fur et à mesure que l'agent croît, il se divise progressivement. La plupart, néanmoins, séparent ces deux moments et considèrent une croissance des agents jusqu'au doublement de leur taille : dans les modèles 2D, le rayon du disque croît d'une valeur R_{\min} à $\sqrt{2}R_{\min}$ (où R_{\min} était le rayon de cet agent juste après sa naissance) pour modéliser un doublement de l'aire [DF00 ; DL01 ; DH05] ; dans les modèles 3D, c'est le volume qui est doublé [GLD05 ; KSO07]. Ces précédents modèles réduisent le cycle cellulaire à une étape de croissance puis une étape de division, c'est-à-dire qu'ils ne détaillent pas

les stades de l'interphase. Un tel détail est cependant abordé dans [SM05], qui propose de distinguer cinq états : phase G1, phase S et G2, phase M, phase G0, phase nécrotique.

Concernant la mitose, certains modèles proposent une division instantanée en deux cellules filles identiques [KSO07]. Cependant, la plupart des approches la modélisent comme une déformation de la cellule mère au cours de laquelle elle a une forme « en 8 » [DF00 ; DL01 ; DH05 ; DHB07 ; DH12 ; Mac+12 ; SM05] ; un tel choix se retrouve dès le milieu des années 1990 [DKM95]. Au cours de cette division, une quantité est conservée : l'aire ou le volume, selon que le modèle est 2D ou 3D ; néanmoins, cette conservation n'est pas toujours strictement respectée : dans [DH05], le volume des agents naissants diminue d'un facteur 0,95 pour modéliser une diminution de la taille des cellules après chaque division.

Notre modèle va lui aussi séparer le cycle cellulaire en une période de croissance et une de division, sans distinguer les différents stades de l'interphase : ainsi, la croissance est continue sauf au moment de la division. Cependant, bien que nous modélisions une situation 2D, la croissance de nos agents s'effectue jusqu'au doublement de leur *volume* initial, c'est-à-dire que le rayon passe de R_{\min} à $\sqrt[3]{2}R_{\min}$. Notre choix, original par rapport aux modèles précédemment cités, se justifie ainsi : les différences de taille entre un agent naissant et un agent prêt à se diviser conviennent mieux aux biologistes ; par ailleurs, cela permet de rendre compte du caractère 3D des objets modélisés. Nous effectuons la division de manière progressive, c'est-à-dire que la cellule mère se déforme en plusieurs étapes, de manière discrète, pour donner deux cellules filles, identiques ; au cours de ce processus, elle a une forme « en 8 ». Nous choisissons aussi de conserver une quantité au cours de la division et, conformément à notre volonté de mimer des agents 3D, la déformation se fait à *volume* constant. Par ailleurs, la durée d'une mitose étant très faible devant celle de l'interphase (de l'ordre de 1 pour 24), nous décidons que, lorsqu'une division se produit, aucune cellule ne croît ; cela consiste à modéliser la mitose comme un phénomène instantané. À l'instar de [SM05], nous supposons que les conséquences de telles hypothèses simplificatrices sont faibles, étant donné la très petite durée d'une mitose dans le cycle cellulaire.

Conformément à la situation biologique étudiée, dans laquelle toutes les cellules ont un égal et permanent accès aux nutriments, l'influence de leur présence n'est pas modélisée ; le besoin et la consommation des nutriments sont abordés dans [SM05] et [DH05].

Choix de l'orientation de division

Comme annoncé en section « [Positionnement biologique et problématiques](#) » page 7, un des objectifs de notre modèle est d'étudier l'influence de l'*orientation de division*. La direction dans laquelle s'oriente l'agent en forme « de 8 » est souvent prise aléatoire [DHB07 ; Mac+12 ; Osb+10 ; Van+09], car il ne s'agit pas d'un phénomène étudié. Cependant d'autres choix existent, selon le modèle ; cette orientation peut ainsi être dans la direction du total des forces [GLD05 ; KSO07] (pour les modèles où l'interaction des agents est gérée par un bilan des forces) ou être déterminée par la forme des agents [And+09] (lorsque celle-ci est prise en compte dans le modèle). Par ailleurs, tous

les agents ne suivent pas nécessairement le même choix d'orientation : dans [RA08], l'orientation est aléatoire pour certains et déterminée par leur forme pour d'autres.

Notre modèle utilise le même choix d'*orientation de division* pour tous les agents ; afin d'explorer l'influence de cette orientation, nous proposons quatre possibilités :

- dans une direction aléatoire, choisie à chaque division ;
- dans une direction fixée à l'avance ;
- dans une direction égale à l'angle de la droite passant par l'origine et l'agent se divisant ;
- dans une direction orthogonale à cette même droite.

Les deux dernières directions sont donc fonction de la position de l'agent au sein de la population. Par ailleurs, en plus du choix de l'orientation, deux cas (exclusifs) sont modélisés au cours de la division, que nous appelons *stratégies* ; notre modèle utilise le même choix de stratégie pour toutes les cellules.

1. Premier cas (dit « stratégie d'orientation libre ») : l'orientation est choisie au tout début de la division, puis est libre d'évoluer au cours de la déformation (lorsque l'agent a une forme « en 8 »), en fonction des interactions avec les agents voisins.
2. Second cas (dit « stratégie d'orientation contrainte ») : l'orientation est choisie au tout début de la division, et l'agent se divisant conserve cette orientation tout au long du processus de déformation « en 8 », jusqu'à ce que la division soit achevée.

Le premier cas est proche de ce qui se fait dans [SM05] et dans [DH12], où la direction initiale de la mitose est aléatoire mais peut changer au cours de la division en fonction de la configuration des voisins. Par ailleurs, distinguer ces deux cas est en accord avec des situations biologiques observées : selon le type de cellules, l'orientation peut être déterminée pendant l'interphase ou au début de la mitose, mais peut aussi être modifiée par rotation au cours de la mitose [MBC11].

Échanges continuels et confrontations mutuelles mathématiques-biologie

La démarche d'interaction entre les mathématiques et la biologie est souvent décrite de la manière linéaire suivante [Mac+10].

1. Des expériences biologiques conduisent à la formulation d'hypothèses, qui sont ensuite intégrées dans les équations d'un modèle mathématique.
2. Des études (théoriques ou numériques) du modèle permettent de comprendre la dynamique du système et de proposer de nouvelles hypothèses pour expliquer les comportements biologiques observés.
3. De nouvelles expériences sont réalisées pour tester ces hypothèses, valider le modèle et en suggérer des améliorations.

En particulier, les modèles agent-centré se prêtent bien à cette démarche [TBP07] : dans le cas de la guérison de blessures de la peau [Wal+04a ; Wal+04b], leurs conclusions ont orienté de nouvelles expériences, qui à leur tour sont venues enrichir le modèle.

Dans notre travail de modélisation, nous réalisons ces différentes étapes mais pas de manière linéaire : l'interaction forte entre les deux disciplines fait que ces échanges se font tout au long du processus de travail. Ainsi :

- nos simulations mathématiques sont réalisées en même temps, et avec (quasi-) le même protocole expérimental, que les expériences biologiques ;
- les résultats des deux approches sont étudiés avec les *mêmes* outils, appelés « indicateurs statistiques » ou « diagnostics » dans la suite ;
- les conditions initiales et les paramètres du modèle sont calibrés sur les données expérimentales.

Une proximité aussi directe est assez rare ; elle se retrouve dans [PVS04], pour un modèle de reconstruction microvasculaire : des images de situations expérimentales sont prises puis numérisées pour être utilisées comme conditions initiales du modèle agent-centré ; les expériences et les simulations sont ensuite conduites simultanément, et les résultats étudiés par les mêmes indicateurs.

Au terme d'une démarche d'interaction forte entre mathématiques et biologie, il arrive que des conclusions du modèle mathématique soient à posteriori observées expérimentalement : c'est le cas pour notre modèle ainsi que, par exemple, dans [Fri+06] pour un modèle d'invasion tumorale.

Organisation du manuscrit

Corps

Outre les chapitres « Introduction générale » et « Conclusion et perspectives », le corps du manuscrit est structuré en cinq chapitres et se termine par une bibliographie commune.

Ces cinq chapitres suivent la chronologie d'une démarche de modélisation mathématique : créer un modèle, en apporter une résolution (numérique), étudier ces résultats, avoir des données expérimentales et les confronter aux résultats du modèle, modifier le modèle pour le rapprocher de la réalité. Leur contenu est résumé ci-dessous.

Chapitre 1 – Principes de modélisation Nous développons un nouveau modèle agent-centré 2D pour étudier des cellules cancéreuses organisées en population cellulaire, qui sont représentées par des disques incompressibles. Ce modèle est original en deux principaux points, qui correspondent à nos problématiques 1 et 2 page 8. D'une part dans l'interaction entre les agents, modélisée par un problème de minimisation d'un potentiel extérieur sous contraintes de non-interpénétration des agents ; cette minimisation détermine le positionnement des agents à tout temps : leur déplacement est donc uniquement engendré par la contrainte de congestion. D'autre part, dans sa prise en compte de l'orientation de division afin d'en faire une étude approfondie. Par ailleurs, la croissance des agents est linéaire en volume et la division se fait par déformation progressive « en 8 », à volume constant.

Chapitre 2 – Implémentation numérique du modèle Nous effectuons une implémentation numérique du modèle établi au chapitre 1. Pour résoudre numériquement, à chaque pas de temps, le problème de minimisation sous contraintes, nous mettons en

place un algorithme inspiré d'une méthode d'Uzawa. Nous programmons ensuite les règles de croissance et de division introduites dans la modélisation, notamment les deux stratégies de division pour explorer l'influence de l'orientation de division. Quant aux valeurs des paramètres de modélisation, elles sont choisies en fonction des données biologiques.

Chapitre 3 – Mise en place des indicateurs statistiques et étude des simulations

Nous effectuons une étude statistique des résultats numériques, qui proviennent de simulations dont le protocole est validé par les biologistes. Pour cela, nous développons un jeu de mesures, appelées « indicateurs statistiques » ou « diagnostics », qui reposent en partie sur une notion de voisinage fournie par un diagramme de Voronoï adapté à notre situation. Ces indicateurs mesurent différents critères morphologiques sur la population cellulaire entière ou sur ses lignages. Ils montrent que les comportements des lignages sont très différents selon que la cellule originelle du lignage est au centre ou en périphérie de la population : les lignages périphériques sont plus nombreux, davantage morcelés et orientés de manière tangentielle à la population.

Chapitre 4 – Expériences biologiques et confrontation avec le modèle Parallèlement aux travaux décrits au chapitre 3, les biologistes réalisent des expériences en suivant un protocole très proche de celui des simulations numériques. Nous participons étroitement au post-traitement de ces données : cela nous permet d'être en mesure de les représenter via le même diagramme de Voronoï adapté que pour l'étude statistique de nos simulations et de leur appliquer les mêmes diagnostics. Les résultats du modèle mathématique guident l'étude des données biologiques et certains sont retrouvés expérimentalement. Cette confrontation entre le modèle mathématique et les expériences biologiques nous amène à formuler l'hypothèse qu'il y a davantage de déplacements au centre de la population dans la réalité que dans notre modèle.

Chapitre 5 – Retour sur le modèle Pour tester l'hypothèse précédemment formulée, nous apportons des modifications aux principes de modélisation. Cela nécessite d'adapter l'implémentation numérique du modèle et de refaire une étude statistique des résultats numériques : elle est bien sûr faite avec les mêmes indicateurs qu'aux chapitres 3 et 4. Cette étude du modèle enrichi à travers le filtre des diagnostics montre que nos modifications sont des améliorations, puisque le modèle mathématique est dorénavant plus proche du modèle expérimental. Nous avons ainsi apporté une première validation, dans le cadre de notre modèle, à l'hypothèse formulée après la confrontation entre mathématiques et biologie.

Pages liminaires, annexe et pages finales

En dehors des entrées habituelles, les pages liminaires contiennent une liste des algorithmes énoncés aux chapitres 2 et 5.

En annexe A figure le protocole des expériences réalisées par les biologistes sur les cultures cellulaires. Elle vient en complément des informations données au chapitre 4.

Les pages finales contiennent quant à elles un glossaire de termes et acronymes biologiques : il contient les expressions composées en *caractères inclinés* dans le manuscrit (ces expressions étant des hyperliens vers le glossaire dans la version numérique). Le glossaire vient en particulier en appui à l'introduction générale et au chapitre 1, pour un lecteur non familier avec ce vocabulaire.

O’neill [jetant un œil sur un livre
extraterrestre] : Mais c’est quoi cette langue
bizarre ?
Carter : Ce sont des maths mon colonel.

Stargate SG1

Chapitre 1

Principes de modélisation

L’objectif de ce chapitre est de construire notre modèle agent-centré, dans le cadre du positionnement biologique et des problématiques données en introduction générale. La résolution numérique de ces équations sera faite au chapitre 2, pour pouvoir étudier les résultats du modèle au chapitre 3.

Après une présentation des grandes lignes du modèle (section 1.1), nous en détaillons les différentes règles de modélisation : le positionnement (section 1.2), la croissance (section 1.3) et la division (section 1.4). Cette démarche suit la chronologie présentée en section « Un nouveau modèle agent-centré de croissance tumorale » page 12.

Sommaire du chapitre

1.1	Mise en place du modèle	22
1.1.1	Choix d’un modèle discret, agent-centré	22
1.1.2	Choix des agents et des règles de comportement	23
	Représentation des agents	23
	Règle de comportement : état d’équilibre (à temps fixé)	24
	Règle de comportement : évolution (en temps)	24
	Bilan de la mise en place du modèle	25
1.2	Loi de positionnement	26
1.2.1	Principe général de la loi de positionnement	26
1.2.2	Valeurs des paramètres et des fonctions de modélisation	27
1.3	Loi de croissance	28
1.3.1	Croissance linéaire en volume	28
1.3.2	Croissance linéaire en volume avec bruit	29
1.4	Loi de division	30
1.4.1	Principe général de la loi de division	30
1.4.2	Processus de division d’une cellule : la transformation	32
1.4.3	Processus de division d’une cellule : l’interaction	35
	Règle n° 1 de positionnement, dite « stratégie d’orientation libre »	36
	Règle n° 2 de positionnement, dite « stratégie d’orientation contrainte »	37
1.4.4	Choix d’orientation	37

1.1 Mise en place du modèle

La construction d'un modèle mathématique, et plus particulièrement l'établissement de ses caractéristiques, nécessite de faire des choix et des hypothèses de modélisation ; en effet, modéliser est par définition traduire en mathématique une vision *simplifiée* de la réalité. Une telle démarche est pilotée par deux lignes directrices :

1. Explorer les problématiques fixées.
2. Trouver un bon équilibre sur la richesse du modèle : s'il est trop complexe alors il est difficilement exploitable ; s'il est trop simpliste alors il ne permet pas d'aboutir à des conclusions pertinentes.

Un modèle est donc établi en vue de répondre à des problématiques, et non pour apporter une description exhaustive de la réalité. Dans la présentation du modèle qui suit, nous nous attachons à souligner ce lien entre hypothèses de modélisation et problématiques visées. Nous rappelons les problématiques 1 et 2, introduites en page 8 dans l'introduction générale.

Problématique 1. Quel est l'impact de la multiplication des cellules sur l'organisation de l'ensemble de la population cellulaire ?

Problématique 2. L'*orientation de division* a-t-elle une influence sur l'évolution de la population cellulaire et sur l'organisation de ses lignages ?

1.1.1 Choix d'un modèle discret, agent-centré

Le phénomène biologique que nous étudions est la croissance d'une population cellulaire. Une première analyse permet de dire que :

- une population cellulaire est un système composé d'un grand nombre de cellules qui s'organisent entre-elles ;
- la croissance de cette population émane de la reproduction des cellules qui la composent.

Ainsi, le comportement des cellules est au cœur de l'objet biologique étudié. À ce titre, une population cellulaire rentre dans le cadre de ce que nous appelons les systèmes complexes, décrits par les propriétés suivantes : il s'agit de systèmes composés d'un grand nombre d'entités, sans leader, mais dans lesquels des phénomènes d'auto-organisation et des émergences de structures sont observés.

Nous décidons de construire un modèle qui est discret, de type agent-centré. Ce choix se justifie par les arguments suivants :

- un modèle discret apporte une souplesse de modélisation. En effet, il permet de définir explicitement les règles de comportement des phénomènes modélisés, ce qui permet d'avoir un contrôle fin des caractéristiques du modèle. Dans l'optique de la création d'un modèle, avoir cette marge de manœuvre est important et offre la possibilité de modifier les caractéristiques de modélisation explicitement ;
- parmi les modèles discrets, nous choisissons un modèle agent-centré plutôt qu'un modèle sur grille. En effet, l'organisation et le déplacement des cellules

sont essentiellement la conséquence de la contrainte de congestion : il est donc important de ne pas se limiter à des déplacements sur une grille préalablement fixée, afin de conserver une liberté de mouvement ;

- les modèles agent-centré sont bien adaptés à la modélisation des systèmes complexes. En effet, un tel modèle focalise la modélisation sur les « agents », qui sont des entités propres munies de règles de fonctionnement et à partir desquelles émane une dynamique collective. Dans notre cas, les « agents » sont les cellules vivantes et le « système complexe » est la population cellulaire.

Le travail de modélisation consiste donc, dorénavant, à choisir les caractéristiques de ces agents ainsi que les règles définissant leur comportement.

1.1.2 Choix des agents et des règles de comportement

Cette section apporte les règles principales du modèle, qui sont ensuite décrites en détail dans les sections 1.2 à 1.4.

Ces règles sont déterminées en fonction :

- des problématiques auxquelles nous nous intéressons ;
- d'une sélection de phénomènes biologiques jugés essentiels pour reproduire le comportement d'une population cellulaire.

Cela nécessite de simplifier la réalité observée, pour en extraire les principes importants : à chaque fois que c'est possible, nous précisons les hypothèses de modélisation faites, en les reliant aux phénomènes biologiques négligés.

Pour explorer notre problématique 1, les cellules vivantes sont étudiées sur des temps assez grands pour que plusieurs *cycles cellulaires* se produisent. Les observations montrent que les cellules peuvent croître et se diviser, tout en restant regroupées — sans interpénétration — avec leurs voisines. Cette analyse rapide indique que, en ce qui concerne les agents, le modèle mathématique doit au moins pouvoir prendre en compte : leur position et leur taille, leur cohésion (entre eux), leur non-interpénétration (deux à deux) et l'augmentation de leur nombre via leur croissance et leur division.

Représentation des agents

La forme d'une cellule vivante est très variable et évolue au cours du temps, en fonction des états de cette cellule et de la population cellulaire : pour simplifier cette variabilité, nous commençons par approcher le bord de la cellule par une sphère. Ensuite, puisque les cellules se développent dans un environnement monocouche 2D, nous allons approcher la variation de forme d'une sphère par la variation de forme d'un cercle. Ainsi, chaque cellule est géométriquement représentée par un disque, dont les coordonnées du centre indiquent sa position et dont le rayon indique sa taille.

Hypothèse de modélisation 1. Les cellules sont considérées comme des entités vides, avec un contour circulaire. Autrement dit, nos cellules sont uniquement constituées de la membrane plasmique et celle-ci est parfaitement ronde pour toutes les cellules. En général, seuls les modèles emboîtés [PT09] (ou multi-échelles) décrivent le détail de phénomènes intracellulaires.

Par ailleurs, nous équipons chaque cellule de deux caractéristiques supplémentaires : son orientation et son lignage. L'orientation est un paramètre qui intervient au moment de la division et qui modélise l'*orientation de division* : elle correspond à la direction dans laquelle partiront les deux cellules filles ; l'introduction de ce paramètre est motivé par notre problématique 2. Le lignage est un paramètre informatif qui est commun à toutes les cellules émanant d'une même cellule mère : à chaque division, les deux cellules filles héritent du même lignage que leur cellule mère. La notion de lignage sera revue, définie et utilisée en section 3.2.1 page 71.

Hypothèse de modélisation 2. La direction de l'orientation du *plan de division*, lié à l'alignement des chromosomes au cours de la mitose, est explicitement donnée par un paramètre de modélisation interne à chaque cellule : son orientation.

Nous adoptons les notations suivantes, qui seront utilisées dans tous les chapitres du manuscrit :

- $t \in \mathbb{R}_+$ le temps ;
- $N(t) \in \mathbb{N}^*$ le nombre de cellules ;
- pour $i \in \llbracket 1 ; N(t) \rrbracket$, l'indice d'une cellule :
 - C_i la cellule numéro i ;
 - $X_i(t) = (x_i(t), y_i(t)) \in \mathbb{R}^2$ les coordonnées de son centre,
 - $R_i(t) \in \mathbb{R}_+^*$ son rayon,
 - $\omega_i(t) \in \mathbb{S}^1 / \{\pm 1\}$ son orientation,
 - $\ell_i \in \mathbb{N}$ son lignage (qui ne dépend pas du temps) ;
- $X(t) = (X_1(t), X_2(t), \dots, X_{N(t)}(t)) \in \mathbb{R}^{2 \times N(t)}$ l'ensemble des positions des cellules.
- $R(t) = (R_1(t), R_2(t), \dots, R_{N(t)}(t)) \in \mathbb{R}^{N(t)}$ l'ensemble des rayons des cellules.

Pour éviter des notations trop lourdes, la dépendance en temps t est souvent omise dans la suite.

Règle de comportement : état d'équilibre (à temps fixé)

Il s'agit de définir la configuration spatiale dans laquelle sont les cellules, à tout temps. Cet état d'équilibre est basé sur les deux observations biologiques principales suivantes : les cellules sont toujours regroupées entre-elles, puisqu'une cellule isolée finit par mourir, et elles ne s'interpénètrent pas.

L'objectif de cette règle de positionnement consiste à calculer la configuration des centres des agents $\{X_1, X_2, \dots, X_N\}$ à chaque temps. Sa description détaillée figure en section 1.2 page 26.

Règle de comportement : évolution (en temps)

Il s'agit de construire les règles de modélisation qui régissent l'évolution de nos agents dans le temps. En vertu de l'analyse rapide en début de section 1.1.2, nous considérons que les phénomènes de croissance et de division sont essentiels pour notre modèle. Afin de les mettre en valeur et de tester s'ils suffisent pour obtenir un comportement proche de la réalité, nous nous limitons à ces deux phénomènes.

Hypothèse de modélisation 3. L'évolution des cellules est uniquement déterminée par ces deux phénomènes, qui vont respectivement constituer la *loi de croissance* et la *loi de division*, décrivant respectivement l'étape d'*interphase* et l'étape de *mitose* d'un *cycle cellulaire*. Cette simplification est fréquente ; elle se retrouve dans [DH12].

La taille des agents étant modélisée par leur rayon, l'objectif de la règle de croissance est, pour tout $i \in \llbracket 1 ; N \rrbracket$, de calculer l'évolution du rayon R_i , pendant la période de croissance de la cellule C_i . L'objectif de la règle de division est de définir quand et comment se déroule la division d'une cellule mère en deux cellules filles ; elle consiste à supprimer une cellule C_i et à créer deux nouvelles cellules, notées C_i et C_{N+1} .

Différentes hypothèses simplificatrices par rapport à la réalité seront sous-jacentes lorsque nous décrirons ces deux règles. Elles sont regroupées et résumées ici.

Hypothèse de modélisation 4. Aucune activité cellulaire (autre que la croissance et la division) ni aucun échange avec l'extérieur n'est modélisé. En particulier, il n'y a aucune influence des nutriments ni de contraintes mécaniques du milieu ambiant et les mécanismes intracellulaires ne sont pas modélisés.

Hypothèse de modélisation 5. La succession des périodes de croissance et de division du cycle cellulaire n'est jamais interrompue. En particulier, les phénomènes biologiques suivants sont négligés : points de contrôle du cycle et stade G0 ; [SM05] les prend en compte.

Ces hypothèses simplificatrices sont cohérentes avec les choix annoncés en début de chapitre et dans l'introduction générale (notamment en section « *Cycle cellulaire et mitose* » page 5), qui sont en particulier :

- considérer les cellules comme des entités vides, en particulier sans *ADN* ni aucun constituant ;
- retenir les quatre premières *caractéristiques fondamentales* des cellules tumorales (autonomie de croissance, insensibilité à la non-prolifération, résistance à l'*apoptose*, répllication à l'infini), ce qui limite toute interruption du cycle cellulaire ;
- étudier des petites populations cellulaires, où toutes les cellules ont accès à la quantité de nutriments et d'oxygène nécessaire pour leur développement, ce qui exclut une sortie temporaire du *cycle cellulaire*.

Remarque 1.1 (Perspectives d'enrichissement de la modélisation). Relâcher l'hypothèse d'exclure le stade G0 (correspondant à une sortie du cycle cellulaire) pourrait consister, d'un point de vue modélisation, à équiper notre cellule d'une caractéristique supplémentaire : son état de prolifération, qui prendrait deux valeurs possibles (« P » pour une cellule proliférante, « Q » pour une cellule *quiescente*). Les règles de comportement (en particulier celle de croissance) seraient alors distinctes selon la valeur de ce nouveau paramètre.

Bilan de la mise en place du modèle

Nous arrivons donc aux caractéristiques synthétiques suivantes du modèle : le modèle est discret, de type agent-centré, et les agents sont les cellules vivantes. Ces

agents sont :

- représentés sous forme circulaire et équipés des caractéristiques suivantes : position du centre, taille du rayon, orientation, lignage ;
- placés, pour tout temps, dans un état d'équilibre défini par la loi de positionnement (modélisant un regroupement sans interpénétration) ;
- soumis, au cours du temps, aux évolutions de la loi de croissance (modélisant l'interphase) et de la loi de division (modélisant la mitose).

1.2 Loi de positionnement

Nous présentons dans cette section la loi qui définit la position des agents. Cette loi est décrite en deux étapes : nous commençons par en donner le cadre général en section 1.2.1 ; puis, en section 1.2.2, nous proposons des choix concrets de fonctions et de paramètres qui respectent ce cadre. Décorrélérer ces deux types d'informations permet de distinguer le cœur du modèle de ses paramètres, ces derniers étant modifiables par la suite sans changer les fondements de la modélisation.

Hypothèse de modélisation 6. Le positionnement des cellules est uniquement déterminé par les deux comportements de regroupement et de non-interpénétration, qui constituent notre loi de positionnement.

1.2.1 Principe général de la loi de positionnement

Nous choisissons de représenter cette loi de positionnement par la minimisation d'un potentiel global sous contraintes de non-recouvrement des agents. Cela se traduit en équations mathématiques de la manière suivante.

Non-recouvrement Pour $i, j \in \llbracket 1; N \rrbracket$, le non-recouvrement des deux cellules C_i et C_j s'exprime comme une contrainte inégalité sur une fonction $\phi_{i,j}: (\mathbb{R}^2)^N \rightarrow \mathbb{R}$, selon :

$$\begin{cases} \phi_{i,j} \leq 0 & \text{quand } C_i \text{ et } C_j \text{ ne se chevauchent pas ou lorsqu'elles sont en contact} \\ & \text{(contrainte satisfaite),} \\ \phi_{i,j} > 0 & \text{quand } C_i \text{ et } C_j \text{ se chevauchent (contrainte violée).} \end{cases} \quad (1.1)$$

Toutes les cellules vérifient la contrainte de non-recouvrement lorsqu'il n'y a aucun chevauchement deux à deux, autrement dit si leurs coordonnées appartiennent à l'ensemble des positions admissibles suivant :

$$\mathcal{A}(t) = \left\{ X \in (\mathbb{R}^2)^N \mid \text{pour tous } i, j \in \llbracket 1; N \rrbracket \text{ avec } i < j, \text{ on a } \phi_{i,j}(X) \leq 0 \right\}, \quad (1.2)$$

où $(\phi_{i,j})_{1 \leq i < j \leq 1}$ sont choisies respectant l'équation (1.1). Ainsi $X \in \mathcal{A}$ modélise la non-interpénétration de toutes les cellules.

Potentiel global Le potentiel global s'exprime comme une fonction $W: (\mathbb{R}^2)^N \rightarrow \mathbb{R}$ telle que

$$W: (X_1, X_2, \dots, X_N) \mapsto \sum_{i=1}^N V(X_i), \quad (1.3)$$

où $V: \mathbb{R}^2 \rightarrow \mathbb{R}$ est une fonction convexe (appelée potentiel des forces extérieures). Le regroupement des cellules est alors modélisé par la minimisation de ce potentiel global W .

Bilan La loi de positionnement, pour combiner ces deux facteurs, s'exprime alors comme un problème de minimisation du potentiel $W(t)$ sous la contrainte d'être dans l'ensemble $\mathcal{A}(t)$, pour tout $t \geq 0$; pour tout temps $t \geq 0$, la position des cellules est donc donnée par le vecteur $X^*(t) \in (\mathbb{R}^2)^N$ qui vérifie

$$X^*(t) = \underset{X(t) \in \mathcal{A}(t)}{\operatorname{argmin}} W(X(t)). \quad (1.4)$$

1.2.2 Valeurs des paramètres et des fonctions de modélisation

Nous donnons maintenant les choix retenus pour les fonctions W et $(\phi_{i,j})_{1 \leq i < j \leq 1}$. Ces choix sont établis pour correspondre aux principes généraux précédemment énoncés et aussi dans la perspective de l'implémentation numérique.

Fonctions contraintes Soit $i, j \in \llbracket 1; N \rrbracket$. Notons $\operatorname{dist}(C_i, C_j)$ la distance euclidienne usuelle sur \mathbb{R}^2 entre les cellules C_i et C_j , de sorte que

$$\operatorname{dist}(C_i, C_j) = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2}. \quad (1.5)$$

La fonction $\phi_{i,j}$ est alors définie par

$$\phi_{i,j}(X) = (R_i + R_j)^2 - \operatorname{dist}(C_i, C_j)^2 \quad \text{pour } X \in (\mathbb{R}^2)^N, \quad (1.6)$$

c'est-à-dire l'opposée de la distance entre le bord des cellules C_i et C_j ; elle vérifie bien la condition exprimée en équation (1.1) page 26.

Fonction potentiel Nous décidons que nos cellules vont vouloir se regrouper, de manière isotrope, autour d'un point fixé arbitrairement, que nous appelons la *position souhaitée*; cette notion est similaire à la notion de *vitesse souhaitée* présentée dans un modèle de mouvement de foule [MV07; MV11]. Nous prenons le point $0 \in \mathbb{R}^2$ qui, par convention, est le centre de la population cellulaire : le potentiel V est donc choisi tel que son minimum est réalisé en 0. Nous choisissons un potentiel quadratique, c'est-à-dire

$$x \mapsto \|x\|_{\mathbb{R}^2}^2 \quad \text{pour } x \in \mathbb{R}^2, \quad (1.7)$$

où $\|\cdot\|_{\mathbb{R}^2}$ est la norme euclidienne usuelle sur \mathbb{R}^2 . Cela donne donc l'expression suivante du potentiel global, notée W_Q :

$$W_Q(X) = \|X\|^2 = \sum_{i=1}^N \|X_i\|_{\mathbb{R}^2}^2 = \sum_{i=1}^N (x_i^2 + y_i^2) \quad \text{pour } X \in (\mathbb{R}^2)^N, \quad (1.8)$$

où $\|\cdot\|$ correspond à la norme euclidienne usuelle sur $(\mathbb{R}^2)^N$.

Remarque 1.2 (Perspectives d'enrichissement de la modélisation). Il est possible de choisir une position souhaitée quelconque, en prenant $V : x \mapsto \|x - p\|^2$, avec $p \in \mathbb{R}^2$; voire aussi une position souhaitée qui dépende de l'état des cellules avec $p : \mathbb{R}^2 \rightarrow \mathbb{R}$ (par exemple le barycentre du système). Une autre perspective d'enrichissement serait d'autoriser un faible recouvrement des cellules, sur une petite couronne, par exemple pour modéliser une déformation [Mac+12] : dans notre modèle, cela consisterait à ajouter une constante additive dans les fonctions $\phi_{i,j}$.

1.3 Loi de croissance

L'évolution en temps de notre modèle consiste, pour chaque cellule, à alterner entre une phase de croissance (où la taille des cellules augmente) et une phase de division (où une cellule se divise en deux cellules — cette seconde phase est discutée en section 1.4). La phase de croissance est donc consécutive à une division et préliminaire à une autre division : elle modélise l'étape d'*interphase* du *cycle cellulaire*.

Hypothèse de modélisation 7. L'interphase est constituée uniquement d'une phase de croissance cellulaire. En particulier, les phénomènes biologiques que sont les stades G1, S et G2 sont regroupés en une seule phase, durant laquelle la cellule n'a qu'un comportement : croître. Cette simplification est très fréquente [DH12 ; KSO07].

Introduisons les paramètres de modélisation suivants :

- R_{\min} est le rayon d'une cellule qui vient de naître ;
- R_{\max} est le rayon d'une cellule qui s'apprête à se diviser, c'est-à-dire la valeur seuil du rayon qui déclenche la phase de division (sa valeur est discutée en section 1.4) ;
- $T_G \in \mathbb{R}_+^*$ est la durée moyenne de la phase de croissance.

La loi de croissance, présentée dans cette section, consiste à déterminer la valeur des rayons $\{R_1, R_2, \dots, R_N\}$, de sorte qu'ils évoluent de R_{\min} à R_{\max} en une durée valant T_G .

1.3.1 Croissance linéaire en volume

La loi de croissance d'une cellule C_i , pour $i \in \llbracket 1; N \rrbracket$, est établie de sorte que le volume d'une sphère de rayon R_i croît linéairement avec le temps ; cela revient à dire que la fonction $t \mapsto R_i^3(t)$ évolue linéairement. Ce choix est motivé par les observations

biologiques et par le fait que, même si nous nous limitons à un environnement deux-dimensionnel pour modéliser le déplacement des cellules vivantes, il s'agit quand même d'objets tridimensionnels.

Hypothèse de modélisation 8. La croissance du volume des cellules au cours de l'*interphase* est supposée linéaire. Cette simplification est souvent effectuée [SM05 ; DH12].

Pour $i \in \llbracket 1 ; N \rrbracket$, le principe d'une croissance linéaire de la fonction $t \mapsto R_i^3(t)$ pour passer de la valeur R_{\min} à la valeur R_{\max} en un temps T_G amène à la loi d'évolution suivante :

$$R_i^3(t) = R_{\min} + \frac{R_{\max}^3 - R_{\min}^3}{T_G} t \quad \text{pour } t \text{ tel que } C_i \text{ est en phase de croissance.} \quad (1.9)$$

Cela signifie notamment que nous choisissons une même vitesse de croissance pour toutes les cellules.

Remarque 1.3. Il est aussi possible d'effectuer une croissance telle que ce soit la surface d'un disque de rayon R_i qui croisse linéairement avec le temps [DF00 ; DL01].

1.3.2 Croissance linéaire en volume avec bruit

La loi de croissance précédente a nécessité de faire comme hypothèse que toutes les cellules réalisent l'*interphase* en une même durée T_G . Nous assouplissons dès maintenant cette hypothèse, en ajoutant de l'aléatoire afin de modéliser la variabilité de la durée de l'interphase.

Nous prenons une variation uniforme autour de T_G : soit $\gamma \in [0; 1[$ et $u_\gamma \sim \mathcal{U}_{[-\gamma; \gamma]}$ (où $\mathcal{U}_{[a; b]}$ désigne la loi uniforme sur $[a; b]$) avec γ un nouveau paramètre de modélisation, dont la valeur est fixée en fonction des données biologiques. Au final, cela donne :

$$R_i^3(t) = R_{\min} + (1 + u_\gamma) \frac{R_{\max}^3 - R_{\min}^3}{T_G} t \quad \text{pour } t \text{ tel que } C_i \text{ est en phase de croissance,} \quad (1.10)$$

de sorte que le rayon R_i peut augmenter plus ou moins fortement au cours du temps. Dans l'optique de l'implémentation numérique, nous pouvons d'ores et déjà exprimer l'équation (1.10) sous la forme

$$R_i(t + dt) = \left(R_i^3(t) + (1 + u_\gamma) \frac{R_{\max}^3 - R_{\min}^3}{T_G} dt \right)^{1/3}. \quad (1.11)$$

Remarque 1.4 (Perspectives d'enrichissement de la modélisation). Intervenir sur le choix de la loi de probabilité suivie par u_γ est un paramètre permettant d'affiner la modélisation, en trouvant des distributions plus proches de la réalité.

Remarque 1.5 (Perspectives d'enrichissement de la modélisation). La loi de croissance pourrait dépendre des cellules. Par exemple, deux populations distinctes de cellules ayant des vitesses de croissance différentes pourraient être modélisées (dans notre

modélisation, cela consisterait à choisir une durée T_G dépendant des cellules). Un autre exemple consisterait à choisir des vitesses de croissance différentes selon la position de la cellule au sein de la population, ce qui permettrait de modéliser une différence d’apport en nutriments.

1.4 Loi de division

Nous présentons dans cette section la loi de division, c’est-à-dire la loi qui détermine l’évolution d’une cellule mère en deux cellules filles. Nous commençons par donner le principe général en section 1.4.1 ; puis, en sections 1.4.2 et 1.4.3, nous détaillons la procédure de division, en la décomposant en deux parties : d’une part la transformation d’une cellule en deux, d’autre part son interaction avec ses voisines ; enfin, en section 1.4.4, nous détaillons les choix d’orientation guidés par la problématique 2.

1.4.1 Principe général de la loi de division

Nous choisissons un processus de division qui se déclenche lorsque le rayon de la cellule atteint une certaine taille, notée R_{\max} (introduite en section 1.3 page 28), et qui consiste en la transformation progressive d’une cellule mère en deux cellules filles. La description de ces deux comportements est la suivante.

Condition de division. Pour $i \in \llbracket 1; N \rrbracket$, la cellule C_i commence sa division dès que $R_i \geq R_{\max}$.

Processus de division. Le processus de division consiste à transformer, de manière discrète, une cellule mère C_i en deux cellules filles ; à chaque étape de ce processus, l’objet déformé a une forme « en 8 » et l’ensemble des cellules est en situation d’équilibre. Nous choisissons comme caractéristiques de modélisation que cette transformation se fait à volume¹ constant et que les deux cellules filles seront deux fois plus petites (en volume) que la cellule mère.

Hypothèse de modélisation 9. Le déclenchement de la *mitose* a lieu dès que la cellule atteint une taille critique, et est uniquement piloté par cette condition : cette taille critique est identique pour toutes les cellules et constante au cours du temps. Elle correspond au doublement de la taille de la cellule, comme dans [DH12 ; GLD05 ; KSO07], ce qui est conforme aux critères transmis par les biologistes.

Hypothèse de modélisation 10. La *mitose* est uniquement constituée de la transformation d’une cellule de taille assez grande en deux cellules de forme strictement identique et symétriques par rapport au plan de division tout au long de cette transformation. En particulier, les différentes étapes de la mitose ne sont pas représentées². Ce processus de transformation est similaire à celui décrit dans [SM05 ; DH12].

1. Nos agents sont des disques dans \mathbb{R}^2 . La terminologie « volume » n’a pas ici de sens physique à proprement parler : il s’agit uniquement de la quantité $\frac{4}{3}\pi R_i^3$.

2. À l’exception, pour explorer notre problématique 2, de l’alignement des chromosomes qui forment le *plan de division* : la direction de l’*orientation de division* est modélisée par le paramètre d’orientation ω_i de la cellule C_i , pour $i \in \llbracket 1; N \rrbracket$.

Hypothèse de modélisation 11. Nous considérons qu’il y a conservation du volume au cours du processus de division. Conserver une quantité est fréquent [SM05 ; DH12] mais nous faisons le choix de conserver le volume pour des agents représentés en 2D par des disques. Ce choix a été discuté en section « Règles de comportement : croissance jusqu’à doublement, puis division par déformation “en 8” » page 15 de l’introduction générale et est justifié en remarque 1.6.

Par ailleurs, nous supposons que :

- il ne peut pas y avoir plusieurs cellules qui réalisent cette transformation simultanément (autrement dit, si deux cellules atteignent la taille critique au même instant t , l’une doit avoir terminé complètement son processus de division avant que l’autre ne l’entame) ;
- l’évolution en temps du modèle est stoppée jusqu’à ce que toutes les cellules ayant atteint la taille R_{\max} à l’instant t aient fini leur division ; en particulier, aucune cellule ne croît lorsqu’une cellule se divise.

Hypothèse de modélisation 12. La *mitose*, dont la durée est environ 24 fois plus petite que celle de l’*interphase* dans les cellules vivantes considérées, est supposée instantanée en temps. Supposer qu’une telle simplification n’a qu’un faible impact a déjà été fait [SM05].

L’hypothèse de conservation du volume permet de déterminer la valeur numérique de R_{\max} . En effet, cela consiste à transformer une cellule de rayon R_{\max} en deux cellules de rayon

$$\frac{R_{\max}}{\sqrt[3]{2}}, \quad (1.12)$$

or nous avons fixé comme paramètre de modélisation le rayon d’une cellule qui vient de naître à R_{\min} , donc cela impose la valeur de R_{\max} :

$$R_{\max} = \sqrt[3]{2}R_{\min}. \quad (1.13)$$

Remarque 1.6 (Conservation du volume). Ce choix de conserver le volume provient d’observations biologiques : conserver l’aire aurait amené à fixer la valeur $\sqrt{2}R_{\min}$ pour R_{\max} et avoir ainsi une plus grande différence de taille entre une cellule naissante et une cellule prête à se diviser ; les biologistes, après observations des simulations numériques (produites au chapitre 2), ont jugé trop grand cet écart.

Pour définir le processus de division de la cellule, nous distinguons sa transformation en elle-même de son interaction avec ses voisines :

- la transformation se fait en plusieurs étapes, illustrées en figure 1.1 page suivante, au cours desquelles la cellule n’a plus une forme circulaire mais une forme « en 8 » ;
- à la fin de chacune de ces étapes, nous imposons aux cellules d’être dans un état d’équilibre, c’est-à-dire de satisfaire à un positionnement sans interpénétration.

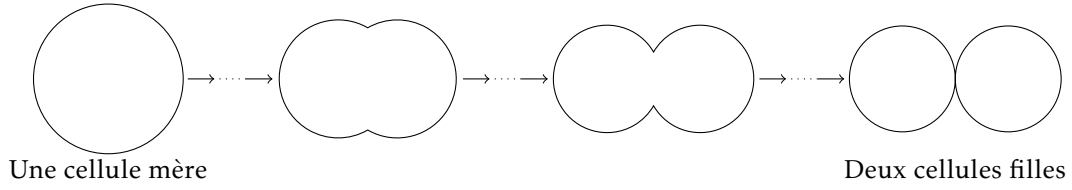


Illustration de la transformation, en plusieurs étapes, d'une cellule mère en deux cellules filles, avec conservation de la taille.

FIGURE 1.1 – Transformation au cours de la division

1.4.2 Processus de division d'une cellule : la transformation

Adoptons les notations suivantes, propres au processus de division en plusieurs étapes d'une cellule :

- $i_0 \in \llbracket 1; N \rrbracket$ l'indice de la cellule mère ;
- R_{i_0} le rayon de la cellule mère lorsque le processus commence ($R_{i_0} \geq R_{\max}$) ;
- ω_{i_0} l'orientation de la cellule mère lorsque le processus commence (dont la valeur est discutée en section 1.4.4 page 37) ;
- $V_{i_0} = \frac{4}{3}\pi R_{i_0}^3$ le volume de la cellule mère lorsque le processus commence, qui correspond à la quantité conservée au cours du processus ;
- N_{div} le nombre d'étapes du processus ;
- $\tau_1, \tau_2, \dots, \tau_{N_{\text{div}}}$ les différentes étapes du processus et τ_0 l'étape d'initialisation, qui correspond à l'état de la cellule mère juste avant de commencer le processus ;
- $(\tilde{x}_+(\tau_k), \tilde{y}_+(\tau_k), \tilde{R}_+(\tau_k), \tilde{\omega}_+(\tau_k), \tilde{\ell}_+(\tau_k))$ et $(\tilde{x}_-(\tau_k), \tilde{y}_-(\tau_k), \tilde{R}_-(\tau_k), \tilde{\omega}_-(\tau_k), \tilde{\ell}_-(\tau_k))$ les caractéristiques des deux cellules naissantes (abscisse, ordonnée, rayon, orientation, lignage) à l'étape τ_k , pour $k \in \llbracket 0; N_{\text{div}} \rrbracket$.

Dans la suite, toute quantité notée avec un tilde est susceptible de varier au cours du processus de division, et sa dépendance en τ_k pour $k \in \llbracket 0; N_{\text{div}} \rrbracket$ n'est pas systématiquement spécifiée (dans le but d'alléger les notations). Toute quantité notée sans tilde fait référence à une quantité non modifiée par le processus de division (comme R_{i_0} , ω_{i_0} et V_{i_0} notamment).

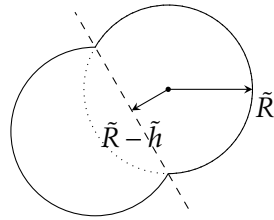
Conformément à nos principes de modélisation, nous devons établir une transformation discrète conservant la quantité $\frac{4}{3}\pi R_{i_0}^3$, appelée volume. Par hypothèse de modélisation, pour tout $k \in \llbracket 0; N_{\text{div}} \rrbracket$,

$$\tilde{R}_+(\tau_k) = \tilde{R}_-(\tau_k) \quad \text{noté } \tilde{R}(\tau_k), \quad (1.14)$$

$$\tilde{\omega}_+(\tau_k) = \tilde{\omega}_-(\tau_k) \quad \text{noté } \tilde{\omega}(\tau_k), \quad (1.15)$$

$$\tilde{\ell}_+(\tau_k) = \tilde{\ell}_-(\tau_k) = \ell_{i_0}. \quad (1.16)$$

Chacune des deux cellules naissantes a la forme décrite en figure 1.2 page suivante, qui correspond à un disque privé d'un arc de disque. Cette forme est modélisée par deux cercles de centre $(\tilde{x}_-, \tilde{y}_-)$ et $(\tilde{x}_+, \tilde{y}_+)$, de rayon \tilde{R} , se superposant ; notons \tilde{h} la hauteur de cet arc de disque. Nous paramétrons la transformation par cette distance \tilde{h} , qui décroît



Configuration des deux cellules naissantes, lors d'une étape du processus de transformation. En tiret est représenté l'axe de division (symbolisant le *plan de division*).

