Mathematical 3D modelling of connective tissues architecture emergence

Pauline Chassonnery

THÈSE

En vue de l'obtention du
DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par l'Université Toulouse 3 - Paul Sabatier

Présentée et soutenue par

Pauline CHASSONNERY

Le 19 décembre 2023

Modélisation mathématique en 3D de l'émergence de l'architecture des tissus conjonctifs

---

Ecole doctorale : **BSB - Biologie, Santé, Biotechnologies**

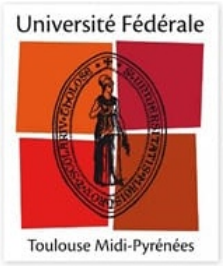
Spécialité : **BIO-INFORMATIQUE, GENOMIQUE ET BIOLOGIE DES SYSTEMES**

Unité de recherche :
RESTORE

Thèse dirigée par
Louis CASTEILLA et Diane PEURICHARD

Jury

Mme Raluca EFTIMIE, Rapporteure
M. Sébastien BENZEKRY, Rapporteur
M. Giacomo DIMARCO, Rapporteur
M. Louis CASTEILLA, Directeur de thèse
Mme Diane PEURICHARD, Co-directrice de thèse
Mme Noélie DAVEZAC, Présidente



THESIS

In order to become

DOCTOR FROM THE UNIVERSITY OF TOULOUSE

Awarded by the Université Toulouse 3 - Paul Sabatier

Presented and defended by

Pauline CHASSONNERY

On December 19th, 2023

Mathematical 3D modelling of connective tissues architecture emergence

---

Doctoral school : **BSB - Biologie, Santé, Biotechnologies**

Academic field : **BIO-INFORMATIQUE, GÉNOMIQUE ET BIOLOGIE DES SYSTÈMES**

Research unit :
RESTORE

Thesis supervised by
Louis CASTEILLA and Diane PEURICHARD

Committee

Mme Raluca EFTIMIE, Referee
M. Sébastien BENZEKRY, Referee
M. Giacomo DIMARCO, Referee
M. Louis CASTEILLA, Supervisor
Mme Diane PEURICHARD, Co-supervisor
Mme Noélie DAVEZAC, President

COLOPHON

Doctoral dissertation entitled “Mathematical 3D modelling of connective tissues architecture emergence”, written by Pauline Chassonnery, completed on December 19th, 2023, typeset with the document preparation system LaTeX.

This thesis has been prepared at

Laboratoire RESTORE

Université Toulouse III – Paul Sabatier
4bis Avenue Hubert Curien
Bâtiment INCERE
31100 Toulouse
France

☎ +33 5 34 60 95 01
Website https://restore-lab.fr/

Laboratoire Jacques-Louis Lions

Sorbonne Université
Campus Pierre et Marie Curie
4 place Jussieu
75005 Paris
France

☎ +33 1 44 27 42 98
Website https://ljll.math.upmc.fr/

It was partially funded by

- the **Institut national de recherche en informatique et en automatique**
- the **Cancer, Ageing & Rejuvenation graduate school**
- the **Alliance Sorbonne Université** through the Emergence project MATHREGEN under grant number S29-05Z101
- the **Agence Nationale de la Recherche** under grant number ANR-22-CE45-0024-01 and ANR-18-EURE-0003 in the framework of the Programme des Investissements d'Avenir.

Parfois, les scientifiques changent d'avis. De nouveaux développements suscitent de nouvelles réflexions. Si cela vous ennuie, pensez aux dommages causés au monde par ceux chez qui de nouveaux développements ne suscitent nulle réflexion d'aucune sorte.

Terry Pratchett, Ian Stewart & Jack Cohen
La Science du Disque-monde

Remerciements

Je tiens en premier lieu à remercier mes directeurs de thèse, Louis Casteilla et Diane Peurichard, qui m'ont accompagnée et soutenue tout au long de ces trois et quelque années. Diane, je reste émerveillée par ta capacité quasi magique à débugger mes codes sans même avoir besoin de les lire. Merci de m'avoir proposé un sujet de recherche aussi passionnant et de m'avoir propulsée, avec tous les filets de sécurité voulus, en dehors du monde des mathématiques droit dans un laboratoire de biologie. Merci Louis de m'avoir réceptionnée à l'autre bout de la trajectoire, avec toute la profondeur de ta réflexion philosophique sur nos champs disciplinaires et nos méthodologies respectives. Apprendre à sortir de ma zone de confort a été un véritable défi et l'un des grands enseignements de cette thèse. Merci énormément à tous les deux pour votre confiance, votre bienveillance, votre disponibilité, pour m'avoir encouragée à l'indépendance tout en sachant me remettre sur les rails lorsque je me perdais dans les détails, et pour tout ce que vous m'avez enseigné : la collaboration interdisciplinaire, la confiance en moi, la prise de recul, la gestion du temps... Je me sens privilégiée d'avoir bénéficié d'un encadrement aussi optimal pour ma thèse, tant sur le plan intellectuel que sur le plan humain.

Merci à toutes celles et ceux qui, chacun à leur manière, ont contribué à faire de cette thèse une expérience formatrice et enrichissante. Je salue en particulier Anne Lorsignol, Jenny Paupert et Marielle Ousset qui m'ont initiée à la biologie et Childérick Sévérac qui a ajouté à notre mélange complexe le liant de la physique. Merci à Laetitia Pieruccioni et Mathieu Vigneau d'avoir exploré avec moi les limites des capacités d'acquisition, de stockage et de traitement d'images du laboratoire Restore. Merci à Pierre Degond pour la sagesse et la hauteur de vue mathématique. Merci à Sinan Haliyo pour son aide cruciale sur le sujet de la visualisation en 3D. Merci aux membres de mon comité de suivi, Annabelle Collin et Patrick Cañadas, pour leur écoute et leurs encouragements qui m'ont redonné confiance lorsque je doutais. Merci à Sébastien Benzekry, Giacomo Dimarco et Raluca Eftimie d'avoir accepté de rapporter ce manuscrit et à Noélie Davezac d'avoir accepté de faire partie du jury.

Je remercie chaleureusement Arièle Rossi et Katia Bailey, membres des services de gestion de l'Université Toulouse III, pour leur implication et leur réactivité face à l'imbroglio administratif qu'a représenté cette thèse à cheval sur deux universités. Merci également à Sylvain Cantaloube, qui m'a tiré d'un sacré guêpier pour mon inscription en première année ; à Claire Jousseaume dont l'aide a été précieuse pour mon passage en commission doctorale ; et à Zoely Rakotomanga-Rajaonah qui a découvert avec moi une toute nouvelle procédure de remboursement des frais de mission.

Un grand merci aux membres des laboratoires Restore et Jacques-Louis Lions, mes deux foyers pendant ces trois dernières années. Du côté des parisiens, citons mes co-bureaux Antoine, Chourouk et Agùstin, dont l'amitié a illuminé mes séjours dans la capitale, ainsi que Pierre, Thomas, Lucas, Nicolai, Matthieu, Elena, María, Guillaume, Charles, Roxanne, Marcel et tant d'autres, pour nos discussions sans queue ni tête autour de la machine à café ou ailleurs. Merci aux toulousains qui m'ont accueillie à bras ouverts : Clémence et Suzy qui m'ont immédiatement prise sous leur aile ; les deux Annes qui m'ont hébergée lors de certains de mes nombreux séjours en Occitanie ; Anastasia qui m'a prêté le soutien émotionnel de ses chats ; Childérick pour les blagues et la bonne humeur permanente ; Clément héraut de la pause déjeuner ; Jean-François éternel Calimero ; Léa si courageuse convertie aux mathématiques ; David pour les discussions sans fin sur l'informatique et la modélisation ; Mélanie, Emmanuel, Haomiao, Jeanne, Miguel et tous les autres, pour votre chaleur et la bonne ambiance que vous cultivez dans notre open space. Merci enfin aux membres du Comité d'Animation Non-Scientifique de Restore qui nous ont

organisé tant de bons moments.

Je tiens à exprimer ma gratitude et ma reconnaissance envers mes professeurs qui, du primaire à la prépa, ont nourri ma soif de connaissances et de compréhension du monde : vous avez été merveilleux. Mention spéciale à M. Larivain qui m'a donné le goût des mathématiques, à M. Jourd'huy qui me l'a redonné après une année difficile et à toute l'équipe pédagogique de la prépa d'Évreux (enthousiaste, passionnée, surmotivée) qui m'a offert mes plus belles années de scolarité !

Merci enfin à ma famille de sang et de cœur qui a toujours été présente à mes côtés. Merci à mes parents pour leur soutien indéfectible, leurs encouragements et leurs conseils. Merci surtout et avant tout de m'avoir offert un foyer et une éducation propices à l'épanouissement. Merci à ma mère de m'avoir poussée à reconsidérer un projet professionnel établi de longue date pour envisager une voie qui m'apparait à présent comme une évidence. Merci à mon oncle Fred pour nos discussions sans fin sur les mathématiques et l'algorithmique, pour les stages de révision intensifs de ces deux matières et pour la relecture exhaustive du chapitre 2 du présent manuscrit. Merci à mon frère pour sa patience, à mes grand-parents pour leur amour, à Claude qui m'a toujours comprise, à Hélène qui partage ma passion pour la Fantasy, à Sarah qui m'a tant donné.

Une pensée particulière pour mon ami Mourad et on-va-dire-mon-oncle François : tout ce qui s'en vient, un jour ou l'autre s'en va.

Pour conclure, merci Jonathan pour toute l'aide que tu m'as apportée durant cette thèse, aussi bien au niveau de la réflexion que de l'écriture, et pour ton soutien précieux dans les moments difficiles. Merci pour toutes ces années partagées et, je l'espère, pour beaucoup d'autres à venir.

Résumé

Dans cette thèse, nous nous interrogeons sur la possibilité que des interactions mécaniques locales simples entre un nombre limité de composants puissent régir l'émergence de l'architecture 3D des tissus biologiques. Pour explorer cette hypothèse, nous développons deux modèles mathématiques. Le premier, ECMmorpho-3D, vise à reproduire un tissu conjonctif non-spécialisé réduit à la matrice extra-cellulaire, c'est à dire à un réseau 3D de fibres interconnectées dynamiquement. Le second, ATmorpho-3D, est obtenu par ajout de cellules sphériques qui apparaissent et croissent spontanément dans ce réseau de fibres afin de modéliser la morphogenèse du tissu adipeux, un tissu conjonctif spécialisé ayant une grande importance sur le plan biomédical.

Pour analyser les données produites par ces deux modèles, nous construisons un outil générique permettant de visualiser en 3D des systèmes composés d'un mélange d'éléments sphériques (cellules) et de bâtonnets (fibres) et de détecter automatiquement dans de tels systèmes des amas d'objets sphériques séparés par des bâtonnets. Cet outil peut également être utilisé pour traiter des images biologiques issues de microscopie en 3D, permettant ainsi une comparaison directe entre les structures *in vivo* et *in silico*.