FIGURE 1.2 – Forme d'une cellule au cours de la division

de $\tilde{h}(\tau_0) = \tilde{R}(\tau_0)$ à $\tilde{h}(\tau_{N_{\text{div}}}) = 0$, pour modéliser le passage de deux demi-disques de rayon $\tilde{R}(\tau_0) = R_{i_0}$ (au début de la division) à deux disques de rayon $\tilde{R}(\tau_{N_{\text{div}}}) = \frac{R_{i_0}}{\sqrt[3]{2}}$ (à la fin de la division). Cela donne, pour $k \in \llbracket 0; N_{\text{div}} \rrbracket$, le paramétrage suivant de la transformation :

$$\tilde{h}(\tau_k) = R_{i_0} - k \frac{R_{i_0}}{N_{\text{div}}}. \quad (1.17)$$

Ainsi, à chaque étape τ_k (pour $k \in \llbracket 0; N_{\text{div}} \rrbracket$) nous devons calculer le rayon et la position des deux disques en forme « de 8 », de sorte que la somme de leur volume reste égal à V_{i_0} .

Valeurs à l'étape τ_0 (initialisation du processus) Cette étape, fictive, décrit l'état de la cellule mère avant que la transformation ne commence (comme représentée en figure 1.1 page 32). Nous avons donc les valeurs suivantes :

$$\tilde{R}(\tau_0) = R_{i_0}, \quad (1.18)$$

$$(\tilde{x}_+, \tilde{y}_+)(\tau_0) = (\tilde{x}_-, \tilde{y}_-)(\tau_0) = (\tilde{x}_c, \tilde{y}_c)(\tau_0) = (x_{i_0}, y_{i_0}), \quad (1.19)$$

$$\tilde{\omega}(\tau_0) = \omega_{i_0}. \quad (1.20)$$

En particulier, ω_{i_0} est un paramètre du modèle, que nous avons la liberté de choisir ; son choix est discuté en section 1.4.4 page 37.

Calcul du rayon à l'étape τ_k Étant donné notre choix de paramétrer les étapes de cette transformation par le pas $\tilde{h}(\tau_k)$, le calcul du rayon dépend uniquement de la valeur de ce pas. Le volume d'un dôme sphérique, tel qu'illustré en figure 1.3 page suivante, est

$$V_{\text{dôme}} = \frac{\pi \tilde{h}}{6} (6\tilde{h}\tilde{R} - 2\tilde{h}^2),$$

d'où nous déduisons qu'une cellule naissante a pour volume

$$\tilde{V}(\tau_k) = \frac{4}{3}\pi\tilde{R}^3 - V_{\text{dôme}} = \frac{\pi}{3} (4\tilde{R}^3 - 3\tilde{h}^2\tilde{R} + \tilde{h}^3).$$

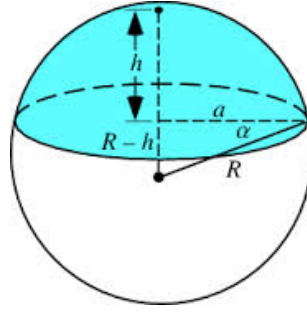


FIGURE 1.3 – Dôme sphérique

Par conservation du volume des deux cellules naissantes, cela donne l'équation sur \tilde{R} suivante :

$$V_{i_0} = \frac{2\pi}{3} (4\tilde{R}^3 - 3\tilde{h}^2\tilde{R} + \tilde{h}^3),$$

qui amène à l'équation

$$\tilde{R}^3 - \frac{3\tilde{h}^2}{4}\tilde{R} + \left(\frac{\tilde{h}^3}{4} - \frac{R_{i_0}^3}{2}\right) = 0, \quad (1.21)$$

dont la solution, pour $\tilde{h} \in [0; R_{i_0}[$, est donnée par

$$\tilde{R} = \sqrt[3]{\frac{-q + \sqrt{\Delta}}{2}} + \sqrt[3]{\frac{-q - \sqrt{\Delta}}{2}}, \quad (1.22)$$

où

$$\begin{aligned} q &= \frac{\tilde{h}^3}{4} - \frac{R_{i_0}^3}{2}, \\ p &= \frac{-3\tilde{h}^2}{4}, \\ \Delta &= q^2 + \frac{4}{27}p^3 = \frac{R_{i_0}^3}{4} (R_{i_0}^3 - \tilde{h}^3) > 0. \end{aligned}$$

Détermination des autres variables (position et orientation) à l'étape τ_k Les cellules naissantes se positionnent le long de la normale à l'axe de division, à égale distance (valant $\tilde{R} - \tilde{h}$, précédemment calculée) de cet axe. Au début de l'étape τ_k , le centre de cette forme « en 8 » ainsi que l'orientation des cellules sont déterminés par les positions à la fin de l'étape τ_{k-1} , d'où

$$\tilde{\omega}(\tau_k) = \begin{cases} \text{atan2}(\tilde{y}_+(\tau_{k-1}) - \tilde{y}_-(\tau_{k-1}), \tilde{x}_+(\tau_{k-1}) - \tilde{x}_-(\tau_{k-1})) & \text{si } k \neq 1, \\ \omega_{i_0} & \text{sinon.} \end{cases} \quad (1.23)$$

et

$$\begin{cases} \tilde{x}_c(\tau_k) = \frac{\tilde{x}_+(\tau_{k-1}) + \tilde{x}_-(\tau_{k-1})}{2}, \\ \tilde{y}_c(\tau_k) = \frac{\tilde{y}_+(\tau_{k-1}) + \tilde{y}_-(\tau_{k-1})}{2}. \end{cases} \quad (1.24)$$

Ces choix déterminent les coordonnées des centres des deux cellules naissantes :

$$\begin{cases} \tilde{x}_\pm(\tau_k) = \tilde{x}_c(\tau_k) \pm (\tilde{R}(\tau_k) - \tilde{h}(\tau_k)) \cos(\tilde{\omega}(\tau_k)), \\ \tilde{y}_\pm(\tau_k) = \tilde{y}_c(\tau_k) \pm (\tilde{R}(\tau_k) - \tilde{h}(\tau_k)) \sin(\tilde{\omega}(\tau_k)). \end{cases} \quad (1.25)$$

Conclusion Le processus de transformation d'une cellule mère en deux cellules filles, indépendamment de toute interaction avec l'extérieur, se fait selon la modélisation suivante :

- étape τ_0 (initialisation), où $\tilde{h}(\tau_0) = R_{i_0}$:
 - leur orientation est donnée en équation (1.20) page 33,
 - leur rayon $\tilde{R}(\tau_0)$ est donné en équation (1.18) page 33,
 - leurs positions initiales sont données en équation (1.19) page 33 ;
- étape τ_k , pour $k \in \llbracket 1; N_{\text{div}} \rrbracket$, où $\tilde{h}(\tau_k)$ est donné en équation (1.17) page 33 :
 - leur orientation est donnée en équation (1.23) page 34,
 - leur rayon $\tilde{R}(\tau_k)$ est donné en équation (1.22) page 34,
 - leurs positions sont données en équation (1.25).

Remarque 1.7. L'orientation $\tilde{\omega}$ n'a aucune raison d'être modifiée au cours de cette transformation. Néanmoins, comme nous allons le voir dans la suite, cette orientation va changer dès que les interactions avec les cellules voisines seront modélisées : c'est pour cette raison que nous précisons le calcul de $\tilde{\omega}(\tau_k)$ pour tout $k \in \llbracket 0; N_{\text{div}} \rrbracket$.

1.4.3 Processus de division d'une cellule : l'interaction

Il s'agit ici de déterminer quelle est l'interaction des cellules naissantes en forme « de 8 » avec leur environnement, c'est-à-dire avec leurs cellules voisines : cela consiste à repositionner l'ensemble des cellules de sorte qu'il n'y ait pas de recouvrement. Nous modélisons cette action en créant une loi de positionnement presque similaire³ à celle décrite en section 1.2, qui s'applique après chaque étape τ_k (pour $k \in \llbracket 1; N_{\text{div}} \rrbracket$) de la transformation définie en section 1.4.2.

Grâce aux hypothèses faites en début de section 1.4, lorsqu'une cellule se divise, nous avons :

- d'une part le temps est gelé (c'est-à-dire qu'aucune cellule ne peut croître) ;
- d'autre part aucune autre cellule ne peut entamer sa division (c'est-à-dire qu'une cellule qui a aussi atteint la taille seuil pour se diviser doit attendre que cette cellule se divisant ait totalement achevé sa division avant de la commencer).

Quand la cellule C_{i_0} se divise, N est incrémenté ; nous notons C_- et C_+ les deux cellules naissantes, i_- et i_+ leur indice (de sorte que $i_+ = i_0$ et $i_- = N$).

3. La différence est qu'il y a une cellule se divisant (donc un agent ayant une forme « en 8 »).

Règle n° 1 de positionnement, dite « stratégie d'orientation libre »

La différence avec la règle de positionnement décrite en section 1.2 page 26 est que, au lieu d'avoir N cellules circulaires, nous avons dorénavant $N - 2$ cellules circulaires et deux qui forment « un 8 ». Ces deux cellules naissantes sont une unique structure : elle peut être vue comme deux cellules ayant le droit de se recouvrir partiellement et telles que ce recouvrement reste identique durant le positionnement (afin que leur forme « en 8 » soit conservée). Nous modélisons cela :

- en supprimant la contrainte inégalité de non-recouvrement entre C_- et C_+ (de sorte qu'elles ne soient plus empêchées de se recouvrir) ;
- en la remplaçant par une contrainte égalité (de sorte que la distance $\text{dist}(C_-, C_+)$ entre ces deux cellules soit constante).

Toutes les autres $\frac{N(N+1)}{2} - 1$ contraintes sont inchangées.

Pour $i, j \in \llbracket 1; N \rrbracket$ tels que $i < j$, nous notons $\tilde{\phi}_{i,j} : (\mathbb{R}^2)^N \rightarrow \mathbb{R}$ les fonctions contraintes, à choisir telles que

$$\left\{ \begin{array}{l} \text{pour } (i, j) = (i_+, i_-), \tilde{\phi}_{i,j}(X) = 0, \\ \text{pour } (i, j) \neq (i_+, i_-) : \\ \quad \left\{ \begin{array}{l} \tilde{\phi}_{i,j}(\tilde{X}) \leq 0 \quad \text{quand } C_i \text{ et } C_j \text{ ne se chevauchent pas ou sont en contact,} \\ \tilde{\phi}_{i,j}(\tilde{X}) > 0 \quad \text{quand } C_i \text{ et } C_j \text{ se chevauchent.} \end{array} \right. \end{array} \right. \quad (1.26)$$

Les contraintes sont respectées lorsque

$$\left\{ \begin{array}{l} \tilde{\phi}_{i,j}(\tilde{X}) = 0 \quad \text{si } (i, j) = (i_+, i_-), \\ \tilde{\phi}_{i,j}(\tilde{X}) \leq 0 \quad \text{sinon.} \end{array} \right. \quad (1.27)$$

Les contraintes sont alors représentées par l'ensemble des positions admissibles suivant :

$$\tilde{\mathcal{A}}(\tau_k) = \left\{ \tilde{X} \in (\mathbb{R}^2)^N \mid \text{pour tous } i, j \in \llbracket 1; N \rrbracket \text{ avec } i < j, \text{ l'équation (1.27) est vérifiée} \right\}, \quad (1.28)$$

Pour $k \in \llbracket 1; N_{\text{div}} \rrbracket$, la position $\tilde{X}(\tau_k)$ de toutes les cellules à la fin de l'étape τ_k est donnée par la résolution du problème de minimisation sous contraintes

$$\tilde{X}^*(\tau_k) = \underset{\tilde{X} \in \tilde{\mathcal{A}}(\tau_k)}{\text{argmin}} W(\tilde{X}(t)). \quad (1.29)$$

Choix des fonctions contraintes Nous choisissons, pour tous $i, j \in \llbracket 1; N \rrbracket$ tels que $i < j$,

$$\tilde{\phi}_{i,j}(\tilde{X}(\tau_k)) = \begin{cases} \left(\tilde{R}_i(\tau_k) + \tilde{R}_j(\tau_k) \right)^2 - \text{dist}(C_i(\tau_k), C_j(\tau_k))^2 & \text{si } (i, j) \neq (i_+, i_-), \\ 4 \left(\tilde{R}(\tau_k) - \tilde{h}(\tau_k) \right)^2 - \text{dist}(C_i(\tau_k), C_j(\tau_k))^2 & \text{si } (i, j) = (i_+, i_-). \end{cases} \quad (1.30)$$

Règle n° 2 de positionnement, dite « stratégie d'orientation contrainte »

Une de nos problématiques est d'étudier la conséquence de l'*orientation de division* des cellules durant la *mitose*. En l'état actuel de la modélisation, une cellule se divisant est libre de se déplacer, sans aucune autre contrainte que de garder une forme « en 8 ». Dans cette 2^e règle, nous modélisons une interaction de positionnement avec les cellules voisines telle que l'orientation des cellules naissantes en forme « de 8 » soit constante au cours de tout le processus de division. Autrement dit, nous imposons que pour tout $k \in \llbracket 1; N_{\text{div}} \rrbracket$, $\tilde{\omega}(\tau_k) = \omega_{i_0}$. Modéliser cela consiste à garder C_-C_+ colinéaire avec ω_{i_0} , c'est-à-dire

$$\det(C_-C_+, \omega_{i_0}) = 0. \quad (1.31)$$

Posons $\tilde{\phi}_{\tilde{\omega}}: (\mathbb{R}^2)^N \rightarrow \mathbb{R}$ une nouvelle fonction contrainte définie par

$$\tilde{\phi}_{\tilde{\omega}} = (\tilde{x}_+ - \tilde{x}_-) \sin(\tilde{\omega}) - (\tilde{y}_+ - \tilde{y}_-) \cos(\tilde{\omega}). \quad (1.32)$$

La condition de colinéarité amène à

$$\tilde{\phi}_{\tilde{\omega}}(\tilde{X}) = 0. \quad (1.33)$$

Cette règle n° 2 d'interaction consiste alors à adopter le positionnement suivant : pour $k \in \llbracket 1; N_{\text{div}} \rrbracket$, la position $\tilde{X}(\tau_k)$ de toutes les cellules à la fin de l'étape τ_k est donnée par la résolution du problème de minimisation sous contraintes

$$\tilde{X}^*(\tau_k) = \underset{\tilde{X} \in \tilde{\mathcal{A}}_{\tilde{\omega}}(\tau_k)}{\operatorname{argmin}} W(\tilde{X}(t)), \quad (1.34)$$

où

$$\tilde{\mathcal{A}}_{\tilde{\omega}}(\tau_k) = \tilde{\mathcal{A}}(\tau_k) \cap \left\{ \tilde{X} \in (\mathbb{R}^2)^N \mid \tilde{\phi}_{\tilde{\omega}}(\tilde{X}) = 0 \right\}, \quad (1.35)$$

avec $\tilde{\omega}$ valant ω_{i_0} , un paramètre de modélisation.

1.4.4 Choix d'orientation

Pour explorer notre problématique 2, nous décidons que l'orientation, notée ω_{i_0} , d'une cellule commençant sa division est un paramètre de modélisation. Dans notre modèle, nous retenons quatre choix pour décider de la direction de cette orientation ω_{i_0} .

1. Aléatoire : au début de chaque division, ω_{i_0} est choisie aléatoirement.
2. Fixe : une même valeur arbitraire est donnée à ω_{i_0} , pour toutes les divisions.
3. Radiale : la direction ω_{i_0} est calculée en fonction de la position relative de la cellule C_{i_0} au sein de la population cellulaire. Notons θ l'angle polaire du point du plan défini par (x_{i_0}, y_{i_0}) ; alors on choisit $\omega_{i_0} = \theta$, autrement dit les deux cellules naissantes vont s'éloigner l'une de l'autre le long de la droite définie en coordonnées polaires par $\{\rho \exp(i\omega_{i_0}), \rho \in \mathbb{R}\}$.
4. Tangentielle : la direction ω_{i_0} est orthogonale à l'orientation radiale.

‘I should have made my meaning plainer,’ said Professor McGonagall, turning at last to look Umbridge directly in the eyes. ‘He has achieved high marks in all Defence Against the Dark Arts tests set by a competent teacher.’

Harry Potter and the Order of the Phoenix

Chapitre 2

Implémentation numérique du modèle

L’objectif de ce chapitre est de développer un algorithme permettant d’effectuer des simulations numériques du modèle présenté au chapitre 1. Ces simulations numériques seront étudiées au chapitre 3, dans le but d’être ensuite comparées aux expériences réalisées par les biologistes sur les cultures cellulaires.

Après une présentation générale de l’algorithme (section 2.1), nous détaillons chacune de ses étapes. Nous commençons par l’étape de positionnement (section 2.2); puis ensuite viennent l’étape de croissance (section 2.3) et l’étape de division (section 2.4); enfin, nous détaillons l’étape d’initialisation et les valeurs numériques données aux paramètres (section 2.5). Cette démarche permet de mettre en place un algorithme associé au modèle.

Sommaire du chapitre

2.1	Structure générale de l’algorithme	40
2.2	Étape de positionnement	40
2.2.1	Algorithme de positionnement	42
	Calcul de l’itération présentée en équation (2.5a)	42
	Critère d’arrêt et signification des paramètres numériques ρ et μ	45
	Bilan de l’algorithme de positionnement	46
2.2.2	Discussion sur les valeurs attribuées aux pas de l’algorithme	47
2.2.3	Améliorations autour de l’algorithme de positionnement	50
2.2.4	Bilan de l’étape de positionnement	51
2.3	Étape de croissance	52
2.4	Étape de division d’une cellule	52
2.4.1	Processus de division à l’étape τ_k pour k fixé	55
2.4.2	Initialisation du processus de division et synthèse de l’étape	58
2.5	Configurations de l’algorithme	59
2.5.1	Étape d’initialisation	59
2.5.2	Valeurs numériques des paramètres	60

2.1 Structure générale de l'algorithme

Nous présentons succinctement dans cette section l'arborescence de l'algorithme. Cela permet de disposer d'une vue globale de l'algorithme avant d'en décrire les différentes parties.

L'algorithme suit les principes de modélisation énoncés au chapitre 1 : ses différentes étapes coïncident avec les lois de modélisation (positionnement, croissance, division). La structure qui en résulte, illustrée par un schéma en figure 2.1 page suivante, est présentée ci-après.

Structure de l'algorithme
<ol style="list-style-type: none"> 1. Choix des paramètres de modélisation et de l'algorithme. 2. Étape d'initialisation. 3. À chaque pas de temps (boucle en temps) : <ol style="list-style-type: none"> (a) étape de croissance ; (b) pour chacune des cellules, les unes après les autres (boucle sur les cellules) : <ol style="list-style-type: none"> (i) test sur la condition de division, (ii) si ce test est positif, alors étape de division ; (c) étape de positionnement.

Ces étapes représentent le cœur de l'algorithme et sont décrites plus en détail dans les sections 2.2 à 2.4 et 2.5.1.

2.2 Étape de positionnement

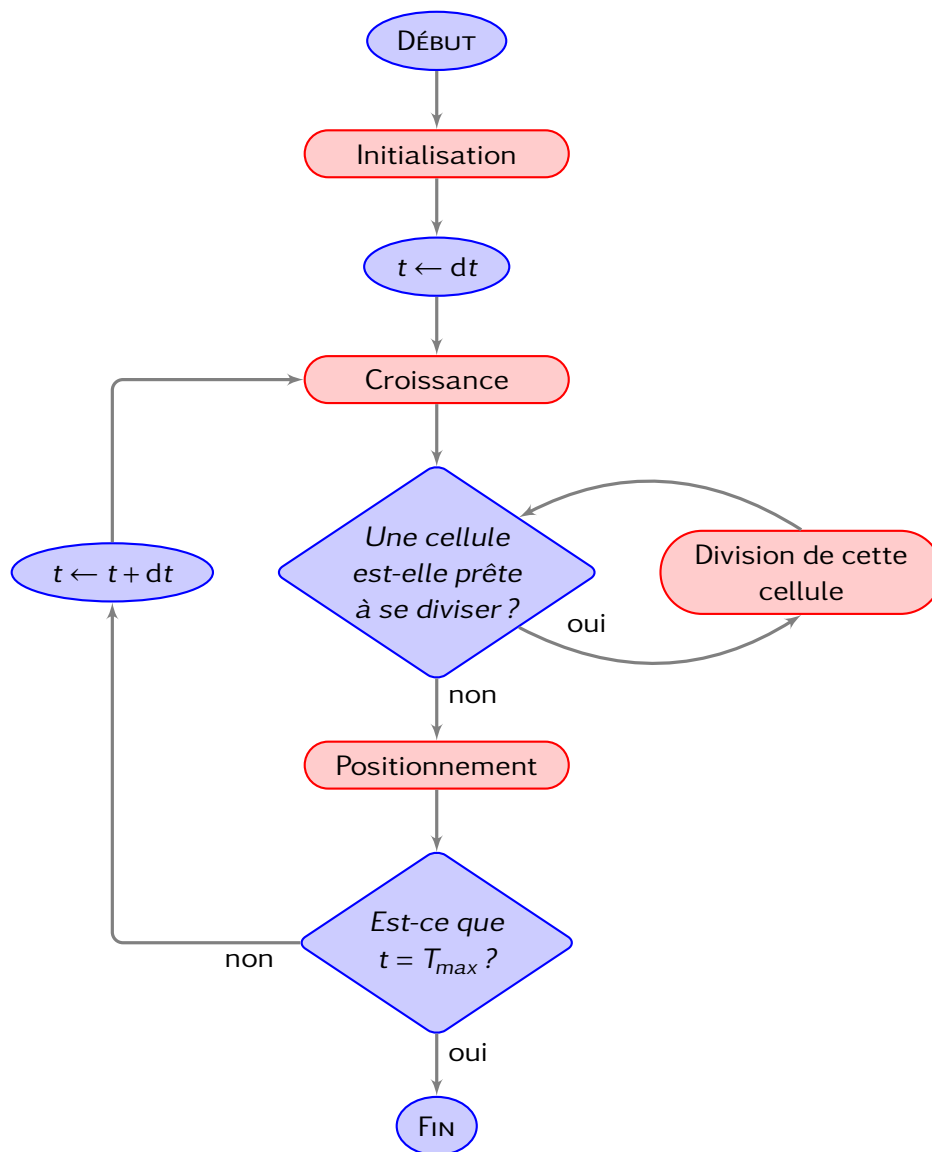
L'objectif de l'étape de positionnement est de trouver une solution au problème de minimisation sous contraintes énoncé en équation (1.4) page 27. Ce problème se réécrit comme suit : chercher $X^* \in (\mathbb{R}^2)^N$ tel que

$$X^* = \underset{\substack{X \in (\mathbb{R}^2)^N \text{ tq} \\ \phi_{i,j}(X) \leq 0}}{\operatorname{argmin}} W(X), \quad (2.1)$$

où W est la fonction à minimiser et $(\phi_{i,j})_{i < j}$ sont les fonctions contraintes ; celles-ci sont définies, pour $X \in (\mathbb{R}^2)^N$ et $R \in \mathbb{R}^N$, par

$$\phi_{i,j}(X) = (R_i + R_j)^2 - \left[(x_j - x_i)^2 + (y_j - y_i)^2 \right], \quad (2.2)$$

$$W = W_Q \text{ où } W_Q(X) = \sum_{i=1}^N (x_i^2 + y_i^2), \quad (2.3)$$




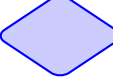

Légende.  est une étape de l'algorithme;  est un test;  est une action mineure. Les paramètres T_{max} et dt sont expliqués en tableaux 2.1 et 2.3 page 60 et page 61.

FIGURE 2.1 – Vision schématique de la totalité de l'algorithme

où nous notons

$$X = \left[\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}, \dots, \begin{pmatrix} x_N \\ y_N \end{pmatrix} \right] \quad \text{et} \quad R = (R_1, R_2, \dots, R_N).$$

Pour résoudre ce problème de minimisation sous contraintes inégalités, nous décidons de mettre en place un algorithme basé sur la méthode d'Uzawa.

2.2.1 Algorithme de positionnement

Soit $M \in \mathbb{N}$ le nombre de fonctions contraintes :

$$M = \frac{N(N-1)}{2}.$$

Pour le problème énoncé en équations (2.1) à (2.3) page 40, la méthode d'Uzawa consiste à trouver un point selle du Lagrangien $\mathcal{L}_Q: (\mathbb{R}^2)^N \times \mathbb{R}^M \rightarrow \mathbb{R}$ défini par

$$\forall (X, \lambda) \in (\mathbb{R}^2)^N \times \mathbb{R}^M, \quad \mathcal{L}_Q(X, \lambda) = W_Q(X) + \sum_{1 \leq i < j \leq N} \lambda_{i,j} \phi_{i,j}(X), \quad (2.4)$$

où les $(\lambda_{i,j})_{i < j}$ sont appelés les multiplicateurs de Lagrange¹. La méthode d'Uzawa et la notion de point selle sont détaillées dans [All12, chap. 10]. Il existe différentes versions de l'algorithme d'Uzawa, dont l'objectif est de calculer une suite d'approximations $(X^{(p)}, \lambda^{(p)})_p$ de sorte que $X^{(p)} \xrightarrow{p \rightarrow +\infty} X^*$. Nous retenons la suivante : à partir des deux éléments quelconques $X^{(0)} \in (\mathbb{R}^2)^N$ et $\lambda^{(0)} \in (\mathbb{R}_+)^M$, nous construisons les suites $(X^{(p)})_p$ et $(\lambda^{(p)})_p$ déterminées, pour $p \geq 0$, par les itérations

$$X^{(p+1)} = X^{(p)} - \rho \nabla \mathcal{L}(X^{(p)}, \lambda^{(p)}), \quad (2.5a)$$

$$\phi_{i,j}^{(p+1)} = \phi_{i,j}(X^{(p+1)}), \quad \text{pour tous } i, j \in \llbracket 1; N \rrbracket \text{ tels que } i < j, \quad (2.5b)$$

$$\lambda_{i,j}^{(p+1)} = \max\left(0, \lambda_{i,j}^{(p)} + \mu \phi_{i,j}^{(p+1)}\right), \quad \text{pour tous } i, j \in \llbracket 1; N \rrbracket \text{ tels que } i < j, \quad (2.5c)$$

où ρ et μ sont des paramètres positifs fixés.

Calcul de l'itération présentée en équation (2.5a)

Pour connaître l'itération sur $X^{(p+1)}$, nous calculons $\nabla \mathcal{L}$. Soit $X \in (\mathbb{R}^2)^N$ et $\lambda \in \mathbb{R}^M$: calculons séparément chacun des deux termes ∇W et $\nabla \phi_{i,j}$ pour $i, j \in \llbracket 1; N \rrbracket$.

1. Ils forment une famille de \mathbb{R}^M mais sont indicés par un couple (i, j) avec $1 \leq i < j \leq N$. Ce choix anticipe le moment où nous les représenterons par une matrice symétrique $N \times N$ (voir équation (2.12) page 44).

— Nous considérons $\nabla W(X)$ comme une matrice de taille $2 \times N$, de sorte que :

$$\nabla W(X) = \left[\left(\begin{array}{c} \frac{\partial W}{\partial x_1} \\ \frac{\partial W}{\partial y_1} \end{array} \right), \left(\begin{array}{c} \frac{\partial W}{\partial x_2} \\ \frac{\partial W}{\partial y_2} \end{array} \right), \dots, \left(\begin{array}{c} \frac{\partial W}{\partial x_N} \\ \frac{\partial W}{\partial y_N} \end{array} \right) \right] (X),$$

d'où

$$\nabla W(X) = 2 \left[\left(\begin{array}{c} x_1 \\ y_1 \end{array} \right), \left(\begin{array}{c} x_2 \\ y_2 \end{array} \right), \dots, \left(\begin{array}{c} x_N \\ y_N \end{array} \right) \right] = 2X.$$

— Soit $i, j \in \llbracket 1; N \rrbracket$ tels que $i < j$; nous avons :

$$\nabla \phi_{i,j}(X) = \left[\left(\begin{array}{c} \frac{\partial \phi_{i,j}}{\partial x_1} \\ \frac{\partial \phi_{i,j}}{\partial y_1} \end{array} \right), \left(\begin{array}{c} \frac{\partial \phi_{i,j}}{\partial x_2} \\ \frac{\partial \phi_{i,j}}{\partial y_2} \end{array} \right), \dots, \left(\begin{array}{c} \frac{\partial \phi_{i,j}}{\partial x_N} \\ \frac{\partial \phi_{i,j}}{\partial y_N} \end{array} \right) \right] (X),$$

or, pour tout $k \in \llbracket 1; N \rrbracket$,

$$\left(\begin{array}{c} \frac{\partial \phi_{i,j}}{\partial x_k} \\ \frac{\partial \phi_{i,j}}{\partial y_k} \end{array} \right) (X) = \begin{cases} 0 & \text{si } k \notin \{i, j\}, \\ 2 \begin{pmatrix} x_j - x_i \\ y_j - y_i \end{pmatrix} & \text{si } k = i, \\ -2 \begin{pmatrix} x_j - x_i \\ y_j - y_i \end{pmatrix} & \text{si } k = j. \end{cases} \quad (2.6)$$

Donc, en notant

$$G_{i,j}(X) = \left[\left(\begin{array}{c} 0 \\ 0 \end{array} \right), \left(\begin{array}{c} 0 \\ 0 \end{array} \right), \dots, \underbrace{\left(\begin{array}{c} x_j - x_i \\ y_j - y_i \end{array} \right), \left(\begin{array}{c} 0 \\ 0 \end{array} \right), \dots, \left(\begin{array}{c} 0 \\ 0 \end{array} \right)}_{\text{position } i}, \underbrace{\left(\begin{array}{c} -x_j + x_i \\ -y_j + y_i \end{array} \right), \left(\begin{array}{c} 0 \\ 0 \end{array} \right), \dots, \left(\begin{array}{c} 0 \\ 0 \end{array} \right)}_{\text{position } j} \right], \quad (2.7)$$

nous obtenons

$$\nabla \phi_{i,j}(X) = 2G_{i,j}(X).$$

Ainsi

$$\nabla \left(\sum_{1 \leq i < j \leq N} \lambda_{i,j} \phi_{i,j}(X) \right) = \sum_{1 \leq i < j \leq N} \lambda_{i,j} \nabla \phi_{i,j}(X) = \sum_{1 \leq i < j \leq N} 2\lambda_{i,j} G_{i,j}(X), \quad (2.8)$$

et cela donne donc

$$\nabla \mathcal{L}(X, \lambda) = 2 \left(X + \sum_{1 \leq i < j \leq N} \lambda_{i,j} G_{i,j}(X) \right). \quad (2.9)$$

Grâce à la forme de $G_{i,j}$, cette quantité peut être simplifiée.

Simplification du terme $\sum_{1 \leq i < j \leq N} \lambda_{i,j} G_{i,j}(X)$. Nous représentons cette quantité sous la forme d'une matrice de taille $2 \times N$, notée A de coefficient générique $[a_{\ell,k}]_{1 \leq \ell \leq 2, 1 \leq k \leq N}$. Soit $k \in \llbracket 1; N \rrbracket$; nous calculons maintenant les termes $a_{1,k}$ et $a_{2,k}$ donnés par :

$$a_{1,k} = \left[\sum_{1 \leq i < j \leq N} \lambda_{i,j} G_{i,j}(X) \right]_{1,k} \quad \text{et} \quad a_{2,k} = \left[\sum_{1 \leq i < j \leq N} \lambda_{i,j} G_{i,j}(X) \right]_{2,k}.$$

En vertu de l'expression de $G_{i,j}$ en équation (2.7) page 43, les seuls facteurs non nuls dans cette somme sur $\{i, j | 1 \leq i < j \leq N\}$ sont les termes en $i = k$ et $j = k$. Ils valent respectivement

$$\sum_{j=k+1}^N \lambda_{k,j} (x_j - x_k) \quad (2.10)$$

et

$$\sum_{i=1}^{k-1} \lambda_{i,k} (-x_k + x_i), \quad (2.11)$$

où, par convention, les sommes sont nulles si l'indice de départ est strictement plus grand que l'indice de fin. Quitte à définir les multiplicateurs λ non plus comme un vecteur de taille M mais comme la partie strictement supérieure d'une matrice de taille $N \times N$, nous pouvons étendre ces multiplicateurs, de $(\lambda_{i,j})_{1 \leq i < j \leq N}$ à $(\lambda_{i,j})_{1 \leq i, j \leq N}$ selon :

$$\begin{cases} \lambda_{i,j} = \lambda_{j,i} & \text{si } i > j, \\ \lambda_{i,j} = 0 & \text{si } i = j. \end{cases} \quad (2.12)$$

Cela permet de réécrire l'équation (2.11) sous la forme

$$\sum_{j=1}^{k-1} \lambda_{k,j} (x_j - x_k),$$

d'où nous déduisons

$$a_{1,k} = \sum_{j=k+1}^N \lambda_{k,j} (x_j - x_k) + \sum_{j=1}^{k-1} \lambda_{k,j} (x_j - x_k).$$

Les termes de ces deux sommes étant nuls pour l'indice $j = k$, nous obtenons donc

$$a_{1,k} = \sum_{j=1}^N \lambda_{k,j} (x_j - x_k). \quad (2.13)$$

Par analogie, nous avons également

$$a_{2,k} = \sum_{j=1}^N \lambda_{k,j} (y_j - y_k). \quad (2.14)$$

Retour à $\nabla \mathcal{L}$ Ainsi, l'expression de $\nabla \mathcal{L}(X, \lambda)$ fournie en équation (2.9) page 43 revient à :

$$\nabla \mathcal{L}(X, \lambda) = 2 \left[\begin{pmatrix} x_1 + a_{1,1} \\ y_1 + a_{2,1} \end{pmatrix}, \dots, \begin{pmatrix} x_k + a_{1,k} \\ y_k + a_{2,k} \end{pmatrix}, \dots, \begin{pmatrix} x_N + a_{1,N} \\ y_N + a_{2,N} \end{pmatrix} \right]. \quad (2.15)$$

L'itération de l'algorithme exprimée en équation (2.5a) page 42 devient alors, pour tout $k \in \llbracket 1; N \rrbracket$,

$$X_k^{(p+1)} = X_k^{(p)} - 2\rho \left(X_k^{(p)} + \sum_{j=1}^N \lambda_{k,j}^{(p)} (X_j^{(p)} - X_k^{(p)}) \right), \quad (2.16)$$

que nous écrivons sous la forme suivante, pour tout $k \in \llbracket 1; N \rrbracket$,

$$X_k^{(p+1)} = (1 - 2\rho) X_k^{(p)} + 2\rho \sum_{j=1}^N \lambda_{k,j}^{(p)} (X_k^{(p)} - X_j^{(p)}). \quad (2.17)$$

Critère d'arrêt et signification des paramètres numériques ρ et μ

Un critère d'arrêt basé sur le contrôle de l'incrément $\|X^{(p+1)} - X^{(p)}\|$ ne donne pas de résultat satisfaisant. Une explication possible est la suivante : étant donné une configuration d'équilibre des cellules, effectuer une même rotation pour toutes les cellules, d'un angle aussi petit que l'on veut, engendre toujours une situation d'équilibre ; or cela amène à des variations des positions $X^{(p+1)}$ et donc à un incrément qui reste élevé, alors que nous sommes sur une position d'équilibre. Pour pallier cela, nous choisissons un critère d'arrêt découplé en deux parties : une sur W et une sur les contraintes $\phi_{i,j}$. Posons, pour $p \geq 0$:

$$\varepsilon_{\phi}^{(p+1)} = \max_{\substack{1 \leq k \leq N \\ 1 \leq \ell \leq k-1}} \left(\phi_{\ell,k}^{(p+1)} \right), \quad (2.18)$$

$$W^{(p+1)} = \sum_{k=1}^N \left[\left(x_k^{(p+1)} \right)^2 + \left(y_k^{(p+1)} \right)^2 \right], \quad (2.19)$$

$$\varepsilon_W^{(p+1)} = \left| \frac{W^{(p+1)} - W^{(p)}}{W^{(p)}} \right|, \quad (2.20)$$

et notons tol_W et tol_{ϕ} deux petites valeurs numériques (les tolérances du critère d'arrêt). Alors le critère d'arrêt consiste à avoir

$$\varepsilon_W^{(p)} < \text{tol}_W \quad \text{et} \quad \varepsilon_{\phi}^{(p)} < \text{tol}_{\phi}. \quad (2.21)$$

Ainsi, nous contrôlons le recouvrement maximal entre deux cellules (par l'inégalité sur tol_ϕ) et la faible variation relative de W (par l'inégalité sur tol_W).

Concernant le paramètre numérique ρ , il est lié au pas de déplacement élémentaire des cellules entre deux itérations de l'algorithme. En effet, l'itération exprimée en équation (2.5a) page 42 s'apparente à une méthode de gradient à pas fixe, de pas ρ .

Quant au paramètre μ , il est lié à la rapidité à laquelle les multiplicateurs de Lagrange λ sont mis à jour lorsque les contraintes deviennent violées ou satisfaites. En effet, si deux cellules C_i et C_j se recouvrent alors $\phi_{i,j}^{(p)} > 0$; donc, en vertu de l'équation (2.5c) page 42, $\lambda_{i,j}^{(p)}$ augmentera d'autant plus que μ est grand; ensuite, la non-nullité de $\lambda_{i,j}^{(p)}$ va avoir un impact sur le calcul de $X^{(p+1)}$.

Les valeurs attribuées à ces deux paramètres sont discutées en section 2.2.2 page suivante et résumées en tableau 2.3 page 61.

Bilan de l'algorithme de positionnement

L'algorithme basé sur les équations (2.5a) à (2.5c) page 42 qui résulte de ces différents calculs est l'algorithme 2.1.

Algorithme 2.1 – Positionnement

Étant donné les pas ρ et μ ; étant donné les tolérances du critère d'arrêt tol_W et tol_ϕ ; étant donné, pour tout $k \in \llbracket 1; N \rrbracket$, les valeurs initiales $X_k^{(0)}$ et $\lambda_k^{(0)}$; étant donné les rayons R .
Tant que le critère d'arrêt

$$\varepsilon_W^{(p)} < \text{tol}_W \quad \text{et} \quad \varepsilon_\phi^{(p)} < \text{tol}_\phi, \quad (2.22)$$

issu de l'équation (2.21) page 45, est non vérifié, boucle sur $p \geq 0$ contenant :

1. Calcul des positions $X^{(p+1)}$, selon les équations (2.5a) et (2.17) page 42 et page 45 : pour tout $k \in \llbracket 1; N \rrbracket$,

$$\begin{cases} x_k^{(p+1)} = (1 - 2\rho)x_k^{(p)} + 2\rho \sum_{j=1}^N \lambda_{k,j}^{(p)} (x_k^{(p)} - x_j^{(p)}), \\ y_k^{(p+1)} = (1 - 2\rho)y_k^{(p)} + 2\rho \sum_{j=1}^N \lambda_{k,j}^{(p)} (y_k^{(p)} - y_j^{(p)}). \end{cases} \quad (2.23)$$

2. Calcul des contraintes, selon l'équation (2.5b) page 42 : pour tous $k \in \llbracket 1; N \rrbracket$ et $\ell \in \llbracket 1; k-1 \rrbracket$,

$$\phi_{\ell,k}^{(p+1)} = (R_\ell + R_k)^2 - \left(x_\ell^{(p+1)} - x_k^{(p+1)} \right)^2 - \left(y_\ell^{(p+1)} - y_k^{(p+1)} \right)^2. \quad (2.24)$$

3. Calcul des multiplicateurs, selon l'équation (2.5c) page 42 : pour tous $k \in \llbracket 1; N \rrbracket$ et $\ell \in \llbracket 1; k-1 \rrbracket$,

$$\lambda_{\ell,k}^{(p+1)} = \max\left(0, \lambda_{\ell,k}^{(p)} + \mu \phi_{\ell,k}^{(p+1)}\right). \quad (2.25)$$

4. Calcul des quantités contrôlant l'arrêt, selon les équations (2.18) à (2.20) page 45 :

$$\begin{cases} \varepsilon_{\phi}^{(p+1)} = \max_{\substack{1 \leq k \leq N \\ 1 \leq \ell \leq k-1}} \left(\phi_{\ell,k}^{(p+1)} \right), \\ W^{(p+1)} = \sum_{k=1}^N \left[\left(x_k^{(p+1)} \right)^2 + \left(y_k^{(p+1)} \right)^2 \right], \\ \varepsilon_W^{(p+1)} = \left| \frac{W^{(p+1)} - W^{(p)}}{W^{(p)}} \right|. \end{cases} \quad (2.26)$$

Remarque 2.1. La quantité $\lambda^{(p)}$ est stockée dans une matrice à diagonale nulle que nous symétrisons, comme indiqué en équation (2.12) page 44, de sorte que l'élément $\lambda_{\ell,k}^{(p)}$ existe pour tout $(k, \ell) \in \llbracket 1; N \rrbracket^2$. Cela simplifie les notations et l'implémentation.

2.2.2 Discussion sur les valeurs attribuées aux pas de l'algorithme

Les valeurs des deux pas ρ et μ , à priori liées entre-elles, influent sur la convergence et sur la vitesse de convergence de l'algorithme. De leurs valeurs dépendent aussi les valeurs des tolérances du critère d'arrêt.

Le choix des valeurs des paramètres s'est fait selon le principe d'essai-erreur : de nombreux tests ont été réalisés et ont consisté à calibrer les *meilleures* valeurs de ρ et μ . Ici, *meilleures* signifie : assez grandes pour éviter un temps de calcul trop élevé mais assez petites pour éviter des déplacements trop importants de cellules². Ce calibrage s'est fait en fonction de N , par paliers. Néanmoins, nous ne sommes pas partis de valeurs quelconques pour ρ , mais de $\rho = \mathcal{O}\left(\frac{1}{N^2}\right)$: ce choix de départ a été guidé par l'analyse suivante, qui consiste à établir la matrice d'itération de la méthode itérative et à regarder la localisation de ses valeurs propres.

Pour tout $k \in \llbracket 1; N \rrbracket$, en vertu de l'équation (2.17) page 45, la coordonnée d'abscisse $x_k^{(p+1)}$ de $X_k^{(p+1)}$ s'écrit :

$$x_k^{(p+1)} = (1 - 2\rho)x_k^{(p)} + 2\rho \sum_{j=1}^N \lambda_{k,j}^{(p)} \left(x_k^{(p)} - x_j^{(p)} \right),$$

2. Les conséquences de ce phénomène sont discutées en section 2.2.3 page 50.

c'est-à-dire

$$x_k^{(p+1)} = \left(1 - 2\rho + 2\rho \sum_{j=1}^N \lambda_{k,j}^{(p)} \right) x_k^{(p)} - 2\rho \sum_{j=1}^N \lambda_{k,j}^{(p)} x_j^{(p)}.$$

Celle-ci peut alors se mettre sous la forme matricielle

$$x^{(p+1)} = \mathcal{I}^{(p)} x^{(p)},$$

en notant $x^{(p)}$ le vecteur des abscisses suivant

$$x^{(p)} = \left(x_1^{(p)}, x_2^{(p)}, \dots, x_N^{(p)} \right)$$

et $\mathcal{I}^{(p)}$ la matrice d'itération symétrique suivante

$$\mathcal{I}^{(p)} = \begin{pmatrix} 1 - 2\rho \left(1 - \Lambda_1^{(p)} \right) & -2\rho \lambda_{1,2}^{(p)} & \cdots & -2\rho \lambda_{1,k}^{(p)} & \cdots & -2\rho \lambda_{1,N}^{(p)} \\ & 1 - 2\rho \left(1 - \Lambda_2^{(p)} \right) & & & & \vdots \\ & & \ddots & & & -2\rho \lambda_{k,N}^{(p)} \\ & & & 1 - 2\rho \left(1 - \Lambda_k^{(p)} \right) & & \vdots \\ & & & & \ddots & -2\rho \lambda_{N-1,N}^{(p)} \\ & & & & & 1 - 2\rho \left(1 - \Lambda_N^{(p)} \right) \end{pmatrix},$$

où, pour tout $k \in \llbracket 1; N \rrbracket$,

$$\Lambda_k^{(p)} = \sum_{j=1}^N \lambda_{k,j}^{(p)}. \quad (2.27)$$

L'utilisation du théorème de Gerschgorin [AD08], qui affirme que toute valeur propre d'une matrice générique $A = [a_{i,j}]_{i,j}$ est située dans l'un au moins des disques de Gerschgorin D_k , définis par

$$D_k = \left\{ z \in \mathbb{C} \left| |z - a_{k,k}| \leq \sum_{j \neq k} |a_{k,j}| \right. \right\}, \quad (2.28)$$

permet de borner les valeurs propres de $\mathcal{I}^{(p)}$. En effet, le centre de chaque D_k est

$$[\mathcal{I}^{(p)}]_{k,k} = 1 - 2\rho + 2\rho \Lambda_k^{(p)} \in \mathbb{R},$$

donc situé sur l'axe réel à droite du point d'abscisse $1 - 2\rho$; et chaque disque passe par ce point $(1 - 2\rho, 0)$, car

$$|(1 - 2\rho) - [\mathcal{I}^{(p)}]_{k,k}| = |2\rho \Lambda_k^{(p)}| = \sum_{j \neq k} [\mathcal{I}^{(p)}]_{k,j}.$$

Il en résulte que la plus grande valeur réelle située dans un disque D_k est

$$1 - 2\rho + 2\left(2\rho\Lambda_k^{(p)}\right),$$

donc toutes les valeurs propres sont inférieures à

$$1 - 2\rho + 4\rho \max_k \left(\Lambda_k^{(p)}\right).$$

Donc, par définition de $\Lambda_k^{(p)}$ et en notant $\lambda_{\max}^{(p)}$ le plus grand coefficient de la matrice $\lambda^{(p)}$, nous avons le majorant suivant pour les valeurs propres :

$$B(\rho, N) \stackrel{\text{déf}}{=} 1 - 2\rho + 4\rho N \lambda_{\max}^{(p)}. \quad (2.29)$$

Notre choix de ρ est dorénavant guidé par le principe suivant : nous cherchons à avoir des valeurs propres plus petites que 1, et donc à avoir un choix de ρ qui rende $B(\rho, N)$ plus petit que 1 (ou le plus proche possible, s'il est plus grand). Considérons deux choix pour ρ :

- posons $\rho_1 = \frac{1}{4\lambda_{\max}^{(p)}N}$, alors $B(\rho_1, N) = 1 - \frac{1}{2\lambda_{\max}^{(p)}N} + 1 = 2 - \frac{1}{2\lambda_{\max}^{(p)}N}$;
- posons $\rho_2 = \frac{1}{4\lambda_{\max}^{(p)}N^2}$, alors $B(\rho_2, N) = 1 - \frac{1}{2\lambda_{\max}^{(p)}N^2} + \frac{1}{N}$.

Il est difficile, sans information supplémentaire (en particulier sur $\lambda_{\max}^{(p)}$), de décider quel choix parmi ρ_1 et ρ_2 est le plus efficace ; néanmoins, le choix ρ_1 amène à un majorant $B(\rho_1, N)$ pour les valeurs propres proche de 2. Après quelques essais numériques, nous avons constaté que les valeurs propres étaient proches de 1,5 pour ρ_1 et proches de 1 (par valeurs supérieures) pour ρ_2 : cela nous a convaincu de choisir

$$\rho = \mathcal{O}\left(\frac{1}{N^2}\right). \quad (2.30)$$

De nombreuses exécutions de l'algorithme ont confirmé, au moins pour $N \in \llbracket 50; 500 \rrbracket$ (ce qui est la zone de variation de N dans laquelle notre étude, au chapitre 3, sera faite), que ce choix pour ρ était convenable.

Remarque 2.2. Par ailleurs, en prenant $\rho < \frac{1}{2}$, nous avons toutes les valeurs propres strictement positives, ce qui est une condition nécessaire de convergence des méthodes de gradient à pas fixe.

Au final, voici les valeurs que nous avons utilisées :

$$\begin{cases} \rho = 3 \times 10^{-4} & \text{pour } 1 \leq N < 100, \\ \rho = 3 \times 10^{-5} & \text{pour } 100 \leq N < 300, \\ \rho = 6 \times 10^{-6} & \text{pour } 300 \leq N \leq 500, \end{cases} \quad (2.31)$$

et

$$\mu = 100.$$

2.2.3 Améliorations autour de l'algorithme de positionnement

En l'état, et après plusieurs exécutions, l'algorithme de positionnement appliqué tel quel nous confronte à différentes difficultés.

Difficulté n° 1 Cette difficulté concerne le temps de calcul qui peut être assez long pour atteindre la précision souhaitée.

Difficulté n° 2 Cette difficulté est liée à l'interprétation des résultats de notre algorithme en terme de modélisation : lorsque des contraintes sont fortement violées, l'algorithme va fortement éloigner toutes les cellules environnantes pour faire de la place puis va ensuite les regrouper. Cependant, deux cellules proches avant l'éloignement ne le seront pas forcément après le regroupement : ce comportement de l'algorithme, même s'il donne bien au final une situation d'équilibre telle que définie dans nos principes de modélisation, n'est absolument pas satisfaisant. En effet, il conduit à une réorganisation locale des cellules, qui correspond à un déplacement discontinu de ces cellules, ce qui n'est pas du tout souhaité. Nous appelons « phénomène d'éjection » ce comportement de l'algorithme.

Réponses à la difficulté n° 1 du temps de calcul Une première réponse, sur le registre du langage de programmation, consiste à migrer le code de Matlab à FORTRAN ; cela a permis de réduire considérablement le temps de calcul. Une deuxième réponse, algorithmique, consiste à appliquer deux fois de suite l'algorithme de positionnement, selon le raisonnement suivant. L'algorithme de positionnement cherche une position d'équilibre parmi toutes celles qui existent : plutôt que d'exiger de cet algorithme de trouver précisément, du premier coup, une position d'équilibre³, nous procédons en deux temps. Ainsi :

1. Nous appliquons l'algorithme de positionnement avec une certaine valeur des pas ρ et μ , qui ne permet pas d'atteindre la précision souhaitée ; relâcher ainsi la valeur des paramètres permet à l'algorithme de s'approcher rapidement d'une position d'équilibre.
2. Ensuite, nous diminuons la valeur du pas ρ et ré-appliquons l'algorithme de positionnement ; cette fois-ci, dans l'objectif d'atteindre précisément une situation d'équilibre de laquelle nous nous étions approchés précédemment.

C'est ce que nous appelons la « dynamique de positionnement » dans la suite.

Réponse à la difficulté n° 2 du « phénomène d'éjection » : l'algorithme de retour Ce phénomène d'éjection est lié à de trop forts recouvrements de cellules ; un problème similaire est présent dans [SM05], où les trop forts recouvrements provoqués par des divisions engendrent une instabilité numérique et sont compensés par un contrôle adaptatif de la taille du pas. Choisir de manière systématique des pas plus petits n'est pas une solution viable, puisqu'elle amplifierait la difficulté n° 1 et court-circuiterait

3. Cela consisterait à appliquer une fois l'algorithme de positionnement avec des paramètres numériques très fins.

le choix des valeurs effectué en section 2.2.2. Nous décidons donc d'implémenter un algorithme supplémentaire *après* la dynamique de positionnement, qui va détecter la présence d'un éventuel trop grand déplacement : le cas échéant, nous l'annulons et nous diminuons le pas ρ , puis nous relançons la dynamique de positionnement. Cet algorithme permet, localement et automatiquement, de diminuer le pas contrôlant le déplacement des cellules en cas de nécessité. Il est décrit en algorithme 2.2, où nous notons $\text{tol}_X > 0$ la valeur seuil du déplacement qui définit ce trop grand déplacement.

Algorithme 2.2 – Retour

Étant donné la position actuelle des cellules et une position précédente, respectivement notées X et \bar{X} ; étant donné le pas ρ et la tolérance du critère d'arrêt tol_W ; étant donné la tolérance de déplacement tol_X :

1. Pour chaque cellule, calcul de la distance parcourue entre les positions X et \bar{X} : pour tout $k \in \llbracket 1 ; N \rrbracket$,

$$d_k = \sqrt{(x_k - \bar{x}_k)^2 + (y_k - \bar{y}_k)^2}. \quad (2.32)$$

2. Évaluation du booléen « $\exists k_0 \in \llbracket 1 ; N \rrbracket, d_{k_0} > \text{tol}_X$ » puis, s'il est vrai :
 - (a) diminution par deux de ρ et de tol_W ;
 - (b) remplacement de X par \bar{X} ;
 - (c) renvoi par l'algorithme de « retour nécessaire ».

2.2.4 Bilan de l'étape de positionnement

Au final, la partie de l'algorithme implémentant le positionnement possède la structure suivante :

Synthèse de l'étape de positionnement

1. Choix ^a des pas ρ et μ et des seuils de tolérance tol_W , tol_ϕ et tol_X .
2. Dynamique de positionnement :
 - (a) algorithme 2.1 de positionnement ;
 - (b) diminution par deux du pas ρ et du seuil de tolérance tol_W de l'algorithme de positionnement ;
 - (c) algorithme 2.1 de positionnement.
3. Algorithme 2.2 de retour (entre les positions actuelles et celles juste avant l'item n° 2) : s'il renvoie « retour nécessaire », alors retour à l'item n° 2.

^a. Voir tableau 2.3 page 61.

2.3 Étape de croissance

L'objectif de l'étape de croissance est d'implémenter la règle de « Croissance linéaire en volume avec bruit », décrite en section 1.3.2 page 29.

L'algorithme de croissance consiste à augmenter le rayon de chaque cellule d'une quantité variable autour d'une valeur moyenne. Cette valeur moyenne est calculée de sorte que le rayon évolue de R_{\min} à R_{\max} en une durée T_G ; quant à la variabilité, elle est contrôlée par une loi uniforme. L'algorithme qui en résulte est l'algorithme 2.3.

Algorithme 2.3 – Croissance

Étant donné les paramètres R_{\min} , R_{\max} , γ et T_G de modélisation ; étant donné le pas de temps dt . Boucle sur les N cellules contenant, pour tout $k \in \llbracket 1 ; N \rrbracket$:

1. Tirage d'un nombre aléatoire u suivant une loi uniforme sur $[-\gamma ; \gamma]$.
2. Changement du rayon R_k de la cellule C_k , selon l'équation (1.11) page 29, en :

$$\left(R_k^3 + (1 + u) \frac{R_{\max}^3 - R_{\min}^3}{T_G} dt \right)^{1/3}. \quad (2.33)$$

L'étape de croissance est uniquement constituée de l'algorithme de croissance. Il en résulte :

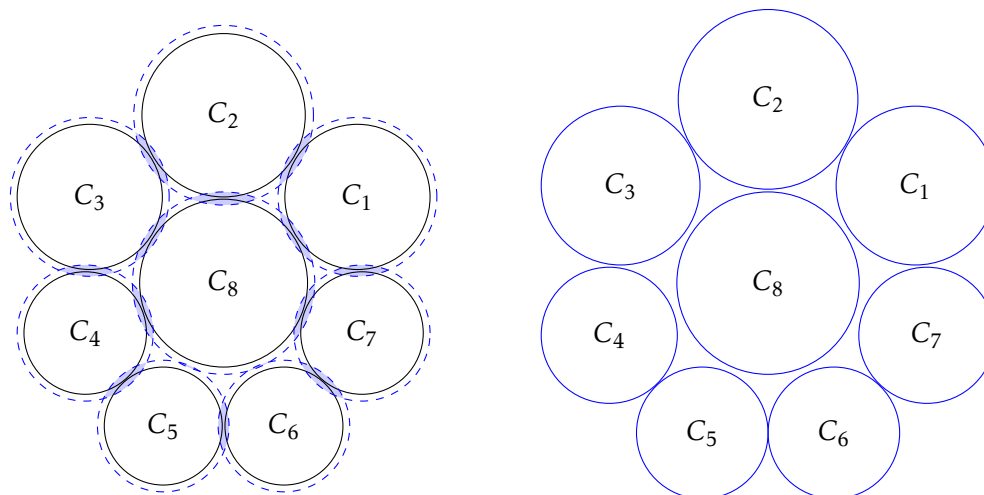
- un nécessaire recouvrement des cellules entre-elles (illustré en figure 2.2 page ci-contre), qui explique le recours à l'étape de positionnement après l'étape de croissance ;
- un éventuel besoin de division, qui explique la présence du test de division après l'étape de croissance, éventuellement suivi de l'étape de division.

2.4 Étape de division d'une cellule

L'objectif de cette section est de décrire l'étape de division telle que définie dans la structure générale de l'algorithme en section 2.1 page 40. Il s'agit donc, au vu des principes de modélisation de la loi de division introduits en section 1.4.1 page 30, d'implémenter le processus de division d'une cellule ayant *préalablement* satisfait la condition de division.

La structure de l'étape de division est illustrée par un schéma en figure 2.3 page 54 ; les rectangles arrondis rouges désignent les actions clés de l'étape de division qui sont discutées dans la suite.

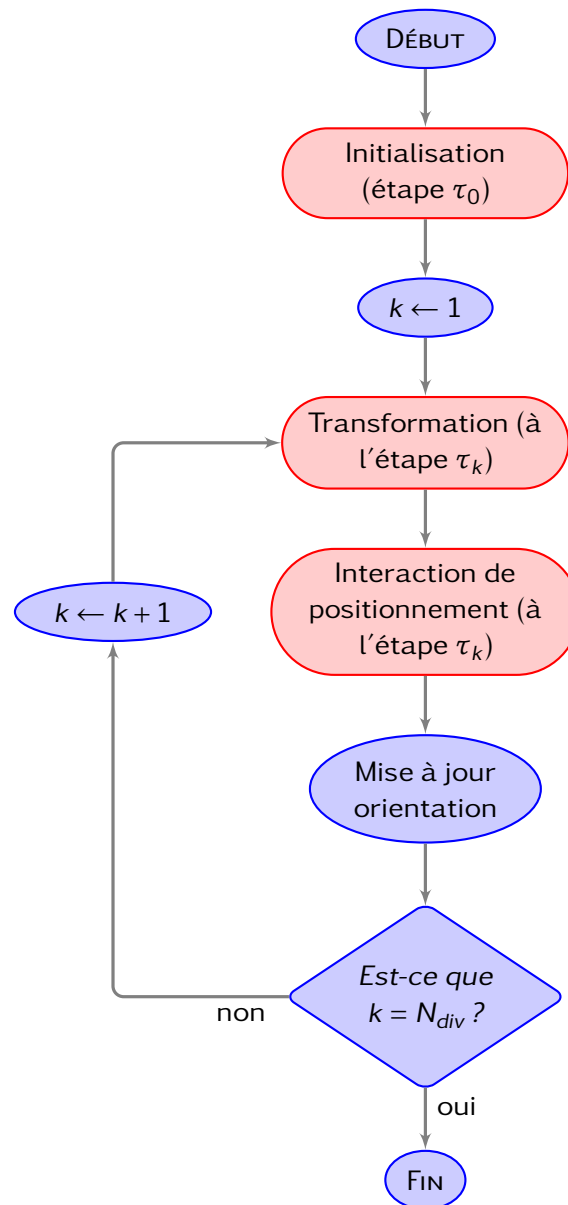
Nous supposons donc être au pas de temps t et avoir une cellule C_{i_0} qui est prête à enclencher sa division, c'est-à-dire telle que $R_{i_0}(t) \geq R_{\max}$. L'étape de division adopte les principes et choix suivants, introduits dans les principes de modélisation au chapitre 1 et adaptés ici.



Légende. Cette illustration graphique décrit ce qui se passe lors d'un pas de temps, en l'absence de division. Nous considérons huit cellules C_1, C_2, \dots, C_8 de taille différente. Nous expliquons schématiquement ce qui se produit lors des étapes de croissance et de positionnement au temps t :

1. En trait noir (à gauche) est représentée une configuration au début du temps t : les cellules sont alors toutes positionnées sans recouvrement.
2. En tiret bleu (à gauche), il s'agit de l'effet de l'étape de croissance : les rayons ont été augmentés. Il en résulte un recouvrement plus ou moins important entre toutes les cellules voisines, représenté par un remplissage bleu : il doit être corrigé.
3. En trait bleu (à droite), il s'agit d'une configuration des cellules après l'étape de positionnement, c'est-à-dire à la fin du temps t . La configuration spatiale des cellules a été revue (via l'algorithme 2.1 de positionnement) pour qu'il n'y ait plus de recouvrement.

FIGURE 2.2 – Illustration de la succession de la croissance et du positionnement




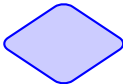

Légende.  est une action clé du processus de division ;  est un test ;  est une action mineure.

FIGURE 2.3 – Vision schématique de l'étape de division

- Nous choisissons pour *toute la durée* de la simulation le même nombre d'étapes N_{div} dans le processus de division.
- Nous choisissons, pour *toute la durée* de la simulation et pour *toutes les cellules*, la même direction d'orientation D_{div} (entier valant 1, 2, 3 ou 4, respectivement pour une direction aléatoire, fixe, radiale ou tangentielle).
- Nous choisissons, pour *toute la durée* de la simulation et pour *toutes les cellules*, la même stratégie d'orientation S_{div} (entier valant 1 ou 2 respectivement pour une stratégie libre ou contrainte) du processus de division.
- Dès que la cellule indicée i_0 commence sa division, alors la valeur de N est incrémentée et cette cellule C_{i_0} disparaît au profit de deux nouvelles cellules, indicées i_0 et N . Il s'agit ici d'un choix de programmation.

Remarque 2.3 (À propos de ce choix de programmation). Un autre choix consisterait à incrémenter N de deux unités et à indexer $N - 1$ et N les deux cellules filles (cela permettrait, en cas de besoin, de garder des informations sur la cellule C_{i_0}). Néanmoins N ne serait alors plus le nombre total de cellules existantes : cet autre choix n'est donc pas retenu.

2.4.1 Processus de division à l'étape τ_k pour k fixé

Comme défini dans le chapitre 1, le processus de division consiste en N_{div} étapes, indicées par τ_k pour $k \in \llbracket 1; N_{\text{div}} \rrbracket$ et paramétrées par $h(\tau_k)$ (donné en équation (1.17) page 33). À chacune de ces étapes, les deux cellules filles effectuent les deux actions clés suivantes, illustrées en figure 2.4 page suivante.

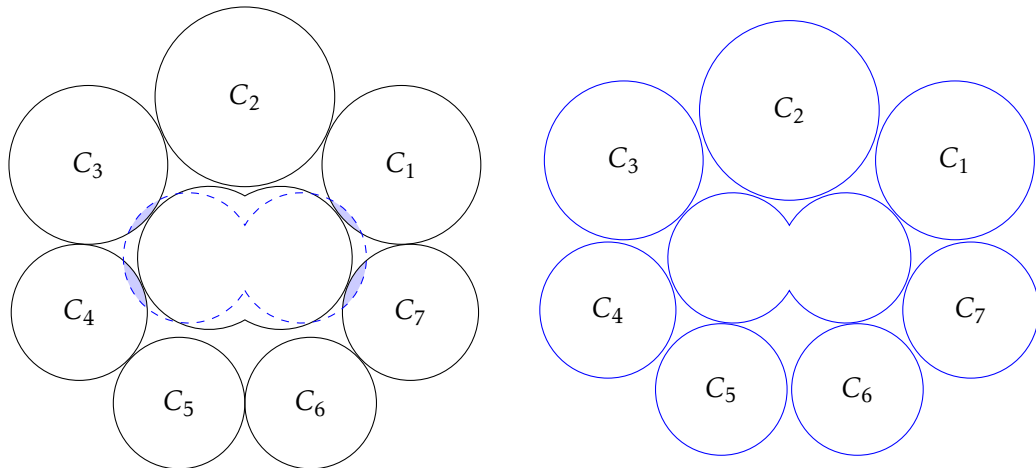
Action 1 : transformation. Elles font un pas de plus dans leur transformation, ce qui les amène à recouvrir partiellement leurs voisines.

Action 2 : interaction de positionnement. Elles interagissent avec les cellules environnantes pour corriger ce recouvrement, via un nouvel algorithme de positionnement.

Détaillons chacune de ces deux actions, en lien avec les principes de modélisation définis en sections 1.4.2 et 1.4.3 page 32 et page 35.

1. L'action n° 1 consiste à calculer le rayon et la position des cellules filles, de sorte qu'elles forment « un 8 » parfaitement symétrique et qu'il y ait conservation du volume ; les formules sont fournies en équations (1.17), (1.22), (1.24) et (1.25) pages 33–35.
2. L'action n° 2 consiste à corriger le recouvrement partiel dû à l'action n° 1 : elle permet à notre système d'être en position d'équilibre⁴. Comme mis en place dans les principes de modélisation, deux stratégies d'interaction sont proposées pour cette 2^e action ; le choix de la stratégie est codé dans la valeur de S_{div} . Voici les deux possibilités :

4. Au sens défini dans la modélisation, c'est-à-dire un regroupement des cellules *sans recouvrement*.



Légende. Cette illustration graphique décrit ce qui se passe au cours du processus de division d'une cellule. Nous considérons sept cellules C_1, C_2, \dots, C_7 de taille différente, qui sont autour d'une cellule ayant entamé son processus de division. Nous expliquons schématiquement ce qui se produit lors du passage d'une étape τ_{k-1} à une étape τ_k du processus de division :

1. En trait noir (à gauche) est représentée une configuration à la fin de l'étape τ_{k-1} : les cellules sont alors toutes positionnées sans recouvrement.
2. En tiret bleu (à gauche), il s'agit de la nouvelle forme des cellules naissantes juste après la 1^{re} action (dite « transformation ») de l'étape τ_k . Il en résulte un recouvrement avec certaines cellules voisines (C_1, C_3, C_4 et C_7), représenté par un remplissage bleu : il doit être corrigé.
3. En trait bleu (à droite), il s'agit d'une configuration des cellules après la 2^e action (dite « interaction de positionnement ») de l'étape τ_k . Les cellules naissantes n'ont pas changé de forme mais la configuration spatiale des cellules a été revue (via un nouvel algorithme de positionnement dit « de positionnement en cours de division ») pour qu'il n'y ait plus de recouvrement.

FIGURE 2.4 – Illustration de la succession des deux actions du processus de division

Stratégie d'orientation libre. L'orientation est libre d'évoluer lors de l'interaction de positionnement ; ω_{i_0} est alors la valeur de l'orientation au début de l'étape τ_1 du processus de division, uniquement.

Stratégie d'orientation contrainte. L'orientation est conservée lors de l'interaction de positionnement ; ω_{i_0} est alors la valeur de l'orientation tout au long du processus de division.

Cette action consiste donc à appliquer un nouvel algorithme de positionnement : il s'agit de l'algorithme 2.4 dit « de positionnement en cours de division ». La différence avec l'algorithme 2.1 page 46 de positionnement classique est qu'il doit dorénavant assurer le regroupement sans interpénétration de $N - 2$ cellules circulaires et de deux cellules naissantes en forme « de 8 ».

Algorithme 2.4 – Positionnement en cours de division

Étant donné les pas ρ et μ ; étant donné les tolérances du critère d'arrêt tol_W et tol_ϕ ; étant donné, pour tout $k \in \llbracket 1; N \rrbracket$, les valeurs initiales $X_k^{(0)}$ et $\lambda_k^{(0)}$; étant donné les rayons R ; étant donné la stratégie de division S_{div} ; étant donné le paramétrage h et l'orientation ω .

Tant que le critère d'arrêt, donné en équation (2.22) page 46, est non vérifié, boucle sur $p \geq 0$ contenant :

1. Calcul des positions $X^{(p+1)}$:
 - (a) pour tout k , $X_k^{(p+1)}$ est donné par l'équation (2.23) page 46 ;
 - (b) si la stratégie est contrainte ($S_{\text{div}} = 2$), alors les positions des deux cellules concernées sont modifiées pour tenir compte de la nouvelle contrainte exprimée en équation (1.32) page 37 ; cela donne :

$$\begin{aligned} x_{i_0}^{(p+1)} &= x_{i_0}^{(p+1)} - \rho(\lambda_1 - \lambda_2) \sin(\omega), \\ x_N^{(p+1)} &= x_N^{(p+1)} + \rho(\lambda_1 + \lambda_2) \sin(\omega), \\ y_{i_0}^{(p+1)} &= y_{i_0}^{(p+1)} + \rho(\lambda_1 - \lambda_2) \cos(\omega), \\ y_N^{(p+1)} &= y_N^{(p+1)} - \rho(\lambda_1 - \lambda_2) \cos(\omega). \end{aligned}$$

2. Calcul des contraintes : il s'agit de permettre aux centres des deux cellules naissantes d'être éloignés de $2(R_{i_0} - h)$ plutôt que de $R_{i_0} + R_N$ (comme illustré en figure 1.2 page 33) :

- (a) pour $\ell = i_0$ et $k = N$ alors

$$\phi_{\ell,k}^{(p+1)} = \left(2(R_{i_0} - h)\right)^2 - \left(x_\ell^{(p+1)} - x_k^{(p+1)}\right)^2 - \left(y_\ell^{(p+1)} - y_k^{(p+1)}\right)^2; \quad (2.34)$$

- (b) pour les autres valeurs de ℓ et k , $\phi_{\ell,k}^{(p+1)}$ est donné en équation (2.24)

page 46 ;

- (c) si la stratégie est contrainte ($S_{\text{div}} = 2$), alors la nouvelle contrainte égalité est exprimée à l'aide de deux contraintes inégalités :

$$\phi_1^{(p+1)} = (x_{i_0} - x_N) \sin(\omega) - (y_{i_0} - y_N) \cos(\omega), \quad (2.35)$$

$$\phi_2^{(p+1)} = -(x_{i_0} - x_N) \sin(\omega) + (y_{i_0} - y_N) \cos(\omega). \quad (2.36)$$

3. Calcul des multiplicateurs :

- (a) pour $\ell = i_0$ et $k = N$ alors $\lambda_{\ell,k}^{(p+1)} = \lambda_{\ell,k}^{(p)} + \mu \phi_{\ell,k}^{(p+1)}$;

- (b) pour les autres valeurs de ℓ et k , $\lambda_{\ell,k}^{(p+1)}$ est donné en équation (2.25) page 47 ;

- (c) si la stratégie est contrainte ($S_{\text{div}} = 2$), alors il y a deux autres multiplicateurs à traiter :

$$\lambda_1^{(p+1)} = \max(0, \lambda_1^{(p)} + \mu \phi_1^{(p+1)}), \quad (2.37)$$

$$\lambda_2^{(p+1)} = \max(0, \lambda_2^{(p)} + \mu \phi_2^{(p+1)}). \quad (2.38)$$

4. Calcul des quantités contrôlant l'arrêt, selon l'équation (2.26) page 47.

2.4.2 Initialisation du processus de division et synthèse de l'étape

L'algorithme d'initialisation consiste à implémenter l'étape τ_0 des principes de modélisation, c'est-à-dire à calculer l'orientation initiale de division et à créer la nouvelle cellule. Il s'agit de l'algorithme 2.5.

Algorithme 2.5 – Initialisation du processus de division

Étant donné la direction D_{div} de division :

1. Calcul de ω_{i_0} en fonction de la valeur de D_{div} .
2. Création de la nouvelle cellule :
 - (a) augmentation du nombre de cellules : N devient $N + 1$;
 - (b) calcul des caractéristiques (position, rayon, orientation, lignage) initiales des cellules naissantes, qui sont celles de la cellule mère ; elles sont fournies en équations (1.16) et (1.18) à (1.20) page 32 et page 33.

Au final, la partie de l'algorithme implémentant l'étape de division (pour une cellule satisfaisant la condition de division) possède la structure suivante :

Synthèse de l'étape de division d'une cellule

Étant donné le nombre d'étapes N_{div} , la direction D_{div} et la stratégie S_{div} de la division ^a ; étant donné une cellule qui a préalablement vérifié la condition de division :

1. Algorithme 2.5 d'initialisation du processus de division.
2. À chaque étape τ_k du processus de division (boucle sur $k \in \llbracket 1 ; N_{\text{div}} \rrbracket$) :
 - (a) transformation ;
 - (b) interaction de positionnement :
 - (i) algorithme 2.4 de positionnement en cours de division,
 - (ii) algorithme 2.2 de retour (entre les positions actuelles et celles juste avant l'item n° 2b) : s'il renvoie « retour nécessaire », alors retour à l'item n° 2b ;
 - (c) si la stratégie est libre ($S_{\text{div}} = 1$), alors mise à jour de l'orientation.

a. Voir tableaux 2.1 et 2.2 page suivante et page 61.

2.5 Configurations de l'algorithme

2.5.1 Étape d'initialisation

L'objectif de l'étape d'initialisation est de fournir les valeurs de départ des différentes caractéristiques des cellules. Il s'agit en particulier :

- de leurs positions, qui doivent respecter une configuration d'équilibre ; elle est déterminée par l'algorithme 2.1 page 46 de positionnement ;
- de leurs lignages, qui doivent être différents deux à deux (afin que chaque cellule initiale soit génératrice de son propre lignage). Soit N_0 le nombre de cellules initiales ; nous choisissons les lignages sous la forme

$$\forall i \in \llbracket 1 ; N_0 \rrbracket, \quad \ell_i = i. \quad (2.39)$$

Au final, la partie de l'algorithme implémentant l'étape d'initialisation possède la structure suivante :

Synthèse de l'étape d'initialisation

Étant donné le nombre initial de cellules N_0 :

1. Choix de positions et de rayons pour N_0 cellules.
2. Choix des lignages, selon l'équation (2.39).
3. Algorithme 2.1 de positionnement.

2.5.2 Valeurs numériques des paramètres

Nous regroupons ici l'ensemble des valeurs numériques données aux paramètres, en distinguant les paramètres de modélisation (tableau 2.1) des paramètres algorithmiques (tableau 2.3 page suivante). Les valeurs indiquées sont celles retenues pour la suite des simulations, en particulier dans les tests statistiques du chapitre 3. Quelques paramètres sont présentés à part (tableau 2.2 page ci-contre) : il s'agit de ceux que nous faisons varier selon les simulations.

Paramètre	Valeur	Signification
N_0	50	Nombre de cellules initiales
T_G	24	Durée moyenne du <i>cycle cellulaire</i>
T_{\max}	72	Durée de la simulation
γ	0,25	Bruit lors de la croissance
R_{\min}	1	Rayon minimal d'une cellule
R_{\max}	$\sqrt[3]{2}$	Valeur seuil du rayon (qui déclenche la division)
N_{div}	8	Nombre d'étapes dans le processus de division progressif

La valeur de N_0 correspond environ au nombre de cellules initiales dans les expériences biologiques présentées au chapitre 4. La valeur de T_G est choisie pour faire coïncider l'unité de temps de l'algorithme avec la grandeur « heure » du monde réel : en effet, les cellules utilisées lors des expériences ont un cycle d'une durée moyenne de 24 h. La durée T_{\max} est alors choisie égale à trois fois cette durée T_G , puisque les expériences ont consisté à observer les populations cellulaires sur trois temps de génération. La durée du cycle des cellules pouvant varier de 18 h à 30 h, γ est choisi en conséquence. Les valeurs de ces quatre premiers paramètres sont donc directement calées sur les observations biologiques. Les trois suivantes dépendent directement des choix de modélisation : la valeur de R_{\min} est fixée à 1 par convention, ce qui fixe celle de R_{\max} en vertu de l'équation (1.13) page 31 ; enfin, la valeur de N_{div} est prise assez petite, tout en évitant de trop gros recouvrements lors de la division.

TABLEAU 2.1 – Valeurs numériques des paramètres de modélisation, données par défaut dans les simulations

Paramètre	Valeurs possibles	Signification
S_{div}	{1, 2}	Stratégie de division
D_{div}	{1, 2, 3, 4}	Direction de division

Les choix pour S_{div} indiquent respectivement une stratégie d'orientation libre ou contrainte. Les choix pour D_{div} indiquent respectivement une direction aléatoire, fixe, radiale ou tangentielle (voir section 1.4.4 page 37).

TABLEAU 2.2 – Valeurs numériques des paramètres de modélisation, à choisir pour chaque simulation

Paramètre	Valeur	Signification
dt	0,25	Pas de temps
ρ	$\mathcal{O}\left(\frac{1}{N^2}\right)$	Pas de l'itération sur les positions
μ	100	Pas de l'itération sur les multiplicateurs
tol_W	$\mathcal{O}\left(\frac{\rho}{10}\right)$	Tolérance du critère d'arrêt sur le potentiel
tol_ϕ	0,2	Tolérance du critère d'arrêt sur le recouvrement
tol_X	2	Tolérance de trop grand déplacement des cellules

La valeur de dt est choisie de sorte que le pas de croissance soit assez petit (pour éviter de trop forts recouvrements). Les différentes valeurs attribuées à ρ sont expliquées en section 2.2.2 page 47 et précisées en équation (2.31) page 49. La valeur de μ a été choisie par essai-erreur. La valeur de tol_W est liée à ρ . La valeur de tol_ϕ est choisie pour correspondre à un très faible recouvrement (de l'ordre de 5×10^{-2} , pour des rayons variant entre $R_{\text{min}} = 1$ et $R_{\text{max}} = 1,26$). La valeur de tol_X permet de détecter des trop grands déplacements de cellules (de l'ordre de deux fois le rayon puisque $R_{\text{min}} = 1$).

TABLEAU 2.3 – Valeurs numériques des paramètres de l'algorithme, données par défaut dans les simulations

Mise en place des indicateurs statistiques et étude des simulations

L'objectif de ce chapitre est d'utiliser l'algorithme mis en place au chapitre 2 pour étudier nos problématiques sur l'organisation des populations cellulaires. Pour réaliser cette étude, nous développons plusieurs indicateurs, appelés « diagnostics », mesurant différents critères sur la population et ses lignages ; afin d'établir des statistiques sur ces diagnostics, ils sont appliqués sur un grand nombre de simulations. Les résultats obtenus seront confrontés aux données expérimentales au chapitre 4.

Dans ce chapitre, nous proposons une représentation des populations cellulaires reposant sur les diagrammes de Voronoï (section 3.1), qui fournit une relation de voisinage entre cellules ; puis nous développons nos diagnostics, en distinguant leur définition de leur implémentation (sections 3.2 et 3.3) et en nous appuyant sur quelques notions de la théorie des graphes ; enfin nous étudions les simulations numériques avec ces diagnostics (section 3.4).

Sommaire du chapitre

3.1 Représentation par diagramme de Voronoï adapté	64
3.1.1 Motivation	64
3.1.2 Rappels sur les diagrammes de Voronoï	65
3.1.3 Modification et utilisation des diagrammes de Voronoï	66
3.2 Définition des diagnostics	70
3.2.1 Motivation des diagnostics	71
Notion de lignages	71
Introduction des diagnostics pour explorer les problématiques	71
3.2.2 Représentation d'une population cellulaire en terme de graphe	72
3.2.3 Définition des diagnostics portant sur la population cellulaire	75
3.2.4 Définition des diagnostics portant sur les lignages	76
3.3 Implémentation des diagnostics	78
3.3.1 Calculs préliminaires	78
3.3.2 Calculs des diagnostics	79

3.4	Protocole des simulations et résultats des diagnostics	81
3.4.1	Protocole d'étude	81
3.4.2	Résultats de chacun des diagnostics	82
	Diagnostic n° 1 (sphéricité de la population)	82
	Diagnostic n° 2 (convexité de la population)	82
	Diagnostic n° 3 (sphéricité d'un lignage)	85
	Diagnostic n° 4 (morcellement d'un lignage)	85
	Diagnostic n° 5 (taille des fragments d'un lignage)	89
	Diagnostic n° 6 (orientation d'un lignage)	90
3.4.3	Conclusions communes des différents résultats	95

3.1 Représentation par diagramme de Voronoï adapté

L'objectif de cette section est d'améliorer la représentation graphique des configurations spatiales de cellules, en s'inspirant des diagrammes de Voronoï. Nous commençons par motiver le besoin de disposer d'une meilleure représentation en section 3.1.1 ; puis nous rappelons ce qu'est un diagramme de Voronoï en section 3.1.2 ; enfin nous décrivons notre nouvelle représentation, dite « par diagramme de Voronoï adapté », en section 3.1.3.

3.1.1 Motivation

La visualisation graphique de ce que produit l'algorithme développé au chapitre 2 est illustrée en figure 3.1 : le départ de la simulation (figure 3.1a) s'effectue avec $N_0 = 50$ cellules, que nous laissons évoluer pendant une durée T_{\max} correspondant à trois temps de génération. Au terme de la simulation (figure 3.1b), il y a donc environ $N = 400$ cellules.

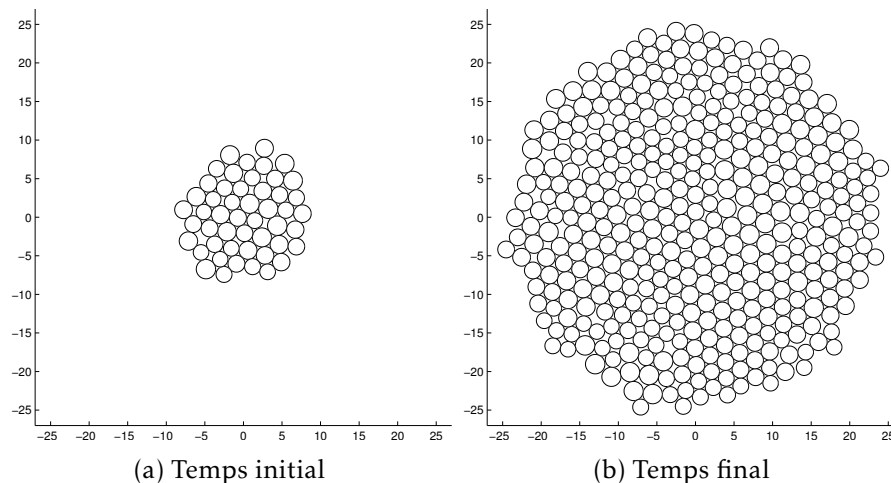


FIGURE 3.1 – Représentation par des disques de configurations issues de l'algorithme

Une telle représentation rend compte du modèle sous-jacent, qui consiste à modéliser les cellules par des disques ne pouvant se recouvrir. Néanmoins, elle admet des limites :

- des interstices sont visibles (suite à la représentation des cellules par des disques), qui n'existent pas dans la réalité ;
- elle ne sera pas adaptée aux données expérimentales, comme nous le verrons au chapitre 4.

Nous proposons une nouvelle représentation, inspirée des diagrammes de Voronoï, qui a le double avantage de corriger les limites visuelles ci-dessus mais aussi de fournir une notion intrinsèque de voisinage, dont nous allons nous servir dans la suite. Les diagrammes de Voronoï ont été considérés dans [Hon78] comme une bonne approximation de la forme des cellules, au moins en 2D ; dans [MPL01 ; SM05], ils sont même directement intégrés au modèle. Quant au calcul de proches voisins par diagramme de Voronoï, il est utilisé dans [Osby+10] pour déterminer des forces d'interaction deux à deux entre cellules.

3.1.2 Rappels sur les diagrammes de Voronoï

Nous utilisons la notion de diagramme de Voronoï en 2D. Nous rappelons ici quelques définitions et propriétés, puis nous donnons des exemples ; des démonstrations et de plus amples informations sont disponibles dans [Oka+09, chap. 2] ou [Aur91]. Soit $\mathcal{S} = \{p_i \mid 1 \leq i \leq n\}$ un ensemble fini de n points de \mathbb{R}^2 , où $n \in \mathbb{N}^*$.

Définition 3.1 (Site de Voronoï). Nous appelons site de Voronoï chacun des points p_i de l'ensemble \mathcal{S} .

Définition 3.2 (Région de Voronoï). Pour $i \in \llbracket 1 ; n \rrbracket$, nous appelons région¹ de Voronoï du site $p_i \in \mathcal{S}$ l'ensemble, noté \mathcal{R}_i , des points de \mathbb{R}^2 suivant :

$$\mathcal{R}_i = \left\{ x \in \mathbb{R}^2 \mid \forall j \neq i, \text{dist}(x, p_i) \leq \text{dist}(x, p_j) \right\}. \quad (3.1)$$

Il s'agit du lieu des points les plus proches de p_i que de n'importe quel autre point de \mathcal{S} .

Définition 3.3 (Diagramme de Voronoï). Nous appelons diagramme de Voronoï de l'ensemble des sites \mathcal{S} l'ensemble des régions \mathcal{R}_i (dont certaines ne sont pas bornées) associées à tous les points p_i de \mathcal{S} :

$$\bigcup_{i=1}^n \mathcal{R}_i. \quad (3.2)$$

Autrement dit, étant donné un ensemble de sites \mathcal{S} , leur diagramme de Voronoï sépare le plan selon la règle du plus proche voisin : chaque point de \mathbb{R}^2 est associé au(x) site(s) dont il est le plus proche.

1. On parle aussi de *cellule* mais pour éviter toute confusion avec les cellules de notre modélisation, nous adoptons le terme de *région*.

Nous terminons en donnant quelques informations supplémentaires sur les diagrammes de Voronoï, qui permettent d’avoir une vision géométrique d’une telle construction. Nous supposons être dans une configuration où $n \geq 3$, où les sites ne sont pas tous alignés² et où il n’y a pas plus de trois points cocirculaires³.

1. L’intersection de deux régions de Voronoï, si elle est non vide, est appelée une arête de Voronoï. Une telle arête, séparant deux régions \mathcal{R}_i et \mathcal{R}_j , appartient à la médiatrice du segment $[p_i p_j]$.
2. Deux sites p_i et p_j sont considérés comme plus proches voisins si leurs régions \mathcal{R}_i et \mathcal{R}_j partagent une arête commune.
3. L’intersection de trois régions de Voronoï, si elle est non vide, est appelée un sommet de Voronoï. Un tel sommet, séparant trois régions \mathcal{R}_i , \mathcal{R}_j et \mathcal{R}_k , est le centre du cercle circonscrit au triangle de sommets p_i , p_j et p_k .
4. La région \mathcal{R}_i est non bornée si et seulement si le point p_i est un sommet (du bord) de l’enveloppe convexe de \mathcal{S} .
5. Une région de Voronoï, si elle est bornée, est un polygone convexe.

Une illustration est proposée en figure 3.2 page ci-contre, où nous représentons des exemples de diagrammes de Voronoï pour $n \in \{4, 5, 6, 9, 12, 20\}$.

3.1.3 Modification et utilisation des diagrammes de Voronoï

Notre objectif est maintenant d’utiliser les diagrammes de Voronoï pour représenter nos populations cellulaires. Comme premier choix de l’ensemble des sites de Voronoï, noté \mathcal{S}_1 , nous prenons l’ensemble des centres des cellules :

$$\mathcal{S}_1 = \{X_i \mid 1 \leq i \leq N\}. \quad (3.3)$$

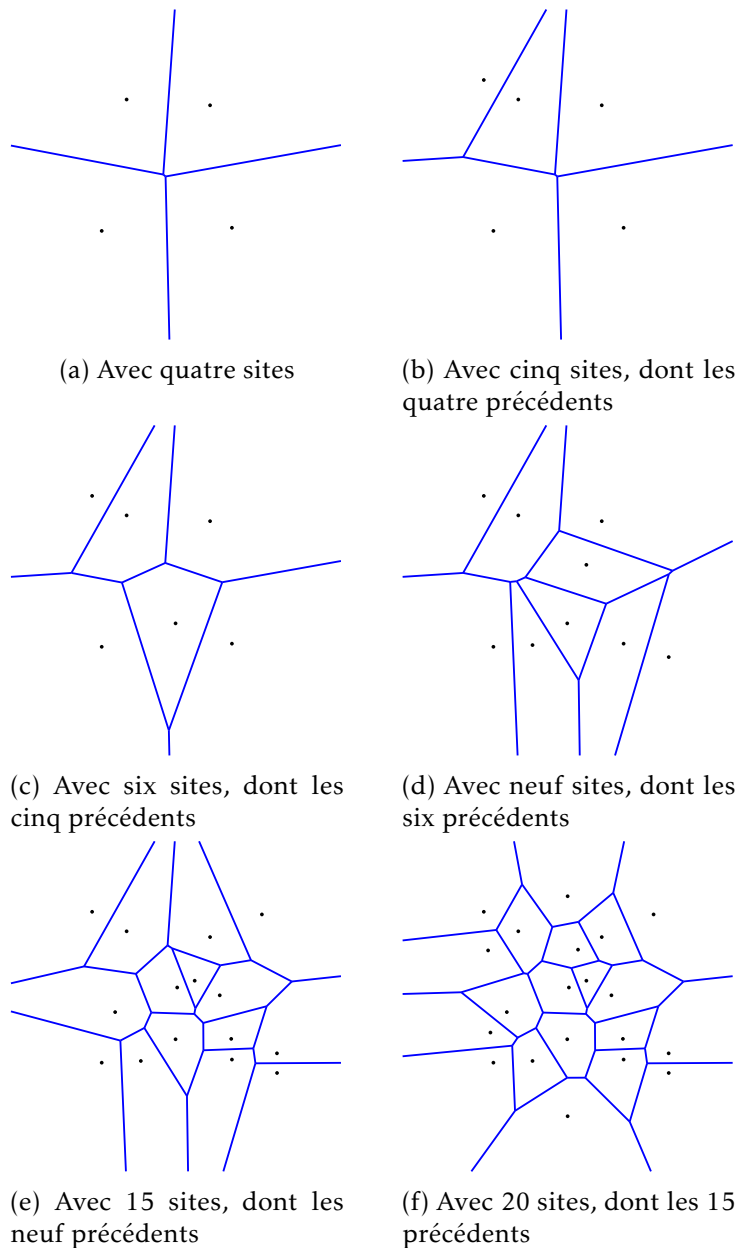
Un tel choix pour les sites amène à une représentation graphique illustrée en figure 3.3a page 69 où, afin de comparer les représentations, nous avons pris la même configuration des cellules qu’en figure 3.1b page 64. Nous pouvons dès à présent souligner un défaut de cette représentation : les régions de Voronoï correspondant à des sites situés à la périphérie de la population sont, soit non bornées, soit de formes trop éloignées par rapport aux régions bornées.

Afin de pallier cela, nous décidons d’ajouter de « faux sites » à la périphérie de notre population, que nous appelons des sites fictifs.