L'étude des structures produites par le modèle ECMmorpho-3D via des simulations numériques montre que ce modèle peut générer spontanément différents types d'architectures, que nous identifions et caractérisons grâce à notre outil d'analyse. Une analyse paramétrique approfondie nous permet d'identifier une variable émergente, le nombre de liens par fibre, qui explique et, dans une certaine mesure, prédit le devenir du système modélisé. Une analyse temporelle révèle que l'échelle de temps caractéristique de ce processus d'auto-organisation est fonction de la vitesse de remodelage du réseau et que tous les systèmes suivent la même trajectoire évolutive.

Enfin, nous utilisons le modèle ATmorpho-3D pour explorer l'influence de cellules sphériques sur l'organisation d'un réseau de fibres dynamique, en prenant comme référence le modèle ECMmorpho-3D. Nous montrons que le nombre de cellules influence l'alignement local des fibres mais pas l'organisation globale du réseau. Par ailleurs, les cellules s'organisent spontanément en amas entourés de feuillets de fibres, dont les caractéristiques morphologiques sont très proches de celles des structures cellulaires *in vivo*. De plus, la distribution des différentes morphologies d'amas cellulaires est similaire dans les systèmes *in silico* et *in vivo*. Ceci suggère que le modèle est capable de produire des morphologies réalistes non seulement à l'échelle d'un amas mais aussi à l'échelle du système entier, en reproduisant les variabilités structurelles observées dans les échantillons biologiques. Une analyse paramétrique révèle que la proportion de chaque morphologie dans un système *in silico* est gouvernée principalement par les capacités de remodelage du réseau de fibres, pointant le rôle essentiel des propriétés de la matrice extra-cellulaire dans l'architecture et le fonctionnement du tissu adipeux (ce qui concorde avec plusieurs constatations biologiques ainsi que des résultats antérieurs en 2D).

Le fait que ces modèles mathématiques très simples puissent générer des structures réalistes corrobore notre hypothèse selon laquelle l'architecture des tissus biologiques pourrait émerger spontanément à partir d'interactions mécaniques locales entre les composants du tissu, indépendamment des phénomènes biologiques complexes se déroulant dans ce tissu. Ce travail ouvre de nombreuses perspectives quant à notre compréhension des principes fondamentaux gouvernant la manière dont l'architecture d'un tissu émerge durant l'organogenèse, est maintenue au cours de la vie et peut être affectée par diverses pathologies. Les applications potentielles vont de l'ingénierie tissulaire à la possibilité de promouvoir la régénération chez les mammifères adultes.

Mots clefs : Biologie mathématique ; Modèles individu-centrés ; Visualisation et segmentation en 3D ; Tissus conjonctifs ; Réseaux de fibres dynamiques en 3D ; Tissus adipeux.

Abstract

In this thesis, we investigate whether simple local mechanical interactions between a reduced set of components could govern the emergence of the 3D architecture of biological tissues. To explore this hypothesis, we develop two mathematical models. The first one, ECMmorpho-3D, aims at reproducing a non-specialised connective tissue and is reduced to the extra-cellular matrix component, that is a 3D dynamically connected fibre network. The second, ATmorpho-3D, is built by adding to this network spherical cells which spontaneously appear and grow in order to mimic the morphogenesis of adipose tissue, a specialised connective tissue with major biomedical importance.

We then construct a unified analysis framework to visualise, segment and quantitatively characterise the fibrous and cellular structures produced by our two models. It constitutes a generic tool for the 3D visualisation of systems composed of a mixture of spherical (cells) and rod-like (fibres) elements and for the automatic detection in such systems of clusters of spherical objects separated by rod-like elements. This tool is also applicable to biological 3D microscopy images, enabling a comparison between *in vivo* and *in silico* structures.

We study the structures produced by the model ECMmorpho-3D by performing numerical simulations. We show that this model is able to spontaneously generate different types of architectures, which we identify and characterise using our analysis framework. An in-depth parametric analysis leads us to identify an intermediate emerging variable, the number of crosslinks per fibre, which explains and partly predicts the fate of the modelled system. A temporal analysis reveals that the characteristic time-scale of the organisation process is a function of the network remodelling speed, and that all systems follow the same, unique evolutionary pathway.

Finally, we use the model ATmorpho-3D to explore the influence of round cells over the organisation of a fibre network, taking as reference the model ECMmorpho-3D. We show that the number of cells can influence the local alignment of the fibres but not the global organisation of the network. On the other hand, the cells inside the network spontaneously organise into clusters with realistic morphological features very close to those of *in vivo* structures, surrounded by sheet-like fibre bundles. Moreover, the distribution of the different morphological types of clusters is similar in *in silico* and *in vivo* systems, suggesting that the model is able to produce realistic morphologies not only on the scale of one cluster but also on the scale of the whole system, reproducing the structural variability observed in biological samples. A parametric analysis reveals that the proportion in which each morphology is present in an *in silico* system is governed mainly by the remodelling characteristic of the fibres, pointing to the essential role of the extra-cellular matrix properties in adipose tissue architecture and function (in agreement with several biological results and previous 2D findings).

The fact that these very simple mathematical models can produce realistic structures supports our hypothesis that biological tissues architecture could emerge spontaneously from local mechanical interactions between the tissue components, independently of the complex biological phenomena taking place around them. This opens many perspectives regarding our understanding of the fundamental principles governing how biological tissue architecture emerges during organogenesis, is maintained throughout life and can be affected by various pathological conditions. Potential applications range from tissue engineering to therapeutic treatment inducing regeneration in adult mammals.

Keywords : Mathematical biology ; Individual-based models ; 3D visualisation and segmentation ; Connective tissues ; 3D dynamical fibre networks ; Adipose tissues.

Table of Contents

List of Figures

List of Acronyms

AT Adipose Tissue

ECM Extra-Cellular Matrix

IBM Individual-Based Model

ATmorpho-2D 2D Individual-Based Model for Adipose Tissue Morphogenesis

ECMmorpho-3D 3D Individual-Based Model for Extra-Cellular Matrix self-organisation

ATmorpho-3D 3D Individual-Based Model for Adipose Tissue architecture emergence

ODE Ordinary Differential Equation

SDE Stochastic Differential Equation

CPU Central Processing Unit

List of Mathematical Notations

Notation	Signification	Mathematical expression
$\mathbb{N}$	the set of natural integers	$\{0,\ 1,\ 2,\ 3,\ ...\}$
$[\![a, b]\!]$	the set of all integers between a and b (included)	$\{a,\ a+1,\ ...,\ b-1,\ b\}$
$\mathbb{R}$	the set of all real numbers	$]-\infty, +\infty[$
$\mathbb{R}^+$	the set of all positive real numbers	$[0, +\infty[$
$\mathbb{R}^{+*}$	the set of all strictly positive real numbers	$]0, +\infty[$
$\mathbb{S}_2^+$	the north hemisphere of the 2-dimensional unit sphere	$\mathbb{S}_2 \cap \{z \geqslant 0\}$
$\perp$	the perpendicular relation	
$\cdot$	the scalar product operator	
$\wedge$	the cross product operator	
$\otimes$	the outer product operator	
$[u]_\wedge$	the matrix form of the cross product by the vector u	$\begin{pmatrix} 0 & -u_z & u_y \\ u_z & 0 & -u_x \\ -u_y & u_x & 0 \end{pmatrix}$
$\lvert x \rvert$	the absolute value of the real x	
$\lvert S \rvert$	the cardinal (i.e. number of elements) of the set S	
$\lVert \cdot \rVert$	the Euclidean norm	
$\mathrm{mean}(S)$	the mean of the set of reals S	$\frac{1}{\lvert S \rvert} \sum_{x \in S} x$
$\mathrm{sign}(x)$	the sign of the real x	$\begin{cases} -1 & \text{if } x < 0, \\ 0 & \text{if } x = 0, \\ +1 & \text{if } x > 0. \end{cases}$
$\lfloor x \rfloor$	the floor of the real x, i.e. the greatest integer lesser than or equal to x	
$\lceil x \rceil$	the ceil of the real x, i.e. the smallest integer greater than or equal to x	
$\underset{x \in S}{\mathrm{argmax}}\, f(x)$	the arguments of the maxima of the function f over the set S, i.e. the elements of S which maximise the value of f	$\left\{ y \in S \mid f(y) = \max_{x \in S}(f(x)) \right\}$

Chapter 1

State-of-the-art : Importance and challenges of 3D biological tissue modelling

Summary of the chapter

1. Motivation : Why model the emergence of 3D architecture in biological tissues ?

Mathematical biology can be defined as the use of tools from applied mathematics to study biological systems. Over the last decades, this field has seen rapid growth due to the conjunction of an increasing volume and complexity of biological data, an increasing availability of computing capability and a lot of efforts to popularise mathematics. Mathematical biology has now emerged as one of the prominent areas of interdisciplinary research [1, 2].

This field is usually divided in two main branches : data processing and modelling. The first consists in applying sophisticated numerical tools to conduct in-depth analysis of large and/or complex datasets such as those derived from genetic sequencing (e.g. single-cell analysis) or three-dimensional tissue imaging (e.g. light-sheet imaging) and render them in an understandable form. The second involves constructing mathematical representations of biological systems and performing computer simulations to investigate the principles governing their structure, development and behaviour. These two branches are highly intertwined, data processing being the two-way link between modelling and reality : complex processing is usually required to compare simulation results to biological data and demonstrate the relevance of a model, while a detailed analysis of biological data provides insights on a system that can in turn be used to improve modelling (see Figure 1.1).

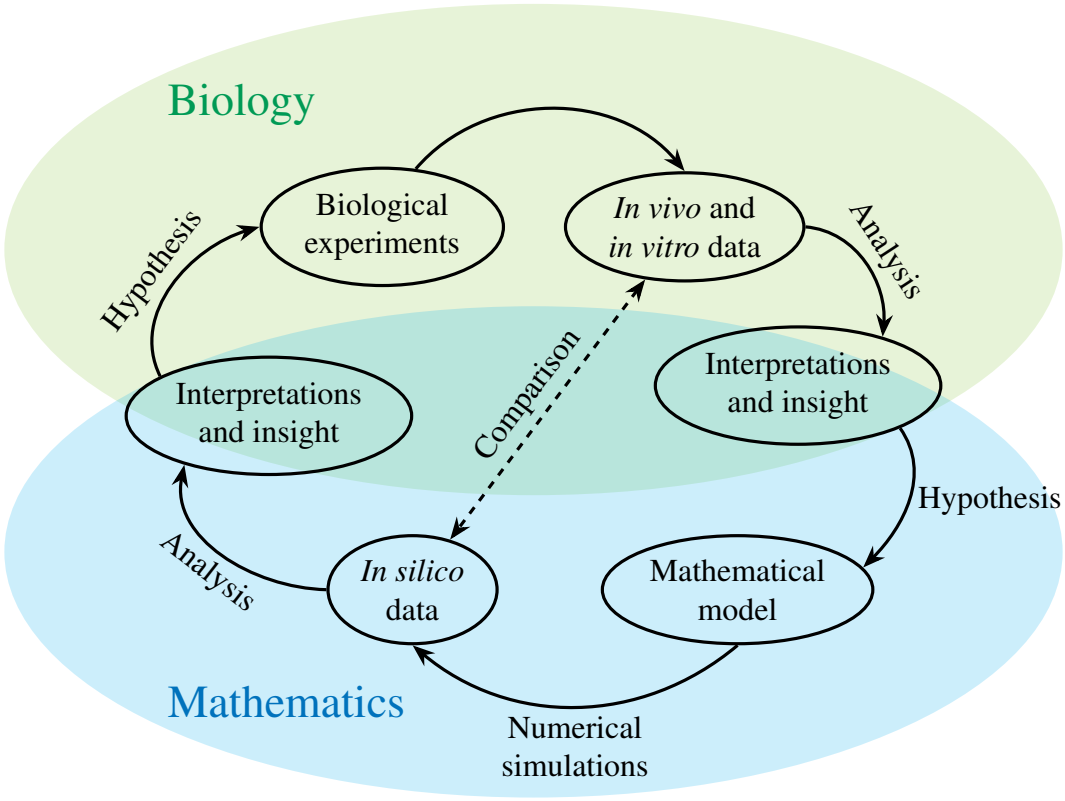


Figure 1.1: The cycle of mathematical biology.