Définition 3.4 (Site réel – Site fictif). Nous appelons site réel un site de Voronoï qui correspond au centre d’une cellule de notre modèle. Nous appelons site fictif un site de Voronoï que nous rajoutons artificiellement aux sites réels (et qui ne correspond donc à aucune cellule).

2. Sinon il n’y aurait pas de sommets.

3. Sinon plus de trois arêtes concourraient en un sommet.



Légende. Les figures représentent le tracé du diagramme de Voronoï pour différents ensembles \mathcal{S} ; à chaque nouvelle figure, nous avons gardé les sites des précédentes figures, afin d'illustrer l'évolution du tracé. Les arêtes sont dessinées en bleu et les sites en noir.

Observations. Sur les dernières figures, nous voyons qu'il existe des régions bornées et d'autres non bornées.

FIGURE 3.2 – Diagrammes de Voronoï pour 4, 5, 6, 9, 15 et 20 points de \mathbb{R}^2

L'intérêt de ces sites fictifs est de les rajouter aux bons endroits, de sorte d'avoir une meilleure représentation graphique.

1. Une première possibilité consiste à placer des sites fictifs le long d'un rectangle englobant les sites réels : cela amène au choix noté \mathcal{S}_2 pour l'ensemble des sites. Soit $\delta > 0$, soit $n \in \mathbb{N}$, soit les quatre points du plan $S_1(\max_i(x_i) + \delta, \max_i(y_i) + \delta)$, $S_2(\min_i(x_i) - \delta, \max_i(y_i) + \delta)$, $S_3(\min_i(x_i) - \delta, \min_i(y_i) - \delta)$ et $S_4(\max_i(x_i) + \delta, \min_i(y_i) - \delta)$ et soit le rectangle formé des quatre sommets $S_1S_2S_3S_4$. Soit n points régulièrement espacés le long des arêtes de ce rectangle, notés F_1, F_2, \dots, F_n ; alors cette possibilité correspond à

$$\mathcal{S}_2 = \{X_i | 1 \leq i \leq N\} \cup \{F_i | 1 \leq i \leq n\}. \quad (3.4)$$

Un tel choix pour les sites est illustré en figure 3.3b page suivante. L'agrandissement d'un facteur δ permet d'éviter que des sites fictifs et réels soient confondus, ce qui pourrait avoir pour conséquence une région de Voronoï non bornée pour un tel site réel. Au final, ce choix des sites fictifs est facile à calculer et permet de borner toutes les régions associées à des sites réels ; cependant, il ne résout pas le problème des régions de formes trop irrégulières au bord : il apparait comme trop grossier.

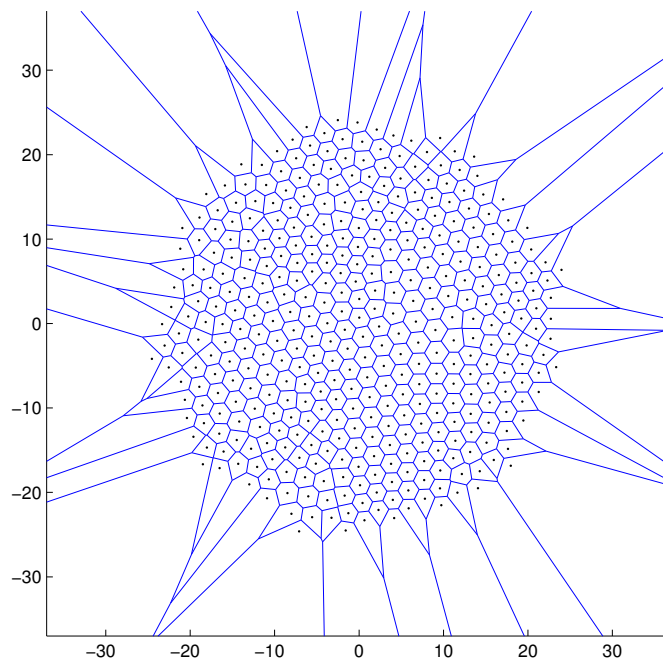
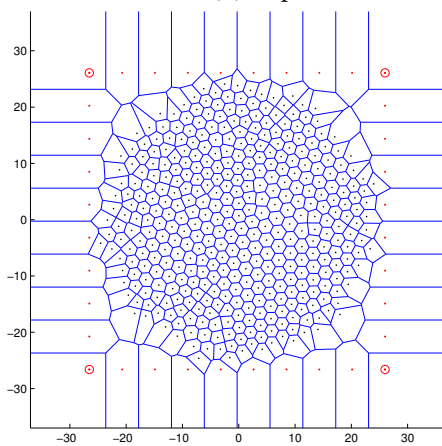
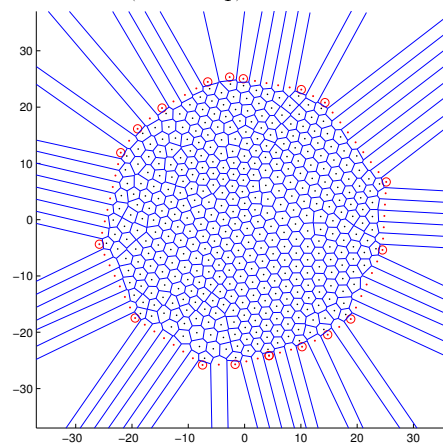
2. Cela nous amène à une deuxième possibilité de placement des sites fictifs, plus précise, qui consiste à les positionner le long d'un polygone beaucoup plus proche des sites réels : nous choisissons pour ce polygone le bord de l'enveloppe convexe des sites réels, légèrement agrandi par homothétie. Cela amène au choix noté \mathcal{S}_3 de l'ensemble des sites. Soit $\delta > 0$, soit $n \in \mathbb{N}$, soit \mathcal{P} le polygone correspondant à l'agrandissement d'un facteur δ du bord de l'enveloppe convexe des sites réels. Soit n points régulièrement espacés le long des arêtes de \mathcal{P} , notés F_1, F_2, \dots, F_n ; alors cette possibilité correspond à

$$\mathcal{S}_3 = \{X_i | 1 \leq i \leq N\} \cup \{F_i | 1 \leq i \leq n\}. \quad (3.5)$$

Un tel choix pour les sites est illustré en figure 3.3c page ci-contre. Cette deuxième possibilité nécessite de calculer une enveloppe convexe mais amène à une représentation beaucoup plus acceptable.

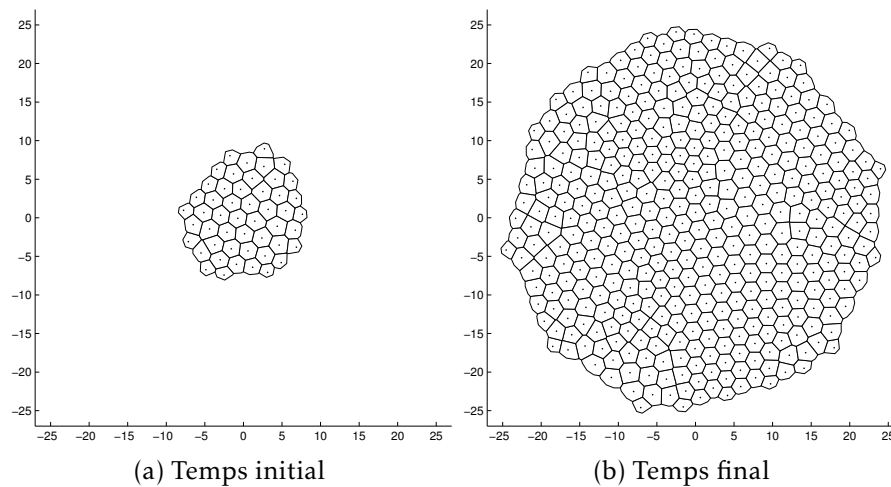
Pour obtenir la représentation de nos populations cellulaires, il ne reste plus qu'à appliquer cette procédure d'ajout de sites fictifs mais à ne dessiner que les régions de Voronoï correspondant à des sites réels : cela est illustré en figure 3.4 page 70, où les configurations des cellules sont les mêmes qu'en figure 3.1 page 64.

Définition 3.5 (Diagramme de Voronoï adapté). Nous appelons diagramme de Voronoï adapté, pour les cellules X_1, X_2, \dots, X_N , le diagramme obtenu en retenant le choix \mathcal{S}_3 exprimé en équation (3.5) pour l'ensemble des sites et en ne traçant que les régions correspondant aux sites réels X_1, X_2, \dots, X_N .

(a) Représentation sans sites fictifs (choix \mathcal{S}_1)(b) Représentation après ajout de sites fictifs le long d'un rectangle englobant les sites réels (choix \mathcal{S}_2)(c) Représentation après ajout de sites fictifs le long d'un polygone englobant les sites réels (choix \mathcal{S}_3)

Légende. Les sites réels (X_1, X_2, \dots, X_N) sont représentés en noir, les sites fictifs (F_1, F_2, \dots, F_n) en rouge, les arêtes des régions en bleu. Les sites fictifs entourés correspondent aux sommets du polygone englobant les sites réels (les autres sites fictifs étant répartis le long des arêtes de ce polygone). Les sites réels sont identiques sur les trois figures, de sorte que seul le choix des sites fictifs change.

FIGURE 3.3 – Représentation par diagramme de Voronoï, avec différents choix de sites fictifs



Légende. Les configurations des cellules y sont les mêmes qu'en figure 3.1 page 64, afin de comparer les deux types de représentation.

FIGURE 3.4 – Représentation par diagramme de Voronoï adapté de configurations issues de l'algorithme

Conclusion de cette section 3.1 Nous avons mis en place une nouvelle représentation, appelée diagramme de Voronoï adapté, qui est basée sur la théorie des diagrammes de Voronoï : elle est systématiquement utilisée dans la suite. Cette représentation a plusieurs avantages :

- elle pallie les défauts visuels de la précédente représentation. En particulier, elle conduit à une vision des cellules sous forme polygonale, bien plus réaliste que des disques, qui correspond mieux aux attentes des biologistes ;
- elle est adaptée aussi bien à la représentation d'une population cellulaire issue des simulations qu'à une population issue des expériences. En effet, contrairement aux simulations qui nous fournissent le centre des cellules, nous verrons par la suite (en section 4.1 page 98) que les données expérimentales ne nous donnent accès qu'au centre du *noyau* des cellules ;
- elle fournit, et cela sera utilisé en définition 3.9 page 74, une relation intrinsèque de voisinage entre les cellules qui sera, elle aussi, bien adaptée à la fois aux simulations et aux expériences.

3.2 Définition des diagnostics

L'objectif de cette section est d'effectuer la mise en place des diagnostics, en expliquant ce qu'ils sont et à quoi ils servent. Nous motivons l'intérêt des diagnostics en section 3.2.1 ; puis nous apportons une vision de la population cellulaire en terme de graphe en section 3.2.2, sur laquelle reposeront les définitions de certains diagnostics ; enfin nous définissons en sections 3.2.3 et 3.2.4 les différents diagnostics, en distinguant

ceux qui portent sur la population de ceux qui portent sur les lignages.

3.2.1 Motivation des diagnostics

Notion de lignages

Notre problématique 2 consiste à étudier l'influence de l'*orientation de division* au moment de la *mitose*. Cette orientation intervenant au moment de la division, elle impacte sur les deux cellules filles ; et lorsque ces deux cellules filles se diviseront, elle impactera sur les quatre cellules « petites-filles ». Par ailleurs, notre problématique 1 portant sur l'impact de la multiplication des cellules, elle est aussi naturellement reliée à la division et au devenir des cellules filles.

Nous sommes donc naturellement amenés à considérer, au temps final, l'ensemble des cellules descendant d'une même cellule ancêtre présente au temps initial ; cette information de filiation est stockée dans la caractéristique « lignage » que nous avons introduite au chapitre 1. En effet, lorsqu'une cellule C_i se divise, nous avons établi que le lignage de chacune de ses deux cellules filles correspond à son lignage ℓ_i : cette information permet donc de caractériser l'appartenance d'une cellule à un même arbre généalogique. Grâce à l'initialisation des lignages proposée en section 2.5.1 page 59, nous pouvons définir récursivement la valeur de ℓ_i pour $i \in \llbracket 1 ; N \rrbracket$.

Définition 3.6 (Lignage d'une cellule). Le lignage ℓ_i d'une cellule C_i est défini par :

- si C_i est présente au temps initial (parmi les N_0 cellules initiales), alors $\ell_i = i$;
- si C_i n'existe pas au temps initial, alors ℓ_i a la même valeur que le lignage de la cellule mère de C_i .

Définition 3.7 (Cellule génératrice (d'un lignage)). Étant donné un lignage, nous appelons cellule génératrice de ce lignage la cellule, présente au temps initial, qui en est à l'origine.

Il s'agit bien de la même notion de lignage que dans la langue courante⁴. La relation « être de même lignage », définie par :

$$C_i \text{ et } C_j \text{ sont de même lignage} \iff \ell_i = \ell_j, \quad (3.6)$$

est donc une relation d'équivalence, qui définit une partition des cellules. Par extension, nous appellerons dans la suite « lignage d'une cellule C_i » l'ensemble de toutes les cellules ayant le même lignage que la cellule C_i .

Introduction des diagnostics pour explorer les problématiques

Un suivi de six lignages au cours du temps est illustré en figure 3.5 page 73. Sur cet exemple, nous voyons que les lignages peuvent être assez différents : certains sont allongés, voire presque alignés, alors que d'autres sont regroupés de manière plus

4. Lignage : groupe de parents, formé d'ascendants et de collatéraux, dont les membres se considèrent comme descendants unilinéaires d'un ancêtre commun.

arrondie ; certains sont d'un seul tenant et d'autres en plusieurs morceaux ; ces morceaux peuvent être proches ou éloignés les uns des autres, voire avec des cellules isolées du reste de leur lignage ; etc. Toutes ces informations qui portent sur le devenir des lignages sont vagues. . . elles doivent donc être mesurées plus précisément, afin d'être étudiées de manière systématique : nous avons besoin de *quantifier ces observations*.

Pour répondre à ce besoin, nous allons développer plusieurs indicateurs qui vont effectuer certaines mesures précises, dans le but d'étudier l'organisation des populations cellulaires. Nous appelons ces indicateurs des « diagnostics » : ils vont avoir pour objectif d'étudier la morphologie des lignages et de la population entière ; pour ce faire, ils devront répondre à différentes contraintes :

- avoir un sens et une interprétation biologiques ;
- être mesurables sur les données issues des simulations numériques ;
- être mesurables sur les données issues des expériences en cultures cellulaires (que nous verrons au chapitre 4).

Les définitions des diagnostics reposent sur des notions de la théorie des graphes, que nous présentons.

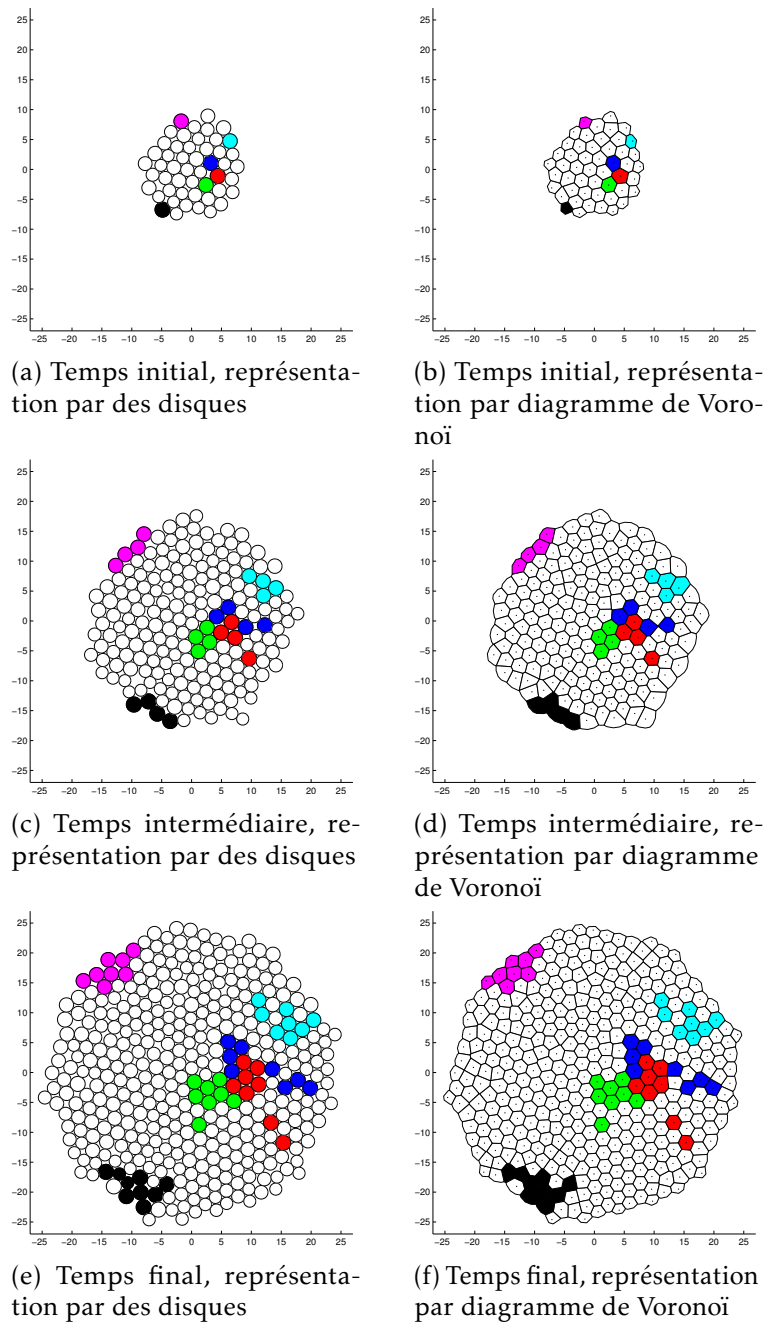
3.2.2 Représentation d'une population cellulaire en terme de graphe

L'objectif de cette section est de donner un sens à la connexité, au périmètre et à l'aire d'une population de cellules, concepts dont nous avons besoin pour définir certains de nos diagnostics à venir. Pour cela, nous nous appuyons sur la théorie des graphes, qui fournit un cadre pour définir rigoureusement ces notions, et nous créons un graphe reproduisant la configuration de la population cellulaire.

Définition 3.8 (Vocabulaire sur les graphes). Nous définissons ici différentes notions de la théorie des graphes.

- Un *graphe* G est un ensemble de points, appelés *sommets*, dont certaines paires sont reliées entre elles par un *lien*⁵. La donnée d'un graphe consiste à avoir ces sommets et ces liens.
- Une *chaîne* dans le graphe G est une suite finie de sommets reliés entre eux par un lien (avec répétition possible de sommets ou de liens).
- Un *cycle* est une chaîne telle que son dernier et son premier termes coïncident.
- Un graphe est dit *connexe* s'il existe une chaîne entre toute paire de sommets.
- Un graphe H est *inclus* dans un graphe G lorsque tous les sommets de H sont des sommets de G et lorsque tous les liens de H sont des liens de G . Dans ce cas, H est appelé un *sous-graphe* de G .
- Une *composante connexe* d'un graphe G est un sous-graphe H connexe qui est maximal pour l'inclusion.
- Le sous-graphe induit du graphe G par un ensemble S de sommets de G est le sous-graphe obtenu en supprimant les sommets de G qui ne sont pas dans S , ainsi que tous les liens incidents à ces sommets supprimés.

5. La terminologie habituelle pour ces liens est *arêtes* ; cependant, pour éviter toute confusion avec les arêtes du diagramme de Voronoï, nous ne l'utilisons pas.



Légende. Ces figures illustrent trois temps (initial sur la première ligne, intermédiaire sur la deuxième, final sur la troisième) d'une seule simulation. Afin de les comparer, nous gardons sur cette figure les deux représentations : sur la colonne de droite, il s'agit de la représentation par diagramme de Voronoï (qui fournit en particulier les relations de voisinage) et sur la colonne de gauche, les cellules sont représentées par des disques.

FIGURE 3.5 – Évolution de lignages au cours du temps

Définition 3.9 (Cellules voisines – Relations de voisinage). Soit $i, j \in \llbracket 1; N \rrbracket$. Deux cellules C_i et C_j sont dites voisines lorsque leurs régions de Voronoï \mathcal{R}_i et \mathcal{R}_j partagent une arête commune.

Définition 3.10 (Graphe des N cellules). L'ensemble des cellules est vu comme un graphe, appelé *graphe des N cellules*, dont les sommets sont les centres des cellules et dont les liens sont définis par les relations de voisinage : il y a un lien entre deux sommets si et seulement s'ils correspondent à deux cellules voisines.

Définition 3.11 (Ensemble connexe de cellules). Soit $p \in \llbracket 1; N \rrbracket$. Soit p cellules notées $C_{i_1}, C_{i_2}, \dots, C_{i_p}$. Cet ensemble de cellules est dit connexe lorsque le sous-graphe (du graphe des N cellules) induit par ces p cellules est un graphe connexe.

Définition 3.12 (Composante connexe d'un lignage). Considérons le sous-graphe (du graphe des N cellules) induit par l'ensemble des sommets correspondant à toutes les cellules d'un même lignage. Une composante connexe d'un lignage est alors définie comme étant une composante connexe de ce sous-graphe induit. Dans la suite, nous parlerons donc de « connexité (resp. composante connexe) d'un lignage » à la place de « connexité (resp. composante connexe) du sous-graphe induit par les sommets constitués de l'ensemble des cellules d'un lignage ».

Définition 3.13 (Cycle cellulaire). Un cycle cellulaire est une suite finie de cellules correspondant à un cycle dans le graphe des N cellules. Autrement dit, pour $p \geq 1$, il s'agit de cellules notées $C_{i_1}, C_{i_2}, \dots, C_{i_p}$ telles que :

$$\text{pour tout } k \in \llbracket 1; p \rrbracket, C_{i_k} \text{ et } C_{i_{k+1}} \text{ sont voisines et } C_{i_{p+1}} = C_{i_1}.$$

Un tel cycle est noté $(C_{i_1}, C_{i_2}, \dots, C_{i_{p+1}})$.

Définition 3.14 (Polygone cellulaire associé à un cycle cellulaire). Soit $p \geq 3$ et soit un cycle cellulaire noté $(C_{i_1}, C_{i_2}, \dots, C_{i_{p+1}})$. Considérons la ligne brisée du plan passant respectivement par les points de coordonnées $X_{i_1}, X_{i_2}, \dots, X_{i_{p+1}}$: il s'agit d'un polygone, que nous appelons polygone cellulaire associé au cycle cellulaire $(C_{i_1}, C_{i_2}, \dots, C_{i_{p+1}})$.

Définition 3.15 (Bord d'un ensemble de cellules connexe – Polygone de bord). Soit un ensemble de cellules connexe. Le bord est défini comme un cycle cellulaire dont le polygone cellulaire associé possède en son intérieur ou sur son bord tous les centres des cellules de l'ensemble connexe considéré. Ce polygone est appelé le polygone de bord.

Grâce à cette notion de bord d'un ensemble de cellules connexe, nous sommes en mesure de définir une aire et un périmètre d'un tel ensemble de cellules.

Définition 3.16 (Périmètre et aire d'un ensemble de cellules connexe). Soit un ensemble de cellules connexe et $p \geq 1$. Considérons le bord de cet ensemble, noté $(C_{i_1}, C_{i_2}, \dots, C_{i_{p+1}})$.

- Le périmètre de cet ensemble de cellules est défini comme le périmètre du polygone cellulaire associé à $(C_{i_1}, C_{i_2}, \dots, C_{i_{p+1}})$; il est donné par

$$\mathcal{P} = \sum_{k=1}^p \text{dist}(C_{i_k}, C_{i_{k+1}}). \quad (3.7)$$

- L'aire de cet ensemble de cellules est définie comme l'aire du polygone cellulaire associé à $(C_{i_1}, C_{i_2}, \dots, C_{i_{p+1}})$; elle est donnée (voir [Bra86]) par

$$\mathcal{A} = \frac{1}{2} \left| \sum_{k=1}^p (x_{i_k} y_{i_{k+1}} - x_{i_{k+1}} y_{i_k}) \right|. \quad (3.8)$$

Définition 3.17 (Enveloppe convexe d'un ensemble de cellules). L'enveloppe convexe d'un ensemble de cellules est définie comme l'enveloppe convexe du nuage de points constitué de l'ensemble des centres de ces cellules.

Remarque 3.1 (Critique et motivation des choix retenus pour définir l'aire et le périmètre). Cette définition de l'aire ne correspond pas à la somme des aires des disques représentant les cellules; de même, le périmètre ne suit pas le bord des disques situés au bord. Nous justifions nos choix par les arguments suivants.

- De telles définitions auraient été plus difficiles à exploiter. En effet, nous nous sommes intéressés dans la suite à comparer l'aire, le périmètre et l'enveloppe convexe d'un ensemble de cellules : à ce titre, il est judicieux de baser ces définitions sur un polygone passant par les sommets des cellules; de plus, il est important que l'aire et le périmètre de l'ensemble des cellules soient mesurés sur le même objet géométrique.
- De telles définitions n'auraient pas nécessairement été plus précises. En effet, les interstices vides observés dans la représentation des cellules par des disques n'auraient pas été comptabilisés dans l'aire de la population.

3.2.3 Définition des diagnostics portant sur la population cellulaire

Diagnostic n° 1 : sphéricité de la population Ce diagnostic est un critère permettant d'évaluer la sphéricité d'une population cellulaire. Il mesure le rapport R_1 entre l'aire et le périmètre de l'ensemble de la population cellulaire, selon :

$$R_1 = \frac{4\pi\mathcal{A}}{\mathcal{P}^2}, \quad (3.9)$$

où \mathcal{P} et \mathcal{A} sont respectivement le périmètre et l'aire de cet ensemble de cellules. Il s'agit d'un indicateur permettant d'évaluer la sphéricité de la population cellulaire, au sens où $R_1 \in [0; 1]$ vaudrait 1 lorsque le bord de la population serait un cercle (en vertu de l'inégalité isopérimétrique). Par ailleurs, il permet aussi de fournir une valeur de référence au diagnostic n° 3 à venir.

Diagnostic n° 2 : convexité de la population Ce diagnostic est un critère permettant d'évaluer la convexité d'une population cellulaire. Il mesure le rapport R_2 entre l'aire \mathcal{A} de l'ensemble des cellules et l'aire $\mathcal{A}_{\text{conv}}$ de son enveloppe convexe, selon :

$$R_2 = \frac{\mathcal{A}}{\mathcal{A}_{\text{conv}}}. \quad (3.10)$$

Il s'agit d'un indicateur permettant d'évaluer la convexité de la population cellulaire, au sens où $R_2 \in [0;1]$ vaut 1 lorsque le bord de la population est associé à un polygone cellulaire convexe.

3.2.4 Définition des diagnostics portant sur les lignages

Diagnostic n° 3 : sphéricité d'un lignage Ce diagnostic est un critère permettant d'évaluer la sphéricité d'une composante connexe de lignage ; il est motivé par le fait d'étudier si ces composantes sont de formes aplaties ou rondes. Il mesure, pour chaque composante connexe (d'au moins deux cellules) d'un lignage, le rapport R_3 entre l'aire et le périmètre cet ensemble connexe, selon une formule similaire R_1 :

$$R_3 = \frac{4\pi\mathcal{A}}{\mathcal{P}^2}, \quad (3.11)$$

où \mathcal{P} et \mathcal{A} sont respectivement le périmètre et l'aire du polygone de bord de la composante considérée. Nous ne prenons pas en compte les composantes constituées d'une seule cellule, où $\mathcal{A} = \mathcal{P} = 0$. En particulier, ce diagnostic détecte des lignages filiformes, pour lesquels l'aire est nulle (puisque le polygone de bord est alors aplati).

Diagnostic n° 4 : morcellement d'un lignage Ce diagnostic permet de mesurer l'éclatement d'un lignage ; il est motivé par le fait d'étudier si les cellules d'un même lignage ont tendance à rester voisines ou bien à s'éparpiller. Ce diagnostic mesure, pour chaque lignage, le nombre R_4 de morceaux en lesquels est séparé le lignage :

$$R_4 = \text{nombre de composantes connexes d'un lignage}. \quad (3.12)$$

En particulier, il permet d'avoir des informations sur l'éclatement d'un lignage.

Diagnostic n° 5 : taille des fragments d'un lignage Ce diagnostic permet d'évaluer la taille d'une composante connexe d'un lignage ; il est motivé par le fait de fournir une information sur les morceaux d'un lignage, complémentaire au diagnostic n° 4. Ce diagnostic mesure, pour chaque composante connexe de chaque lignage, la taille de cette composante :

$$R_5 = \text{nombre de cellules dans une composante connexe d'un lignage}. \quad (3.13)$$

En particulier, il permet de savoir s'il y a davantage de grosses ou de petites composantes.

Diagnostic n° 6 : orientation d'un lignage Ce diagnostic permet d'évaluer dans quelle direction, et avec quelle intensité, un lignage est orienté ; il est motivé par le fait d'étudier l'impact de l'orientation du plan de division au moment de la division cellulaire sur l'orientation globale des lignages. Ce diagnostic mesure, pour chaque lignage, la direction principale dans laquelle il est orienté ; nous calculons pour cela les valeurs propres et les vecteurs propres de la matrice d'inertie des cellules composant ce lignage.

Définition 3.18 (Matrice d'inertie). Soit $p \geq 2$ et, pour $k \in \llbracket 1 ; p \rrbracket$, $X_{i_k} = (x_{i_k}, y_{i_k}) \in \mathbb{R}^2$. Notons $X_G = (x_G, y_G)$ le barycentre de ces points : $X_G = \sum_{k=1}^p \frac{X_{i_k}}{p}$. La matrice d'inertie du nuage de points $(X_{i_k})_{1 \leq k \leq p}$ est définie par la matrice symétrique :

$$\mathcal{E} = \frac{1}{p} \sum_{k=1}^p (X_{i_k} - X_G)(X_{i_k} - X_G)^T. \quad (3.14)$$

Remarque 3.2 (Interprétation géométrique de la matrice d'inertie \mathcal{E}). Si le barycentre est à l'origine, c'est-à-dire dans le cas $X_G = (0, 0)$, alors on est ramené à

$$\mathcal{E} = \frac{1}{p} \sum_{k=1}^p \begin{pmatrix} x_{i_k}^2 & x_{i_k} y_{i_k} \\ x_{i_k} y_{i_k} & y_{i_k}^2 \end{pmatrix}. \quad (3.15)$$

L'ensemble des points X tels que $X\mathcal{E}^{-1}X^T = 1$ est alors une ellipse de centre 0, dont le grand axe est porté par le vecteur propre associé à la plus grande valeur propre de \mathcal{E} .

En supposant que les points X_{i_k} ne sont pas alignés, cette matrice \mathcal{E} est symétrique définie positive, donc diagonalisable à valeurs propres strictement positives. Notons $0 < \lambda_1 \leq \lambda_2$ ses valeurs propres, associées aux vecteurs propres v_1 et v_2 . Ce diagnostic mesure l'écart entre l'orientation principale θ_2 de l'ellipse, fournie par v_2 , et l'angle polaire θ_0 du barycentre du lignage ($\theta_0, \theta_2 \in [-\pi; \pi]$) décalé de $\frac{\pi}{2}$:

$$R_6 = \theta_2 - \theta_0 - \frac{\pi}{2}, \text{ à valeurs dans } \left[-\frac{\pi}{2}; \frac{\pi}{2}\right] \text{ (car angle orienté entre deux droites)}. \quad (3.16)$$

Par ailleurs, grâce aux valeurs propres, il est possible de calculer l'excentricité de l'ellipse, afin de mesurer l'intensité de l'allongement dans la direction fournie par v_2 :

$$\tilde{R}_6 = \sqrt{1 - \frac{\lambda_1}{\lambda_2}}. \quad (3.17)$$

Une illustration est proposée en figure 3.7 page 80.

Autre diagnostic Un autre diagnostic qui mesure le déplacement entre, d'une part la position d'une cellule au temps initial et, d'autre part, la position du centre de gravité de son lignage en temps final a été utilisé. Néanmoins ses résultats n'ont pas amené de conclusion pertinente : nous n'en discutons pas davantage.

3.3 Implémentation des diagnostics

L'objectif de cette section est de fournir l'implémentation algorithmique des diagnostics, sans rentrer dans les détails de programmation. Afin de clarifier la présentation et d'identifier les besoins des différents diagnostics, nous procédons en deux temps : tout d'abord les calculs communs sur lesquels reposent plusieurs diagnostics, en section 3.3.1 ; puis les calculs des diagnostics à proprement parler, en section 3.3.2.

3.3.1 Calculs préliminaires

Il s'agit de calculs venant en amont des diagnostics. Comme nous l'avons indiqué précédemment, les relations de voisinage entre les cellules sont définies par le diagramme de Voronoï adapté, qu'il faut calculer. Cela permet ensuite de pouvoir calculer le polygone de bord, qui doit renvoyer les sommets des cellules du bord, dans l'ordre de parcours.

Diagramme de Voronoï adapté. Cette fonction consiste à choisir les sites puis à utiliser les programmes internes de Matlab pour le calcul du diagramme.

1. Définition des sites du diagramme de Voronoï. Elle se fait selon le choix exprimé en équation (3.5) page 68. Cela nécessite de calculer l'enveloppe convexe de la population, de l'agrandir par dilatation, puis de choisir des points uniformément répartis le long de chacune des arêtes de son enveloppe. L'ensemble des sites est composé des centres des cellules ainsi que de ces nouveaux points.
2. Calcul de la triangulation de Delaunay puis du diagramme de Voronoï correspondant au choix de ces sites.

Une illustration a déjà été introduite en figure 3.4 page 70.

Relations de voisinage. Cette fonction crée une matrice \mathcal{V} symétrique de taille $N \times N$ à coefficients dans $\{0, 1\}$, de diagonale nulle, telle que

$$[\mathcal{V}]_{i,j} = \begin{cases} 1 & \text{si } \mathcal{R}_i \text{ et } \mathcal{R}_j \text{ partagent une arête commune,} \\ 0 & \text{sinon.} \end{cases} \quad (3.18)$$

Elle repose sur le diagramme de Voronoï adapté.

Polygone de bord. Cette fonction calcule le bord d'un ensemble connexe de cellules, sous forme de cycle cellulaire $(C_{i_1}, C_{i_2}, \dots, C_{i_{p+1}})$ qui contient les sommets dans l'ordre de parcours. Elle suit l'algorithme suivant, qui parcourt le bord dans le sens trigonométrique :

1. Choix de la première cellule. Nous prenons celle qui est la plus à droite ; autrement dit, soit i_1 tel que :

$$x_{i_1} = \max_{1 \leq i \leq N} x_i.$$

2. Choix de la deuxième cellule. Parmi l'ensemble des voisines de C_{i_1} (données par la fonction de relations de voisinage), nous choisissons la cellule faisant un angle de vecteur orienté minimal avec la verticale ; autrement dit, soit i_2 tel que :

$$(\overrightarrow{C_{i_1}M_0}, \overrightarrow{C_{i_1}C_{i_2}}) = \min_{\substack{i \text{ tel que} \\ [\mathcal{V}]_{i,i_1}=1}} (\overrightarrow{C_{i_1}M_0}, \overrightarrow{C_{i_1}C_i}),$$

où M_0 est défini comme le point du plan d'abscisse x_{i_1} et d'ordonnée $y_{i_1} + 1$.

3. Choix des cellules suivantes. Supposons $C_{i_1}, C_{i_2}, \dots, C_{i_k}$ trouvées : parmi l'ensemble des voisines de C_{i_k} (pour $i_k \geq 2$) nous choisissons la cellule la plus à droite par rapport à la direction d'où nous venons ; autrement dit, soit i_{k+1} tel que

$$(\overrightarrow{C_{i_k}C_{i_{k-1}}}, \overrightarrow{C_{i_k}C_{i_{k+1}}}) = \min_{\substack{i \text{ tel que} \\ [\mathcal{V}]_{i,i_k}=1}} (\overrightarrow{C_{i_k}C_{i_{k-1}}}, \overrightarrow{C_{i_k}C_i}),$$

où l'angle $(\overrightarrow{C_{i_k}C_{i_{k-1}}}, \overrightarrow{C_{i_k}C_i})$ est défini dans l'intervalle $]0, 2\pi]$ (afin que le choix $i_{k+1} = i_{k-1}$ corresponde à un angle de 2π et non à un angle de zéro, ce qui en fait un choix possible mais en dernier ressort ⁶).

4. Critère d'arrêt. Notre algorithme s'arrête lorsque nous sommes revenus au point de départ, c'est-à-dire dès que $i_{k+1} = i_1$; nous posons alors $p = k$ et le bord est défini par $(C_{i_1}, C_{i_2}, \dots, C_{i_p}, C_{i_{p+1}})$.

Cette fonction utilise les relations de voisinage, et donc le diagramme de Voronoï adapté. Une illustration est proposée en figure 3.6 page suivante.

Choix des lignages suivis : lignage central, lignage périphérique. Cette fonction consiste à choisir deux lignages : un dont la cellule génératrice était au plus proche du centre de la population et un dont la cellule génératrice était au plus loin (c'est-à-dire sur le bord de la population cellulaire). Nous parlons respectivement de lignage central et de lignage périphérique.

Définition 3.19 (Zone centrale – Zone périphérique). La zone centrale (respectivement périphérique) fait référence au suivi d'un lignage central (resp. périphérique).

3.3.2 Calculs des diagnostics

Diagnostics n° 1 et 2 En amont de ces deux diagnostics, nous calculons le diagramme de Voronoï adapté (afin d'obtenir les relations de voisinage) puis le polygone de bord de toute la population. Une fois obtenu ce polygone, nous calculons son périmètre \mathcal{P} et son aire \mathcal{A} avec les formules fournies en équations (3.7) et (3.8) page 75. Nous calculons l'enveloppe convexe de la population, pour en déduire l'aire $\mathcal{A}_{\text{conv}}$. Cela permet de calculer R_1 et R_2 .

6. Le polygone de bord est alors croisé, puisque deux de ses cotés au moins sont superposés.

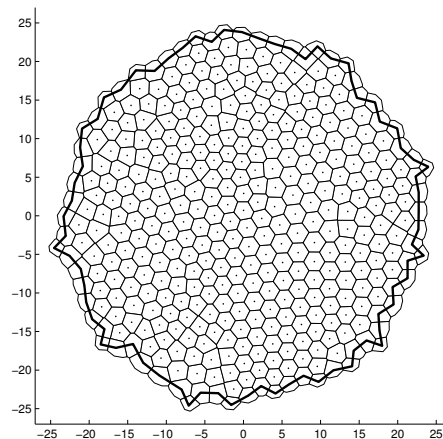
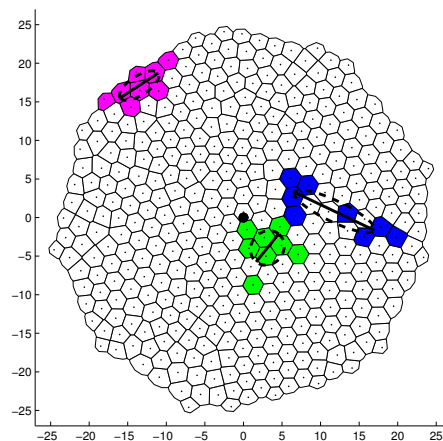


FIGURE 3.6 – Représentation du polygone de bord pour une population cellulaire

Diagnostics n° 3, 4 et 5 En amont de ces trois diagnostics, nous calculons le diagramme de Voronoï adapté (afin d’obtenir les relations de voisinage de toute la population). Puis nous calculons les composantes connexes de chaque lignage, afin d’identifier le nombre et le contenu de chacune de ces composantes. Cela fournit directement le nombre (R_4) de composantes connexes des lignages suivis et leur taille (R_5). Ensuite, toujours pour les lignages suivis, nous calculons le polygone de bord de chacune de leurs composantes connexes (de deux cellules au moins), afin d’obtenir leur périmètre et leur aire (R_3).

Diagnostics n° 6 Pour les lignages suivis, nous calculons la matrice d’inertie \mathcal{E} donnée en équation (3.14) page 77 ; ensuite nous en extrayons les valeurs et vecteurs propres afin de calculer R_6 et \tilde{R}_6 . Une illustration est proposée en figure 3.7.



Légende. La direction de chaque lignage est donnée par l’orientation du grand axe de l’ellipse tracée représentant sa matrice d’inertie.

FIGURE 3.7 – Représentation de la direction principale de trois lignages

3.4 Protocole des simulations et résultats des diagnostics

L'objectif de cette section est de présenter l'ensemble des résultats obtenus grâce à chacun des diagnostics précédemment mis en place ; afin d'étudier si des comportements particuliers émergent, nous appliquons ces diagnostics sur un grand nombre de simulations. Nous présentons en section 3.4.1 le protocole suivi pour cette étude ; puis nous détaillons en section 3.4.2, diagnostic par diagnostics, les résultats obtenus ; enfin, nous terminons en section 3.4.3 par une conclusion qui rassemble et confronte les résultats de chacun des diagnostics.

3.4.1 Protocole d'étude

Dans l'algorithme développé au chapitre 2, nous avons laissé deux paramètres à choisir ultérieurement pour chaque simulation (voir tableau 2.2 page 61) : la stratégie de division S_{div} (parmi « libre » ou « contrainte ») et la direction d'orientation D_{div} (parmi « aléatoire », « constante », « radiale » ou « tangentielle »). Nous rappelons que ce choix ne varie pas au cours d'une simulation, c'est-à-dire que toutes les divisions d'une même simulation se font selon la même stratégie et la même orientation. Cela amène donc à huit possibilités de division, que nous appelons *situations*. Les autres valeurs des paramètres de modélisation sont celles indiquées en tableau 2.1 page 60 ; nous rappelons que ces valeurs sont choisies en complète cohérence avec les expériences biologiques menées en parallèle des simulations, et dont nous présenterons les résultats au chapitre 4. Le protocole des simulations consiste à effectuer 100 simulations dans chacune de ces huit situations.

Une fois les 800 simulations obtenues, notre protocole d'analyse par les diagnostics consiste à calculer les valeurs des indicateurs $R_1, R_2, \dots, R_6, \tilde{R}_6$:

1. Pour les diagnostics concernant la population entière (diagnostics n^{os} 1 et 2), nous calculons leur valeur sur la population entière.
2. Pour les diagnostics concernant les lignages, nous décidons pour chaque simulation d'extraire deux lignages (un lignage central et un lignage périphérique, selon la procédure de choix des lignages décrite en section 3.3.1 page 78) et d'appliquer les diagnostics n^{os} 3 à 6 à ces deux lignages uniquement. Le choix de distinguer lignages périphériques et lignages centraux a été motivé par les premières observations de quelques diagnostics, qui mettaient en lumière une forte disparité entre ces deux types de lignages. Comme nous le verrons par la suite, cette disparité se confirme sur tous les diagnostics, et est même observée sur les expériences biologiques.

Dans ce qui suit, le terme *situation* fait référence à une stratégie et une direction de division données ; le terme *zone* fait référence à la zone centrale ou à la zone périphérique pour les lignages. Par ailleurs, la direction de division « constante » ne donnant pas de résultats significativement différents et étant éloignée de la réalité de notre modèle biologique, nous ne la présentons pas dans les résultats qui suivent.

3.4.2 Résultats de chacun des diagnostics

Nous présentons dans cette section les résultats de tous les diagnostics, les uns après les autres, en suivant l'ordre dans lequel ils ont été définis.

Diagnostic n° 1 (sphéricité de la population)

La première lecture est, qu'en stratégie libre comme en stratégie contrainte, pour les trois directions, la distribution de R_1 à un profil gaussien, de moyenne située autour de 0,7 ; ces distributions sont illustrées en figures 3.8a et 3.8b page suivante. Cela permet de dire que, quelles que soient la stratégie et la direction, le profil global de la population cellulaire observée à travers ce diagnostic, est très similaire. En particulier, ce résultat tend à montrer que les situations de division testées n'ont qu'un impact très limité sur la forme globale de la population ; cela signifie que la loi de positionnement a un impact beaucoup plus fort sur la forme globale de la population que la loi de division.

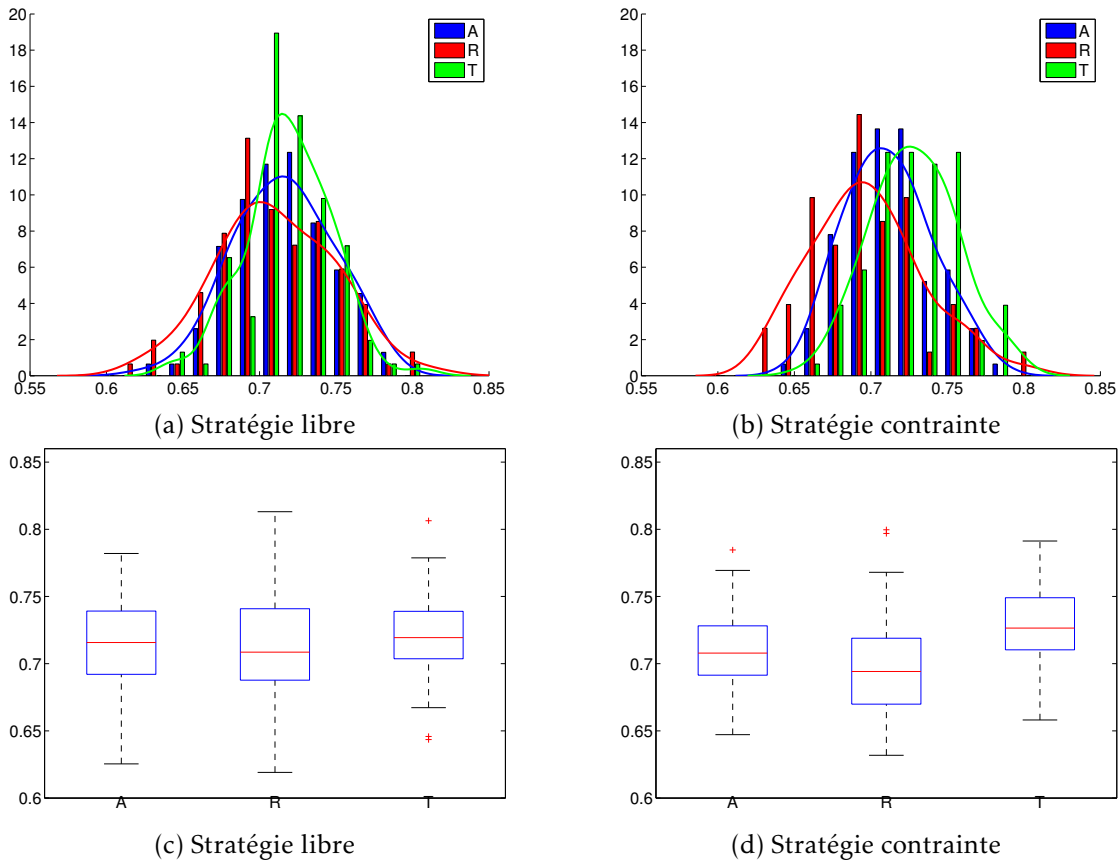
En deuxième lecture, nous pouvons illustrer quelques différences entre ces situations : en comparant les trois directions dans chaque stratégie, nous constatons que les différences entre les trois sont davantage prononcées en stratégie contrainte qu'en stratégie libre (cela est particulièrement visible sur une représentation graphique par diagramme en boîte, présentée en figures 3.8c et 3.8d page ci-contre). En particulier, une direction tangentielle à stratégie contrainte amène à une valeur distinctement plus élevée qu'une direction radiale. Ces deux observations peuvent s'interpréter de manière cohérente avec les choix de modélisation :

- une stratégie libre laisse l'orientation évoluer au cours de la division, alors qu'une stratégie contrainte la maintient fixée : il est alors cohérent que les différences entre les trois directions ressortent davantage en stratégie contrainte ;
- une cellule située au bord et se divisant radialement va apporter une irrégularité au bord plus forte qu'une cellule se divisant tangentiellement : cette irrégularité au bord va augmenter le périmètre de la population, sans vraiment en changer l'aire, et donc diminuer la valeur de R_1 .

Diagnostic n° 2 (convexité de la population)

Les observations que nous pouvons tirer de ce diagnostic sont en tout point cohérentes avec celles du diagnostic n° 1.

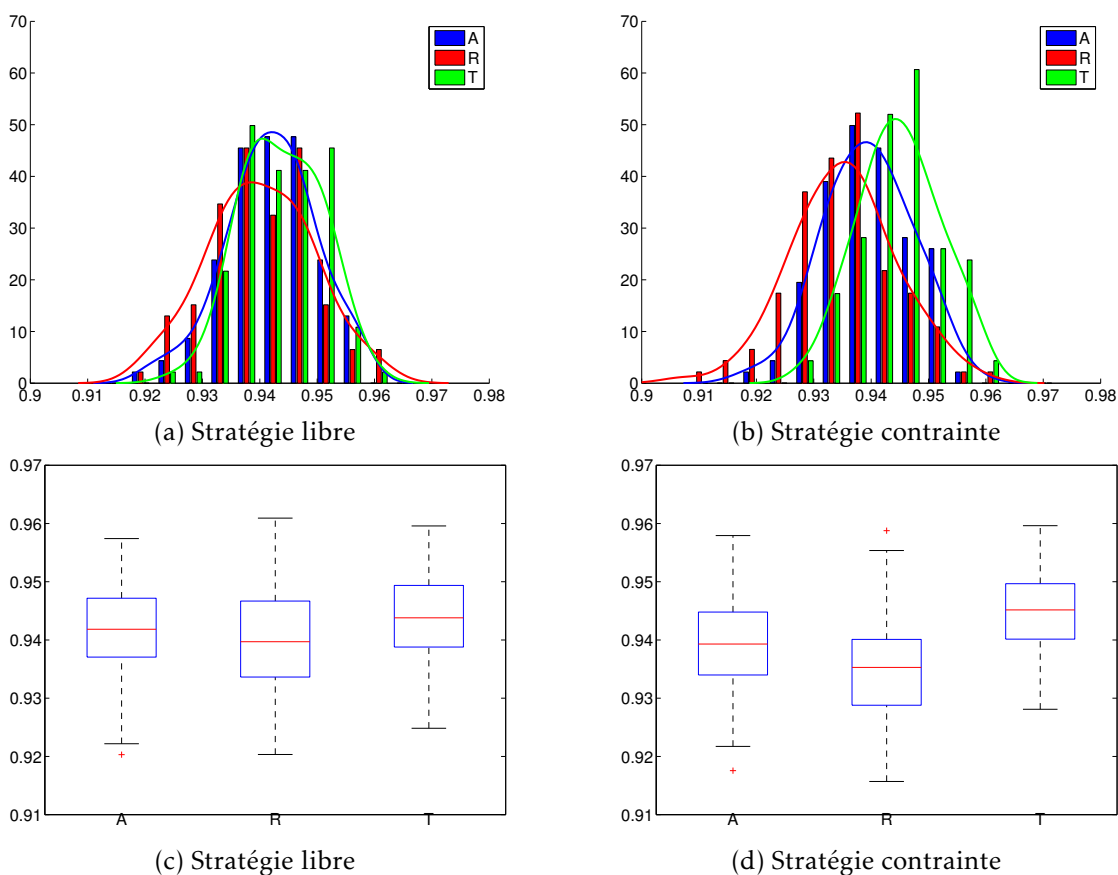
Quelles que soient la stratégie et la direction, le profil global de la population cellulaire observée à travers ce diagnostic est très similaire, autour d'une moyenne de 0,94, comme illustré en figures 3.9a et 3.9b page 84. Si nous regardons plus en détail les différentes situations, nous retrouvons que les différences entre les trois directions sont davantage prononcées en stratégie contrainte que libre, comme illustré en figures 3.9c et 3.9d page 84 ; autrement dit, le choix de la direction en stratégie libre n'a quasiment pas de conséquence. Par ailleurs, ces différences illustrent une plus forte valeur de R_2 pour la direction tangentielle que pour la direction radiale, ce que nous interprétons comme suit : la forme au bord est plus régulière lorsque la direction est tangentielle.



Légende. Répartition de R_1 , pour les trois directions, en stratégies libre pour (a), (c) et contrainte pour (b), (d). Les figures (a) et (b) contiennent chacune une superposition de trois histogrammes normalisés en aire, correspondant respectivement à une direction aléatoire (A – bleu), radiale (R – rouge) et tangentielle (T – vert). Les figures (c) et (d) contiennent chacune trois diagrammes en boîte correspondant respectivement à une direction aléatoire, radiale et tangentielle.

Observations. Les moyennes respectives en stratégie libre sont 0,715, 0,710 et 0,719 alors qu'en stratégie contrainte, elles sont 0,711, 0,697 et 0,728. En stratégie libre, les trois rectangles se chevauchent en grande partie alors qu'en stratégie contrainte, et particulièrement entre les directions radiale et tangentielle, ce chevauchement est beaucoup moins prononcé.

FIGURE 3.8 – Diagnostic n° 1 sur le rapport aire versus périmètre de la population (R_1), pour les simulations



Légende. Répartition de R_2 , pour les trois directions, en stratégies libre pour (a), (c) et contrainte pour (b), (d). Les figures (a) et (b) contiennent chacune une superposition de trois histogrammes normalisés en aire, correspondant respectivement à une direction aléatoire (A – bleu), radiale (R – rouge) et tangentielle (T – vert). Les figures (c) et (d) contiennent chacune trois diagrammes en boîte correspondant respectivement à une direction aléatoire, radiale et tangentielle.

Observations. Les moyennes respectives en stratégie libre sont 0,942, 0,940 et 0,944 alors qu'en stratégie contrainte, elles sont 0,939, 0,935 et 0,945. En stratégie libre, les trois rectangles se chevauchent en grande partie alors qu'en stratégie contrainte, et particulièrement entre les directions radiale et tangentielle, ce chevauchement est quasiment inexistant.

FIGURE 3.9 – Diagnostic n° 2 sur le rapport aire versus aire de l'enveloppe convexe de la population (R_2), pour les simulations

Diagnostic n° 3 (sphéricité d'un lignage)

Ce diagnostic mesure un rapport aire versus périmètre, tel que défini en équation (3.11) page 76 : notre objectif est d'étudier la filamentarité des lignages. Pour cela, nous traçons la répartition des composantes connexes de lignage (en proportion), en fonction de ce ratio R_3 . Ce premier diagnostic qui porte sur les lignages est l'occasion de détecter la disparité entre zone centrale et zone périphérique sur la morphologie des lignages.

Observations et conclusions principales Afin d'appuyer la disparité centre versus périphérie, nous regroupons nos données par zone, comme présenté en figure 3.10 page suivante. Ces histogrammes mettent en valeur que la proportion de petites valeurs de R_3 est nettement plus forte en zone périphérique qu'en zone centrale ; en particulier, quasiment la moitié (43 %) des composantes sont à $R_3 = 0$ en zone périphérique alors que seul un tiers (30 %) l'est en zone centrale. Nous pouvons en conclure que les lignages ont davantage tendance à rester agglomérés en zone centrale qu'en zone périphérique, ce qui peut s'interpréter par notre modélisation : comme une cellule est poussée par la croissance de toutes les cellules situées entre elle et le centre de la population, une cellule située en zone périphérique subira davantage ces poussées, qui l'empêcheront de rester groupée avec son lignage.

Observations et conclusions secondaires Toutes les données, situation par situation, sont représentées en figure 3.12 page 87. Un élément à noter est que la sphéricité des lignages est très faible, au regard de la sphéricité de la population entière fournie par R_1 ; en effet, la valeur moyenne de R_1 est proche de 0,7 alors que celle de R_3 est inférieure à 0,2.

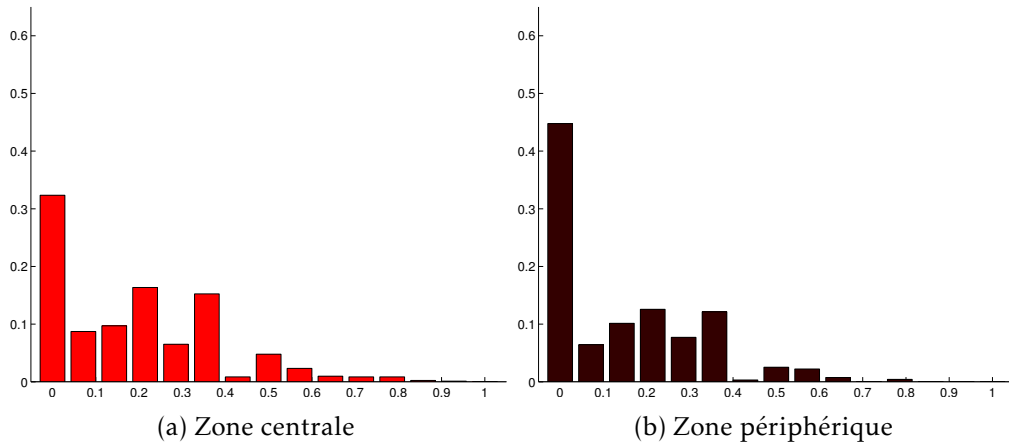
Diagnostic n° 4 (morcellement d'un lignage)

Ce diagnostic mesure le nombre de composantes connexes par lignage : notre objectif est d'étudier le morcellement des lignages. Pour cela, nous traçons un histogramme de la répartition des lignages (en proportion) en fonction de leur nombre de composantes connexes R_4 .

Là encore, comme pour le diagnostic précédent, le principal résultat que nous pouvons tirer est qu'une forte disparité se produit entre les lignages centraux et les lignages périphériques.

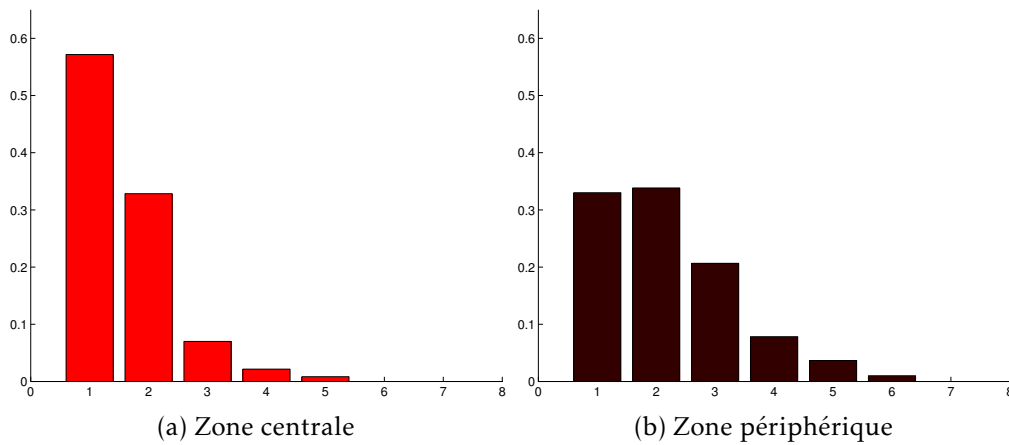
Observations et conclusions principales Afin d'appuyer la disparité centre versus périphérie, nous regroupons nos données par zone, comme présenté en figure 3.11 page suivante. Ces histogrammes mettent en valeur que :

- quasiment tous les lignages sont entre un et quatre (respectivement cinq) morceaux en zone centrale (resp. périphérique) ;
- en zone centrale, 57 % des lignages sont en un seul morceau et 33 % sont en deux morceaux, ce qui regroupe la quasi totalité (90 %) des lignages ;



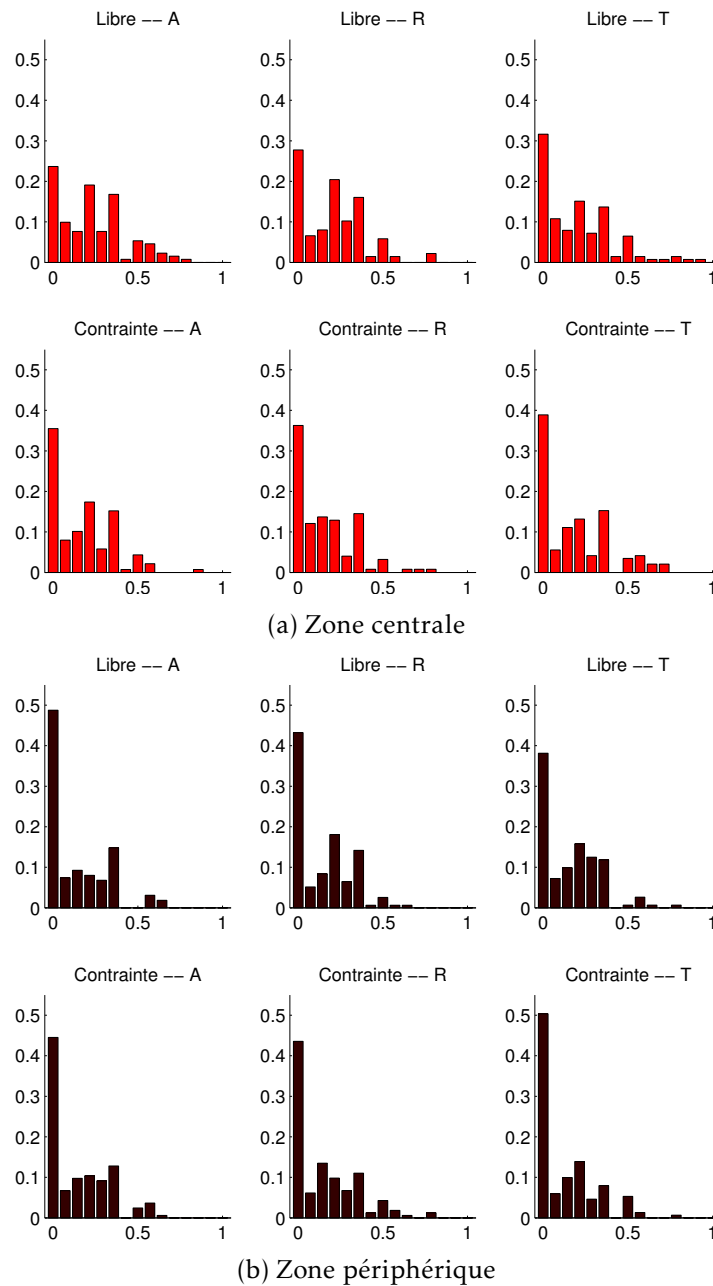
Légende. En abscisses figurent les valeurs de R_3 et en ordonnées la proportion de composantes connexes ; au centre (a) et en périphérie (b), toutes situations confondues.
Observations. Les composantes à $R_3 = 0$ sont 43 % en zone périphérique et seulement 30 % en zone centrale. Les valeurs de R_3 sont par ailleurs beaucoup plus faibles que celle de R_1 .

FIGURE 3.10 – Diagnostic n° 3 sur le rapport aire versus périmètre par composante connexe de lignage (R_3), pour les simulations



Légende. En abscisses figurent les valeurs de R_4 et en ordonnées la proportion de lignages ; au centre (a) et en périphérie (b), toutes situations confondues.
Observations. Les lignages à une ou deux composantes représentent $57 + 33 = 90\%$ des lignages centraux alors que cette proportion tombe à $33 + 34 = 67\%$ en zone périphérique (et qu'elle n'y est pas atteinte en ajoutant les lignages à trois composantes).

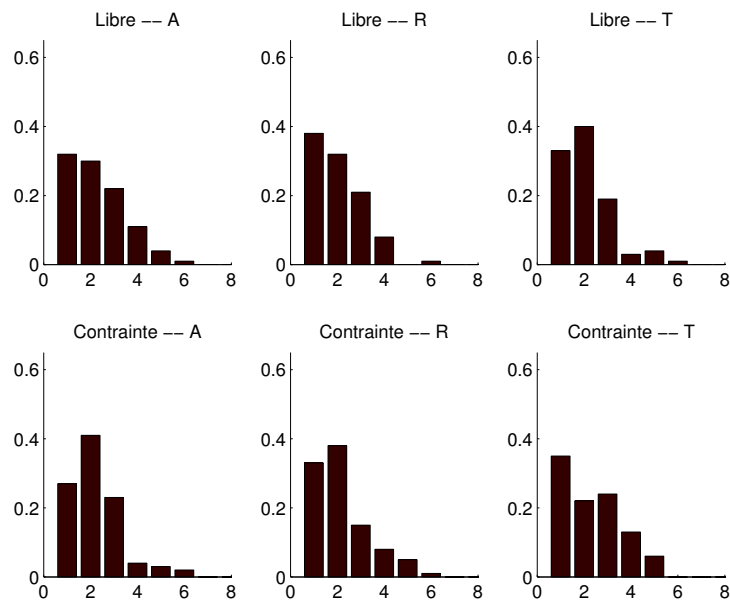
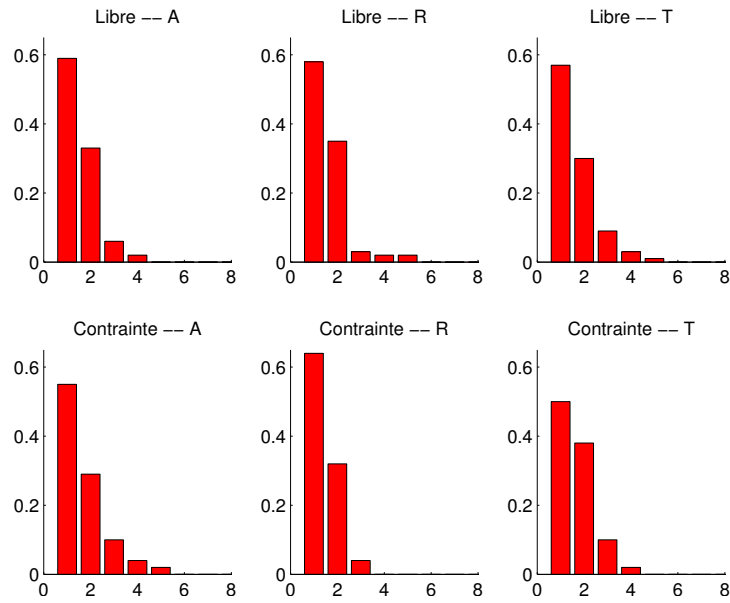
FIGURE 3.11 – Diagnostic n° 4 sur le nombre de composantes connexes par lignage (R_4), pour les simulations



Légende. En abscisses figurent les valeurs de R_3 et en ordonnées la proportion de composantes connexes ; au centre (a) et en périphérie (b). Les légendes A, R et T signifient respectivement une direction aléatoire, radiale et tangentielle.

Observations. Par rapport au diagnostic n° 1, qui mesure le même rapport mais pour la population, la répartition est différente et fortement concentrée sur des faibles valeurs.

FIGURE 3.12 – Diagnostic n° 3 sur le rapport aire versus périmètre par composante connexe de lignage (R_3), pour chacune des six situations des simulations



Légende. En abscisses figurent les valeurs de R_4 et en ordonnées la proportion de lignages ; au centre (a) et en périphérie (b). Les légendes A, R et T signifient respectivement une direction aléatoire, radiale et tangentielle.

Observations. À direction fixée, la différence entre les deux stratégies est peu visible en zone centrale, alors qu'elle l'est davantage en zone périphérique.

FIGURE 3.13 – Diagnostic n° 4 sur le nombre de composantes connexes par lignage (R_4), pour chacune des six situations des simulations

- en zone périphérique, ces proportions sont respectivement de 33 % et 34 %, ce qui regroupe 67 % des lignages.

Nous pouvons conclure que cela appuie le résultat principal du diagnostic n° 3 sur la disparité entre lignages centraux et lignages périphériques. Ensuite, nous pouvons affirmer que les lignages sont davantage morcelées en zone périphérique qu'en zone centrale, ce qui est cohérent avec notre interprétation suivante de la modélisation : les cellules d'un même lignage situé à la périphérie, de par la croissance de toutes les cellules situées entre elles et le centre de la population, se retrouvent séparées les unes des autres.

Observations et conclusions secondaires Les données complètes, situation par situation, sont représentées en figure 3.13 page 88. Nous pouvons remarquer que deux cas se distinguent des autres : en zone centrale, la situation contrainte radiale n'a quasiment que des lignages en un ou deux morceaux ; en zone périphérique, la situation contrainte tangentielle possède beaucoup de lignages en quatre ou cinq fragments. Une explication possible, commune à ces deux cas, est la suivante : après une division tangentielle en zone périphérique, la croissance de la population cellulaire rend très probable qu'une tierce cellule vienne s'intercaler entre ces deux cellules sœurs ; ce même phénomène d'intercalation d'une tierce cellule est peu probable dans le cas d'une division radiale en zone centrale, où le lignage peut s'étendre radialement.

Diagnostic n° 5 (taille des fragments d'un lignage)

Ce diagnostic mesure le nombre de cellules par composante connexe de lignage : notre objectif est de mettre en valeur les plus grosses composantes (celles contenant huit cellules ou davantage⁷) et les plus petites (celles avec une ou deux cellules). Pour cela, nous traçons un histogramme de la répartition des composantes connexes (en pourcentage cumulé) en fonction de leur nombre de cellules R_5 ; ensuite, nous regardons l'évolution de ce pourcentage entre cinq zones (de zéro à une cellule, de une à deux, de deux à sept, de sept à huit, de huit et plus), ce qui consiste à tracer les droites d'interpolation des extrémités ces cinq zones. Cela permet de mettre en valeur l'augmentation du nombre de petites et de grosses composantes.

Observations et conclusions principales Afin d'appuyer la disparité centre versus périphérie, là encore nous regroupons nos données par zone, comme présenté en figure 3.14 page 91. Ces histogrammes mettent en valeur que :

- en zone périphérique, il y a un pourcentage très fort (presque 50 %) de petites composantes connexes et un pourcentage faible (16 %) de grosses composantes ;
- en zone centrale, il y a un pourcentage fort (presque 30 %) de petites composantes connexes et un pourcentage de même ordre (34 %) de grosses composantes ;

7. Comme nous simulons la croissance des populations cellulaires sur une durée moyenne de trois temps de génération, il y a en moyenne huit cellules dans chaque lignage à la fin.

- en zone centrale et en zone périphérique, le pourcentage de lignages de taille moyenne (trois à sept cellules) est quasiment identique (autour de 7,5%), ce qui appuie que les deux observations précédentes correspondent vraiment à un basculement entre petites et grosses composantes.

Nous pouvons conclure que ces résultats viennent confirmer la disparité centre versus périphérie et compléter ceux du diagnostic n° 4. Ils permettent d'affirmer que les composantes connexes en zone périphérique, qui sont en nombre plus important qu'en zone centrale d'après le diagnostic n° 4, correspondent à des très petites composantes (une à deux cellules) et non à des composantes de taille moyenne. Autrement dit, davantage de cellules se retrouvent isolées de leur lignage.

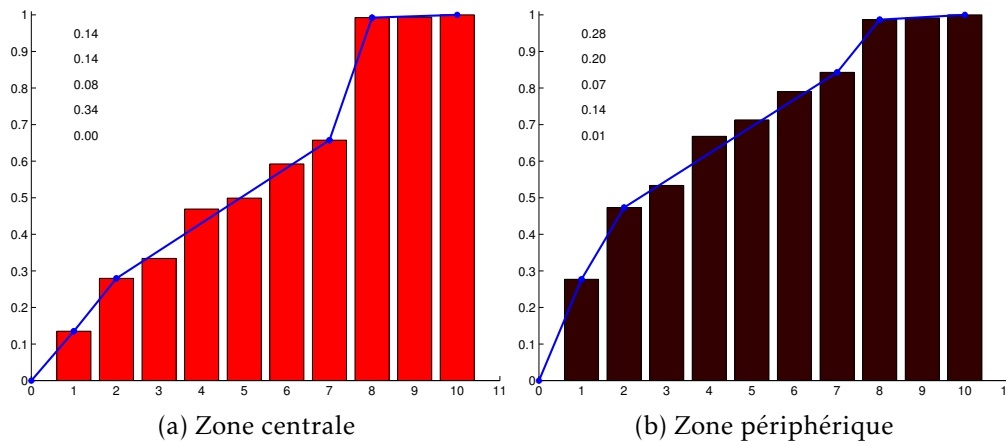
Observations et conclusions secondaires Les données complètes, situation par situation, sont représentées en figure 3.16 page 92. Nous pouvons remarquer qu'en stratégie libre les trois directions donnent des résultats très similaires, alors qu'en stratégie contrainte des différences deviennent davantage visibles. Deux cas sont remarquables :

- la situation contrainte radiale, qui possède beaucoup de grosses composantes en zone centrale. Cela est cohérent avec le diagnostic n° 4 et peut s'interpréter comme suit : dans une telle situation, un lignage va pouvoir s'étendre de manière radiale. En zone périphérique, cette situation diffère moins des autres directions. Une interprétation est, qu'en zone périphérique, la loi de positionnement s'impose fortement sur la loi de division : un lignage se divisant radialement au bord de la population ne va pas se positionner radialement mais va se rabattre le long du bord ;
- la situation contrainte tangentielle qui, en zone périphérique, possède beaucoup de composantes d'une cellule. Là encore, cela peut s'interpréter avec le modèle : de par la poussée de toutes les cellules présentes entre elles et le centre de la population, deux cellules s'étant divisées tangentiellement risquent de voir une cellule s'intercaler en travers d'elles.

Diagnostic n° 6 (orientation d'un lignage)

Ce diagnostic mesure la direction principale d'orientation (R_6) d'un lignage, ainsi que son intensité (\tilde{R}_6) : notre objectif est de mettre en valeur une orientation particulière des lignages. Pour cela, nous traçons un histogramme de la répartition des lignages (en pourcentage cumulé) en fonction de leur écart à la direction tangentielle, c'est-à-dire en fonction de $|R_6|$. Ainsi, une concentration de l'histogramme proche de 0 indique une orientation tangentielle et une concentration proche de $\frac{\pi}{2}$ indique une orientation radiale.

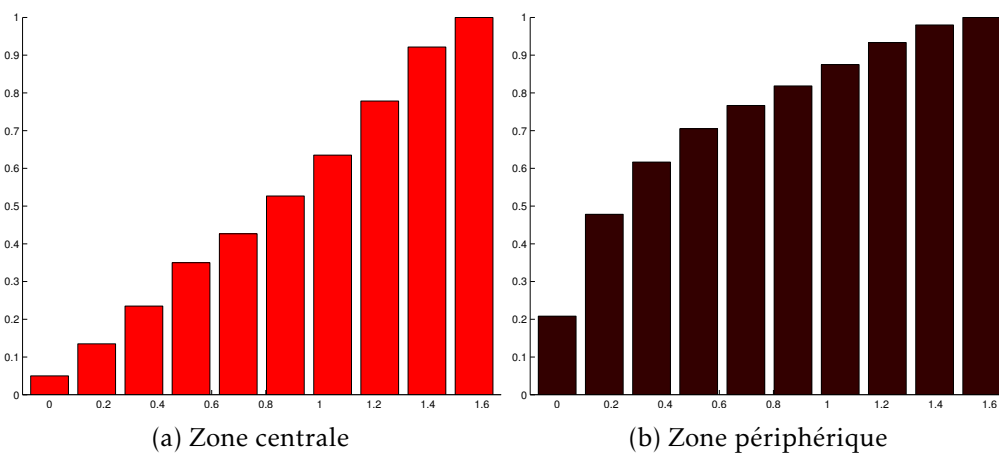
Observations et conclusions principales Afin d'appuyer la disparité centre versus périphérie, là encore nous regroupons nos données par zone, comme présenté en figure 3.15 page ci-contre. Ces histogrammes mettent en valeur que :



Légende. En abscisses figurent les valeurs de R_5 et en ordonnées la proportion cumulée de composantes connexes ; au centre (a) et en périphérie (b), toutes situations confondues. Les pentes des cinq droites bleues sont indiquées dans le graphique.

Observations. Les composantes à huit cellules sont 34 % des composantes en périphérie mais deux fois moins (16 %) au centre. Comme la proportion de composantes possédant trois à sept cellules est similaire (7,5 %), la différence est donc sur celles d'une à deux cellules, qui sont moins d'un tiers (28 %) au centre et la moitié en périphérie (48 %).

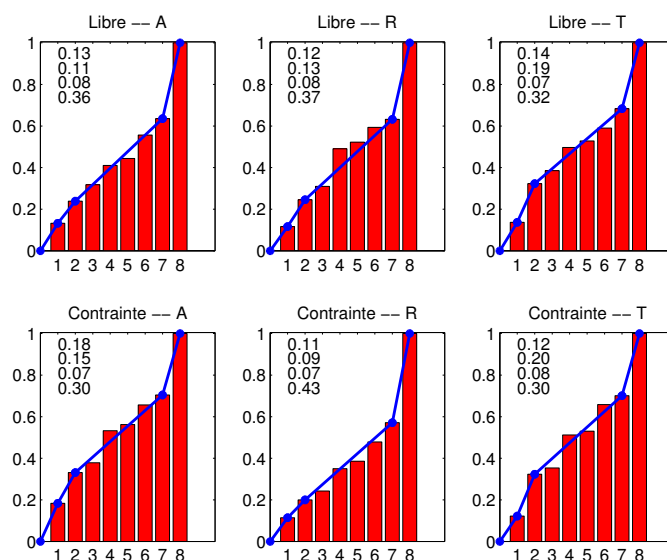
FIGURE 3.14 – Diagnostic n° 5 sur le nombre de cellules par composante connexe de lignage (R_5), pour les simulations



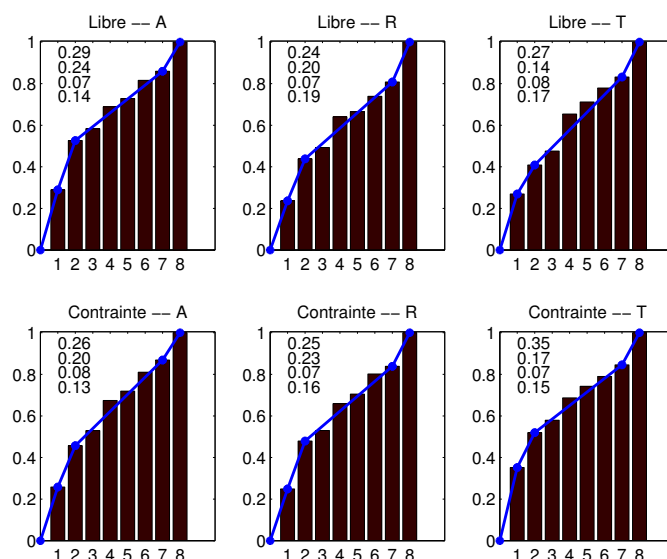
Légende. En abscisses figurent les valeurs de $|R_6|$ et en ordonnées la proportion cumulée de lignages ; au centre (a) et en périphérie (b), toutes situations confondues. L'axe des abscisses est divisé en dix morceaux, couvrant les plages angulaires suivantes (exprimées en degrés) : $[0; 5]$, $[5; 15]$, ..., $[75; 85]$, $[85; 90]$.

Observations. En périphérie, 48 % des lignages (contre 14 % au centre) sont dans les deux premières colonnes, c'est-à-dire dans un écart de 0° à 15° de la direction tangentielle.

FIGURE 3.15 – Diagnostic n° 6 sur l'écart de l'orientation des lignages à la direction tangentielle ($|R_6|$), pour les simulations



(a) Zone centrale

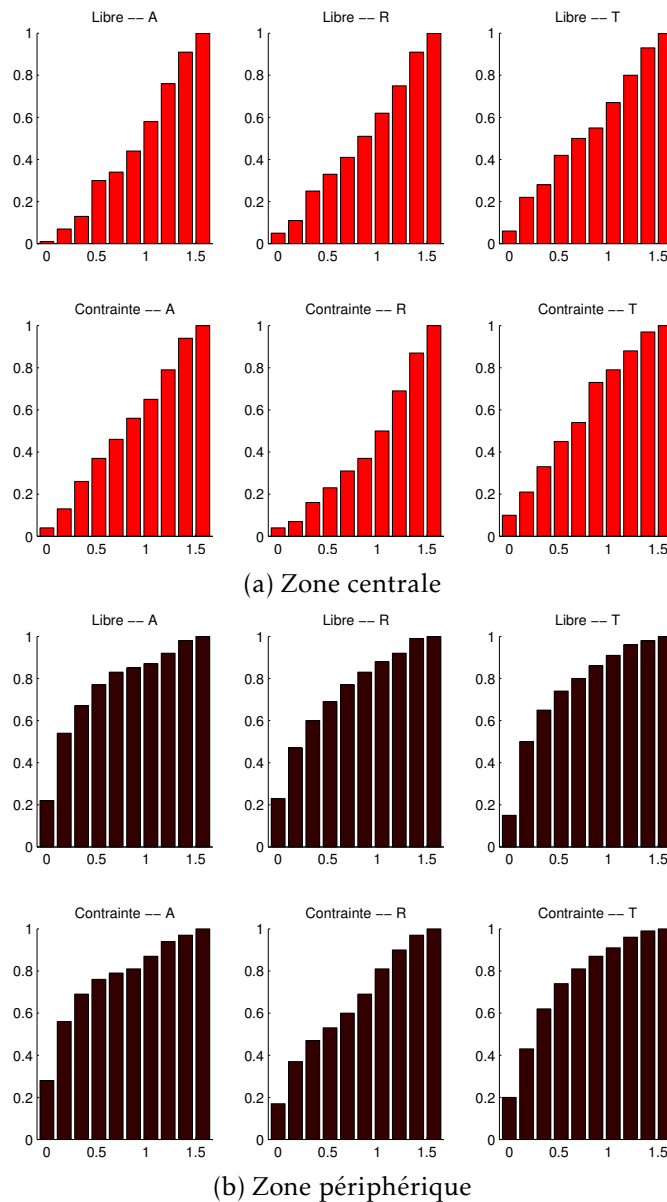


(b) Zone périphérique

Légende. Similaire à la figure 3.14 page 91, sauf que (a) et (b) sont pour les six situations. Les A, R et T signifient respectivement une direction aléatoire, radiale et tangentielle.

Observations. À direction fixée, la différence entre les deux stratégies est peu visible. Par contre, deux situations à stratégie contrainte ont un profil particulier : le cas radial en zone centrale à une plus forte proportion de composantes à huit cellules que les deux autres cas centraux ; le cas tangential en zone périphérique à une plus forte proportion de composantes à une cellules que les deux autres cas périphériques.

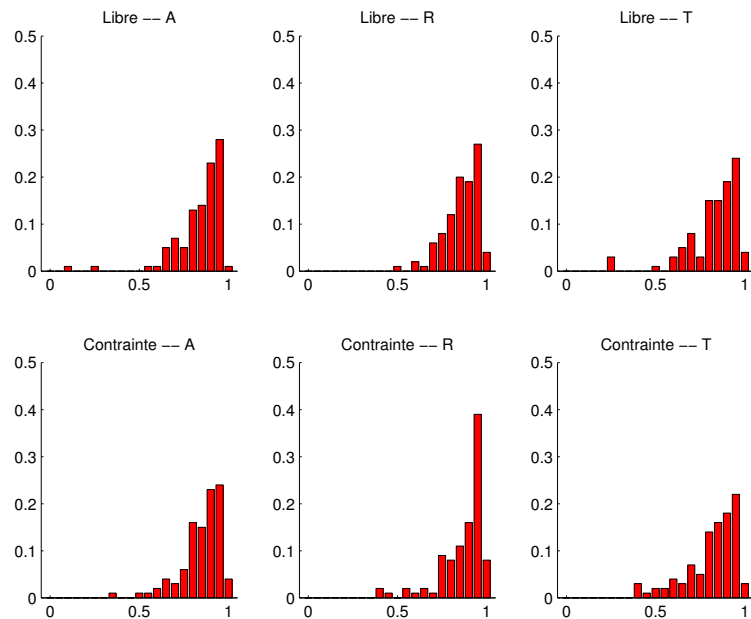
FIGURE 3.16 – Diagnostic n° 5 sur le nombre de cellules par composante connexe de lignage (R_5), pour chacune des six situations des simulations



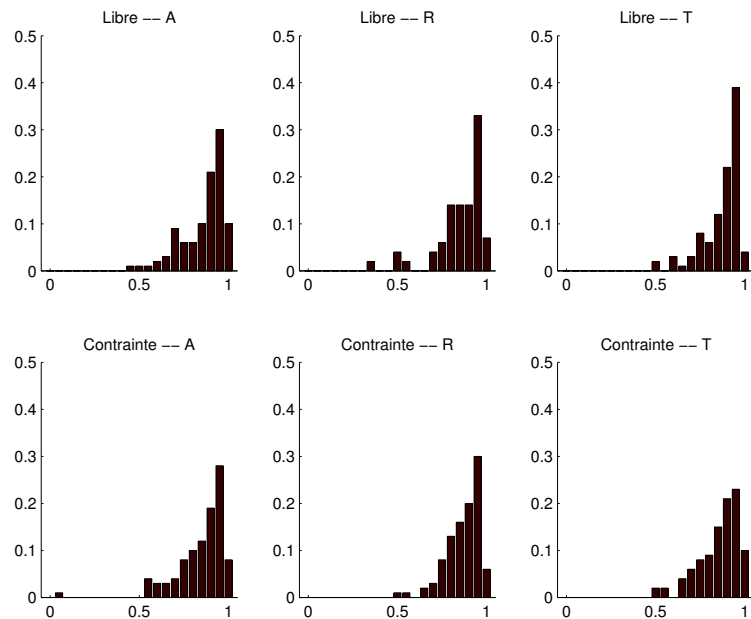
Légende. Similaire à la figure 3.15 page 91, sauf que (a) et (b) sont pour les six situations. Les A, R et T signifient respectivement une direction aléatoire, radiale et tangentielle.

Observations. En stratégie libre, il y a peu de différences entre les trois directions, que ce soit au centre ou en périphérie. La situation contrainte radiale se distingue des autres : au centre, elle est radialement orientée et, en périphérie, son orientation tangentielle est moins prononcée.

FIGURE 3.17 – Diagnostic n° 6 sur l'écart de l'orientation des lignages à la direction tangentielle ($|R_6|$), pour chacune des six situations des simulations



(a) Zone centrale



(b) Zone périphérique

Légende. En abscisses figurent les valeurs de \tilde{R}_6 et en ordonnées la proportion de lignages ; au centre (a) et en périphérie (b). Les légendes A, R et T signifient respectivement une direction aléatoire, radiale et tangentielle.

FIGURE 3.18 – Diagnostic n° 6 sur l'excentricité des lignages (\tilde{R}_6), pour chacune des six situations des simulations

- en zone périphérique, il y a un pourcentage très fort (presque 50 %) de lignages dont l'orientation fait, avec la direction tangentielle, un angle inférieur à 15° (la plage $[0^\circ; 15^\circ]$ représente 16,7 % de l'intervalle d'appartenance de $|R_6|$, qui est $[0^\circ; 90^\circ]$);
- en zone centrale, aucune orientation ne semble privilégiée.

Nous pouvons conclure que ces résultats viennent appuyer la disparité centre versus périphérie. Ils permettent d'affirmer qu'en zone périphérique, l'orientation tangentielle est fortement privilégiée par les lignages. Cela est cohérent avec l'interprétation suivante de la modélisation : des lignages situés en bord de la population vont, pour maintenir la forme arrondie de la population, se répartir de manière tangentielle.

Observations et conclusions secondaires Les données complètes, situation par situation, sont représentées en figure 3.17 page 93 pour l'orientation. Nous pouvons remarquer qu'en stratégie libre les trois directions donnent des résultats très similaires, alors qu'en stratégie contrainte des différences deviennent davantage visibles. Ensuite, la situation contrainte radiale est un cas qui se distingue :

- en zone centrale, une orientation radiale se distingue pour beaucoup de lignages (en effet, 50 % des lignages ont une orientation faisant un angle inférieur à 25° avec la direction radiale);
- en zone périphérique, bien que tangentiellement orientée comme les autres directions, cette orientation est moins prononcée (en effet, en direction tangentielle ou aléatoire, les trois-quarts des lignages font un écart à l'orientation tangentielle de moins de 35° , alors que cette proportion n'est atteinte en direction radiale qu'en considérant les lignages jusqu'à un écartement de 60°).

Les données complètes, situation par situation, sont représentées en figure 3.18 page 94 pour l'excentricité mais nous n'en tirons aucune information particulière.

3.4.3 Conclusions communes des différents résultats

Concernant les lignages, le principal enseignement de nos diagnostics est qu'une forte disparité se produit entre les lignages situés en zone centrale et ceux situés en zone périphérique : ce constat est visible sur tous les diagnostics qui traitent des lignages (n^{os} 3 à 6). Cette disparité s'illustre par :

1. De plus nombreux et de plus petits fragments de lignages en zone périphérique. Cela peut s'expliquer par le fait que plus une cellule est éloignée du centre, plus elle est poussée par la croissance des cellules de la population et donc plus la probabilité qu'elle se retrouve écartée de ses voisines (par d'autres cellules s'intercalant entre-elles) est grande ; cette explication impliquerait des lignages peu morcelés lors d'une division radiale, ce qui est notre dernière conclusion ci-dessous.
2. Une orientation tangentielle en zone périphérique. Une interprétation possible est qu'en zone périphérique la loi de positionnement, en ce qui concerne l'orientation des lignages, prend le pas sur la loi de division. Aucune direction ne

semble privilégiée en zone centrale, sauf dans la situation contrainte radiale : cela laisse supposer qu'en zone centrale, la loi de division s'impose sur la loi de positionnement lors d'une direction radiale de division.

Par ailleurs, concernant la stratégie libre de division, les diagnostics n^{os} 1, 2, 5 et 6 mettent en valeur peu de différences entre les trois directions proposées. Cela tend à conclure que la stratégie libre a très peu d'impact sur la forme globale comme sur les lignages, contrairement à la stratégie contrainte. Ce phénomène est cohérent avec la modélisation, puisque la stratégie libre consiste à laisser l'orientation de la cellule mère évoluer au cours du processus de déformation. C'est uniquement en stratégie contrainte que les diagnostics n^o 1 et 2 détectent des différences entre les directions radiales et tangentiels, qui tendent à conclure que le bord est plus régulier pour cette 2^e direction.

Enfin, la conjonction des diagnostics n^{os} 4, 5 et 6 permet de préciser que la situation contrainte radiale en zone centrale est souvent remarquable : elle est radialement orientée et peu morcelée, avec de gros fragments ; notre interprétation est que, dans cette situation, les lignages peuvent s'étendre et se développer radialement, sans en être empêchés par d'autres phénomènes.

An ancestor of mine maintained that if you eliminate the impossible, whatever remains — however improbable — must be the truth.

Spock (Leonard Nimoy), *Star Trek VI*

Chapitre 4

Expériences biologiques et confrontation avec le modèle

L'objectif de ce chapitre est de fournir une confrontation entre le modèle mathématique et les expériences biologiques, pour les comparer. Cette confrontation permet d'évaluer si les résultats obtenus après l'étude statistique des simulations (au chapitre 3) se retrouvent dans les cultures cellulaires des biologistes. Elle permet aussi d'apporter des pistes d'amélioration de la modélisation, qui seront exploitées au chapitre 5. Afin d'obtenir une confrontation la plus pertinente possible, les protocoles des simulations et des expériences sont conçus pour être très proches ; de plus, nous proposons d'étudier les expériences avec précisément les mêmes outils que pour les simulations : les diagnostics établis en section 3.2.