Mathematical models offer a number of advantages over biological experiments :

- One can control all their input parameters freely.
- They allow to isolate a specific process or system from its intricate overall biological setting, which simplifies its study and makes it easier to understand.

- Because they are built in a general framework, they are often not specific to the system or process for which they were initially designed but can be applied to many other cases (e.g. molecular diffusion and population migration can be described by the same equations).

- By merely modifying a few parameters, they can be used to test several hypothesis and explore various scenarios, making *in silico* experiments cheaper and quicker to conduct than *in vivo* or *in vitro* ones.

- *In silico* experiments are also less risky (no use of toxic reagent, no risk of accidental release of bacteria, etc.) and raise fewer ethical issues (no animal testing). Although they do not altogether replace biological experiments, the information they provide allow experimenters to focus on promising avenues.

However, one of the most interesting properties of mathematical models is their ability to explore emerging phenomena. "Emergence" is defined as the fact that simple local rules generate complex overall structures in an unpredictable way [3] : that is, there is a causal link between the rules and the structures, but we are not able to determine which structure will emerge from a given set of rules apart by actually enacting those rules and looking at what happens. Writing a mathematical model and computing its solution(s) is a way of doing that. Thereafter, depending on the nature of the gap between the local rules and the observed structures, it may be possible to construct a logical narrative linking the two.

Emerging phenomena appear in many fields, but especially in biology [4, 5]. In a sense, it can be said that all biological systems exhibit emerging behaviours. For example, it seems that the architecture of a biological tissue is mainly due to local mechanical interactions, but how these interactions result in a given architecture is not well understood. The question of the architecture of a tissue is not a trivial matter since it is now well known that, in biology, architecture dictates function : every biological architecture fulfils one or more functions, and every biological function is performed by a specific architecture. This architecture/function relationship is a fundamental concept in biology [6, 7]. Of course, it also applies to many systems outside the field of biology, but not as systematically.

Furthermore, tissue architecture must be considered in all its 3D complexity. Until recently, the only available descriptions of biological tissue were essentially 2D, because tissue-scale imaging techniques only produced 2D images. However, improvements in 3D imaging techniques have shown that, in many cases, the 3D architecture was significantly more complex than biologists had thought based on 2D data. This encourages the development of 3D mathematical models and 3D data processing techniques to study these tissues.

In this thesis, we will investigate the hypothesis that the 3D architecture of biological tissues, and more specifically connective tissues, could be driven by simple mechanical interactions between the Extra-Cellular Matrix (ECM) and the surrounding cells.

2. Existing mathematical models for biological tissues

Before designing our own 3D model for the emergence of biological tissue architecture, we will give a quick overview of the two main classes of mathematical models that can be found in the literature for modelling biological tissue and discuss their pros and cons depending on the modelled system and the purpose of the model. A comprehensive review of the different types of mathematical models and methods used in biology can be found in [1].

2.1. Microscopic scale : Individual-Based Models

The principle behind Individual-Based Models (IBMs) is to represent a system by describing each of its components, called individuals or agents (hence the alternative name of Agent-Based Models), and the rules governing their actions and interactions. This is why IBMs are said to be microscopic models, even though the actual size of the components may be very large depending on the system modelled. Examples cover various scientific fields and range from molecular interactions inside a cell to flocking birds to clusters of stellar bodies :

- In atomic or molecular models, the agents would be protons, neutrons and electrons.
- In cellular models, the agents would be organelles and molecules (RNA, proteins, chemical signal, etc).
- In tissue or organ models, the agents would be various types of cells and molecules.
- In population models, the agents would be living beings.
- In astronomical models, the agents would be stars, planets, black-holes, etc.

The description of an agent consists in a list of features such as its size, mass, shape, position, speed, orientation, angular momentum, age, sex, infectious status, opinion, etc. As implied above, the same model can contain multiple types of agents, each type having its own list of features.

The second key element of an IBM is the set of rules describing how all these features vary over time. The choice of these rules is based on observations of real systems but also on the context and hypotheses underlying the model : a phenomenon that is prominent in one context may be negligible in another. All these rules are usually formalised as functions whose variables can be various features of the different agents, as well as external influences such as random processes, and can be continuous or discrete in time.

IBMs are the most widely used models in mathematical biology [8,9]. They are very simple to design and their description is usually quite intuitive. They are also very flexible and can take into account multiple aspects of a problem on multiple scales, discrete local properties and local random processes that do not average out over time or on a global scale.

Because of their design, IBMs are especially adapted to study the emergence of complex collective properties and behaviours from simple individual ones, including biological tissues self-organisation. This topic has given rise to numerous models of interacting cells and/or fibres. Some focus on cellular interactions in an environment where the ECM is either absent or acts as an external force, for instance when studying tumour growth [10]. Others focus on the properties of fibre networks where beams with various degrees of flexibility are connected by crosslinks [11–16]. Lastly, interactions between cells and fibres can involve one way or reciprocal mechanical interactions [17], contact guidance (the fibres providing directional information for cell motion) [18] or crosslink remodelling through cell action [19, 20].

2.2. Macroscopic scale : Continuous Models

Continuous models, on the other hand, simplify systems made up of a large number of components by only taking into account their average features and interactions, and do not explicitly model the individual components. They rely on the theory of the propagation of chaos which states that, as the number of components in a system tends to infinity, the behaviour of each component will become independent of

the rest of the system : this makes looking at each component individually meaningless to understand the global features. For example, pressure and temperature are macroscopic features emerging from the random motion of many particles, and the exact trajectory of one particle is totally irrelevant.

This is why continuous models are said to be macroscopic : they only look at the big (macroscopic) picture and neglect the granularity of the system. The prerequisite for this class of models is that the modelled system must be composed of a very high number of components and show small changes on the length scale of a single component.

Contrarily to IBMs where the features are spatially discrete (with one value per component), in continuous models the features are continuous in space, meaning that the variables representing them are scalar fields or vector fields instead of scalars or vectors. They can be roughly separated into two types : those that express the average of individual components features (e.g. average size, average speed or average age) and those that emerge from the components collective actions (e.g. temperature, pressure or viscosity). As for IBMs, the rules governing the features temporal evolution are formalised as functions of these features, and their choice depends both on empirical observations and on the context and hypotheses of the model. They are often derived from conservation relations such as the law of conservation of mass or the law of conservation of energy.

[21] gives numerous examples of continuous models from different biological fields : population models for single or multiple interacting species, reaction-diffusion models for chemotaxis, propagation models for infectious diseases or parasite invasions, etc.

In the more specific case of biological tissues, continuous models can be used to describe the behaviour of large cohorts of cells [22] and of large swaths of fibrous materials. For example, the ECM can be represented as a poroelastic environment [23], a two-phase viscous fluid or an active polar gel [24], while a strained muscle or sinew can be modelled as an assembly of undulating fibres hooked between two edges and subjected to tensile load [25].

2.3. Link between the two scales

For the most part, the distinction between an IBM and a continuous model is a matter of resolution : do the user want or need a resolution high enough to distinguish the individual components of a system, and up to which limit ? After all, everything is made of atoms and elementary particles, so the most accurate way to model any system would be to describe it down to the last electron... except that in most cases this is neither feasible nor relevant.

For an IBM, the computational cost of a numerical simulation scales with the number of agents modelled, which itself often scales with the size of the system through a quadratic (in 2D) or cubic (in 3D) law. On the other hand, the computational cost of a continuous model is independent of the size of the simulated system, enabling the study of large spatial and/or temporal scales. Moreover, continuous models are easier to analyse theoretically, which allows to predict some of their properties without going through numerical simulations. Their main downside is of course their lesser resolution, that is the loss of information at the microscopic level, which can be problematic.

The choice of a class of model is thus mostly a matter of evaluating the resolution needed for each part of the system studied and balancing them with the technical constraints induced by the global scale of the study. This can lead to mixed models where some components are modelled by individual agents and other by continuous fields. For example, [26] presents a model for the organisation of ECM fibres into aligned structures in which the ECM is represented as a (continuous) vector field whose direction is locally modified by the motion of (discrete) fibroblasts.

As stated above, continuous models are mainly used for large-scale systems (compared to the size

of a typical component) with little need for a microscopic resolution, whereas IBMs are used when microscopic accuracy is necessary for the purpose of the study, even if it means reducing the scale of the system studied. Difficulties arise when one need to study a large-scale system up to a microscopic resolution. For instance, studying the architecture of a biological tissue requires a resolution at the cellular level on the scale of the whole tissue. This technical constraint is all the more acute given that we want to study the 3D architecture.

This is one of the reasons why most of the computational models developed thus far for the study of biological tissue architecture are two-dimensional [11, 12, 14–17, 27–31]. Few studies have been conducted on 3D models [13, 19, 32–36], although these are expected to yield different, more realistic results than 2D ones since they better mimic biological architectures themselves immersed in 3D environments. Another reason is the lack of high quality data in 3D : this impeded the development of 3D mathematical models because their results could not have been validated against biological data. However, the recent emergence of high resolution 3D imaging technologies and their wider availability has removed this obstacle.

The usual practice for dealing with such a dilemma can be summarised as follows : try to design an IBM with a manageable cost, and if not possible design a continuous model as accurate as possible. The first step includes optimising the algorithmic implementation of the model, but also reducing its intrinsic complexity by neglecting or simplifying some phenomena in a first approach, even if it makes it less realistic. The simplified model is then used as a stepping stone : if it produces completely false results, then perhaps its founding hypotheses are wrong and should be abandoned altogether or one of the neglected phenomena was predominant. If it produces results accurate enough for our needs, then the phenomena neglected were not strictly necessary to answer the original question and hence should not be included according to Occam's razor. If it produces relatively good, but not accurate enough results, then some of the neglected phenomena may need to be integrated in the model, and the additional numerical cost is justified by the expectancy of better results.

An example of this strategy is to precede a 3D IBM by a 2D version : this reduced model, although less realistic, would also be much less costly and could serve as a concept proof, guiding the development of its 3D successor. Of course, this is not systematic : sometimes the adjunction of the third dimension does not increase the computational cost so much (e.g. with sparse systems such as in astronomy) and sometimes, or rather often, the adjunction of the third dimension greatly changes the fundamental properties of a system (e.g. a random walk on a 2D grid will almost surely pass again through its starting point, but a random walk on a 3D grid only have a 34% chance to do so). The success (or failure) of a 2D model is thus not always a reliable indicator for 3D.

Concerning biological tissue architecture, the computational gains of using a 2D (instead of 3D) model is very high, but the relevance of the 2D approach is questionable, e.g. because it is unlikely that linear elements such as the ECM fibres have similar behaviour in 2D and 3D environments. On the other hand, until recently there was a lack of 3D biological data to which the results of a 3D model may have been confronted for validation and calibration. Hence, for tissues such as the adipose tissue, whose 3D architecture was believed to be qualitatively equivalent to its 2D representation, the 2D approach is natural. And indeed, the results obtained by Peurichard and al. with their 2D model were strikingly impressive [17, 37] (see section 4.3 for a summary). However, the 3D architecture of adipose tissue have recently been shown to be much more complex than what was suggested by the observation of 2D data. The combination of these two facts justifies the investment of human and computational resources in the development of a 3D model for adipose tissue architecture derived from [17], and thus leads to the present thesis.

The second step may be enacted even if the numerical cost of the IBM is manageable. It consists in deriving a continuous model from an IBM through a change of scale in space and time [38, 39] : this process is called a micro-macro (or mean-field) derivation. The results of the ensuing continuous model can be compared to that of its parent IBM to see if the emerging properties and structures were preserved through the derivation. The continuous model will then either replace the IBM (if the cost of the IBM is intractable and/or if the continuous model is accurate enough to answer the initial issue) or be used alongside it to quickly perform large scale simulations that will later be refined by slower, more detailed small-scale simulations. Examples of micro-macro derivations in the context of biological tissues modelling can be found in [27–31, 40–42].

Rigorously establishing the link between the microscopic and macroscopic scales is the main aim of the kinetic theory [1, 21, 38, 39]. However, when modelling complex biological systems, it is not uncommon for classical techniques to fail due to the nature of the agent interactions, namely : the lack of conservation relations, the existence of multiple equilibria and symmetry-breaking phase transitions and the appearance of correlations preventing the propagation of chaos.

3. Data processing in 3D

As mentioned before, the other main branch of mathematical biology besides modelling is data processing. In the last decades, both the volume and complexity of biological data have increased significantly. As a result, data processing is more and more often becoming a bottleneck for biological research [43].

In this section, we will discuss two aspects of data processing that are relevant to our work and are particularly complex in a 3D setting : data rendering (or visualisation) and structures segmentation. The development of an appropriate 3D data processing framework represented an important part of this thesis and will be detailed in chapter 3.

3.1. Data visualisation

First, a little terminology. We say that data is rendered in 3D, or that a visualisation is three-dimensional, when the resulting visual is itself three-dimensional. However, these three dimensions are not in any way compelled to represent the three spatial dimensions. In the same way that a 2D visual can either be a photography or a graphic showing the dependance between two variables (e.g. medication dose and efficiency), a 3D visual can either be a spatial representation of the (3D) system under study or a graphic representing the dependance between multiple variables. For the sake of clarity, we will speak of images (whether they are 2D or 3D) in the first case and of graphics in the second.

Data visualisation is sometimes treated as an optional step mostly aimed at aesthetics, but it is in fact a crucial tool for transforming data into knowledge [44, 45]. It has three main uses : (i) checking for errors or misconceptions when developing and implementing a new analysis protocol, (ii) ascertain that an analysis protocol is appropriate for the data under study, and (iii) conveying knowledge to fellow researchers and the general public.

While points (i) and (iii) are fairly well known, point (ii) tends to be overlooked. Let us give two practical examples :

- Concerning numerical quantification, the relevance of a quantifier such as the mean depends on how the data are distributed. If the visual shows that the data points form a regular point cloud around their mean value, then the mean (together with the standard deviation) is indeed a good

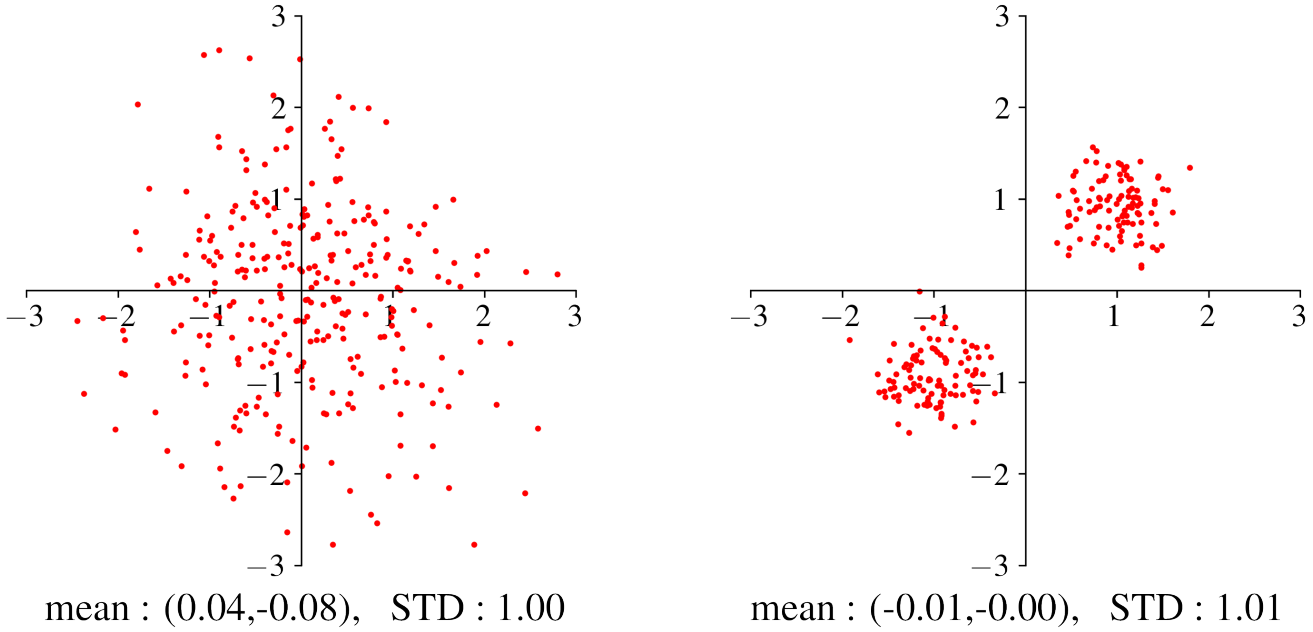


Figure 1.2: Why visualisation is necessary to interpret data. Example of two point clouds with different distributions but similar mean and standard deviation (STD). The data on the left was sampled from a gaussian distribution centred on $(0, 0)$ with scale 1, while the data on the right was sampled from two gaussian distributions of respective centres $(-1, -1)$ and $(1, 1)$, and identical scale 0.3.

quantifier. But if the visual shows the presence of two distinct clusters of points, then the mean is probably not appropriate and can even be misleading. An illustration is shown in Figure 1.2.

- Concerning objects segmentation, the segmentation protocols usually involve a number of parameters for which the appropriate values may not be obvious or deducible from empirical knowledge. Hence, the only way to gauge the influence of these parameters and to determine their appropriate values is to visualise the result of the segmentation to check whether or not it corresponds to the expected result. This question is addressed in more details in the next section, where it is illustrated with Figure 1.5.

It is important to understand that visualisation is a double-edged sword : a well done visual can be a very powerful tool for interpreting data, but a badly done one can lead to biased interpretations. This is even more true in the case of 3D visuals, because most visualisation devices (piece of paper, computer screen) are themselves 2D : a good 3D visual is thus very difficult to produce. The recent development and increased availability of technologies such as virtual reality [46,47] or 3D printing [48,49] are making this less of an obstacle, but do not remove it entirely.

For this reason, mathematicians tend to discourage the use of 3D visuals when they are not strictly necessary. However, there are indeed numerous cases where 3D rendering is necessary, particularly when the object being studied is itself 3D : if a system unfolds in a 3D space, it is unlikely that 2D visuals will be enough to get a good idea of its layout (especially for non-specialists). This is the case for most biological systems, which are fundamentally three-dimensional and can only be understood as such. Figure 1.3 gives an example of both a 3D graphic and a 3D image, together with their 2D counterparts.