Dans ce chapitre, nous expliquons quelles sont les informations extraites des données expérimentales (section 4.1) ; puis nous étudions les expériences grâce aux mêmes diagnostics que ceux établis pour l'étude des simulations (section 4.2) ; enfin nous comparons les résultats ainsi obtenus à ceux du modèle (section 4.3). Cette démarche permet d'établir une confrontation entre le modèle et les expériences.

Sommaire du chapitre

4.1	Utilisation des données expérimentales	98
4.1.1	Post-traitement des données expérimentales	98
4.1.2	Préparation des données expérimentales post-traitées	99
4.2	Application des diagnostics aux données expérimentales post-traitées . . .	104
4.2.1	Protocole d'analyse	104
4.2.2	Utilisation des diagnostics sur les données expérimentales	105
4.2.3	Observations et conclusions des diagnostics	109
4.3	Confrontation avec les résultats des simulations	112
4.3.1	Comparaison entre les résultats des simulations et les résultats des expériences	112
4.3.2	Bilan de la modélisation grâce à la confrontation	115

4.1 Utilisation des données expérimentales

Les données expérimentales sont des films montrant l'évolution de cellules sur plusieurs jours. Ces cellules miment le développement et l'organisation de cellules cancéreuses s'organisant en population cellulaire : c'est ce que nous appelons le « modèle biologique ».

Les données expérimentales sont post-traitées en section 4.1.1, afin d'en extraire des informations pertinentes ; ensuite, en section 4.1.2, nous choisissons quelles informations extraire et nous illustrons leur utilisation, en représentant les situations initiale et finale des expériences par les mêmes outils que pour les simulations.

L'objectif est de disposer de toutes les données nécessaires afin de pouvoir, en section 4.2, appliquer les diagnostics sur les expériences.

4.1.1 Post-traitement des données expérimentales

Les données expérimentales brutes sont obtenues de la manière suivante : des îlots de cellules d'adénocarcinome colique, de la lignée HCT116, sont filmés pendant trois jours au vidéomicroscope. Ces cellules ont pour caractéristiques :

- d'être adhérentes au support ;
- d'avoir été génétiquement modifiées afin d'exprimer une protéine associée à l'*ADN*, l'histone H₂B, fluorescente.

La protéine fluorescente permet de visualiser le noyau de chaque cellule de la population : cela rend possible, lors du post-traitement des données brutes, la segmentation et le suivi des noyaux au cours du temps ; nous insistons sur le fait que cette fluorescence détecte les *noyaux* des cellules, et non les cellules elles-mêmes. Le détail du protocole expérimental figure en annexe A page 145.

Une fois ces données expérimentales obtenues, le post-traitement consiste à effectuer le travail suivant, sur chaque film.

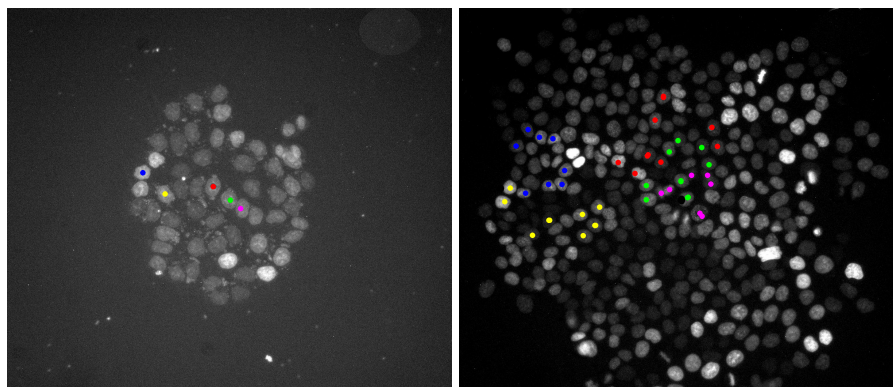
Suivi de lignages. Des cellules sont choisies à l'instant initial et leur filiation est manuellement suivie à chaque temps. Ce suivi consiste à marquer les cellules du lignage, sur *chaque* image du film (une illustration est proposée en figure 4.1 page ci-contre). Le choix des cellules est expliqué en remarque 4.1 page suivante.

Segmentation. Un programme (sur MatLab) est utilisé pour réaliser la segmentation de chaque noyau par rapport au niveau de fluorescence. Ce programme génère des images à partir desquelles la segmentation est vérifiée manuellement : lorsqu'une erreur est détectée, elle est corrigée à la main. La segmentation n'est faite que sur l'image du temps initial et sur celle du temps final. Un résultat de segmentation est proposé en figure 4.2 page 100.

L'objectif du post-traitement est que chaque agent de la population soit clairement identifié et qu'il soit possible d'extraire les informations utiles à la confrontation du modèle et des expériences¹. Ce travail a été réalisé en collaboration étroite avec l'ITAV :

1. Comme nous allons le détailler en section 4.1.2 page suivante, il s'agit des informations permettant d'appliquer les diagnostics sur les expériences biologiques ; en particulier, les seules informations à extraire

le résultat, toujours illustré avec la même expérience, est présenté en figure 4.3 page suivante.



(a) Temps initial

(b) Temps final

Légende. Chaque gomme colorée a été ajoutée à la main dans un logiciel de traitement et d'analyse d'images.

Observations. Les noyaux sont parfois collés les uns aux autres et souvent d'intensité de fluorescence différente : cela va rendre difficile la segmentation.

FIGURE 4.1 – Expérience : suivi de certains lignages

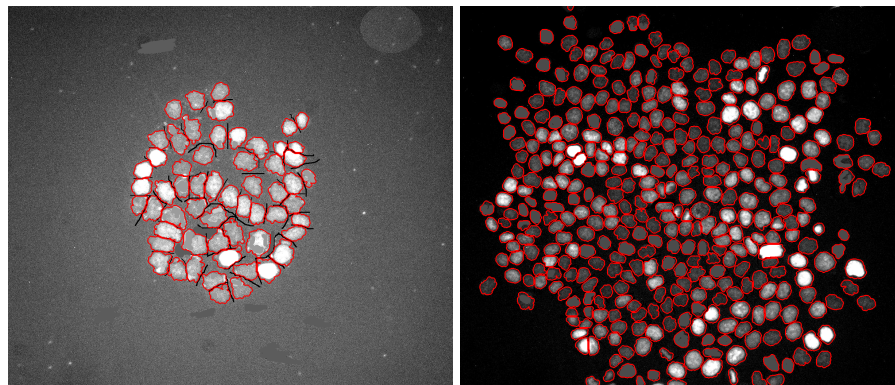
Remarque 4.1 (Choix, au temps initial, des cellules dont le lignage est suivi). Ce choix des cellules à l'instant initial a, dans un premier temps, été guidé par la facilité de suivi. Dans un deuxième temps (après les premières exploitations des résultats), il a été guidé par la volonté de suivre des lignages centraux et périphériques (la signification précise de ces deux types de lignage, basée sur la représentation des populations cellulaires par diagramme de Voronoï adapté, figure en définition 4.2 page 105). Ainsi, le choix des lignages pertinents à suivre est fait en fonction de la représentation graphique (mathématique), qui elle-même n'est accessible qu'après une première exploitation des données expérimentales post-traitées : cet exemple est une illustration des allers-retours entre mathématiques et biologie, réalisables grâce à l'interaction entre les deux disciplines.

4.1.2 Préparation des données expérimentales post-traitées

Cette préparation consiste à sélectionner et à extraire, parmi les données expérimentales post-traitées, des informations ciblées : celles permettant d'appliquer les diagnostics sur les expériences.

Définition 4.1 (Région d'intérêt). Nous appelons région d'intérêt chacune des zones détectées lors de la segmentation au cours du traitement d'image.

concernent le temps initial et le temps final, raison pour laquelle la segmentation n'est faite qu'à ces deux temps.

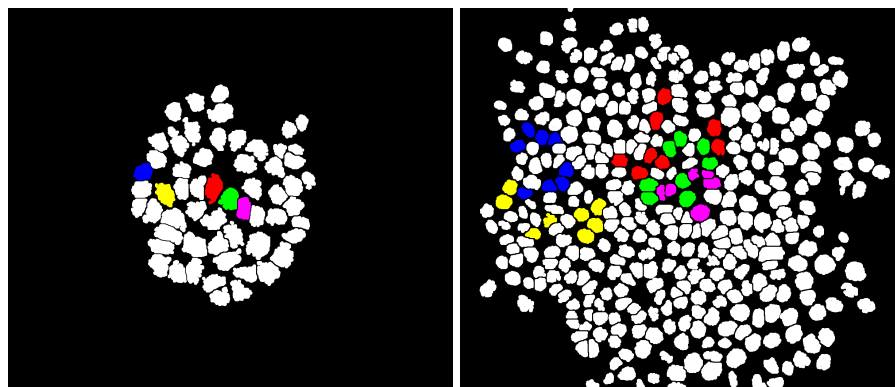


(a) Temps initial

(b) Temps final

Légende. Les traits noirs et les zones grisées correspondent à des corrections de l'image, effectuées à la main dans un logiciel de traitement et d'analyse d'images, pour corriger les erreurs de segmentation.

FIGURE 4.2 – Expérience : délimitation des régions détectées par la segmentation



(a) Temps initial

(b) Temps final

Légende. Les pixels correspondant à un noyau sont en blanc et les autres en noir. Pour l'illustration graphique, les noyaux appartenant aux lignages suivis sont colorés.

FIGURE 4.3 – Expérience : représentation après la segmentation

Une fois la segmentation correctement et totalement effectuée, chaque région d'intérêt correspond au noyau d'une cellule. Nous extrayons alors, pour chacune de ces régions d'intérêt :

1. L'abscisse de son barycentre.
2. L'ordonnée de son barycentre.
3. Son équivalent diamètre (il s'agit du diamètre du disque qui a une surface de même aire que cette région d'intérêt).
4. Sa couleur (afin d'avoir la même coloration dans toute représentation graphique basée sur ces données extraites).
5. Un entier indiquant son lignage (entier non nul pour chacun des lignages suivis et 0 pour *tous* les autres lignages non suivis ²).

Pour illustrer l'utilisation de cette extraction, nous commençons par effectuer une représentation graphique des instants initial et final de chaque expérience, à l'instar de ce que nous avons fait pour les simulations en section 3.1 page 64. Cela remplit plusieurs objectifs :

- s'assurer, grâce à une représentation visuelle, que nous avons bien exporté toutes les données nécessaires ;
- disposer d'une *même* représentation graphique pour les simulations et pour les expériences, qui donne des résultats pertinents et acceptables dans les deux cas.

Dans un premier temps, nous effectuons une représentation par des disques, comme nous le proposons pour les simulations en figure 3.1 page 64 : le résultat est présenté en figure 4.4 page suivante. Cette figure illustre qu'une telle représentation, à partir des informations extraites, n'est pas du tout satisfaisante ; en particulier, deux critiques peuvent être émises : d'énormes interstices sont visibles et des disques se superposent. Cela :

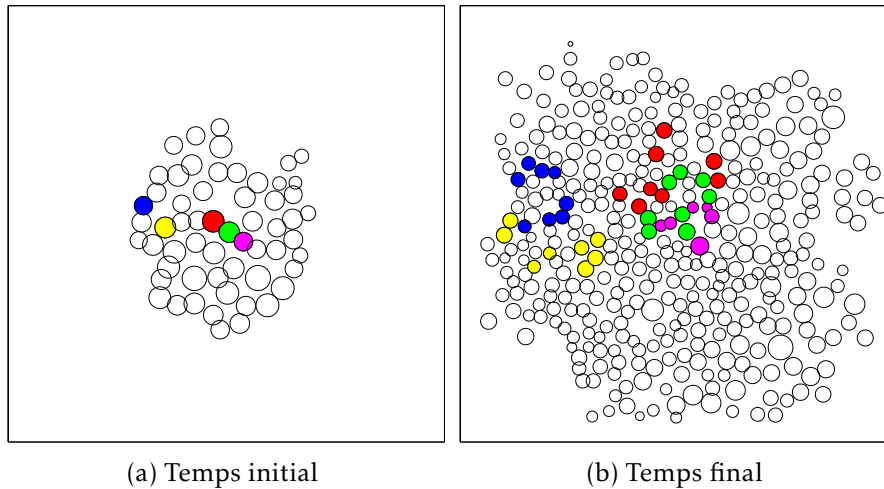
- n'est pas acceptable, en particulier au regard du phénomène biologique de regroupement sans interpénétration qui est au cœur de notre modélisation ;
- ne correspond pas du tout à une configuration biologique observée. En effet, il s'agit uniquement d'un artefact de la représentation, par des disques, d'objets n'ayant pas une forme circulaire.

Ce comportement était annoncé en section 3.1.1 page 64, où il motivé l'utilisation d'une meilleure représentation graphique.

Afin d'améliorer la visualisation des données recueillies, nous proposons deux autres représentations graphiques :

1. Une représentation des noyaux par des points, comme présenté en figure 4.5 page 103. Un tel choix se libère des deux critiques faites sur la représentation par disques mais ne rend plus compte de la taille de chaque cellule. Par ailleurs, les relations de voisinage y sont difficiles à visualiser.

2. En effet, nous avons vu en section 4.1.1 page 98 que nous ne suivions pas tous les lignages lors du post-traitement des données expérimentales.



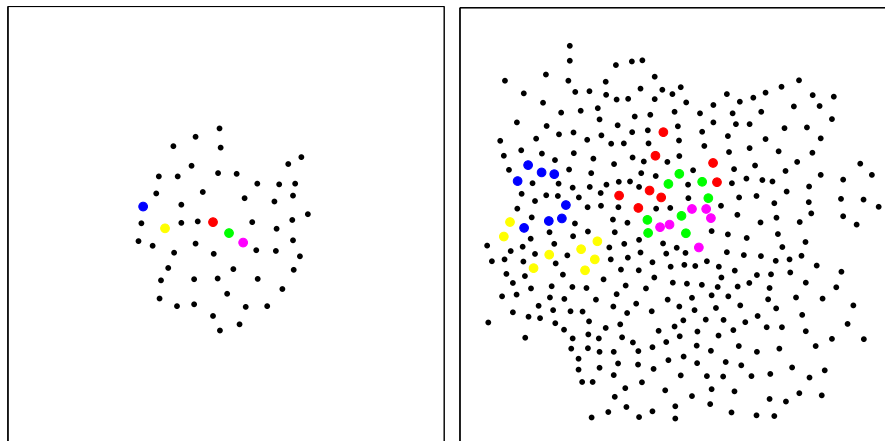
(a) Temps initial (b) Temps final
Légende. Les disques colorés correspondent aux cellules dont le lignage est suivi.
Observations. Des disques se chevauchent et il y a trop d'interstices vides.

FIGURE 4.4 – Expérience : représentation par disques des noyaux

2. Une représentation de la population cellulaire par diagramme de Voronoï adapté (voir définition 3.5 page 68), comme présenté en figure 4.6 page suivante : chaque point représente le barycentre d'un noyau et chaque région de Voronoï représente une cellule. Ce choix a le double avantage de :
 - se libérer des deux critiques précédentes ; en effet, de part la construction même du diagramme de Voronoï, tout point du plan appartient à une région de Voronoï (il n'y a donc aucun interstice) et les intérieurs de deux régions de Voronoï sont d'intersection vide (il n'y a donc aucune interpénétration) ;
 - de fournir une représentation visuelle bien plus proche de la réalité ; en effet :
 - les régions de Voronoï sont de formes polygonales, ce qui correspond davantage (que des disques) à la forme de cellules ; une telle proximité de représentation est déjà signalée dans [Hon78 ; SM05],
 - les régions de Voronoï sont de taille et de forme assez variables, ce qui rend compte de la diversité d'une population cellulaire,
 - les noyaux et les cellules y sont représentés, ce qui permet de bien distinguer les deux³ et de visualiser que le noyau n'est pas nécessairement au centre de la cellule.

Remarque 4.2. Nous pouvons remarquer que ces deux nouveaux choix n'utilisent pas le paramètre « équivalent diamètre » que nous extrayons. Cela n'est pas surprenant : pour une région d'intérêt ayant une forme très éloignée d'un disque, cette information est peu pertinente.

3. Nous rappelons que dans la modélisation nos agents représentent des *cellules* et que, dans les expériences, ce sont les *noyaux des cellules* qui sont détectés par la fluorescence.



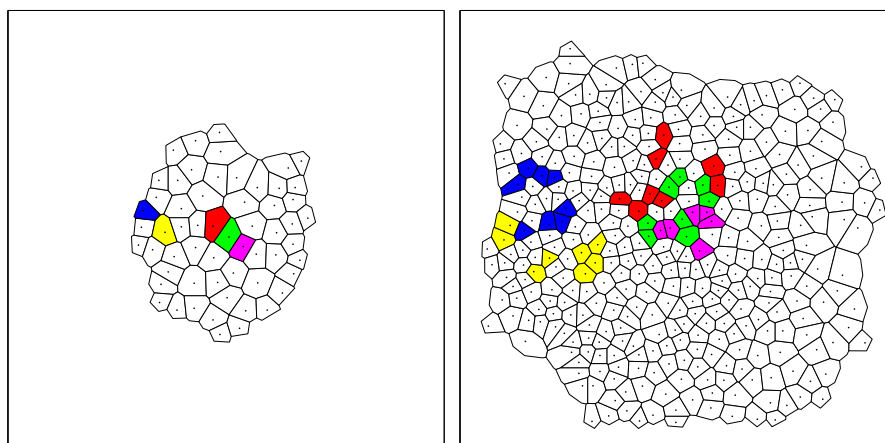
(a) Temps initial

(b) Temps final

Légende. Les points colorés correspondent aux cellules dont le lignage est suivi.

Observations. La taille des cellules et leurs relations de voisinage ne sont pas visualisables.

FIGURE 4.5 – Expérience : représentation par points des (centres des) noyaux



(a) Temps initial

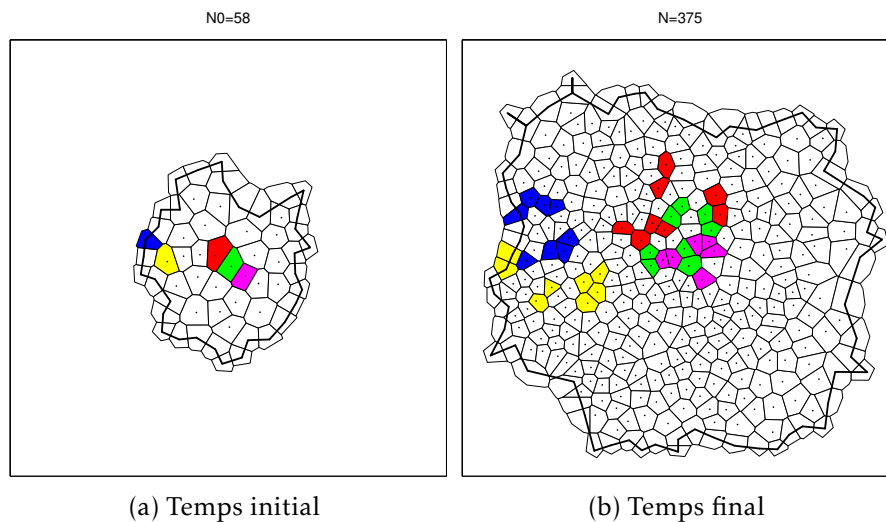
(b) Temps final

Légende. Les régions de Voronoï représentent les cellules ; celles qui sont colorées correspondent aux cellules dont le lignage est suivi. Les points, qui sont les sites du diagramme de Voronoï adapté, correspondent aux centres des noyaux.

Observations. Cette représentation s'approche davantage de la réalité, notamment concernant la forme des cellules ; elle permet aussi de visualiser les relations de voisinage. Cela en fait une représentation pleinement satisfaisante, tant du point de vue biologique que mathématique.

FIGURE 4.6 – Expérience : représentation par diagramme de Voronoï adapté de la population cellulaire

Visualisation du bord d'une population cellulaire Le bord d'une population cellulaire est déterminé par le polygone de bord (voir définition 3.15 page 74), calculé après représentation par diagramme de Voronoï adapté de cette population ; cela est illustré en figure 4.7. En particulier, c'est à partir de ce bord qu'est défini, en section 4.2.1, un lignage biologique périphérique ou central : sa détermination a donc un impact sur le choix des lignages suivis.



Légende. Le polygone tracé, qui englobe la population cellulaire, est le polygone de bord.

Observations. Au temps initial, il y a parmi les cellules suivies : une qui est sur le bord (la bleue), une qui est voisine du bord (la jaune) et trois qui sont davantage éloignées du bord (la rouge, la verte et la mauve).

FIGURE 4.7 – Expérience : représentation du polygone de bord de la population

4.2 Application des diagnostics aux données expérimentales post-traitées

L'objectif de cette section est de se servir des diagnostics qui ont été utilisés lors de l'analyse des simulations, afin d'appliquer exactement les mêmes sur les données issues des expériences.

4.2.1 Protocole d'analyse

Comme nous venons de le voir, les informations extraites lors du post-traitement des données expérimentales sont suffisantes pour représenter, via notre représentation par diagramme de Voronoï adapté, les instants initial et final d'une expérience. Cela

signifie que nous sommes placés dans une configuration matérielle où nous avons toutes les informations nécessaires pour appliquer les diagnostics aux données expérimentales.

Il reste à décider, pour les diagnostics concernant les lignages, sur quels lignages nous les appliquons. Dans le but d'étudier si la disparité centre versus périphérie observée sur les simulations existe aussi sur les expériences, nous avons besoin de définir ce que sont un lignage central et un lignage périphérique. Afin de coïncider au plus près avec les choix effectués lors de l'analyse des simulations, nous posons les définitions suivantes.

Définition 4.2 (Lignage biologique périphérique – Lignage biologique central). Un lignage issu des expériences biologiques est défini comme :

- périphérique lorsque sa cellule génératrice, au temps initial, est située sur le bord de la population cellulaire ;
- central lorsque sa cellule génératrice, au temps initial, est située à au moins deux assises cellulaires du bord⁴ de la population cellulaire.

Une illustration du polygone de bord et de lignages centraux et périphériques est proposé en figure 4.7 page 104. L'éloignement au bord, en terme d'assises cellulaires, est défini par les relations de voisinage données par le diagramme de Voronoï adapté.

Ces deux définitions ont pour conséquence que certains lignages initialement suivis lors du post-traitement des données expérimentales ne seront pas utilisés au cours de l'analyse avec les diagnostics, puisqu'ils ne correspondent à un lignage ni périphérique ni central. Au final, nous avons l'échantillon suivant de données :

- 14 expériences exploitables (taille de l'échantillon pour les diagnostics n^{os} 1 et 2) ;
- 22 lignages centraux et 54 lignages périphériques (taille des échantillons pour les diagnostics n^{os} 4 et 6) ;
- 53 composantes connexes de lignage en zone centrale et 115 en zone périphérique (taille des échantillons pour les diagnostics n^{os} 3 et 5).

4.2.2 Utilisation des diagnostics sur les données expérimentales

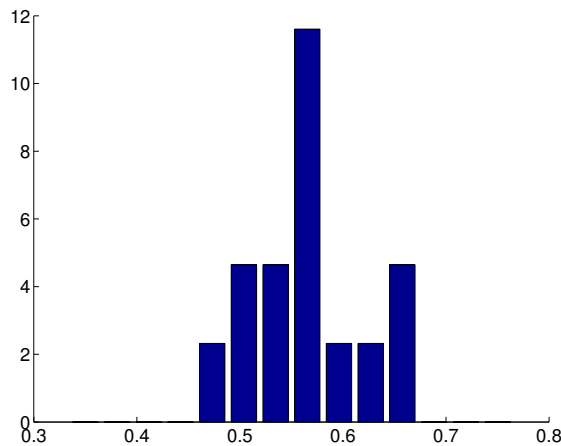
Dans cette section, chaque diagnostic introduit en section 3.2 page 70 est appliqué sur les données expérimentales et est accompagné d'un bref rappel. Les conclusions et les interprétations sont l'objet de la section 4.2.3.

Diagnostic n° 1 (sphéricité de la population) Il mesure un rapport aire versus périmètre pour l'ensemble de la population, tel que défini par R_1 en équation (3.9) page 75 :

$$R_1 = \frac{4\pi A}{\mathcal{P}^2}.$$

Notre objectif est d'étudier la sphéricité de la population et de fournir une valeur de référence au diagnostic n° 3. Nous représentons en figure 4.8 page suivante la répartition de R_1 (normalisée en aire).

4. Être située à deux assises cellulaires du bord signifie être éloignée du bord par deux cellules ; autrement dit : il y a au moins une cellule entre cette cellule et une cellule située au bord.



Légende. En abscisses figurent les valeurs de R_1 et en ordonnées les proportions (normalisées en aire) d'expériences, sur un échantillon de taille 14.

Observations. Les valeurs sont regroupées autour d'une moyenne de 0,56.

FIGURE 4.8 – Diagnostic n° 1 sur le rapport aire versus périmètre de la population (R_1), pour les expériences

Diagnostic n° 2 (convexité de la population) Il mesure un rapport aire versus aire de l'enveloppe convexe pour l'ensemble de la population, tel que défini par R_2 en équation (3.10) page 76 :

$$R_2 = \frac{\mathcal{A}}{\mathcal{A}_{\text{conv}}}.$$

Notre objectif est d'étudier la convexité de la population. Nous représentons en figure 4.9 page ci-contre la répartition de R_2 (normalisée en aire).

Diagnostic n° 3 (sphéricité d'un lignage) Il mesure un rapport aire versus périmètre par composante connexe de lignage, tel que défini par R_3 en équation (3.11) page 76 :

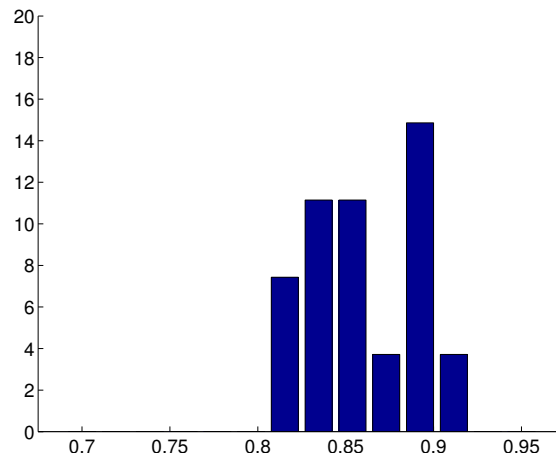
$$R_3 = \frac{4\pi\mathcal{A}}{\mathcal{P}^2}.$$

Notre objectif est d'étudier la filamentarité des lignages. Nous représentons en figure 4.10 page ci-contre la répartition de R_3 (en pourcentage).

Diagnostic n° 4 (morcellement d'un lignage) Il mesure le nombre de composantes connexes par lignage, tel que défini par R_4 en équation (3.12) page 76 :

$$R_4 = \text{nombre de composantes connexes d'un lignage.}$$

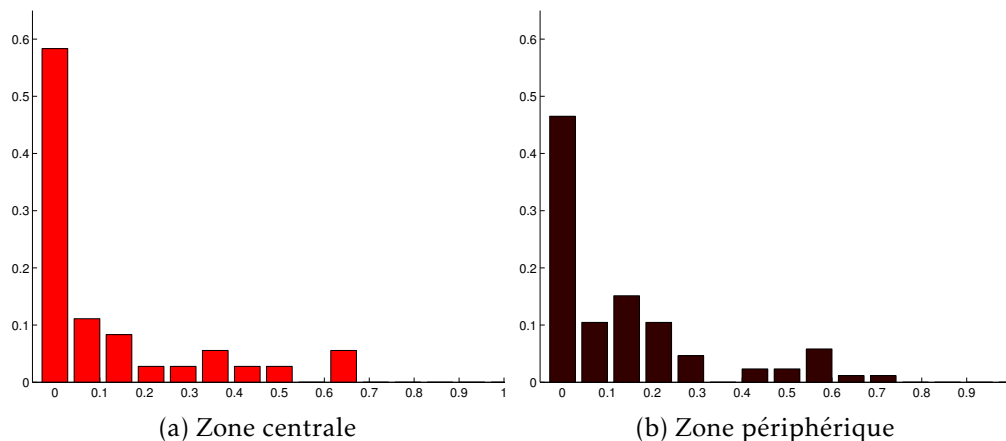
Notre objectif est d'étudier le morcellement des lignages. Nous représentons en figure 4.11 page 108 la répartition de R_4 (en pourcentage).



Légende. En abscisses figurent les valeurs de R_2 et en ordonnées les proportions (normalisées en aire) d'expériences, sur un échantillon de taille 14.

Observations. Les valeurs sont regroupées autour d'une moyenne de 0,86.

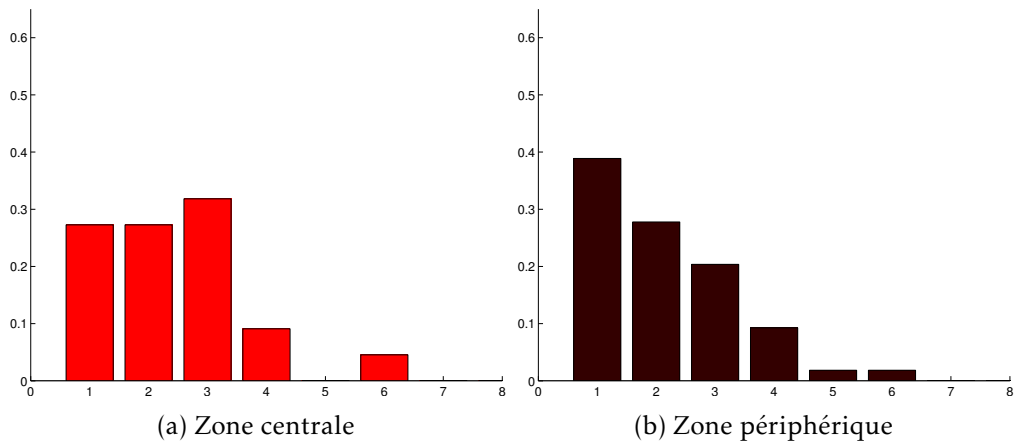
FIGURE 4.9 – Diagnostic n° 2 sur le rapport aire versus aire de l'enveloppe convexe de la population (R_2), pour les expériences



Légende. En abscisses figurent les valeurs de R_3 et en ordonnées les proportions de composantes connexes ; au centre pour (a) et en périphérie pour (b), sur 53 et 115 composantes connexes respectivement.

Observations. Les composantes à $R_3 = 0$ représentent 58 % des composantes en zone centrale alors qu'elles ne sont que 47 % en zone périphérique. Les valeurs de R_3 sont de moyenne 0,12, ce qui est faible.

FIGURE 4.10 – Diagnostic n° 3 sur le rapport aire versus périmètre par composante connexe de lignage (R_3), pour les expériences



Légende. En abscisses figurent les valeurs de R_4 et en ordonnées les proportions de lignages ; au centre pour (a) et en périphérie pour (b), sur 22 et 54 lignages respectivement.

Observations. Les lignages a une ou deux composantes représentent 55 % des lignages centraux et 67 % des lignages périphériques. Les proportions de lignages a une, deux, trois ou quatre composantes sont respectivement de 27%, 27%, 32% et 9% en zone centrale alors qu'elles suivent une décroissance exponentielle en zone périphérique (39 %, 28 %, 20 % et 9 %, pour une moyenne de 2,13).

FIGURE 4.11 – Diagnostic n° 4 sur le nombre de composantes connexes par lignage (R_4), pour les expériences

Diagnostic n° 5 (taille des fragments d'un lignage) Il mesure le nombre de cellules par composante connexe de lignage, tel que défini par R_5 en équation (3.13) page 76 :

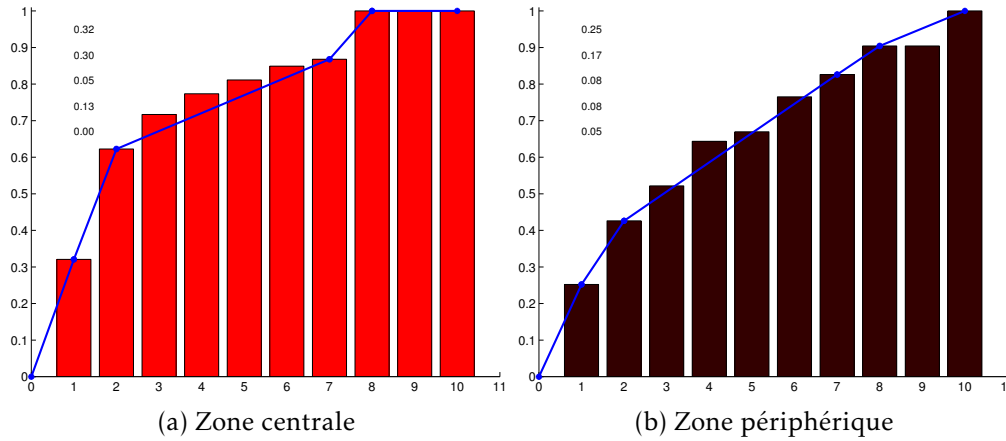
R_5 = nombre de cellules dans une composante connexe d'un lignage.

Notre objectif est de mettre en valeur les plus grosses composantes et les plus petites. Nous représentons en figure 4.12 page ci-contre la répartition de R_5 (en pourcentage cumulé) ; pour étudier l'évolution de cette répartition sur cinq zones (de zéro à une cellule, de une à deux, de deux à sept, de sept à huit, de huit et plus), nous traçons les droites d'interpolation aux extrémités de ces cinq zones.

Diagnostic n° 6 (orientation d'un lignage) Il mesure la direction principale d'orientation d'un lignage, telle que définie par R_6 en équation (3.16) page 77 :

$$R_6 = \theta_2 - \theta_0 - \frac{\pi}{2}, \text{ à valeurs dans } \left[-\frac{\pi}{2}; \frac{\pi}{2} \right],$$

où θ_2 est l'angle défini par le vecteur propre correspondant à la plus grande des deux valeurs propres de la matrice d'inertie du lignage et θ_0 est l'angle polaire du barycentre du lignage. Notre objectif est de mettre en valeur une orientation particulière des



Légende. En abscisses figurent les valeurs de R_5 et en ordonnées les proportions cumulées de composantes connexes ; au centre pour (a) et en périphérie pour (b), sur 53 et 115 composantes respectivement. Les pentes des cinq droites bleues sont indiquées dans le graphique.

Observations. Les composantes connexes d'une ou deux cellules représentent 62 % des composantes au centre alors que cette proportion est de 42 % en périphérie. En zone centrale, il n'existe pas de composante au-delà de huit cellules, alors qu'il y en a 10 % en zone périphérique.

FIGURE 4.12 – Diagnostic n° 5 sur le nombre de cellules par composante connexe de lignage (R_5), pour les expériences

lignages. Nous représentons en figure 4.13 page suivante la répartition de $|R_6|$, c'est-à-dire l'écart à la direction tangentielle. Ce diagnostic mesure aussi l'intensité des orientations, telle que définie par \tilde{R}_6 en équation (3.17) page 77 :

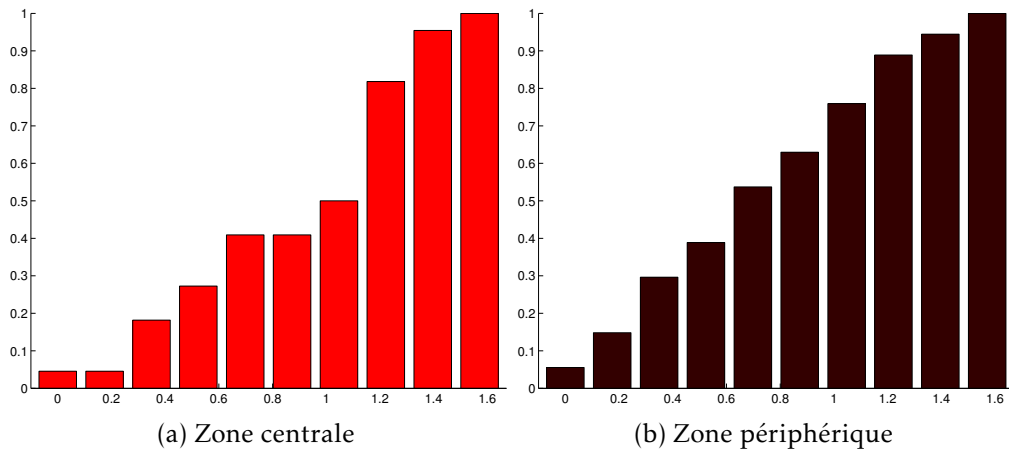
$$\tilde{R}_6 = \sqrt{1 - \frac{\lambda_1}{\lambda_2}}, \quad (4.1)$$

où λ_1, λ_2 sont les valeurs propres de la matrice d'inertie du lignage (avec $\lambda_1 \leq \lambda_2$). Nous représentons en figure 4.14 page suivante la répartition de \tilde{R}_6 (en pourcentage).

4.2.3 Observations et conclusions des diagnostics

Avec les diagnostics n°s 1 et 3, nous observons que la sphéricité des lignages est très faible au regard de celle de la population entière ; en effet, la valeur moyenne de R_1 vaut 0,56 alors que R_3 est proche de 0,12. Par ailleurs, les lignages à $R_3 = 0$ représentent une plus grande proportion en zone centrale qu'en zone périphérique. Nous en concluons que les lignages sont très filamenteux, qu'une disparité entre zones périphérique et centrale se dégage et que les lignages centraux sont plus filamenteux que les lignages périphériques.

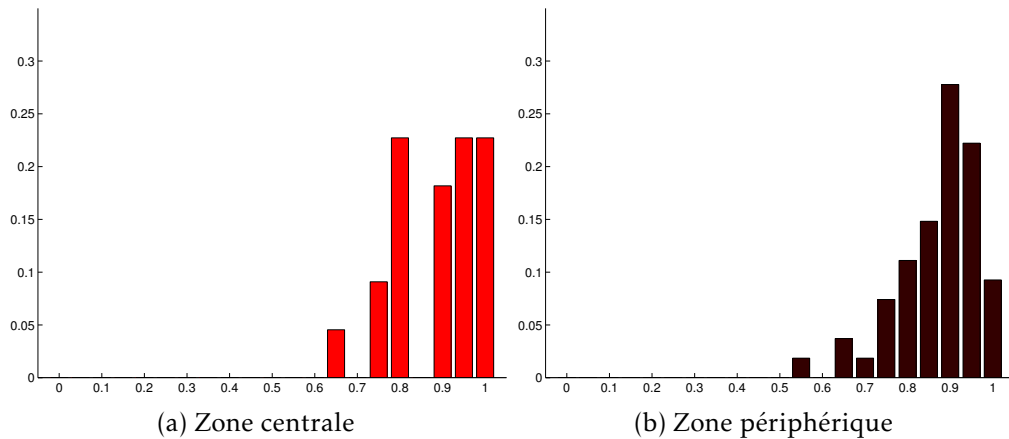
Sur le diagnostic n° 4, nous observons une forte différence entre zone centrale et



Légende. En abscisses figurent les valeurs de $|R_6|$ et en ordonnées les proportions cumulées de lignages ; au centre pour (a) et en périphérie pour (b), sur 22 et 54 lignages respectivement. Les valeurs proches de 0 (respectivement $\frac{\pi}{2}$) correspondent à une orientation tangentielle (resp. radiale).

Observations. En zone centrale, 50 % des lignages sont dans la zone $[65^\circ; 90^\circ]$, c'est-à-dire faisant un écart inférieur à 25° par rapport à la direction radiale. En zone périphérique, la répartition est équilibrée sur tout l'intervalle.

FIGURE 4.13 – Diagnostic n° 6 sur l'écart de l'orientation des lignages à la direction tangentielle ($|R_6|$), pour les expériences



Légende. En abscisses figurent les valeurs de \tilde{R}_6 et en ordonnées les proportions de lignages ; au centre pour (a) et en périphérie pour (b), sur 22 et 54 lignages respectivement. Les valeurs proches de 0 correspondent à des ellipses proches de cercles.

Observations. Les valeurs sont dans plus de 90 % des cas au-dessus de 0,7. (L'excentricité d'une ellipse dont le grand axe est deux fois plus grand que le petit axe est de $\frac{\sqrt{3}}{2} \approx 0,86$.)

FIGURE 4.14 – Diagnostic n° 6 sur l'excentricité des lignages (\tilde{R}_6), pour les expériences

zone périphérique concernant la distribution des lignages séparés en un, deux ou trois fragments : en zone centrale, la proportion de lignages dans ces trois cas est sensiblement identique alors qu'en zone périphérique elle est fortement décroissante. Par ailleurs, les lignages en un ou deux fragments représentent la moitié des lignages en zone centrale alors que cette proportion atteint les deux-tiers en zone périphérique. Nous en concluons qu'il existe une forte disparité entre les deux zones : elle s'exprime par un plus fort morcellement des lignages centraux que des lignages périphériques.

Sur le diagnostic n° 5, nous observons une forte différence entre zone centrale et zone périphérique concernant les composantes constituées d'une ou deux cellules : elles représentent quasiment les deux-tiers en zone centrale alors que cette proportion est de 40 % en zone périphérique. Par ailleurs, il n'existe pas de composante contenant plus de huit cellules en zone centrale, alors que de telles grosses composantes représentent 10 % des lignages périphériques. Nous en concluons qu'il existe une forte disparité entre zones centrale et périphérique : les fragments des lignages centraux sont beaucoup plus petits.

Sur le diagnostic n° 6, nous observons qu'en zone centrale la moitié des lignages ont une orientation proche de la direction radiale, alors qu'en zone périphérique aucune direction ne semble privilégiée. Toutefois, il est à noter une remarque limitante pour l'interprétation de l'orientation en zone périphérique : les populations cellulaires en cultures n'étant pas aussi sphériques que dans les simulations (ce qui est confirmé par les valeurs du diagnostic n° 1), un lignage peut se retrouver le long du bord sans être tangentiellement orienté. Nous ne notons aucune observation particulière concernant l'intensité des orientations. Nous en concluons qu'il existe une forte disparité entre zones centrale et périphérique : les lignages centraux sont radialement orientés alors qu'aucune orientation privilégiée n'apparaît pour les lignages périphériques.

Conclusions communes des diagnostics et interprétation biologique Nous pouvons conclure, à travers le filtre des quatre diagnostics (n°s 3 à 6) étudiant les lignages, que nous avons mis en valeur une disparité entre les lignages centraux et périphériques. Cette disparité s'illustre par :

1. De plus nombreux et de plus petits fragments en zone centrale. Autrement dit, les descendants d'une cellule qui se trouvait initialement proche du centre de la population cellulaire se retrouvent davantage éparpillés (que ceux issus d'une cellule se trouvant initialement à sa périphérie).
2. Une orientation radiale en zone centrale et aucune orientation particulière en zone périphérique. Autrement dit, les lignages issus de cellules proches du centre ont tendance à s'aligner en direction du bord, alors que ceux issus de la périphérie ne présentent aucune orientation.

Par ailleurs, en comparant les diagnostics n°s 1 et 3, nous pouvons conclure que les composantes connexes des lignages sont très peu sphériques par rapport à la population entière.

4.3 Confrontation avec les résultats des simulations

Nous pouvons maintenant effectuer une confrontation entre les résultats des simulations et des expériences : c'est l'objectif de cette section. En effet, tout ce que nous avons construit jusqu'à présent l'a été dans le but de faire cette confrontation :

- les simulations numériques et les expériences réalisées par les biologistes ont été conçues pour être très proches ;
- les diagnostics ont été définis et implémentés pour être appliqués aux résultats des simulations (voir section 3.4 page 81) et des expériences (voir section 4.2 page 104).

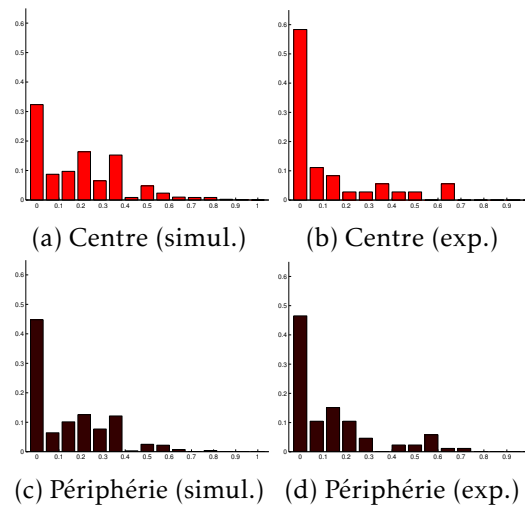
4.3.1 Comparaison entre les résultats des simulations et les résultats des expériences

Concernant les deux diagnostics n^{os} 1 et 2 qui étudient la population cellulaire entière, la comparaison est identique : les valeurs moyennes expérimentales sont bien plus faibles que celles des simulations. Nous l'interprétons par le fait que les populations cellulaires sont beaucoup moins régulières dans le monde biologique : en particulier, le bord est moins lisse et la forme est moins ronde.

Concernant le diagnostic n^o 6 sur l'orientation des lignages, la comparaison est illustrée en figure 4.18 page 114. Un point commun entre simulations et expériences est la disparité centre versus périphérie. Cependant, cette disparité s'exprime différemment : dans les expériences, aucune orientation tangentielle n'apparaît en zone périphérique, alors qu'une orientation radiale (qui pourrait coïncider avec la situation contrainte radiale des simulations, voir figure 3.17a page 93) est visible.