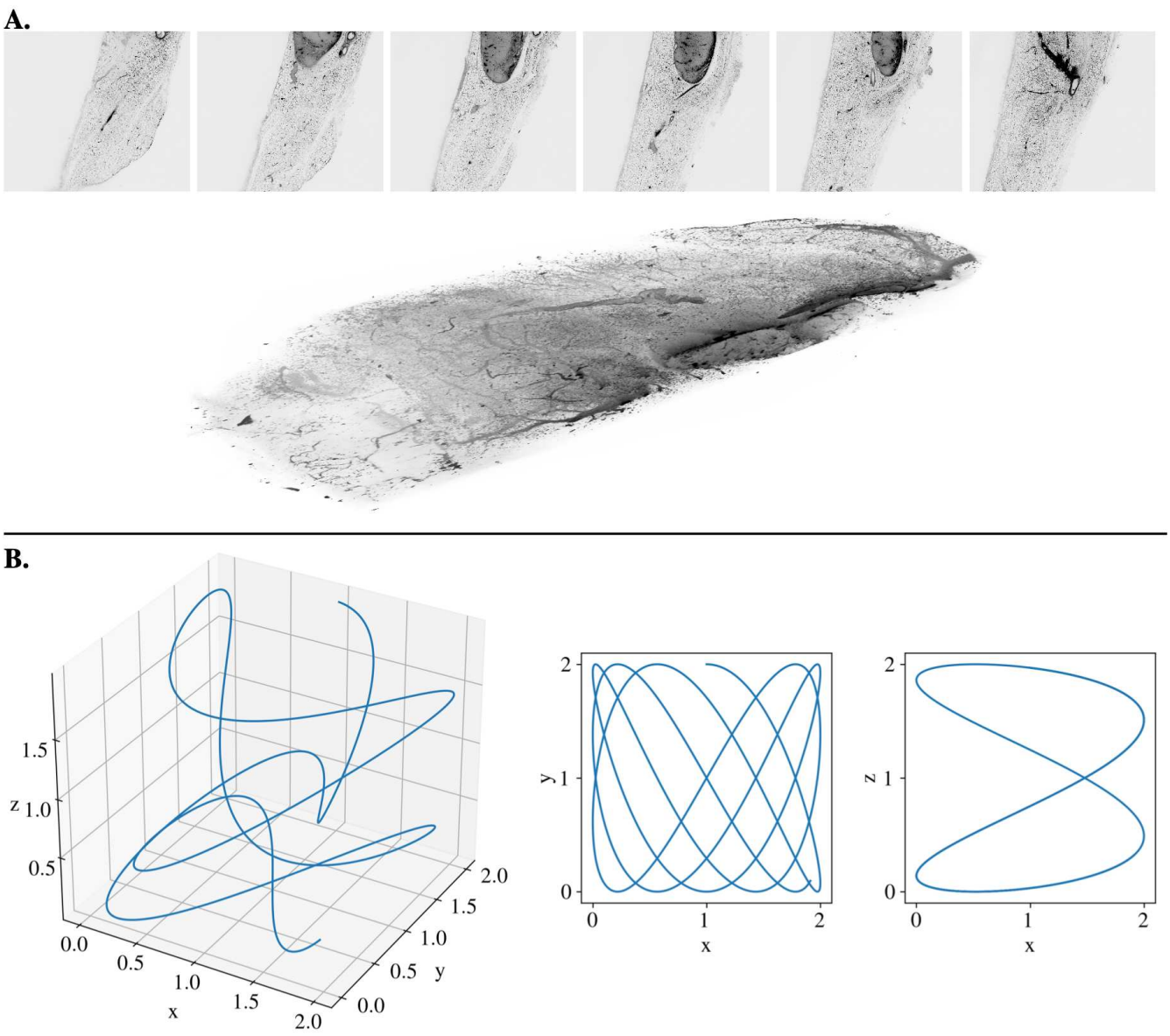


Figure 1.3: Why 3D visualisation is sometimes necessary : 3D structures can be difficult to grasp from 2D visuals. **A :** Images of a large portion of a mouse subcutaneous adipose tissue depot, acquired at the Restore Institute using light-sheet imaging on a lectin-stained tissue. The 2D slices (top row) are combined using the Paraview software to produce the 3D view (bottom row). **B :** Example of a 3D parametric curve (on the left) and of two of its planar projections (on the right).

3.2. Structures segmentation

In its most general meaning, structure segmentation (or data clustering) is the process of partitioning data into multiple subsets, so that each subset forms a coherent set designated as a structure or object. More commonly, the term refers to the same process applied to images. The process in itself extracts a lot of informations from the data and facilitates further analysis (e.g. by comparing the subsets between them), but also allows to create visuals that are more meaningful and easier to interpret. Structure segmentation is a wide-spread problem which have given rise to a major branch of data analysis.

It must be noted that segmenting structures, and creating words to designate them, is the very basis of

the spontaneous human approach to understanding our environment. It is also the basis of the scientific approach. Humans continuously and unconsciously perform structure segmentation, e.g. to identify from their field of vision meaningful structures such as roads, walkways, buildings, pedestrians and so on. This same process is at work when a doctor identify bones, organs or cancer tumours in CT scans. When studying tissue architecture, biologists will try to identify (and then proceed to name) basic architectural units such as planes of fibres or clusters of cells (see sections 4.1 and 4.2 below).

Automatic structure segmentation (performed by a computer rather than a human) covers distinct issues depending on whether the task could have been carried out manually or not : the gains and obstacles are different. First of all, it is easier to evaluate the performance of an algorithm if we can compare its results to what a human would have made. Otherwise, ascertaining the validity of the segmentation performed by an algorithm can be very difficult.

When automating a process that can be performed by a human, the main gains to be expected are increased speed (and consequently the capacity to treat much larger volume of data) and increased accuracy or reduced bias (although algorithms can "inherit" the human bias of their developers). Consider the example of cancer tumours : there is lot of important informations that a doctor can extract from a CT scan, such as the tumour size and shape, its proximity with a sensitive organ, its solid or infiltrating status... To acquire these information, the doctors have to separate the tumour from the rest of the image, something they do easily thanks to their training and experience. However, doctors usually segment only one (or at best a few) 2D slices of a 3D scan, because processing every slices is tedious - but it allows to compute the volume of the tumour, which is a more accurate quantifier of its scale and potential dangerousness than its average diameter on a few slices. Moreover, different doctors will not segment a tumour in the exact same way, which makes it difficult to compare data acquired by different individuals in a medical survey. Automating the segmentation with a dedicated software will thus provide many advantages.

When the segmentation can not be performed by a human, something which often happens with 3D data, then the gain is obvious : the algorithm will make possible what was not before.

In any cases, a segmentation algorithm may uncover previously unknown structures, opening the way to new interpretations and hypotheses. Many methods have been developed to solve the problem of data segmentation. We list below some of the most widely used.

Model-based segmentation assumes that the structures of interest have, or tend to have, a specific shape. Therefore, one can look for groups of points that fit this shape and evaluate the significance of such fit (e.g. the probability of it being due to a random distribution) [50]. An example is the RANSAC (RANdom SAmple Consensus) algorithm [51] which can be summarised as follows :

1. Consider a dataset containing N points and a shape that can be fully defined from n points (with $n \ll N$);
2. Randomly select n points in the dataset and draw the corresponding shape;
3. Count how many points of the dataset fit this shape according to a preset tolerance threshold;
4. Repeat steps (2) and (3) a preset number of times, each time retaining the best fit (i.e. the fit containing the most points);
5. Finally, check if the best fit is valid, e.g. if it contains more points than what could have been obtained with a random gaussian dataset of N points.

An illustration is displayed in Figure 1.4 : it can be seen that the classic linear regression method is not appropriate, because the data does not contain a single linear structure but two. The RANSAC

method will detect the largest of these structures (by number of points) and, if run again on the remaining unfitted points, will also detect the smallest.

One of the advantages of RANSAC is that it can detect structures among a large number of outliers (it is thus robust to noise). Another is that the user does not have to predict the number of structures present in the data : running RANSAC recursively until the structures detected are no longer valid automatically determines this number. One of its downsides is that it is a single model method (it can search for lines or circles, but not for both at once), so that if more than one type of shapes (or models) coexist in the data, RANSAC may fail to detect either of them.

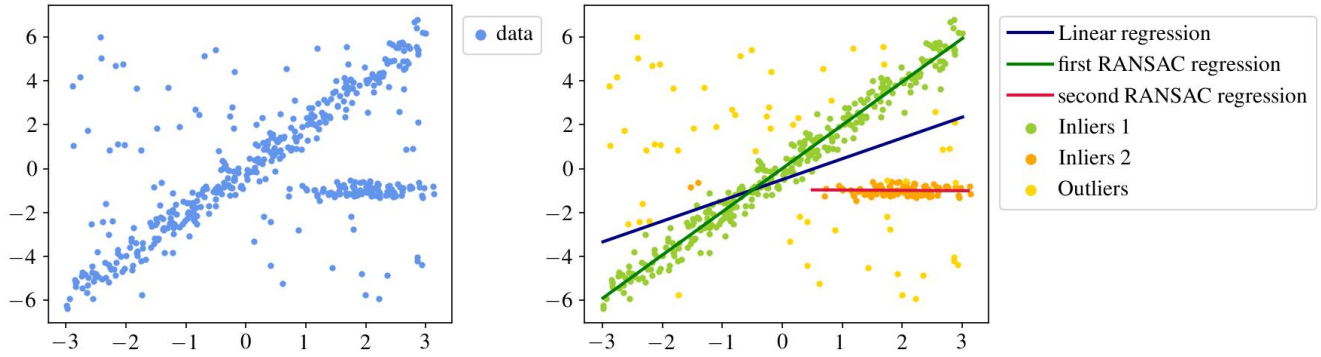


Figure 1.4: Example of model-based segmentation (here line detection) in a dataset using RANSAC. **Left :** raw data. **Right :** data fitted either by a linear regression (in blue) or two successive RANSAC regressions (respectively in green and red).

Centroid-based or k-means clustering consists in partitioning a dataset into an chosen number of clusters in a way that minimises the distance between each point and the centre of its assigned cluster [52]. As for connectivity-based clustering, here the notion of distance can involve many parameters other than spatial distance. This can be seen as choosing the optimal locations for a chosen number of sources (the centre of the clusters) to supply specific spots (the data points). However, this seemingly simple problem is extremely hard to solve and algorithms usually resort to heuristic approaches which may converge towards a local optimum instead of the global solution. For example, the classic k-mean algorithm [53] can be summarised as follows :

1. Consider a dataset that should be partitioned into k groups;
2. Pick k cluster centres;
3. Assign each data point to the cluster whose centre is closest;
4. Recompute each cluster centre as the mean of all the data point assigned to that cluster;
5. Repeat steps (3) and (4) until the process converges (i.e. no points change cluster anymore).

The main downside of this algorithm is that the quality of the solution reached is highly dependent on the choice of the number of clusters and of the initial cluster centres. On the other hand, it is guaranteed to converge and has a comparatively very small computational cost, which makes it easy to apply to large datasets. Moreover, its shortcomings can be improved by various optimisation procedures, such as running it with different values of k and retaining the best segmentation (based on a quality-assessment of the clusters), or using prior knowledge of the dataset to select the initial cluster centres. See first column of Figure 1.5 for an illustration.

Connectivity-based or hierarchical clustering consists in grouping data points together based on their respective distance, producing a hierarchy of clusters that gradually merge together at increasing connecting distances [54]. The final partition requires the user to chose either the connecting distance or the number of clusters to produce.

Note that the definition of the distance function can include many parameters apart from purely spatial distance, such as pixel colour and intensity (for an image). These methods can detect structures with very different shapes and sizes, but are sensible to outliers which will either show up as autonomous clusters or cause other clusters to merge by artificially connecting them (see second column of Figure 1.5 for an illustration).

Density-based segmentation assumes that the structures of interest define areas of higher density than the remainder of the dataset. Therefore, all points in the same structure can be linked by a densely packed path, i.e. a series of points that are close to each other and have a dense neighbourhood [55]. This is called "density-reachability". In contrast to many other methods, density-based segmentation allows the simultaneous segmentation of various arbitrary shapes. The most widely used method is DBSCAN [56], which has a low computational cost. It only requires the user to choose a connecting distance and a minimal density to connect, not to predict the number of structures in the dataset.

The main drawback of these methods is that dense structures are assumed to be separated by sparser areas, so that nearby or overlapping structures can not be distinguished (see third column of Figure 1.5 for an illustration).

Spectral graph clustering consists in segmenting a dataset based on the eigenvectors of its similarity matrix [57, 58]. It can be summarised as follows :

1. Measure the "similarity" or "affinity" between each pair of data points and compile these measures in the form of a matrix;
2. Compute the eigenvalues of this matrix and retain the k largest;
3. Map the original data points to $\mathbb{R}^k$ using the k associated eigenvectors;
4. Group the resulting vectors based on their largest components (i.e. to which eigenvector does the original data point contribute the most).

Note that some variants of this algorithm automatically determine the optimal number of clusters k from the distribution of the eigenvalues, for example by maximising the drop (called eigengap) between the last retained eigenvalue and the first rejected one. See fourth column of Figure 1.5 for an illustration.

Watershed transformation is an image segmentation method based on the geographical concept of water catchment basins [59]. The image is seen as a topographic surface where water placed on a "high" pixel will run downhill towards a local minimum. Pixels draining to the same point define a catchment basin, and each catchment basin is a separate structure. This is the method we finally retained to segment our data; it will be described more fully in chapter 3.

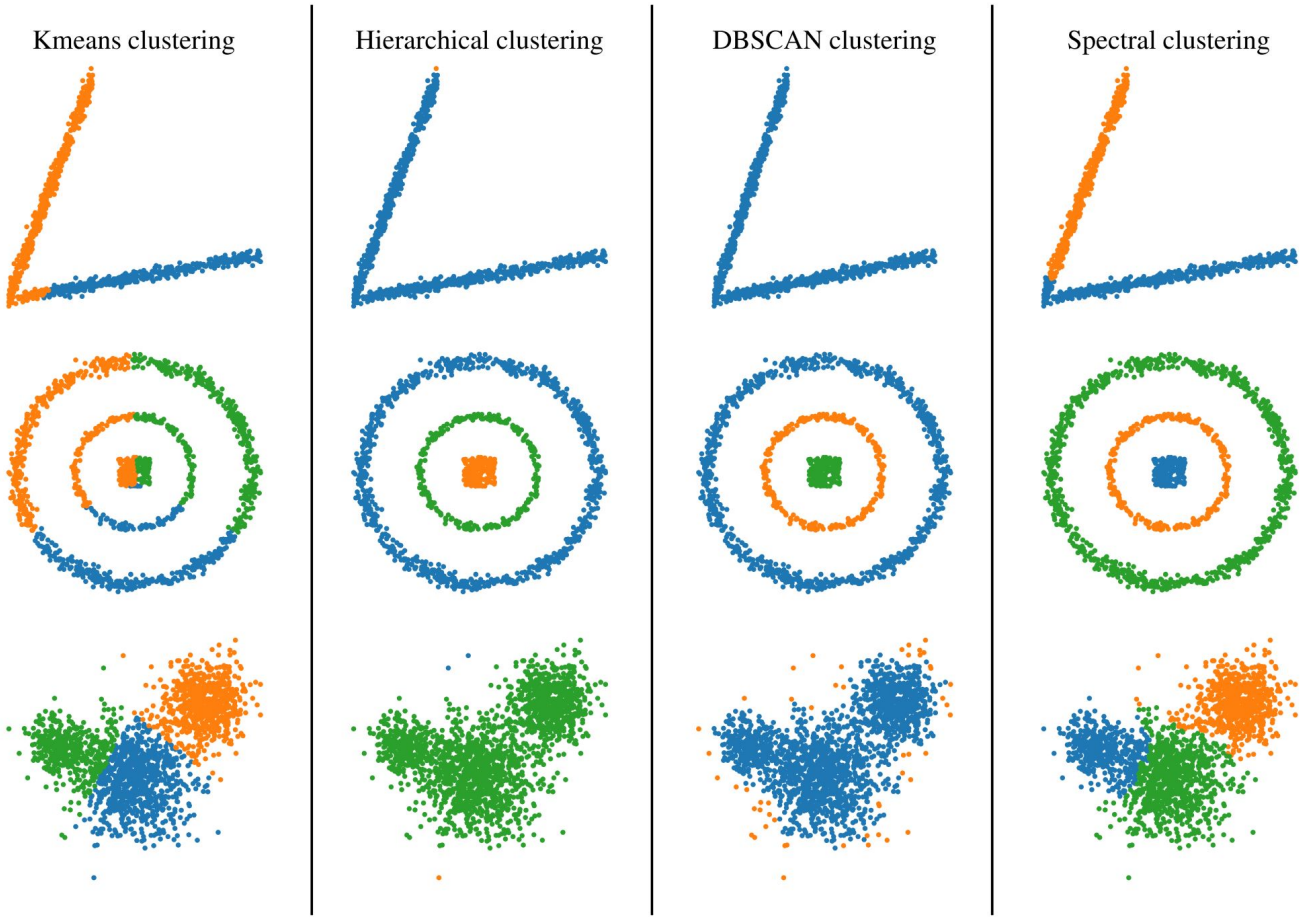


Figure 1.5: Illustration of the results produced by various segmentation methods on some typical test data. **From left to right :** k-means clustering with user-defined number of clusters (2 for the first case, 3 for the second and third cases); Agglomerative hierarchical clustering with user-defined number of clusters (2 for the first case, 3 for the second and third cases); DBSCAN density-based clustering with automatic computation of the number of clusters; Spectral clustering with automatic computation of the number of clusters from the eigengap.

4. Biological tissue architecture supports their functions

In this section, we will give a brief overview of the fundamental principles underlying biological tissues architecture. Given that an adequate architecture is mandatory for the efficient physiological function of any organ, it is crucial to understand how this architecture emerges during the organogenesis and is maintained throughout life.

4.1. Connective tissues : importance of the extra-cellular matrix

Animal biological tissues are classified into four main types [60] : (i) muscle tissues that ensure the individual's mobility, (ii) nervous tissues that ensure communication inside the body and between the organism and its environment, (iii) epithelial tissues that form a protective barrier between the organism and its external environment (covering e.g. the skin, the digestive track or the airways) and (iv) connective tissues. The latter are the most abundant, accounting for approximately two thirds of the body volume.

As the term implies, "connective tissue" is a generic name given to several body tissues that connect, support and bind together other tissues. They are composed of stromal cells, fibres whose arrangement form the Extra-Cellular Matrix (ECM) and interstitial fluid [61]. The proportions of these three components (cells, ECM and fluid) vary greatly from one connective tissue to another. Connective tissues thus form a very heterogeneous family ranging from blood (a fluid nearly deprived of ECM) to soft adipose tissue to highly stiff bone tissue (due to a calcified ECM).

Connective tissues constitute the basic brick of all organs. They provide mechanical and biochemical support for the functional cells, sustain and connect together the various parts of an organ and the various organs of the body, but also transport nutrients and wastes, store energy or mineral and participate in immune defence [62]. Depending on their role, they are usually subdivided into connective tissues proper (loose or dense) and specialised connective tissues (comprising blood, cartilage, bones and adipose tissue).

Due to their great diversity, connective tissues can be affected by numerous pathologies, ranging from surgical conditions (e.g. tendon tears and bony fractures) to autoimmune diseases (e.g. scleroderma and rheumatoid arthritis) and genetic disorders (e.g. joint hyper-mobility syndrome and brittle bone disease) [61, 62]. Moreover, their dysfunctions have been shown to be involved in several pathologies such as fibrosis and cancer.

While all tissues contain ECM, connective tissues hold the greatest part, enabling their mechanical and architectural roles [63]. The ECM is a three-dimensional network of macromolecules (divided into fibrous proteins and proteoglycans) connected by various types of enzymatic and non-enzymatic crosslinking agents (among which glycoproteins) [64, 65]. Collagen fibres are the most abundant component, accounting for around 90% of the ECM, and can represent up to 30% of the total protein mass in vertebrates. To date, more than 20 types of collagens have been identified in vertebrates [66]. They can self-assemble but also bind to other ECM components, and provide tensile strength to tissues [67, 68]. Collagens are associated with elastin, the main component of elastic fibres : it is a fibrous, insoluble, hydrophobic and extremely stable protein which provides recoil to tissues. A third major type of ECM fibre is the fibronectin, which forms a mechanical link between the ECM and the cells and plays a key role in cell adhesion and migration. Proteoglycans on the other hand are structural macromolecules composed of a protein core covalently linked to one or more glycosaminoglycans (a type of carbohydrate chain). They are very hydrophilic and, accordingly, form a hydrated gel which fills the ECM interstitial space and enables tissues to withstand high compressive forces.

The composition and spatial structure of the ECM varies over time in response to various cells activity : ECM components are continuously synthesised and spatially reorganised, mainly by fibroblasts, and at the same time degraded by a number of enzymes [69]. Their mechanical properties and crosslinks can also be altered. Moreover, crosslinks can unbind spontaneously or under tension, which leads to viscoplastic material responses such as softening and tension relaxation [70].

This dynamic nature of the ECM has been shown to be crucial to maintain tissue integrity and function.

The mechanical support provided by the ECM is twofold : it operates as a buffer against traction and compression loads, and it creates sheet-like deposits on which functional cells can rest. This latter point means that the ECM is indeed the skeleton on which the spatial organisation of a tissue is built.

Moreover, beyond the purely physical question of resistance to tensile stress, the mechanical properties (e.g. stretching and bending ability) of the ECM have a deep impact on the tissue function : they participate in the regulation of many cellular processes such as cell migration, cell differentiation, cell proliferation, cell death and gene expression (the last point being not yet fully understood). These proper-

ties can vary by several orders of magnitude depending on the network architecture, which itself depends on many factors such as the relative proportion of collagen and elastin fibres (e.g. a 9 to 1 ratio in bones). Overall, collagens fibres are mainly involved in rigidity, elastin fibres in elasticity and proteoglycans in viscosity.

As a whole, understanding the fundamental principles behind the ECM three-dimensional architecture is crucial to comprehend many physiological and pathological conditions (e.g. degenerative diseases or tumour invasion [71]). Pathologies may affect the nature and/or synthesis of the ECM (collagenopathies, fibrosis), or its remodelling (arthrosis, cancer, fibrosis) [72, 73]. For example, ageing is characterised by decreasing secretion of elastin fibres and increasing secretion of crosslinking molecules [74, 75], a combination leading to a stiffer ECM with less remodelling ability [71]. As a result, the ECM becomes less organised and more fragmented, hence weakening tissue integrity and strength [76, 77].

4.2. The adipose tissue : a specialised connective tissue with biomedical importance

The Adipose Tissue (AT) is one of the four types of specialised connective tissues. For decades, it has been studied only for its role energy storage. However, with the discovery of the leptin in the 1990s, our vision of this tissue changed considerably. Indeed we know now that adipose tissue, through its endocrine activity, interacts with the rest of the body and contributes to many physiological functions, notably thermoregulation, lipid homeostasis, immune responses and reproduction. This explains why the slightest dysfunction of AT can have major pathological consequences.

Cellular composition and physiological roles

The main functional cell type of AT is the adipocyte or fat cell, whose primary function is to absorb carbohydrates and lipids from the blood to store them into their cytoplasm as triglycerides droplets, and to release glycerol and free fatty acids into blood circulation as the result of triglycerides hydrolysis. Adipocytes thus contribute to regulate energy availability within the body according to metabolic needs. The whole process is finely controlled by hormones such as insulin (to which adipocytes react by increasing their storage and decreasing their release) or adrenalin (which has the opposite effect) and also by innervation from the autonomous nervous system. In addition, adipocytes release numerous specific proteins (e.g. leptin, ghrelin, adiponectin, cytokines) involved in energy balance [78, 79], appetite regulation [80], immune function [81], reproduction, etc.