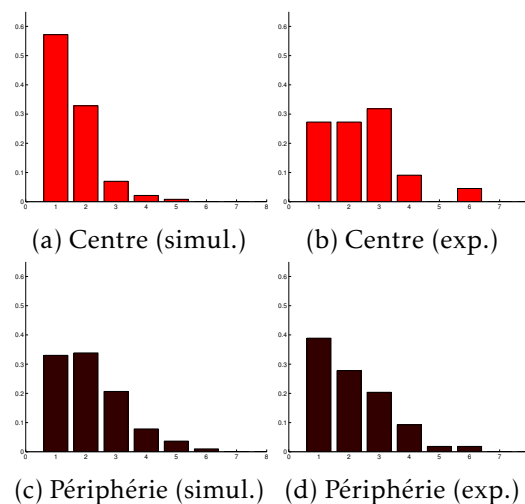
Concernant les trois autres diagnostics qui s'appliquent aux lignages, la comparaison est illustrée en figures 4.15 à 4.17 page suivante et page 114. Un point commun aux expériences et aux simulations est que les résultats sont très différents entre zone centrale et zone périphérique : cela est très net sur les diagnostics n^{os} 4 et 5. Cependant, cette disparité est inversée : sur le modèle mathématique, les lignages périphériques sont plus filamenteux, plus morcelés et en plus petits fragments que les lignages centraux, alors que c'est l'inverse en biologie. Par ailleurs, un autre point commun à ces trois diagnostics est, qu'au centre, les résultats entre les simulations et les expériences sont assez éloignés, en comparaison des résultats en périphérie : nous en concluons que, au travers le filtre des diagnostics, notre modèle est meilleur en zone périphérique qu'en zone centrale.

Bilan En conclusion, nous pouvons affirmer qu'un point commun entre les simulations et les expériences est la disparité entre zones centrale et périphérique : celle-ci a été mise en valeur sur les simulations puis, ensuite, elle a été confirmée sur les expériences. En revanche, un apport de la confrontation mathématiques-biologie est que cette disparité est inversée : dans les expériences, c'est en zone centrale que les lignages sont davantage fragmentés et filamenteux. Un dernier élément important de cette confrontation, toujours concernant ces deux zones, est que le modèle est davantage proche des expériences en zone périphérique que centrale.



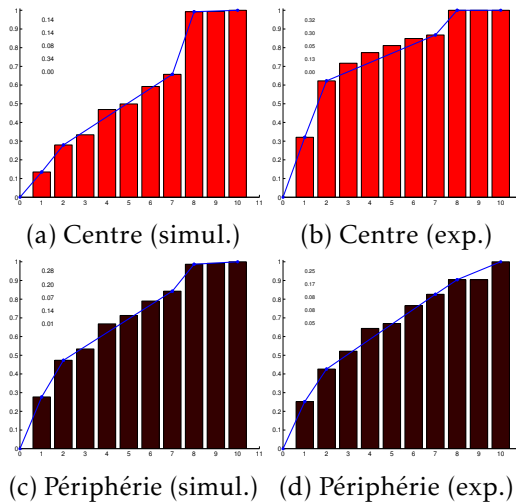
Légende. Reproduction des figures 3.10 et 4.10 page 86 et page 107, aux mêmes échelles.
Observations. Il y a une disparité centre versus périphérie, dans les deux cas. Cependant, elle est inversée : pour les expériences, c'est au centre que le pic en 0 est plus grand (lignages filamenteux). De plus, en zone centrale, les résultats du modèle (a) et des expériences (b) sont assez différents.

FIGURE 4.15 – Diagnostic n° 3 : comparaison entre les simulations et les expériences



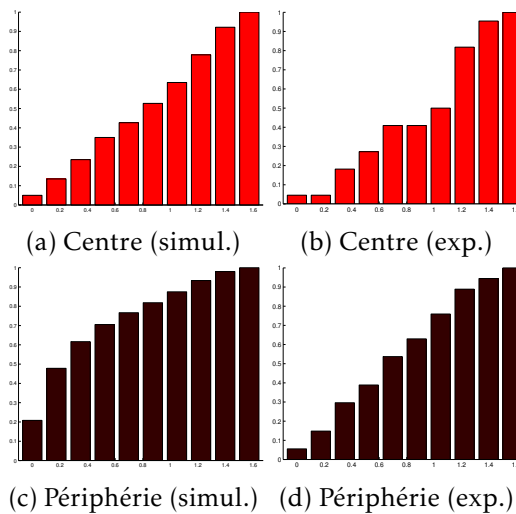
Légende. Reproduction des figures 3.11 et 4.11 page 86 et page 108, aux mêmes échelles.
Observations. Il y a une disparité centre versus périphérie, dans les deux cas. Cependant, elle est inversée : pour les simulations, c'est au centre que les valeurs en 1 et 2 sont davantage élevées (faible morcellement). De plus, en zone centrale, les résultats du modèle (a) et des expériences (b) sont assez différents.

FIGURE 4.16 – Diagnostic n° 4 : comparaison entre les simulations et les expériences



Légende. Reproduction des figures 3.14 et 4.12 page 91 et page 109, aux mêmes échelles.
Observations. Il y a une disparité centre versus périphérie, dans les deux cas. Cependant, elle est inversée : pour les expériences, c'est au centre que l'augmentation sur les valeurs 1 et 2 est la plus forte (petits fragments). De plus, en zone centrale, les résultats du modèle (a) et des expériences (b) sont assez différents.

FIGURE 4.17 – Diagnostic n° 5 : comparaison entre les simulations et les expériences



Légende. Reproduction des figures 3.15 et 4.13 page 91 et page 110, aux mêmes échelles.
Observations. Il y a une disparité centre versus périphérie, dans les deux cas ; cependant, elle est différente. Pour les expériences, l'orientation est radiale (b) ou quelconque (d) ; pour les simulations, elle est quelconque (a) ou tangentielle (c).

FIGURE 4.18 – Diagnostic n° 6 : comparaison entre les simulations et les expériences

4.3.2 Bilan de la modélisation grâce à la confrontation

La confrontation entre les expériences et les simulations a mis en valeur des similitudes ainsi que des différences : cela signifie que notre modèle est proche de la réalité mais pas assez. Concernant l'étude de la problématique 1 sur l'organisation de la population : nous avons montré que le devenir et la morphologie d'un lignage étaient fortement liés à la position, dans la population, de la cellule génératrice de ce lignage (plus précisément, selon qu'elle était au centre ou à la périphérie de la population cellulaire). Quant à l'étude de la problématique 2 sur l'orientation, elle n'a pas établi de similitude entre le modèle et les expériences.

Concernant le modèle, elle a mis en valeur un fort éloignement des résultats entre simulations et expériences en zone centrale, en particulier pour le diagnostic n° 4. Une possible explication serait qu'il existe une composante de mouvement chez les cellules centrales : par exemple, pour les pousser à atteindre la zone périphérique. Ce déplacement perturberait l'organisation des cellules, qui est uniquement pilotée par le regroupement dans le modèle mathématique actuel, et rendrait ces lignages plus morcelés et filamenteux.

Nous formulons donc l'hypothèse suivante :

il y a davantage de déplacements en zone centrale qu'en zone périphérique.

L'objet du chapitre suivant est d'adapter la modélisation afin de tester cette hypothèse.

Chapitre 5

Retour sur le modèle

L'objectif de ce chapitre est d'enrichir le modèle, tout en lui assurant d'être ancré au plus proche de la réalité. Les améliorations de la modélisation s'appuient sur les résultats de la confrontation entre le modèle mathématique et les expériences biologiques. Ces améliorations sont guidées par l'exploration de notre hypothèse formulée à la fin du chapitre 4 et visent à rapprocher le modèle des expériences.

Dans ce chapitre, nous commençons par proposer deux améliorations au modèle du chapitre 1, cohérentes avec notre hypothèse et avec les conclusions de la confrontation du chapitre 4 (section 5.1); puis nous apportons à l'algorithme du chapitre 2 les modifications nécessaires pour prendre en compte cet enrichissement de la modélisation (section 5.2); enfin nous étudions, à travers le filtre des diagnostics du chapitre 3, les résultats des nouvelles simulations et nous les comparons aux expériences réalisées par les biologistes (section 5.3). Cette démarche constitue un retour sur le modèle.

Sommaire du chapitre

5.1 Améliorations des principes de modélisation	118
5.1.1 Ajout d'une permutation entre cellules	118
5.1.2 Choix par zone du potentiel global W	119
5.2 Adaptation de l'algorithme	120
5.2.1 Étape de positionnement	122
Résolution du problème de minimisation avec le potentiel linéaire	122
Algorithme de positionnement modifié et bilan de l'étape	123
5.2.2 Étape de division	125
Interaction de positionnement modifié	125
Décision de permuter	126
Bilan de l'étape de division d'une cellule	126
5.2.3 Étape de permutation	128
5.2.4 Valeurs numériques des nouveaux paramètres de modélisation	128
5.3 Étude du modèle enrichi avec le filtre des diagnostics	129
5.3.1 Impacts de chacune des deux améliorations	129
5.3.2 Impacts des deux améliorations combinées et conclusions	132

5.1 Améliorations des principes de modélisation

L'objectif de cette section est de proposer deux améliorations du modèle : l'ajout d'une permutation entre deux cellules pour modéliser un déplacement discret, qui figure en section 5.1.1 ; la possibilité de choisir entre deux valeurs du potentiel W au cours de l'évolution de la population cellulaire, qui est détaillée en section 5.1.2.

5.1.1 Ajout d'une permutation entre cellules

Une première amélioration vise à rendre les lignages davantage fragmentés en zone centrale, puisque nous avons conclu qu'ils ne l'étaient pas assez. Elle consiste à ajouter une composante de déplacement dans la modélisation : nous la choisissons sous la forme d'une permutation entre une cellule naissante appartenant à un lignage proche du centre et une de ses voisines.

Ce terme de déplacement de cellules peut avoir pour origine deux phénomènes différents :

- une composante migratoire (une cellule se déplace) ;
- une composante mécanique (une cellule, de par les pressions exercées par ses voisines, peut être déformée jusqu'à avoir une forme lui permettant d'être déplacée).

Comme les cellules sont des disques incompressibles dans le modèle mathématique, elles ne peuvent subir de déformation¹ : la permutation est une manière de modéliser, dans le modèle discret, un déplacement par déformation. La décision de déclencher cette permutation juste après la division est justifiée par le fait qu'il s'agit du moment où une cellule est la plus petite, ce qui facilite son déplacement au sein d'une population cellulaire. Enfin, le choix d'autoriser ce déplacement uniquement aux cellules appartenant à un lignage proche du centre est motivé de la manière suivante : la contrainte de congestion étant plus forte au centre de la population, les cellules y sont plus enclines à s'agiter pour mieux se positionner.

Définition 5.1 (Lignage intérieur). Nous appelons lignage intérieur un lignage issu d'une cellule proche du centre. Le choix de modélisation retenu pour la proximité au centre est la distance euclidienne à 0 : pour $i \in \llbracket 1; N \rrbracket$, une cellule X_i est proche du centre lorsque

$$\text{dist}(0, X_i) \leq C_{\text{int}} \max_{j \in \llbracket 1; N \rrbracket} (\text{dist}(0, X_j)), \quad (5.1)$$

avec $C_{\text{int}} \in [0; 1]$ un paramètre de modélisation à choisir.

La règle de permutation est la suivante : après une division dans un lignage intérieur, une des deux cellules filles a la possibilité de permuter avec une de ses voisines ; cette décision de permuter est modélisée par une probabilité qui suit une loi de Bernoulli de paramètre p . Pour $i, j \in \llbracket 1; N \rrbracket$, en notant C_{i_0} une cellule venant de se diviser (respectivement C_j une de ses voisines), de centre X_{i_0} (resp. X_j), de rayon R_{i_0} (resp. R_j) et de lignage ℓ_{i_0} (resp. ℓ_j), cela revient à : si ℓ_{i_0} est un lignage intérieur alors, avec probabilité p , nous échangeons les valeurs de X_{i_0} et de X_j ; sinon, il n'y a pas de permutation.

1. Excepté au cours du processus de division.

5.1.2 Choix par zone du potentiel global W

La deuxième amélioration vise à rendre les lignages moins fragmentés en zone périphérique, ainsi qu'à affaiblir la dynamique de positionnement². Elle consiste à appliquer aux cellules un potentiel global moins fort que le potentiel quadratique : nous choisissons un potentiel linéaire, qui n'est appliqué qu'à partir d'une certaine taille de la population cellulaire et uniquement à des cellules éloignées du centre.

Cet affaiblissement du potentiel peut avoir pour origine que des cellules subissent un confinement moins fort que d'autres. Le choix d'appliquer ce potentiel uniquement à des cellules éloignées du centre vise à rendre compte d'une plus faible compression de la population cellulaire dans les zones proches du bord. Quant à la décision de ne pas appliquer ce potentiel lorsque la population cellulaire est trop petite, il est motivé ainsi : dans des petites populations, les cellules ont une forte volonté de rester agglomérées, pour survivre.

Définition 5.2 (Zone éloignée – Cellule éloignée du centre). Pour $i \in \llbracket 1; N \rrbracket$, nous appelons cellule éloignée du centre une cellule X_i telle que

$$\text{dist}(0, X_i) \geq C_L \max_{j \in \llbracket 1; N \rrbracket} (\text{dist}(0, X_j)), \quad (5.2)$$

avec $C_L \in [0; 1]$ un paramètre de modélisation à choisir. La modélisation de l'éloignement est la même que celle introduite en définition 5.1 page 118. Nous appelons zone éloignée l'ensemble des cellules éloignées du centre.

La règle de positionnement modifié est la suivante : dès que le nombre de cellules a atteint une valeur critique, notée N_s , un potentiel linéaire est utilisé pour les cellules de la zone éloignée et le potentiel quadratique usuel est utilisé pour les autres cellules. Autrement dit, le principe général de la loi de positionnement est toujours le problème de minimisation sous contraintes

$$X^* = \underset{\substack{X \in (\mathbb{R}^2)^N \text{ tq} \\ \phi_{i,j}(X) \leq 0}}{\text{argmin}} W(X), \quad (5.3)$$

où les contraintes $(\phi_{i,j})_{i < j}$ restent définies, pour $X \in (\mathbb{R}^2)^N$ et $R \in \mathbb{R}^N$, par

$$\phi_{i,j}(X) = (R_i + R_j)^2 - \left[(x_j - x_i)^2 + (y_j - y_i)^2 \right], \quad (5.4)$$

mais où le potentiel W est choisi comme suit :

- si $N \geq N_s$ (c'est-à-dire le nombre de cellules a atteint une valeur critique), alors :
 - si une cellule est dans la zone éloignée, alors le choix pour W est le potentiel

2. Nous rappelons que, dans nos principes de modélisation, le regroupement des cellules est assuré par une minimisation sous contraintes d'un potentiel global de confinement ; ce potentiel est jusqu'à présent choisi quadratique.

linéaire $W_L : (\mathbb{R}^2)^N \rightarrow \mathbb{R}$, défini par

$$W_L(X) = \sum_{i=1}^N (|x_i| + |y_i|) \quad \text{pour } X \in (\mathbb{R}^2)^N, \quad (5.5)$$

— sinon le choix pour W est le potentiel quadratique W_Q introduit au chapitre 1, en équation (1.8) page 28, défini par

$$W_Q(X) = \sum_{i=1}^N (x_i^2 + y_i^2) \quad \text{pour } X \in (\mathbb{R}^2)^N; \quad (5.6)$$

— sinon le choix pour W reste le potentiel quadratique W_Q .

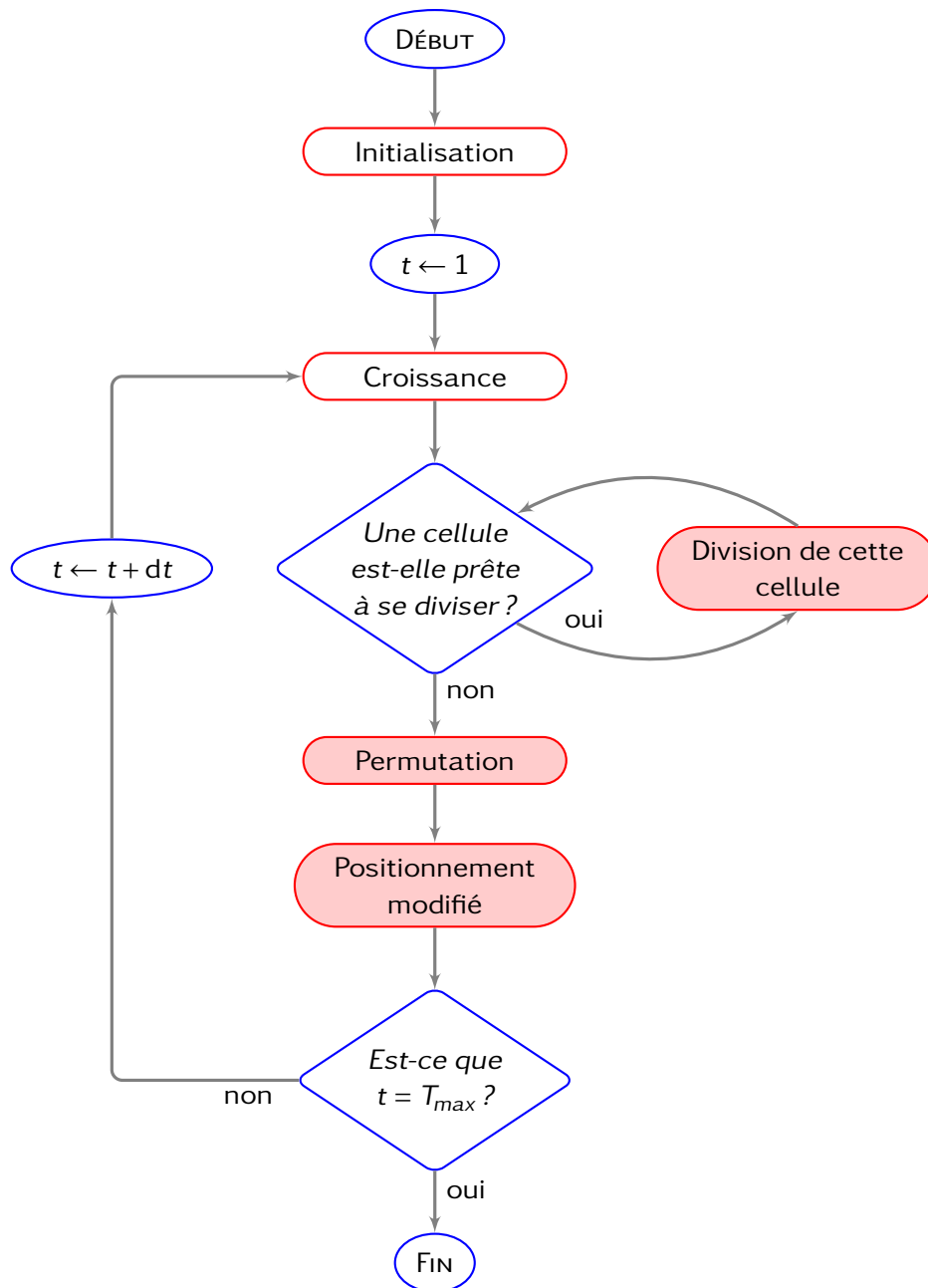
5.2 Adaptation de l'algorithme

L'objectif de cette section est de disposer d'un algorithme prenant en compte les améliorations apportées à la modélisation en section 5.1, afin de pouvoir effectuer de nouvelles simulations en section 5.3.

Les changements apportés à l'algorithme, sous forme d'ajouts ou de modifications, apparaissent *sous cet aspect* dans la structure générale présentée ci-après. Cette structure est illustrée par un schéma en figure 5.1 page suivante.

Structure de l'algorithme
<ol style="list-style-type: none"> 1. Choix des paramètres de modélisation et de l'algorithme. 2. Étape d'initialisation. 3. À chaque pas de temps (boucle en temps) : <ol style="list-style-type: none"> (a) étape de croissance ; (b) pour chacune des cellules, les unes après les autres (boucle sur les cellules) : <ol style="list-style-type: none"> (i) test sur la condition de division, (ii) si ce test est positif, alors <i>étape de division</i> ; (c) <i>étape de permutation</i> ; (d) <i>étape de positionnement</i>.

Nous détaillons chacun de ces changements, en les présentant dans l'ordre de dépendance des uns par rapport aux autres. L'étape de positionnement est expliquée en section 5.2.1, puis vient l'étape de division en section 5.2.2 (qui fait intervenir un algorithme de positionnement) et enfin l'étape de permutation (amorcée au cours de l'étape de division) est décrite en section 5.2.3.



Légende. Les nœuds identiques à ceux présents en figure 2.1 page 41, c'est-à-dire inchangés par les améliorations du modèle, sont sur fond blanc. Les changements apparaissent sur fond coloré.

FIGURE 5.1 – Vision schématique de la totalité de l'algorithme pour le modèle enrichi

5.2.1 Étape de positionnement

Par rapport à l'étape de positionnement décrite en section 2.2 page 40, les améliorations de la modélisation nécessitent d'apporter une modification concernant le changement de potentiel. En effet, dans la résolution du problème de minimisation présenté en équation (5.3) page 119, la fonction potentiel W peut désormais être égale à W_L ou à W_Q .

Le cas de la résolution avec le potentiel quadratique W_Q ne changeant pas, nous détaillons uniquement le cas du potentiel linéaire W_L .

Résolution du problème de minimisation avec le potentiel linéaire

Il s'agit d'adapter notre algorithme basé sur la méthode d'Uzawa, afin de tenir compte du nouveau potentiel. Dans les calculs ci-dessous, nous suivons le même raisonnement qu'en section 2.2.1 page 42.

Le Lagrangien associé au choix $W = W_L$, noté $\mathcal{L}_L: (\mathbb{R}^2)^N \times \mathbb{R}^M \rightarrow \mathbb{R}$, est défini par

$$\forall (X, \lambda) \in (\mathbb{R}^2)^N \times \mathbb{R}^M, \quad \mathcal{L}_L(X, \lambda) = W_L(X) + \sum_{1 \leq i < j \leq N} \lambda_{i,j} \phi_{i,j}(X). \quad (5.7)$$

À partir des deux éléments quelconques $X^{(0)} \in (\mathbb{R}^2)^N$ et $\lambda^{(0)} \in (\mathbb{R}_+)^M$, nous construisons les suites $(X^{(p)})_p$ et $(\lambda^{(p)})_p$, de sorte que $X^{(p)} \xrightarrow{p \rightarrow +\infty} X^*$. Ces suites sont déterminées, pour $p \geq 0$, par les itérations

$$X^{(p+1)} = X^{(p)} - \rho \nabla \mathcal{L}(X^{(p)}, \lambda^{(p)}), \quad (5.8a)$$

$$\phi_{i,j}^{(p+1)} = \phi_{i,j}(X^{(p+1)}), \text{ pour tous } i, j \in \llbracket 1; N \rrbracket \text{ tels que } i < j, \quad (5.8b)$$

$$\lambda_{i,j}^{(p+1)} = \max\left(0, \lambda_{i,j}^{(p)} + \mu \phi_{i,j}^{(p+1)}\right), \text{ pour tous } i, j \in \llbracket 1; N \rrbracket \text{ tels que } i < j, \quad (5.8c)$$

où ρ et μ sont des paramètres positifs fixés. Adoptons les deux notations suivantes :

$$X = \left[\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}, \dots, \begin{pmatrix} x_N \\ y_N \end{pmatrix} \right] \quad \text{et} \quad |X| = \left[\begin{pmatrix} |x_1| \\ |y_1| \end{pmatrix}, \begin{pmatrix} |x_2| \\ |y_2| \end{pmatrix}, \dots, \begin{pmatrix} |x_N| \\ |y_N| \end{pmatrix} \right]. \quad (5.9)$$

Nous avons ensuite, en notant $x \mapsto \text{sgn } x$ la fonction signe d'un réel, le résultat :

$$\nabla W_L(X) = \left[\begin{pmatrix} \text{sgn}(x_1) \\ \text{sgn}(y_1) \end{pmatrix}, \begin{pmatrix} \text{sgn}(x_2) \\ \text{sgn}(y_2) \end{pmatrix}, \dots, \begin{pmatrix} \text{sgn}(x_N) \\ \text{sgn}(y_N) \end{pmatrix} \right]. \quad (5.10)$$

Cela permet, en reprenant les calculs ayant mené au résultat présenté en équation (2.15) page 45, d'obtenir l'expression de $\nabla \mathcal{L}_L(X, \lambda)$ suivante :

$$\nabla \mathcal{L}(X, \lambda) = 2 \left[\begin{pmatrix} \text{sgn}(x_1) + a_{1,1} \\ \text{sgn}(y_1) + a_{2,1} \end{pmatrix}, \dots, \begin{pmatrix} \text{sgn}(x_k) + a_{1,k} \\ \text{sgn}(y_k) + a_{2,k} \end{pmatrix}, \dots, \begin{pmatrix} \text{sgn}(x_N) + a_{1,N} \\ \text{sgn}(y_N) + a_{2,N} \end{pmatrix} \right], \quad (5.11)$$

où, pour $k \in \llbracket 1; N \rrbracket$, les $a_{1,k}$ et $a_{2,k}$ donnés en équations (2.13) et (2.14) page 44 et page 45 valent

$$a_{1,k} = \sum_{j=1}^N \lambda_{k,j} (x_j - x_k) \quad \text{et} \quad a_{2,k} = \sum_{j=1}^N \lambda_{k,j} (y_j - y_k).$$

L'itération exprimée en équation (5.8a) devient alors, pour tout $k \in \llbracket 1; N \rrbracket$,

$$X_k^{(p+1)} = \left(1 - \frac{\rho}{|X_k^{(p)}|} \right) X_k^{(p)} + 2\rho \sum_{j=1}^N \lambda_{k,j}^{(p)} (X_k^{(p)} - X_j^{(p)}). \quad (5.12)$$

Cette équation (5.12) est à mettre en regard du résultat obtenu en équation (2.17) page 45 pour le potentiel quadratique W_Q , qui était :

$$X_k^{(p+1)} = (1 - 2\rho)X_k^{(p)} + 2\rho \sum_{j=1}^N \lambda_{k,j}^{(p)} (X_k^{(p)} - X_j^{(p)}).$$

Remarque 5.1 (Constance du choix du potentiel au cours de l'algorithme de positionnement). Il est à noter que, au cours de l'algorithme de positionnement modifié, la position des cellules change : il se peut alors qu'une cellule, située initialement proche de la frontière entre la zone soumise à W_L et celle soumise à W_Q , change de zone. Nous choisissons néanmoins, pour chaque cellule, de ne pas changer le choix entre W_L et W_Q au cours de la résolution numérique du problème de minimisation.

Algorithme de positionnement modifié et bilan de l'étape

L'implémentation de ces changements figure en algorithme 5.1, dit « de positionnement modifié ». Afin de ne pas en alourdir le contenu, nous ne récrivons pas les détails non affectés par ces changements et nous renvoyons à l'algorithme 2.1 page 46 du chapitre 2.

Algorithme 5.1 – Positionnement modifié

Étant donné les pas ρ et μ ; étant donné les tolérances du critère d'arrêt tol_W et tol_ϕ ; étant donné, pour tout $k \in \llbracket 1; N \rrbracket$, les valeurs initiales $X_k^{(0)}$ et $\lambda_k^{(0)}$; étant donné les rayons R ; étant donné les paramètres C_L et N_s liés au choix du potentiel W_L .

Tant que le critère d'arrêt, identique à celui de l'algorithme 2.1 page 46, est non vérifié, boucle sur $p \geq 0$ contenant :

1. Calcul des positions $X^{(p+1)}$ en fonction du choix du potentiel W (selon C_L et N_s) : pour tout $k \in \llbracket 1; N \rrbracket$, $X_k^{(p+1)}$ est donné par

(a) l'équation (2.17) page 45 si $W = W_Q$:

$$\begin{cases} x_k^{(p+1)} = (1 - 2\rho)x_k^{(p)} + 2\rho \sum_{j=1}^N \lambda_{k,j}^{(p)} (x_k^{(p)} - x_j^{(p)}), \\ y_k^{(p+1)} = (1 - 2\rho)y_k^{(p)} + 2\rho \sum_{j=1}^N \lambda_{k,j}^{(p)} (y_k^{(p)} - y_j^{(p)}). \end{cases} \quad (5.13)$$

(b) l'équation (5.12) page 123 si $W = W_L$:

$$\begin{cases} x_k^{(p+1)} = x_k^{(p)} - \rho \frac{x_k^{(p)}}{|x_k^{(p)}|} + 2\rho \sum_{j=1}^N \lambda_{k,j}^{(p)} (x_k^{(p)} - x_j^{(p)}), \\ y_k^{(p+1)} = y_k^{(p)} - \rho \frac{y_k^{(p)}}{|y_k^{(p)}|} + 2\rho \sum_{j=1}^N \lambda_{k,j}^{(p)} (y_k^{(p)} - y_j^{(p)}). \end{cases} \quad (5.14)$$

2. Calcul des contraintes, comme en algorithme 2.1.
3. Calcul des multiplicateurs, comme en algorithme 2.1.
4. Calcul des quantités contrôlant l'arrêt, comme en algorithme 2.1.

Par ailleurs, nous conservons les améliorations établies en section 2.2.3 page 50 autour de l'algorithme de positionnement. Elles consistaient à utiliser deux fois de suite cet algorithme, en augmentant la précision demandée entre les deux, et à recourir à l'algorithme 2.2 page 51 de retour pour adapter automatiquement la taille du pas ρ .

Au final, la partie de l'algorithme implémentant le positionnement modifié possède la structure suivante :

Synthèse de l'étape de positionnement

1. Choix ^a des pas ρ et μ et des seuils de tolérance tol_W , tol_ϕ et tol_X .
2. Dynamique de positionnement :
 - (a) algorithme 5.1 de positionnement modifié ;
 - (b) diminution par deux du pas ρ et du seuil de tolérance tol_W de l'algorithme de positionnement modifié ;
 - (c) algorithme 5.1 de positionnement modifié.
3. Algorithme 2.2 de retour (entre les positions actuelles et celles juste avant l'item n° 2) : s'il renvoie « retour nécessaire », alors retour à l'item n° 2.

^a. Voir tableau 2.3 page 61.

5.2.2 Étape de division

Par rapport à l'étape de division décrite en section 2.4 page 52, les améliorations de la modélisation nécessitent d'apporter deux modifications :

1. Une concernant le changement de potentiel, puisque l'action n° 2 (appelée « interaction de positionnement ») de chaque étape du processus de division fait appel à un algorithme de positionnement.
2. Une concernant la permutation puisque, d'après la règle de permutation, la décision de permuter est prise juste après une division.

La structure de l'étape de division qui en résulte est illustrée en figure 5.2 page 127.

Interaction de positionnement modifié

Il s'agit d'effectuer un changement au niveau de l'étape τ_k (pour k fixé dans $\llbracket 1; N_{\text{div}} \rrbracket$) du processus de division, décrite en section 2.4.1 page 55. En effet, l'algorithme 2.4 de positionnement en cours de division y est utilisé : il doit être adapté pour tenir compte du choix entre le potentiel linéaire W_L et le potentiel quadratique W_Q . Cette adaptation est similaire à ce qui est fait en section 5.2.1 pour adapter l'étape de positionnement au choix des deux potentiels.

L'implémentation de ces changements figure en algorithme 5.2, dit « de positionnement modifié en cours de division ». Afin de ne pas en alourdir le contenu, nous ne récrivons pas les détails non affectés par ces changements et nous renvoyons à l'algorithme 2.4 page 57 du chapitre 2.

Algorithme 5.2 – Positionnement modifié en cours de division

Étant donné les pas ρ et μ ; étant donné les tolérances du critère d'arrêt tol_W et tol_ϕ ; étant donné, pour tout $k \in \llbracket 1; N \rrbracket$, les valeurs initiales $X_k^{(0)}$ et $\lambda_k^{(0)}$; étant donné les rayons R ; étant donné la stratégie de division S_{div} ; étant donné le paramétrage h et l'orientation ω .

Tant que le critère d'arrêt, identique à celui de l'algorithme 2.4 page 57, est non vérifié, boucle sur $p \geq 0$ contenant :

1. Calcul des positions $X^{(p+1)}$ en fonction du choix du potentiel W :
 - (a) pour tout $k \in \llbracket 1; N \rrbracket$, $X_k^{(p+1)}$ est donnée par
 - (i) l'équation (5.13) page 124 si $W = W_Q$,
 - (ii) l'équation (5.14) page 124 si $W = W_L$;
 - (b) si la stratégie est contrainte ($S_{\text{div}} = 2$), alors les positions des deux cellules concernées sont modifiées pour tenir compte de la nouvelle contrainte, comme en algorithme 2.4.
2. Calcul des contraintes, comme en algorithme 2.4.
3. Calcul des multiplicateurs, comme en algorithme 2.4.
4. Calcul des quantités contrôlant l'arrêt, comme en algorithme 2.4.

Décision de permuter

Il s'agit d'ajouter, après le processus de division, la prise de décision sur la permutation³. Conformément aux choix de modélisation, une permutation ne peut se produire que sur une cellule qui vient de naître : c'est donc à la fin de l'étape de division que nous décidons d'implémenter cette décision.

L'implémentation de cette action figure en algorithme 5.3.

Algorithme 5.3 – Prise de décision sur la permutation

Étant donné les paramètres p et C_{int} liés à la permutation ; étant donné une cellule, notée C_{i_0} , qui vient d'achever son processus de division :

- si ℓ_{i_0} est un lignage intérieur (déterminé par C_{int}), alors on tire un nombre entre 0 et 1 :
 - si ce nombre est plus grand que p alors la décision de permuter est négative,
 - sinon la décision est positive ;
- sinon, la décision est négative.

Bilan de l'étape de division d'une cellule

Au final, la partie de l'algorithme implémentant l'étape de division (pour une cellule satisfaisant la condition de division) possède la structure suivante :

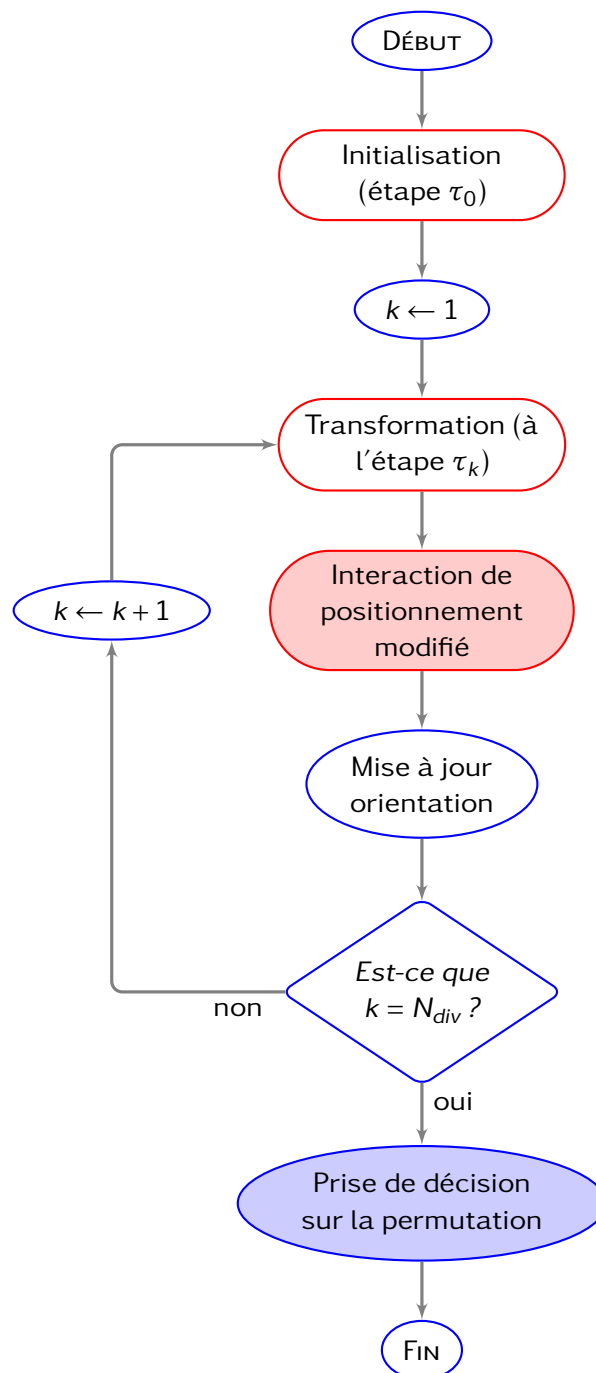
Synthèse de l'étape de division d'une cellule

Étant donné le nombre d'étapes N_{div} , la direction D_{div} et la stratégie S_{div} de la division^a ; étant donné une cellule qui a préalablement vérifié la condition de division :

1. Algorithme 2.5 d'initialisation du processus de division.
2. À chaque étape τ_k du processus de division (boucle sur $k \in \llbracket 1 ; N_{\text{div}} \rrbracket$) :
 - (a) transformation ;
 - (b) interaction de positionnement :
 - (i) algorithme 5.2 de positionnement modifié en cours de division,
 - (ii) algorithme 2.2 de retour (entre les positions actuelles et celles juste avant l'item n° 2b) : s'il renvoie « retour nécessaire », alors retour à l'item n° 2b ;
 - (c) si la stratégie est libre ($S_{\text{div}} = 1$), alors mise à jour de l'orientation.
3. Algorithme 5.3 de prise de décision sur la permutation.

a. Voir tableaux 2.1 et 2.2 page 60 et page 61.

3. La permutation en elle-même est réalisée après l'étape de division (voir section 5.2.3 page 128).



Légende. Les nœuds identiques à ceux présents en figure 2.3 page 54, c'est-à-dire inchangés par les améliorations du modèle, sont sur fond blanc. Les changements apparaissent sur fond coloré.

FIGURE 5.2 – Vision schématique de l'étape de division pour le modèle enrichi

5.2.3 Étape de permutation

Cette étape est nouvelle et succincte : il s'agit, conformément aux principes de modélisation énoncés en section 5.1.1 page 118, de réaliser le phénomène de permutation pour les cellules ayant préalablement pris la décision de le réaliser. Pour chacune de ces cellules, l'algorithme de permutation consiste dans un premier temps à choisir avec quelle cellule permuter, puis ensuite à réaliser cette permutation.

Le détail de l'implémentation est proposé en algorithme 5.4.

Algorithme 5.4 – Permutation

Étant donné une cellule, notée C_{i_0} , ayant préalablement satisfait le test de permutation (au cours de l'étape de division) :

1. Choix de la cellule avec qui réaliser la permutation :
 - (a) calcul des cellules proches de C_{i_0} ;
 - (b) sélection, au hasard, d'une cellule parmi ces proches, notée C_j .
2. Échange des valeurs de x_{i_0} et x_j , ainsi que celles de y_{i_0} et y_j .

Au final, la partie de l'algorithme implémentant l'étape de permutation consiste à faire une boucle, sur l'ensemble des cellules prêtes à permuter, qui contient l'algorithme de permutation.

5.2.4 Valeurs numériques des nouveaux paramètres de modélisation

Nous regroupons ici les valeurs numériques données aux paramètres nouveaux par rapport à la section 2.5.2 page 60 : il s'agit uniquement de paramètres de modélisation, donnés en tableau 5.1.

Paramètre	Valeur	Signification
p	0,2	Probabilité de succès pour la permutation
C_{int}	0,5	Délimitation des lignages intérieurs (voir page 118)
N_s	200	Seuil du nombre de cellules pour le potentiel linéaire
C_L	0,5	Délimitation de la zone éloignée (voir page 119)

La valeur de C_{int} est choisie pour que la zone intérieure coïncide avec la définition d'un lignage biologique central (voir définition 4.2 page 105). Différents essais sont effectués pour choisir les valeurs de p , N_s et C_L : les valeurs retenues sont celles qui rendent les résultats très proches des résultats expérimentaux (voir sections 5.3.1 et 5.3.2 page ci-contre et page 132).

TABLEAU 5.1 – Valeurs numériques des paramètres de modélisation (du modèle enrichi), données par défaut dans les simulations

5.3 Étude du modèle enrichi avec le filtre des diagnostics

L'objectif de cette section est d'utiliser le nouvel algorithme résultant des modifications présentées en section 5.2, afin d'étudier les enrichissements apportés aux principes de modélisation en section 5.1. Cette étude consiste à effectuer de nouvelles simulations numériques et à leur appliquer nos diagnostics du chapitre 3 concernant les lignages.

5.3.1 Impacts de chacune des deux améliorations

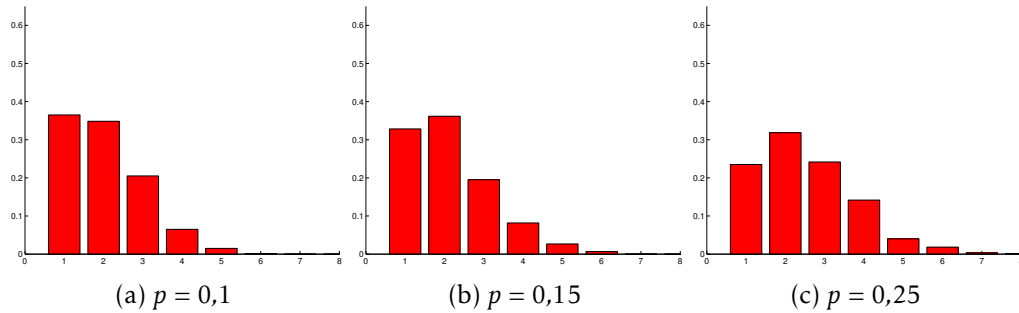
Effet de la permutation Tout d'abord, comme nous pouvions nous y attendre, si la permutation est autorisée pour toute la population cellulaire⁴ alors, en ce qui concerne la zone périphérique, nous nous éloignons des données expérimentales. Nous définissons donc la zone intérieure avec $C_{\text{int}} = \frac{1}{2}$, ce qui limite à certaines cellules seulement la possibilité de permuter. Nous regardons alors l'évolution des diagnostics n° 4 et 5 dans cette même zone : les résultats pour trois valeurs de p sont proposés en figures 5.3 et 5.4 page suivante. Ils illustrent que plus p est grand, plus les lignages sont morcelés et en plus petits fragments : il s'agit bien de l'objectif que nous cherchions à atteindre. Nous observons aussi des changements en zone périphérique, mais ceux-ci sont assez faibles. Enfin, concernant l'orientation des lignages (diagnostic n° 6), il n'y a pas de changement remarquable (voir figure 5.5 page suivante).

Remarque 5.2. Dans la présentation des graphiques, nous nous limitons à trois valeurs pour p ($p = 0,1$, $p = 0,15$ et $p = 0,25$) : elles illustrent les changements dus à la permutation tout en restant proches des résultats expérimentaux.

Effet du changement de potentiel Là aussi nous regardons l'évolution des diagnostics n° 4 et 5, mais cette fois-ci en zone périphérique : c'est en effet dans cette zone que nous voulons étudier l'impact de cette amélioration. Nous définissons la zone éloignée avec $C_L = \frac{1}{2}$: les résultats pour deux valeurs de N_s ($N_s = 100$ et $N_s = 200$) sont proposés en figures 5.6 et 5.7 page 131. Ils illustrent que plus N_s diminue, moins les lignages sont morcelés et en petits fragments : il s'agit bien d'un rapprochement vers les données expérimentales, ce qui nous cherchions à obtenir. Ce changement de potentiel a aussi un effet en zone centrale, où les lignages deviennent eux aussi moins morcelés, ce que nous ne souhaitons pas. Enfin, concernant l'orientation des lignages (diagnostic n° 6), il n'y a là encore pas de changement remarquable (voir figure 5.8 page 131).

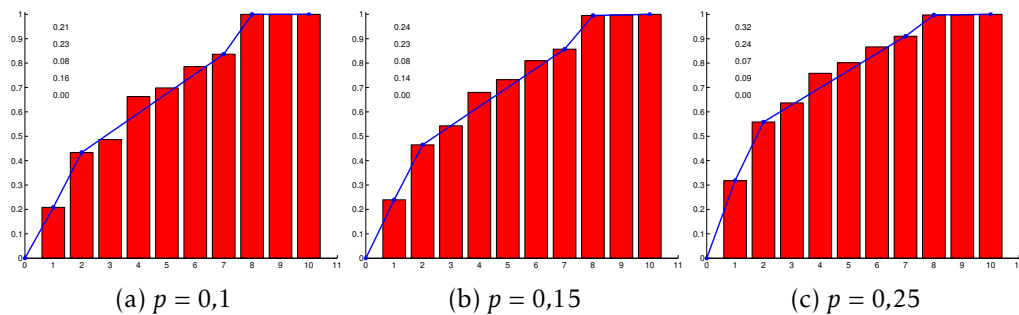
Bilan Chacune de ces deux améliorations permet de rapprocher le modèle des données expérimentales mais, uniquement en zone centrale pour l'une (la permutation) et uniquement en zone périphérique pour l'autre (le changement de potentiel). Ainsi, utiliser une seule de ces améliorations ne suffit pas : nous allons donc combiner les deux.