Adipocytes represent around 50% of all cells in AT and make up for most of its volume. The remainder (called the stromal vascular fraction) is made of different cell types : adipocyte precursors (or pre-adipocytes), mesenchymal stem cells, fibroblasts, vascular endothelial cells (which make up blood vessel walls) and a large variety of immune cells (e.g. macrophages, lymphocytes, eosinophils or mast cells) [82].

Two types of adipocytes, white and brown, are generally described in literature, allowing a rough classification of AT depots according to their "colour". White adipocytes specialise in energy storage [83], but also displays major endocrine and immunological functions [84]. They contain a single lipid droplet that can swell until it represents $\approx 90\%$ of their volume, giving the cell a spherical shape and a very large diameter (up to 100 µm) [85].

Brown adipocytes are smaller than white adipocytes (typically between 15 µm and 50 µm in diameter) [85] and contain several small lipid droplets. They also contain more numerous and larger mitochondria [86]. Moreover, these mitochondria exhibit in their inner membrane a uniquely high concentration of thermogenin (also called uncoupling protein 1 or UCP-1) [87, 88], a protein that interferes with

the classic ATP energy-production process to generate heat. Brown adipocytes thus specialise in adaptive non-shivering thermogenesis, a fundamental process in maintaining body temperature [89].

However, this cell dichotomy (and consequently the dichotomy of AT depots) is not so strict. Indeed, since the 1990s an intermediate type of adipocyte, called beige or brite (for "brown in white"), have been described [90–92]. These adipocytes have characteristics from both white and brown adipocytes and display phenotypic plasticity : in resting conditions they are closer to white adipocytes, but they can "brown" (i.e. acquire a brown phenotype with numerous mitochondria and small lipid droplets supporting a thermogenic activity) through cold exposure [93, 94], diet [95] or exercise [96] and "whiten" again in a warm environment [97, 98], thus suggesting a real cell plasticity.

In addition, marrow [99] and pink [100] adipocytes (found in bone marrow and mammary glands respectively) have also been described with characteristic of both white and brown adipocytes, further complicating the landscape of the "adipocytes field".

Anatomy of adipose depots

White AT depots (defined as containing white adipocytes) represent the largest volume of AT in most mammals, including humans. They can be found as a more or less developed subcutaneous layer throughout the entire body and around major blood vessels and organs (e.g. heart, digestive tract, kidney and eye-balls). This visceral fat is mostly located in the abdominal cavity (it is sometimes called abdominal fat) and protects organs against mechanical impacts. The subcutaneous layer on the other hand provides thermal insulation for the body and contains very large adipocytes.

Brown AT depots (defined as containing brown adipocytes) are mostly present in infant (which are nearly deprived of subcutaneous AT and have not yet acquired diet-induced shivering) and/or hibernating mammals. Although it was long thought that these depots disappeared after birth, their presence has recently been demonstrated in the abdominal and thoracic cavities of adult humans, where they can represent 4.3% of all AT [101, 102]. These depots are more vascularised and innervated by the sympathetic nervous system than white ones [89]. Their exact nature, brown or beige, is still debated.

One of the characteristics of all these adipose depots is their extraordinary plasticity. Indeed, they are capable of massive expansion or reduction in response to changes in energy balance (from 5% to 60% of body mass). The reversibility of these changes depends on their amplitude. In the case of energy excess, AT depots expand through the swelling of white adipocytes, a phenomenon called hypertrophy. It also causes beige adipocytes to whiten [95, 103]. When existing adipocytes are no longer able to absorb excess energy, this hypertrophic phase is followed by an hyperplasia (differentiation of pre-adipocytes into new adipocytes) which results in a quasi-permanent increase of the number of adipocytes in the body [104]. In the case of an energy deficit, the reduction of the AT size is mainly due to adipocyte shrinkage following triglycerides hydrolysis, with no major change in adipocyte number.

Finally, despite their distinct locations, specific vascular and nerve supplies and distinctive metabolic abilities, all AT depots are now considered to form a single large soft "organ" [103, 105, 106].

Architecture of adipose tissues

The first description of adipose tissue architecture was made by Wassermann in 1965 on the basis of 2D photonic microscopy images. It evidenced the existence of ellipsoidal clusters of adipocytes (called lobules) surrounded, but not totally separated, by well organised sheets of ECM (called fascia), each of them irrigated by a tree-like vascular system [107]. Usually, mature white AT depots contain large, compact lobules densely packed together, while brown AT depots contain smaller and better separated lobules. Nevertheless, since Wassermann's observations most of the studies concerning AT have focused on its cellular composition and anatomical location rather than its architecture. Studies on tissular organ-

isation of AT depots remain descriptive without investigating its origin. The question of the emergence of AT architecture is therefore entirely neglected.

For several years, Pr. Casteilla's team investigated adipose tissue architecture, confirming the existence of strong heterogeneity between depots (some have a clear lobular architecture and others do not) and from one area to another inside the same depot. Indeed, exploration of the whole subcutaneous AT in mice have evidenced a strong regionalisation of the lobular architecture, associated with different browning ability of the adipocytes [108]. Adipocytes in the inner region display high browning ability and are organised in a network of interconnected lobules, while the periphery is a shapeless, unstructured layer of very large adipocytes with poor browning ability [108]. Moreover, image segmentation of autofluorescence or vascularisation signals revealed that the lobules have an unexpectedly complex 3D shape with several ellipsoidal lobes and many digitations [109], far from the grain of rice or olive pit observed on 2D sections. In addition, the poly-lobed units can be grouped, on the basis of their connectivity, into two clusters harbouring distinct molecular profiles and thermogenic ability [109].

These results highlight the importance of the architecture/function relationship.

Development of adipose tissues

AT is the last tissue to be formed during embryonic development. An intrascapular depot of brown adipose tissue appears just before birth (between embryonic days 18 and 19 out of 21 in mice [110, 111]) so as to protect the newborn against the thermal shock of the birth. Brown adipose depots decline as the individual ages, but do not disappear completely. White adipose tissue appears just after birth (first three postnatal days in mice) and continue to develop throughout life.

Figure 1.6 illustrates the successive development stages of subcutaneous AT in mice.

- At birth or postnatal day 0 (P0), the depot is made of small, rounded, well-delineated and rather spaced clusters of small adipocytes (see Figure 1.6.P0). Higher magnification allows to distinguish the significant space between lobules and the tree-like organisation of the vasculature.

- Following birth, the lobules begin to swell and merge together, likely under the conjugated effects of adipocytes growth (hypertrophy) and pre-adipocytes differentiation (hyperplasia). The inter-lobular spacing thus diminishes and the boundaries between lobules are less well-defined (see Figure 1.6.P1).

- Several days later, the lobules are so densely packed that they are hard to distinguish from each other (see Figure 1.6.P3) even with higher magnification. The tree-like organisation of the vasculature also becomes less apparent.

The fact that the lobular architecture is present from the start is questioning. However, as previously mentioned, the questions of how and why this specific spacial structure emerges have not been studied.

Obesity and adipose tissue dysfunctions

As mentioned above, white adipocytes participate in long-term energy storage by accumulating energy intake exceeding daily metabolic needs and releasing it during periods of food shortage or intense physical exercises. Brown adipocytes, for their part, participate in the control of energy expenditure by storing and burning energy excess. All adipose tissue depots thus play a key role in the control of energy balance, and any AT dysfunction will necessarily lead to the disruption of energy homeostasis and the development of miscellaneous diseases such as obesity.

Obesity is defined as an abnormal or excessive accumulation of fat (body mass index greater that 30), which is detrimental to health. According to the World Health Organisation (WHO), the prevalence

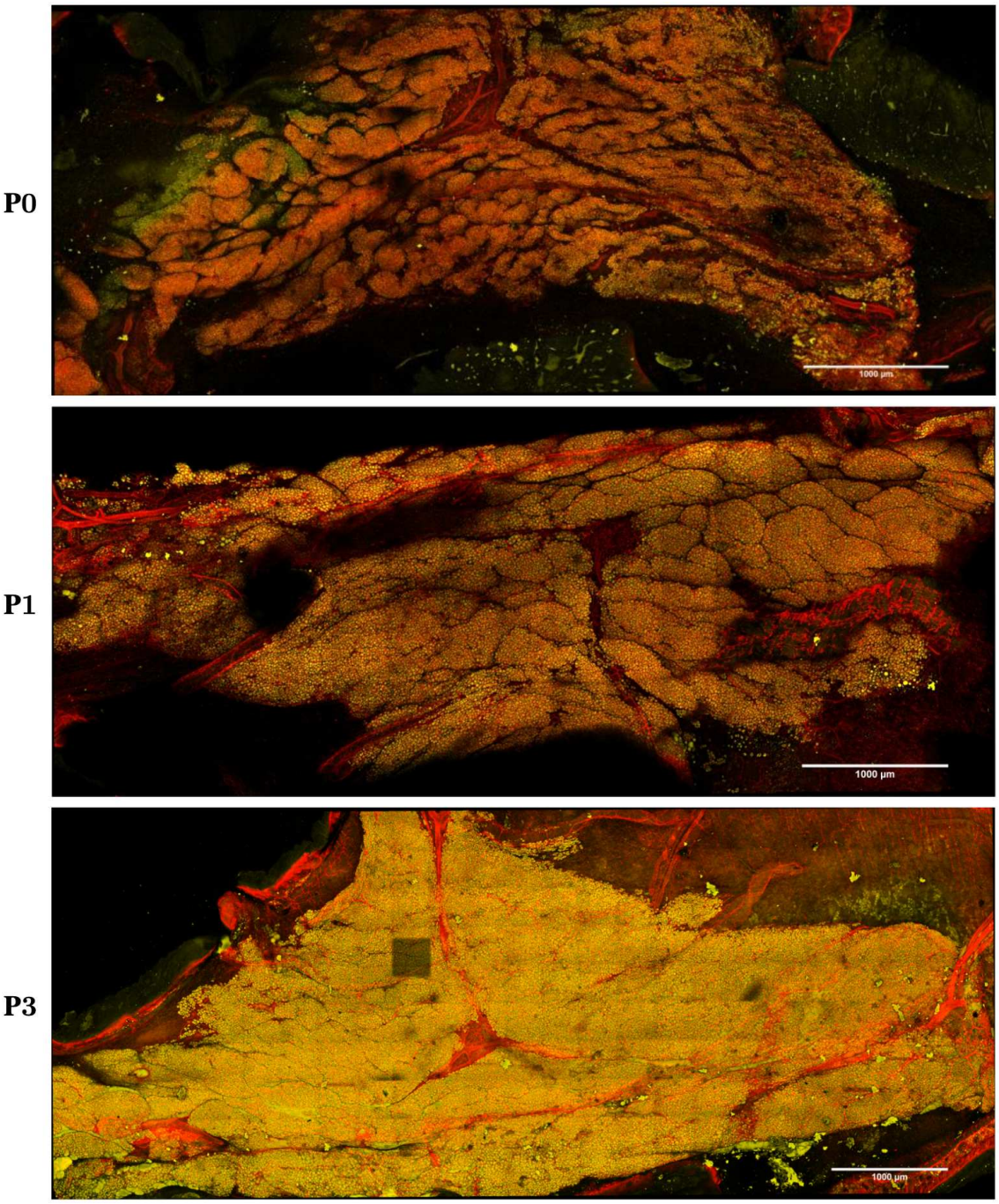


Figure 1.6: Images of a subcutaneous adipose tissue depot in juvenile mice, with adipocytes in yellow (bodipy staining) and vasculature in red (lectin staining). The white scale bars at the bottom right correspond to 1000 µm. **From top to bottom :** birth (P0), first day after birth (P1) and third day after birth (P3).

of obesity has nearly tripled since 1975, affecting 13% of all adults (over 650 million people) in 2016 while 39% (over 1.9 billion people) were overweight, with strong geographical disparities [112, 113]. Obesity is considered as the main risk factor for many metabolic diseases such as type 2 diabetes, arterial hypertension, atherosclerosis, fat liver disease and chronic kidney disease. It is also associated with an increased risk of cancer (including breast, uterine and liver cancers) or neurodegenerative diseases. Obesity is therefore responsible for at least 2.8 million deaths per year, making it the fifth leading cause of death according to the WHO [113].