4. C'est-à-dire si $C_{\text{int}} = 1$.



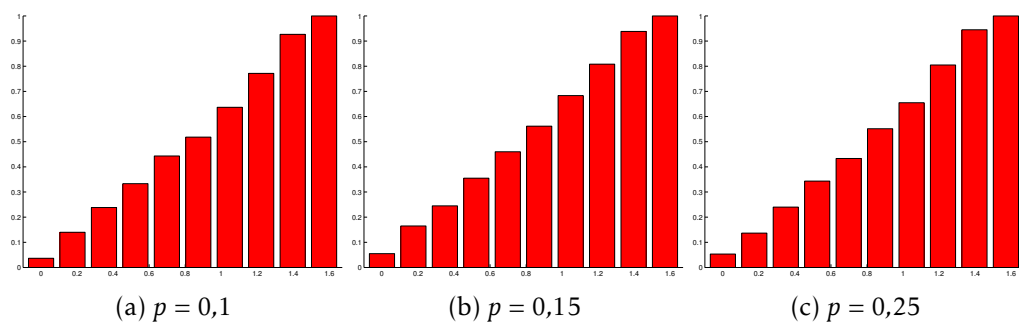
Observations. Plus p augmente et plus la répartition des valeurs de l'histogramme se décale vers la droite : les lignages sont davantage séparés en plusieurs fragments.

FIGURE 5.3 – Diagnostic n° 4 sur le morcellement d'un lignage, pour le modèle enrichi par l'ajout de permutation : en zone centrale avec différentes valeurs de la probabilité p .



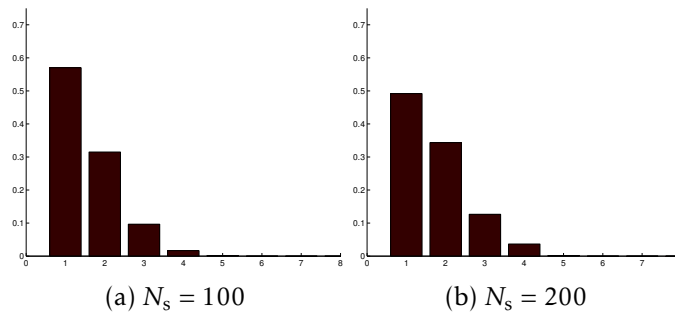
Observations. Plus p augmente et plus les valeurs sont concentrées en début d'histogramme : cela signifie qu'il y a de plus en plus de petites composantes.

FIGURE 5.4 – Diagnostic n° 5 sur la taille des fragments d'un lignage, pour le modèle enrichi par l'ajout de permutation : en zone centrale avec différentes valeurs de la probabilité p .



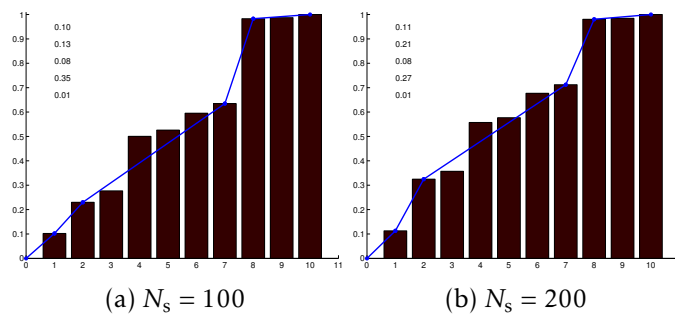
Observations. Il n'y a pas de différence notable lorsque p change de valeur.

FIGURE 5.5 – Diagnostic n° 6 sur l'orientation d'un lignage, pour le modèle enrichi par l'ajout de permutation : zone centrale pour différentes valeurs de la probabilité p .



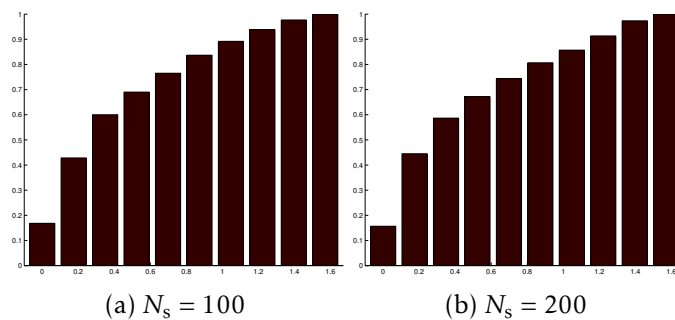
Observations. Plus N_s diminue et plus la répartition des valeurs de l'histogramme se décale vers la gauche : les lignages sont moins morcelés.

FIGURE 5.6 – Diagnostic n° 4 sur le morcellement d'un lignage, pour le modèle enrichi par changement de potentiel : évolution en zone périphérique pour différentes valeurs de la probabilité N_s .



Observations. Plus N_s diminue et plus les valeurs sont concentrées en fin d'histogramme : cela signifie qu'il y a de plus en plus de grosses composantes.

FIGURE 5.7 – Diagnostic n° 5 sur la taille des fragments d'un lignage, pour le modèle enrichi par changement de potentiel : évolution en zone périphérique pour différentes valeurs de la probabilité N_s .



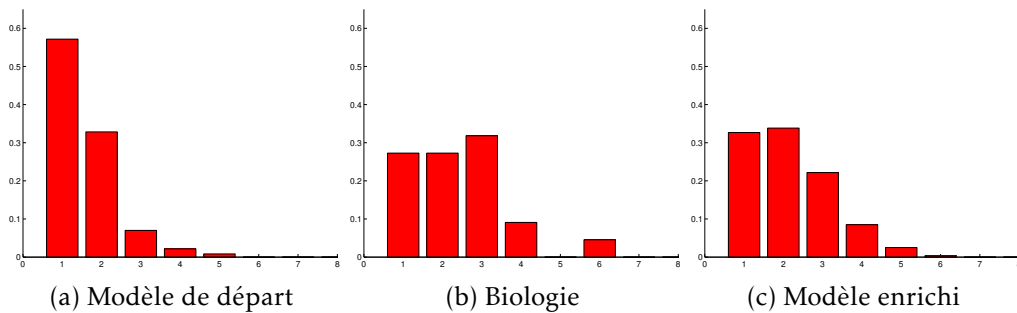
Observations. Il n'y a pas de différence notable lorsque N_s change de valeur.

FIGURE 5.8 – Diagnostic n° 6 sur l'orientation d'un lignage, pour le modèle enrichi par changement de potentiel : évolution en zone périphérique pour différentes valeurs de la probabilité N_s .

5.3.2 Impacts des deux améliorations combinées et conclusions

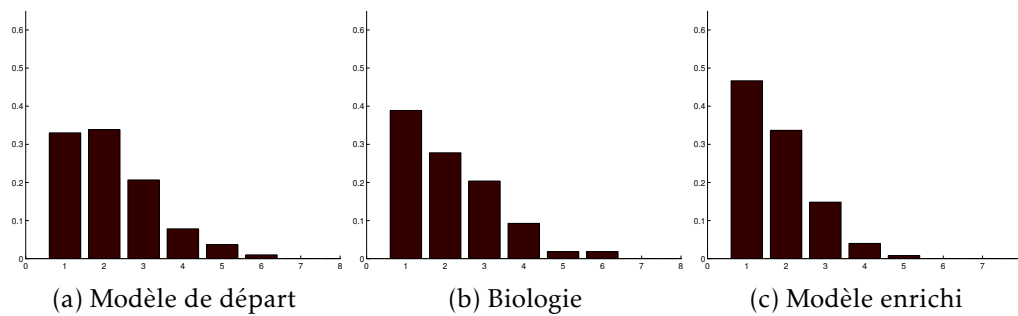
L'effet conjugué des deux améliorations produit une bien meilleure correspondance entre les résultats du modèle et les résultats expérimentaux, que ce soit en zone centrale ou en zone périphérique. Après différents essais, effectués dans le but de se rapprocher des données expérimentales, nous proposons de retenir les valeurs suivantes : $p = 0,2$ et $N_s = 200$. Les résultats des diagnostics n^{os} 4 et 5 dans les zones centrale et périphérique sont en figures 5.9 à 5.12 de la présente page et page suivante. Dans ces figures, le modèle issu du chapitre 1 est appelé « modèle de départ » et ce même modèle modifié par les améliorations de la section 5.1 est dénommé « modèle enrichi ».

Conclusion Nous concluons que ces enrichissements, guidés par la confrontation aux expériences biologiques, nous ont permis d'explorer notre hypothèse. Ils se révèlent être une amélioration du modèle, car le morcellement et la taille des lignages dans les simulations mathématiques obtenues avec le modèle enrichi coïncident dorénavant mieux avec ceux des données expérimentales : il s'agit d'une première validation du modèle enrichi, qui est plus proche des résultats biologiques que le modèle de base.



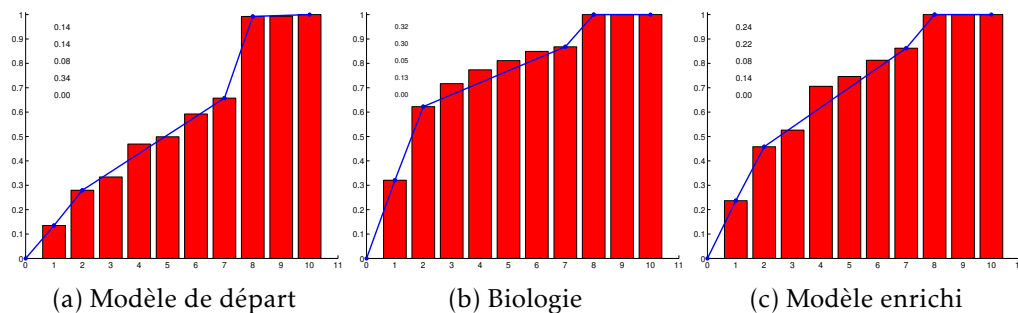
Observations. Le modèle enrichi (c) est plus proche des données expérimentales (b) que ne l'est le modèle de départ (a).

FIGURE 5.9 – Diagnostic n^o 4 sur le morcellement d'un lignage : comparaison entre le modèle enrichi, le modèle de départ et les données expérimentales, en zone centrale



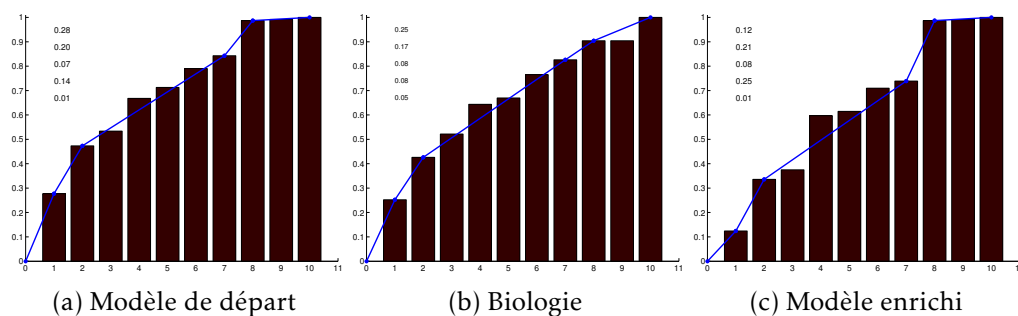
Observations. Le modèle enrichi (c) est plus proche des données expérimentales (b) que ne l'est le modèle de départ (a).

FIGURE 5.10 – Diagnostic n° 4 sur le morcellement d'un lignage : comparaison entre le modèle enrichi, le modèle de départ et les données expérimentales, en zone périphérique



Observations. Le modèle enrichi (c) est plus proche des données expérimentales (b) que ne l'est le modèle de départ (a).

FIGURE 5.11 – Diagnostic n° 5 sur la taille des fragments d'un lignage : comparaison entre le modèle enrichi, le modèle de départ et les données expérimentales, en zone centrale



Observations. Le modèle enrichi (c) est plus proche des données expérimentales (b) que ne l'est le modèle de départ (a).

FIGURE 5.12 – Diagnostic n° 5 sur la taille des fragments d'un lignage : comparaison entre le modèle enrichi, le modèle de départ et les données expérimentales, en zone périphérique

Kathryn Janeway: Is there coffee where we're going?
Lady Q: Bottomless cups of it.
Kathryn Janeway: Thank God.
Lady Q: You're welcome.

Before Dishonor, Star Trek TNG novel

Conclusion et perspectives

Dans ce court chapitre, nous commençons par énoncer une conclusion aux travaux réalisés dans cette thèse ; nous proposons ensuite quelques perspectives, pour motiver une utilisation future de ces travaux.

Conclusion

Un nouveau modèle agent-centré pour décrire la croissance de petites populations cellulaires tumorales a été établi. Son originalité repose sur la modélisation des interactions entre les agents, via un problème de minimisation sous contraintes d'un potentiel, ainsi que sur l'étude de la problématique d'orientation. Cette dernière explore un processus peu connu des biologistes et rarement abordé dans les modèles mathématiques.

Une résolution numérique a été développée, essentiellement inspirée d'une méthode d'Uzawa. Ensuite, un jeu de mesures a été mis en place pour explorer nos problématiques et quantifier nos observations sur les populations et les lignages : appliqué aux résultats des nombreuses simulations effectuées, il a montré une forte disparité entre les lignages centraux et périphériques ; il a aussi montré que les lignages périphériques sont plus nombreux, davantage morcelés et tangentiellement orientés.

Ce même jeu de mesures a été appliqué aux résultats des expériences biologiques, dont le protocole et celui des simulations ont été choisis extrêmement proches ; pour ce faire, nous avons participé activement au post-traitement de ces résultats. La disparité central versus périphérique a alors été retrouvée expérimentalement, comportement qui n'était pas connu des biologistes. Une confrontation avec le modèle mathématique a été réalisée et a mis en valeur des différences, qui nous ont conduit à formuler l'hypothèse suivante : il y a, en zone centrale, plus de déplacements de cellules dans la réalité que dans notre modèle. Pour tester cette hypothèse, nous avons modifié quelques points du modèle en ajoutant une permutation entre agents et un deuxième potentiel. Ce modèle enrichi, évalué au moyen du même jeu de mesures que précédemment, se rapproche des résultats obtenus par les biologistes sur les cultures cellulaires : il s'agit donc d'un premier pas dans la validation de notre hypothèse et plus généralement dans la validation du modèle mathématique établi.

Perspectives

Les perspectives que nous proposons ici sont axées sur la continuité de l'interaction entre les mathématiques et la biologie. Elles reflètent des travaux en cours, effectués par l'équipe IP3D de l'ITAV.

Le passage à la 3D est la première perspective d'amélioration, qui permettrait de modéliser l'évolution de sphéroïdes. Par rapport au contenu du manuscrit, les principales difficultés se situeraient dans un premier temps au niveau des diagnostics, pour les convertir au cas tridimensionnel (à noter que les diagrammes de Voronoï s'étendent naturellement à la 3D, donc ce qui repose uniquement sur ces diagrammes est généralisable). Dans un second temps, elles seraient au niveau de l'obtention, et donc ensuite de l'étude, des données expérimentales ; en effet, le suivi des lignages que nous avons effectué fut manuel : un tel traitement d'analyse d'image en 3D nécessiterait d'être automatisé.

Une autre perspective d'amélioration, orientée vers une étude plus approfondie de l'*orientation de division*, consisterait à effectuer un suivi systématique de ces orientations tout au long de la prolifération de la population cellulaire : autant cela semblerait aisément accessible sur le modèle mathématique autant, là encore, nous serions limités par l'exploitation des résultats biologiques. En effet, la segmentation des images a été réalisée, en grande partie manuellement, uniquement aux temps initial et final : la faire à tous les temps nécessiterait d'être automatisée.

Ces deux premières perspectives trouvent leurs difficultés essentiellement dans la confrontation à des données expérimentales. Une troisième perspective irait en sens inverse : il s'agirait de modifier le modèle mathématique, pour explorer de nouvelles problématiques sur lesquelles les biologistes disposent de données expérimentales. Une possibilité serait d'étudier des populations cellulaires évoluant dans des environnements contraints ; en effet, les tumeurs en croissance évoluent dans des milieux de rigidité variable et peuvent rencontrer des tissus extrêmement rigides, comme le tissu osseux. Nous proposons deux situations :

- lorsqu'il y a des obstacles dans le domaine, par exemple quelques objets ponctuels que les populations sont obligées de contourner au cours de leur prolifération [Aou+14] ;
- lorsque le domaine lui-même est confiné, par exemple s'il s'agit de gouttières dans lesquelles prolifèrent les populations [Des+13].

Il est à noter que ces deux exemples s'expriment en terme de *contraintes* : le cœur de notre modélisation des interactions étant un problème de minimisation sous contraintes, elle semblerait fort adaptée à l'exploration de ces phénomènes, et ainsi permettre d'« avancer vers l'inconnu » !

Bibliographie

- [ACR07] Alexander RA ANDERSON, Mark AJ CHAPLAIN et Katarzyna A REJNIAK. *Single-cell-based models in biology and medicine*. Springer, 2007 (cf. p. 9, 12).
- [AD08] Luca AMODEI et Jean-Pierre DEDIEU. *Analyse numérique matricielle*. Dunod, 2008. ISBN : 978-2-10-052085-5 (cf. p. 48).
- [All12] Grégoire ALLAIRE. *Analyse numérique et optimisation. Une introduction à la modélisation mathématique et à la simulation numérique*. 2^e éd. Éditions de l'École polytechnique, 2012 (cf. p. 15, 42).
- [AM04] Robyn P ARAUJO et DL Sean McELWAIN. « A history of the study of solid tumour growth : the contribution of mathematical modelling ». In : *Bulletin of Mathematical Biology* 66.5 (2004), p. 1039–1091 (cf. p. 8).
- [AM06] Robyn P ARAUJO et DL Sean McELWAIN. « The role of mechanical host–tumour interactions in the collapse of tumour blood vessels and tumour growth dynamics ». In : *Journal of Theoretical Biology* 238.4 (2006), p. 817–827 (cf. p. 11).
- [AMD05] Chaitanya ATHALE, Yuri MANSURY et Thomas S DEISBOECK. « Simulating the impact of a molecular ‘decision-process’ on cellular phenotype and multicellular patterns in brain tumors ». In : *Journal of Theoretical Biology* 233.4 (2005), p. 469–481 (cf. p. 11).
- [And+09] Alexander RA ANDERSON et al. « Microenvironment driven invasion : a multiscale multimodel investigation ». In : *Journal of Mathematical Biology* 58.4 (2009), p. 579–624 (cf. p. 12, 16).
- [And05] Alexander RA ANDERSON. « A hybrid mathematical model of solid tumour invasion : the importance of cell adhesion ». In : *Mathematical Medicine and Biology* 22.2 (2005), p. 163–186 (cf. p. 12).
- [Aou+14] Laurène AOUN et al. « Microdevice arrays of high aspect ratio poly(dimethyl-siloxane) pillars for the investigation of multicellular tumour spheroid mechanical properties ». In : *Lab on a Chip* 14.13 (2014), p. 2344–2353 (cf. p. 136).
- [Aub+06] Marine AUBERT et al. « A cellular automaton model for the migration of glioma cells ». In : *Physical Biology* 3.2 (2006), p. 93 (cf. p. 9).

- [Aub+08] Marine AUBERT et al. « A model for glioma cell migration on collagen and astrocytes ». In : *Journal of The Royal Society Interface* 5.18 (2008), p. 75–83 (cf. p. 9).
- [Aur91] Franz AURENHAMMER. « Voronoi diagrams – A survey of a fundamental geometric data structure ». In : *ACM Computing Surveys (CSUR)* 23.3 (1991), p. 345–405 (cf. p. 65).
- [BC10] Denis BITOUZÉ et Jean-Côme CHARPENTIER. *LATEX, l'essentiel*. 1^{re} éd. Pearson Education France, 2010. ISBN : 978-2-7440-7451-6 (cf. p. xviii).
- [BD09] Helen BYRNE et Dirk DRASDO. « Individual-based and continuum models of growing cell populations : a comparison ». In : *Journal of Mathematical Biology* 58.4-5 (2009), p. 657–687 (cf. p. 11).
- [BH11] Mina J BISSELL et William C HINES. « Why don't we get more cancer ? A proposed role of the microenvironment in restraining cancer progression ». In : *Nature Medicine* 17.3 (2011), p. 320–329 (cf. p. 5).
- [BJJ09] Amy L BAUER, Trachette L JACKSON et Yi JIANG. « Topography of extracellular matrix mediates vascular morphogenesis and migration speeds in angiogenesis ». In : *PLoS Computational Biology* 5.7 (2009), e1000445 (cf. p. 10).
- [BM07] Tilo BEYER et Michael MEYER-HERMANN. « Modeling emergent tissue organization involving high-speed migrating cells in a flow equilibrium ». In : *Physical Review E* 76.2 (2007), p. 021929 (cf. p. 13, 14).
- [Bon02] Eric BONABEAU. « Agent-based modeling : Methods and techniques for simulating human systems ». In : *Proceedings of the National Academy of Sciences of the United States of America* 99.Suppl 3 (2002), p. 7280–7287 (cf. p. 13).
- [BR06] Daniel G BROWN et Derek T ROBINSON. « Effects of heterogeneity in residential preferences on an agent-based model of urban sprawl ». In : *Ecology and Society* 11.1 (2006), p. 46 (cf. p. 13).
- [Bra86] Bart BRADEN. « The surveyor's area formula ». In : *The College Mathematics Journal* 17.4 (1986), p. 326–337 (cf. p. 75).
- [Bre+10] Didier BRESCH et al. « Computational modeling of solid tumor growth : the avascular stage ». In : *SIAM Journal on Scientific Computing* 32.4 (2010), p. 2321–2344 (cf. p. 11).
- [Byr10] Helen M BYRNE. « Dissecting cancer through mathematics : from the cell to the animal model ». In : *Nature Reviews Cancer* 10.3 (2010), p. 221–230 (cf. p. 8).
- [DD05] Andreas DEUTSCH et Sabine DORMANN. *Cellular automaton modeling of biological pattern formation*. FASEB, 2005 (cf. p. 9).

- [Des+13] Annaïck DESMAISON et al. « Mechanical stress impairs mitosis progression in multi-cellular tumor spheroids ». In : *PloS one* 8.12 (2013), e80447 (cf. p. 136).
- [DF00] Dirk DRASDO et Gabor FORGACS. « Modeling the interplay of generic and genetic mechanisms in cleavage, blastulation, and gastrulation ». In : *Developmental Dynamics* 219.2 (2000), p. 182–191 (cf. p. 10, 13, 15, 16, 29).
- [DH05] Dirk DRASDO et Stefan HÖHME. « A single-cell-based model of tumor growth *in vitro* : monolayers and spheroids ». In : *Physical Biology* 2.3 (2005), p. 133 (cf. p. 10, 11, 13–16).
- [DH12] Dirk DRASDO et Stefan HOEHME. « Modeling the impact of granular embedding media, and pulling versus pushing cells on growing cell clones ». In : *New Journal of Physics* 14.5 (2012), p. 055025 (cf. p. 13, 15–17, 25, 28–31).
- [DHB07] Dirk DRASDO, Stefan HOEHME et Michael BLOCK. « On the role of physics in the growth and pattern formation of multi-cellular systems : What can we learn from individual-cell based models ? » In : *Journal of Statistical Physics* 128.1-2 (2007), p. 287–345 (cf. p. 10, 15, 16).
- [DKM95] Dirk DRASDO, R KREE et JS McCASKILL. « Monte Carlo approach to tissue-cell populations ». In : *Physical Review E* 52.6 (1995), p. 6635 (cf. p. 16).
- [DL01] Dirk DRASDO et Markus LOEFFLER. « Individual-based models to growth and folding in one-layered tissues : intestinal crypts and early development ». In : *Nonlinear Analysis : Theory, Methods & Applications* 47.1 (2001), p. 245–256 (cf. p. 13, 15, 16, 29).
- [Fin+11] Jenny FINK et al. « External forces control mitotic spindle positioning ». In : *Nature Cell Biology* 13.7 (2011), p. 771–778 (cf. p. 7).
- [Fri+06] Hermann B FRIEBOES et al. « An integrated computational/experimental model of tumor invasion ». In : *Cancer Research* 66.3 (2006), p. 1597–1604 (cf. p. 18).
- [GA07] Philip GERLEE et Alexander RA ANDERSON. « An evolutionary hybrid cellular automaton model of solid tumour growth ». In : *Journal of Theoretical Biology* 246.4 (2007), p. 583–603 (cf. p. 9).
- [GG92] François GRANER et James A GLAZIER. « Simulation of biological cell sorting using a two-dimensional extended Potts model ». In : *Physical Review Letters* 69.13 (1992), p. 2013 (cf. p. 9).
- [GG93] James A GLAZIER et François GRANER. « Simulation of the differential adhesion driven rearrangement of biological cells ». In : *Physical Review E* 47.3 (1993), p. 2128 (cf. p. 9).
- [GHA09] Jörg GALLE, Martin HOFFMANN et Gabriela AUST. « From single cells to tissue architecture – a bottom-up approach to modelling the spatio-temporal organisation of complex multi-cellular systems ». In : *Journal of Mathematical Biology* 58.1-2 (2009), p. 261–283 (cf. p. 10, 14).

- [Gib+11] William T GIBSON et al. « Control of the mitotic cleavage plane by local epithelial topology ». In : *Cell* 144.3 (2011), p. 427–438 (cf. p. 7).
- [GLD05] Jörg GALLE, Markus LOEFFLER et Dirk DRASDO. « Modeling the effect of de-regulated proliferation and apoptosis on the growth dynamics of epithelial cell populations in vitro ». In : *Biophysical Journal* 88.1 (2005), p. 62–75 (cf. p. 10, 13, 15, 16, 30).
- [Gos+06] John D GOSS-CUSTARD et al. « Test of a behavior-based individual-based model : response of shorebird mortality to habitat loss ». In : *Ecological Applications* 16.6 (2006), p. 2215–2222 (cf. p. 13).
- [GR13] Volker GRIMM et Steven F RAILSBACK. *Individual-based modeling and ecology*. Princeton University Press, 2013 (cf. p. 13).
- [Heg+11] Rachel A HEGEMANN et al. « Geographical influences of an emerging network of gang rivalries ». In : *Physica A : Statistical Mechanics and its Applications* 390.21 (2011), p. 3894–3914 (cf. p. 13).
- [Hil28] Archibald V HILL. « The diffusion of oxygen and lactic acid through tissues ». In : *Proceedings of the Royal Society of London. Series B, Containing Papers of a Biological Character* (1928), p. 39–96 (cf. p. 8).
- [Hon78] Hisao HONDA. « Description of cellular patterns by Dirichlet domains : The two-dimensional case ». In : *Journal of Theoretical Biology* 72.3 (1978), p. 523–543 (cf. p. 65, 102).
- [HW00] Douglas HANAHAN et Robert A WEINBERG. « The hallmarks of cancer ». In : *Cell* 100.1 (2000), p. 57–70 (cf. p. 2).
- [HW11] Douglas HANAHAN et Robert A WEINBERG. « Hallmarks of cancer : the next generation ». In : *Cell* 144.5 (2011), p. 646–674 (cf. p. 3).
- [Kan+00] Anuraag R KANSAL et al. « Simulated brain tumor growth dynamics using a three-dimensional cellular automaton ». In : *Journal of Theoretical Biology* 203.4 (2000), p. 367–382 (cf. p. 9).
- [KHM06] Jos KÄFER, Paulien HOGEWEG et Athanasius FM MARÉE. « Moving forward moving backward : directional sorting of chemotactic cells due to size and adhesion differences ». In : *PLoS Computational Biology* 2.6 (2006), e56 (cf. p. 9).
- [KRA12] Yoonseok KAM, Katarzyna A REJNIAK et Alexander RA ANDERSON. « Cellular modeling of cancer invasion : integration of in silico and in vitro approaches ». In : *Journal of Cellular Physiology* 227.2 (2012), p. 431–438 (cf. p. 10).
- [KSO07] Yangjin KIM, Magdalena A STOLARSKA et Hans G OTHMER. « A hybrid model for tumor spheroid growth in vitro I : theoretical development and early results ». In : *Mathematical Models and Methods in Applied Sciences* 17 (2007), p. 1773 (cf. p. 12, 14–16, 28, 30).

- [LF05] Terry LECHLER et Elaine FUCHS. « Asymmetric cell divisions promote stratification and differentiation of mammalian skin ». In : *Nature* 437.7056 (2005), p. 275–280 (cf. p. 7).
- [Mac+10] Paul MACKLIN et al. « Agent-based modeling of ductal carcinoma in situ : application to patient-specific breast cancer modeling ». In : *Computational Biology*. Springer, 2010, p. 77–111 (cf. p. 8, 10, 13, 14, 17).
- [Mac+12] P. MACKLIN et al. « Patient-calibrated agent-based modelling of ductal carcinoma in situ (DCIS) : from microscopic measurements to macroscopic predictions of clinical progression ». In : *Journal of Theoretical Biology* 301 (2012), p. 122–140 (cf. p. 15, 16, 28).
- [May32] William V MAYNEORD. « On a law of growth of Jensen’s rat sarcoma ». In : *The American Journal of Cancer* 16.4 (1932), p. 841–846 (cf. p. 8).
- [MBC11] Nicolas MINC, David BURGESS et Fred CHANG. « Influence of cell geometry on division-plane positioning ». In : *Cell* 144.3 (2011), p. 414–426 (cf. p. 7, 17).
- [Mei03] Laurent MEIJER. « Le cycle de division cellulaire et sa régulation ». In : *Oncologie* 5.7/8 (2003), p. 311–326 (cf. p. 7).
- [MN09] Charles M MACAL et Michael J NORTH. « Agent-based modeling and simulation ». In : *Winter Simulation Conference*. Winter Simulation Conference. 2009, p. 86–98 (cf. p. 13).
- [MPL01] Frank A MEINEKE, Christopher S POTTEN et Markus LOEFFLER. « Cell migration and organization in the intestinal crypt using a lattice-free model ». In : *Cell Proliferation* 34.4 (2001), p. 253–266 (cf. p. 10, 65).
- [MV07] Bertrand MAURY et Juliette VENEL. « Un modèle de mouvements de foule ». In : *ESAIM : proceedings*. T. 18. EDP Sciences. 2007, p. 143–152 (cf. p. 15, 27).
- [MV11] Bertrand MAURY et Juliette VENEL. « A discrete contact model for crowd motion ». In : *ESAIM : Mathematical Modelling and Numerical Analysis* 45.01 (2011), p. 145–168 (cf. p. 15, 27).
- [Nor+10] Kerri-Ann NORTON et al. « A 2D mechanistic model of breast ductal carcinoma *in situ* (DCIS) morphology and progression ». In : *Journal of Theoretical Biology* 263.4 (2010), p. 393–406 (cf. p. 10).
- [Oka+09] Atsuyuki OKABE et al. *Spatial tessellations : concepts and applications of Voronoi diagrams*. T. 501. John Wiley & Sons, 2009 (cf. p. 65).
- [Osborne+10] James M OSBORNE et al. « A hybrid approach to multi-scale modelling of cancer ». In : *Philosophical Transactions of the Royal Society A : Mathematical, Physical and Engineering Sciences* 368.1930 (2010), p. 5013–5028 (cf. p. 11, 14, 16, 65).

- [Pal08] Eirikur PALSSON. « A 3-D model used to explore how cell adhesion and stiffness affect cell sorting and movement in multicellular systems ». In : *Journal of Theoretical Biology* 254.1 (2008), p. 1–13 (cf. p. 10, 14, 15).
- [Pat+01] Aalpen A PATEL et al. « A cellular automaton model of early tumor growth and invasion : the effects of native tissue vascularity and increased anaerobic tumor metabolism ». In : *Journal of Theoretical Biology* 213.3 (2001), p. 315–331 (cf. p. 12).
- [Peu] Diane PEURICHARD. « Simple mechanical cues explain adipose tissue organization ». Thèse de doctorat. Université Toulouse III. En prép. (cf. p. 14).
- [PO00] Eirikur PALSSON et Hans G OTHMER. « A model for individual and collective cell movement in *Dictyostelium discoideum* ». In : *Proceedings of the National Academy of Sciences* 97.19 (2000), p. 10448–10453 (cf. p. 10, 14, 15).
- [PT09] Luigi PREZIOSI et Andrea TOSIN. « Multiphase and multiscale trends in cancer modelling ». In : *Mathematical Modelling of Natural Phenomena* 4.03 (2009), p. 1–11 (cf. p. 11, 23).
- [PT11] Jillian C PEASE et Jennifer S TIRNAUER. « Mitotic spindle misorientation in cancer – out of alignment and into the fire ». In : *Journal of Cell Science* 124.7 (2011), p. 1007–1016 (cf. p. 7).
- [PVS04] Shayn M PEIRCE, Eric J VAN GIESON et Thomas C SKALAK. « Multicellular simulation predicts microvascular patterning and in silico tissue assembly ». In : *The FASEB journal* 18.6 (2004), p. 731–733 (cf. p. 18).
- [Qua+08] Vito QUARANTA et al. « Invasion emerges from cancer cell adaptation to competitive microenvironments : quantitative predictions from multiscale mathematical models ». In : *Seminars in Cancer Biology*. T. 18. 5. Elsevier. 2008, p. 338–348 (cf. p. 9, 10, 12).
- [RA08] Katarzyna A REJNIAK et Alexander RA ANDERSON. « A computational study of the development of epithelial acini : I. Sufficient conditions for the formation of a hollow structure ». In : *Bulletin of Mathematical Biology* 70.3 (2008), p. 677–712 (cf. p. 17).
- [Ram+08] Ignacio RAMIS-CONDE et al. « Modeling the influence of the E-cadherin- β -catenin pathway in cancer cell invasion : a multiscale approach ». In : *Biophysical Journal* 95.1 (2008), p. 155–165 (cf. p. 11).
- [Ram+09] Ignacio RAMIS-CONDE et al. « Multi-scale modelling of cancer cell intravasation : the role of cadherins in metastasis ». In : *Physical Biology* 6.1 (2009), p. 016008 (cf. p. 11).
- [RCM07] Tiina ROOSE, S Jonathan CHAPMAN et Philip K MAINI. « Mathematical models of avascular tumor growth ». In : *Siam Review* 49.2 (2007), p. 179–208 (cf. p. 9, 11, 13).

- [Rej07] Katarzyna A REJNIAK. « An immersed boundary framework for modelling the growth of individual cells : an application to the early tumour development ». In : *Journal of Theoretical Biology* 247.1 (2007), p. 186–204 (cf. p. 10).
- [RM10] Katarzyna A REJNIAK et Lisa J McCAWLEY. « Current trends in mathematical modeling of tumor-microenvironment interactions : a survey of tools and applications ». In : *Experimental Biology and Medicine* 235.4 (2010), p. 411–423 (cf. p. 10).
- [SC01] Jonathan A SHERRATT et Mark AJ CHAPLAIN. « A new mathematical model for avascular tumour growth ». In : *Journal of Mathematical Biology* 43.4 (2001), p. 291–312 (cf. p. 11).
- [Shi+09] Abbas SHIRINIFARD et al. « 3D multi-cell simulation of tumor growth and angiogenesis ». In : *PloS one* 4.10 (2009), e7190 (cf. p. 10).
- [SM05] Gernot SCHALLER et Michael MEYER-HERMANN. « Multicellular tumor spheroid in an off-lattice Voronoi-Delaunay cell model ». In : *Physical Review E* 71.5 (2005), p. 051910 (cf. p. 10, 11, 13, 14, 16, 17, 25, 29–31, 50, 65, 102).
- [SM06] Gernot SCHALLER et Michael MEYER-HERMANN. « Continuum versus discrete model : a comparison for multicellular tumour spheroids ». In : *Philosophical Transactions of the Royal Society A : Mathematical, Physical and Engineering Sciences* 364.1843 (2006), p. 1443–1464 (cf. p. 11).
- [SP12] Marco SCIANNA et Luigi PREZIOSI. « Multiscale developments of the cellular Potts model ». In : *Multiscale Modeling & Simulation* 10.2 (2012), p. 342–382 (cf. p. 9, 10).
- [SP13] Marco SCIANNA et Luigi PREZIOSI. « A Cellular Potts Model simulating cell migration on and in matrix environments ». In : *Mathematical Biosciences and Engineering* 10.1 (2013), p. 235–261 (cf. p. 9, 10).
- [TB06] Manuel THÉRY et Michel BORNENS. « Cell shape and cell division ». In : *Current Opinion in Cell Biology* 18.6 (2006), p. 648–657 (cf. p. 7).
- [TBP07] Bryan C THORNE, Alexander M BAILEY et Shayn M PEIRCE. « Combining experiments with multi-cell agent-based modeling to study biological tissue patterning ». In : *Briefings in Bioinformatics* 8.4 (2007), p. 245–257 (cf. p. 13, 17).
- [Van+09] Ingeborg MM VAN LEEUWEN et al. « An integrative computational model for intestinal tissue renewal ». In : *Cell Proliferation* 42.5 (2009), p. 617–636 (cf. p. 10, 16).
- [VVB03] Katrien VERMEULEN, Dirk R VAN BOCKSTAELE et Zwi N BERNEMAN. « The cell cycle : a review of regulation, deregulation and therapeutic targets in cancer ». In : *Cell Proliferation* 36.3 (2003), p. 131–149 (cf. p. 7).

- [Wal+04a] DC WALKER et al. « Agent-based computational modeling of wounded epithelial cell monolayers ». In : *NanoBioscience, IEEE Transactions on* 3.3 (2004), p. 153–163 (cf. p. 17).
- [Wal+04b] DC WALKER et al. « The epitheliome : agent-based modelling of the social behaviour of cells ». In : *Biosystems* 76.1 (2004), p. 89–100 (cf. p. 17).
- [WSS07] Arwen C WILCOCK, Jason R SWEDLOW et Kate G STOREY. « Mitotic spindle orientation distinguishes stem cell and terminal modes of neuron production in the early spinal cord ». In : *Development* 134.10 (2007), p. 1943–1954 (cf. p. 7).
- [ZAD07] Le ZHANG, Chaitanya A ATHALE et Thomas S DEISBOECK. « Development of a three-dimensional multiscale agent-based tumor model : simulating gene-protein interaction profiles, cell phenotypes and multicellular patterns in brain cancer ». In : *Journal of Theoretical Biology* 244.1 (2007), p. 96–107 (cf. p. 11).

Protocole expérimental

Les expériences ont été réalisées sur des cellules d'adénocarcinome colique de la lignée HCT116 modifiées par transduction lentivirale pour exprimer une histone H₂B fusionnée à la protéine fluorescente mCherry ; cela permet de visualiser en microscopie de fluorescence les noyaux des cellules. Les cellules sont ensemencées dans des chambres de culture (Lab-Tek, Dutscher) à une densité de 7500 cellules/cm² dans du milieu OPTI-MEM supplémenté de 3 % de sérum de veau fœtal (SVF) et de pénicilline/streptomycine. Après 48 h, la densité cellulaire choisie permet d'obtenir sur le fond de la chambre de culture des îlots d'environ 50 cellules, isolés les uns des autres, ce qui permet de pouvoir suivre leur évolution individuelle en microscopie inversée de fluorescence en temps réel. Avant de réaliser les acquisitions, des groupes de cellules isolés sont sélectionnés et le fond de la chambre de culture traité manuellement, afin d'éviter que des cellules voisines puissent rejoindre le groupe principal.

Les acquisitions sont réalisées sur un microscope Axiovert (Zeiss) équipé d'une caméra CoolSnap HQ (Roper Scientific) et piloté par le logiciel Metavue. À chaque acquisition, plusieurs groupes individuels sont suivis parallèlement par vidéomicroscopie (1 image toutes les 10 min). Les images sont traitées en utilisant les logiciels Métamorph et Image J avant d'être analysées.

Glossaire

Ce glossaire est un recueil de quelques termes et acronymes biologiques : il s'agit des termes composés *en caractères inclinés* au cours du manuscrit. Les numéros de page composés en gras dans ce glossaire renvoient à une description plus exhaustive.

A | C | D | G | I | M | O | P | Q

A

ADN acide désoxyribonucléique. Molécule, présente dans toutes les cellules, qui porte l'information génétique et constitue le génome de l'être vivant. [2](#), [3](#), [6](#), [7](#), [25](#), [98](#)

angiogenèse Développement de nouveaux vaisseaux sanguins à partir du système sanguin préexistant ; ils permettent notamment d'irriguer les tumeurs cancéreuses. [3](#), [4](#)

apoptose Mécanisme de mort cellulaire programmée, c'est-à-dire processus d'auto-destruction de la cellule, en réponse à un signal intérieur ou extérieur. [3](#), [25](#)

C

caractéristiques fondamentales (d'une cellule cancéreuse) Il s'agit de propriétés supposées nécessaires pour qu'une cellule soit considérée comme cancéreuse. [2](#), [4](#), [5](#), [25](#)

cycle cellulaire Suite d'évènements que vit une cellule : il commence par la naissance de la cellule et se termine par sa division. [5](#), [6](#), [7](#), [11](#), [15](#), [23](#), [25](#), [28](#), [60](#)

D

différenciation (cellulaire) Processus par lequel une cellule se spécialise. Une cellule souche est une cellule indifférenciée capable de se spécialiser par différenciation. [6](#), [7](#)

G

génome Ensemble du matériel génétique, c'est-à-dire des molécules d'ADN, d'une cellule. [4](#)

I

interphase Partie du cycle cellulaire comprenant la croissance et la préparation de la division cellulaire. [6](#), [11](#), [15](#), [25](#), [28](#), [29](#), [31](#)

M

métastase Foyer secondaire de cellules cancéreuses. Il résulte de la dissémination dans un tissu de cellules cancéreuses appartenant à une tumeur initiale, localisée dans un autre organe du même organisme. [3](#), [4](#)

mitose Partie du cycle cellulaire comprenant la division de la cellule en deux cellules filles. [6](#), [11](#), [15](#), [25](#), [30](#), [31](#), [37](#), [71](#)

O

orientation de division Orientation dans laquelle, au cours de la mitose, s'éloignent les deux lots de chromosomes issus de la réplication de l'ADN. Cette orientation détermine la direction dans laquelle la cellule mère va se déformer en deux cellules filles. [7](#), [8](#), [12](#), [16](#), [17](#), [22](#), [24](#), [30](#), [37](#), [71](#), [136](#)

P

plan de division Plan dans lequel s'alignent les paires de chromosomes au début de la mitose, avant leur séparation en deux lots identiques. Ce plan détermine l'orientation de division. [6](#), [7](#), [24](#), [30](#), [33](#)

Q

quiescente En état de quiescence, c'est-à-dire en phase de repos. Cette état, aussi appelé stade G₀, correspond à une sortie (temporaire) du cycle cellulaire pendant laquelle la cellule ne peut ni croître ni se diviser. [6](#), [11](#), [12](#), [25](#)

ÉTUDE DE LA CROISSANCE TUMORALE VIA LA MODÉLISATION AGENT-CENTRÉ DU COMPORTEMENT COLLECTIF DES CELLULES AU SEIN D'UNE POPULATION CELLULAIRE

Résumé

L'objectif de cette thèse de mathématiques appliquées à la biologie est de comprendre des mécanismes fondamentaux impliqués dans la croissance de tumeurs cancéreuses non vascularisées.

Un nouveau modèle discret 2D de type agent-centré et son implémentation numérique sont établis pour étudier les liens entre, d'une part, la croissance et la division des cellules et, d'autre part, l'organisation de la population cellulaire ; ce modèle explore en particulier l'influence de l'orientation de la division cellulaire sur la population et ses lignages. Les cellules sont des disques se regroupant de façon isotrope, sans chevauchement : à chaque instant, leur position est obtenue par la résolution d'un problème de minimisation sous contraintes. La méthode utilisée pour l'implémentation numérique de cette minimisation repose sur l'algorithme d'Uzawa.

Des indicateurs statistiques ainsi qu'un procédé de visualisation des populations cellulaires sont mis en place, pour étudier à la fois les résultats mathématiques *in silico* et les résultats biologiques *in vitro* : ils mesurent des caractéristiques sur le comportement de la population et des différentes lignées qui la composent. Il en ressort une forte disparité entre les lignées situées au centre de la population et celles situées en périphérie. L'interaction mathématiques-biologie, inscrite au cœur de notre démarche, permet de réaliser une confrontation du modèle aux expériences : elle en apporte une première validation mais met en valeur des différences. Suivant une hypothèse formulée pour se rapprocher des expériences, un modèle enrichi est proposé : ses résultats sont en meilleur accord avec les expériences.

Mots-clés : modélisation agent-centré, minimisation sous contraintes, confrontation entre modèle et expériences, croissance tumorale, orientation de la division cellulaire, lignages cellulaires

AGENT-BASED MODELING OF THE COLLECTIVE BEHAVIOR OF CELLS IN A CELL POPULATION, TO STUDY TUMOR GROWTH

Abstract

This thesis in Mathematics applied to Biology aims at understanding key behaviors involved in avascular tumor growth.

A new 2D discrete model, agent-based, and its numerical implementation are introduced to study the relation between, on the one hand, cell growth and division and, on the other hand, the organization of the cell population. This model is specially designed to investigate the impact of the orientation during division on the population and on its lineages. Cells are represented as disks grouping together in an isotropic way, without overlapping themselves: at each time, their positions are given by the solution of a constrained minimization problem. This numerical resolution uses an algorithm based on the Uzawa method.

Some statistical indicators and a visualisation process of cell populations are set up, to study both the mathematical results (*in silico*) and the biological ones (*in vitro*): they measure distinctive information on the behavior of the population and its lineages. A large disparity between central and peripheral lineages is highlighted. The interaction between Mathematics and Biology, placed at the heart of our approach, enables us to confront the modeling with the experiments: this brings a first step toward validation but also shows some differences. We then state a hypothesis, the exploration of which leads to an improved model: its results are a better match to the experiments.

Keywords: agent-based modeling, constrained minimization, comparison between modeling and experiments, tumor growth, orientation of the cell division axis, cell lineages

Institut de mathématiques de Toulouse (IMT) – UMR 5219

Université Toulouse III – 118, route de Narbonne – 31062 TOULOUSE Cedex 9