Excessive weight in itself exerts on the body a major mechanical strain with many detrimental consequences, but is at first protective against metabolic diseases, because it prevents excess lipids from being stored in cells that are not suitable for this purpose. In most cases of obesity, despite its expansion the AT is no longer capable of efficiently storing lipids or expending energy. This leads to ectopic lipid deposition, for example in the liver or muscles, which in turn cause lipotoxicity, insulin resistance and inflammatory state development. This is why body mass index alone does not accurately reflect obesity-related risks : individuals who are able to greatly increase their amount of functional AT are clinically healthier than comparatively lighter individuals with dysfunctional AT.

In addition, histological studies of AT from obese induced mice or obese patients have revealed the existence of fibrosis [114–116]. This excessive ECM deposition must exert mechanical constrains that inhibit adipocyte hypertrophy, thereby contributing to AT dysfunction [117–119].

Taking into account that the architecture plays a key role in the function of the tissue, a better understanding of the rules governing the development and preservation of AT architecture is thus essential in the fight against the epidemics of obesity-associated diseases.

Conclusion

The AT is a good starting point for a study on the emergence of biological tissues architecture for at least three reasons. First, it has a relatively simple architecture, meaning that it can reasonably be expected to be easy to reproduce with a mathematical model. Second, lobular architectures can be found in many other biological tissues such as the liver (hepatic lobules), lungs (pulmonary alveoli) and pituitary gland, making AT a generalist object of study. Third, this connective tissue plays a key physiological role and is thus involved in diseases with a high prevalence. Considering the importance of the architecture/function relationship, a mathematical model providing a better understanding of how AT architecture emerges should have a positive impact on the medical field.

4.3. Description of a 2D model for adipose tissue self-organisation

We will now give a detailed example of mathematical model for adipose tissue self-organisation : the two-dimensional individual-based model for adipose tissue morphogenesis and regeneration created by Peurichard et al. [17, 37] (subsequently abbreviated as ATmorpho-2D). We will describe the model construction and give a brief summary of the results obtained with it. This model was validated against biological data and served as the basis for our 3D models for the self-organisation of connective tissues (see chapter 2).

The goal of the model ATmorpho-2D was to reproduce the emergence of the overall architecture of mature white AT and to identify the parameters driving the type of structure produced. The starting hypothesis was that the structures observed in white AT (clusters of adipocytes wrapped in thin sheets of well-organised extra-cellular matrix) could emerge merely from mechanical interactions between the adipocytes and the ECM fibres (hence the need of a microscopic model). This hypothesis has been

validated in [17], where the authors showed that their model produces architectures that compare quantitatively well to experimental observations, and that, counter-intuitively, modelling the vascular system was not necessary.

In [37], the authors showed that this relatively simple model can also be used to explain the way adipose tissue is reconstructed after injury, leading either to regeneration (the tissue regains its original functionalities) or scarring (the new tissue is not fully functional) depending on a few key parameters.

The main features of the model ATmorpho-2D are summarised in Figure 1.7. In the following discussion, we consider a finite, rectangular spatial domain $\Omega \subset \mathbb{R}^2$ with periodic border. Although the present text draws heavily from [17,37], a number of notations were changed so as to harmonise with the description of our own models.

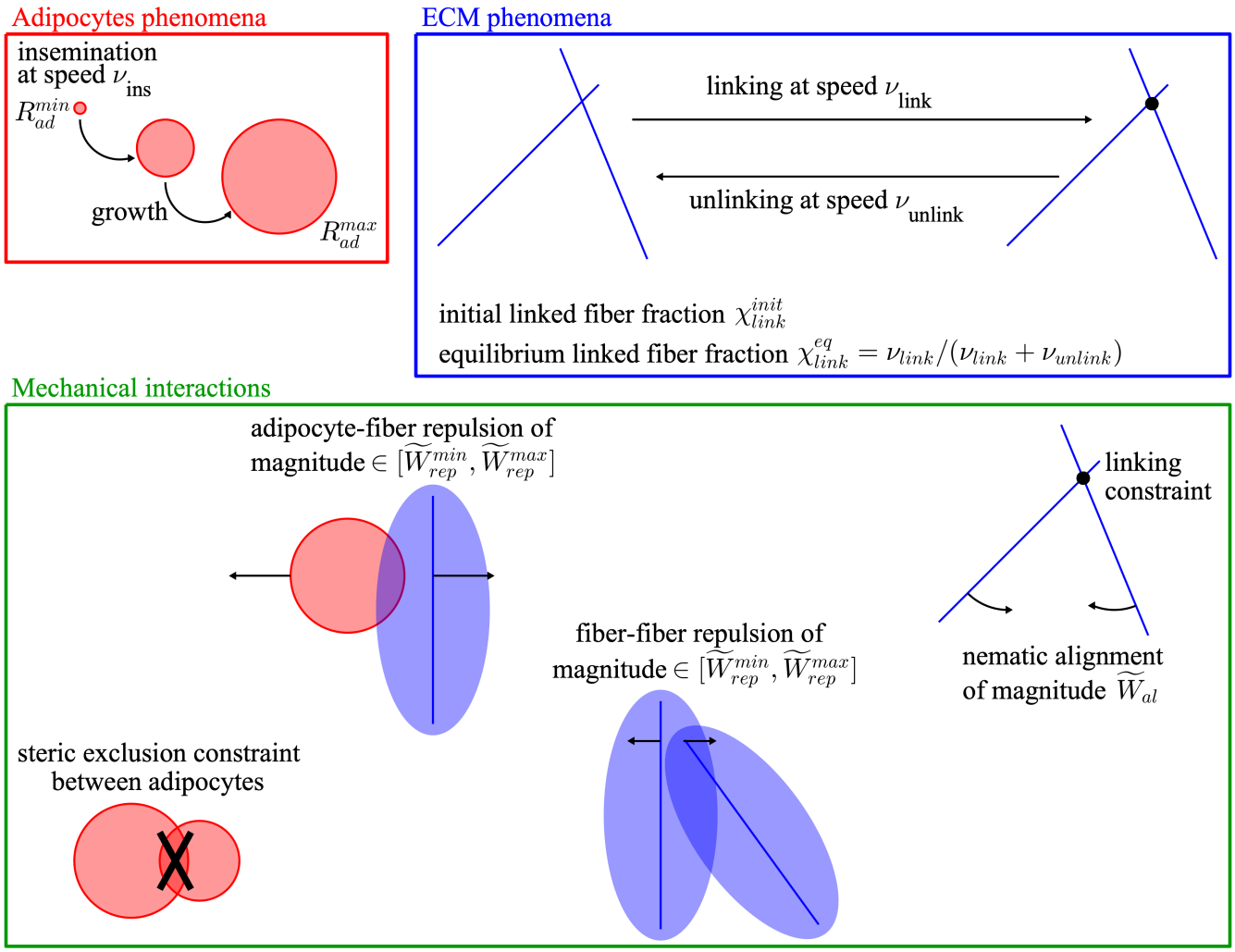


Figure 1.7: Scheme of all the biological processes and mechanical interactions included in the model ATmorpho-2D [17].

4.3.1. Description of the model

Agents

It is assumed that, in white AT, the agents contributing the most to the global mechanical balance are the extra-cellular matrix, which provides mechanical support to the whole tissue, and the adipocytes,

which are bigger and more numerous than all other types of cells. Both these other cells and the vascularisation network are therefore not represented by specific agents, although their biological functions are not neglected.

The ECM is modelled as a network of identical rigid line segments that can spontaneously link together. The linking mechanism account both for actual biological crosslink between ECM fibres and for the complex mechanical properties of said fibres : several "fibre units" linked end-to-end can be seen as a single ECM fibre that can stretch (by the addition of a new unit), shorten (by breaking an existing link) and bend (by having two linked units rotate in relation to each other). Since no limits were imposed to the number of links per fibre unit, the fibres can also branch out and adopt complex geometries. Please note that the agents considered in the model are the fibre units, hereafter refer to merely as fibres for the sake of simplicity. They are characterised by the position of their centre $Y_k(t) \in \Omega$ and by the angle $\theta_k(t) \in]-\pi/2, \pi/2]$ that they make with a reference axis. Their total number is denoted N_{fib}.

Adipocytes being round cells, they are modelled by disks of centre $X_i(t) \in \Omega$ and radius $R_i(t) \in \mathbb{R}^{+*}$, with $1 \leqslant i \leqslant N_{ad}(t)$ and $N_{ad}(t)$ the number of adipocytes in the system.

These data are collected in two sets : $\mathcal{F}(t) = \{(Y_k(t), \theta_k(t)),\ k \in [\![1, N_{fib}]\!]\}$ for the ECM fibres and $\mathcal{A}(t) = \{(X_i(t), R_i(t)),\ i \in [\![1, N_{ad}(t)]\!]\}$ for the adipocytes.

Fibre crosslinking

Crosslink creation is modelled by a random Poisson process of intensity ν_{link} between any pair of intersecting fibres, and crosslink breakage by a Poisson process of intensity ν_{unlink}. The set of all crosslinked fibre pairs (k, m) at time t is denoted $\mathcal{L}(t)$. The ratio $\chi_{link} = \dfrac{\nu_{link}}{\nu_{link} + \nu_{unlink}}$, called the equilibrium linked fibre fraction, corresponds to the proportion of linked fibres among all pairs of intersecting fibres at dynamical equilibrium.

As long as two fibres remain linked, they must keep intersecting at the same point. This set of constraints is written as :

$$\Phi_{link}(\mathcal{F}, \mathcal{L}) \coloneqq \left(Y_k + l^{link}_{k,m}\omega(\theta_k) - Y_m - l^{link}_{m,k}\omega(\theta_m)\right)_{(k,m)\in\mathcal{L}} = \mathbf{0}, \tag{1.1}$$

with $l^{link}_{k,m}$ the position of the crosslink on fibre k, that is the distance between the centre of fibre k and the intersection point with fibre m at the time of the crosslink creation.

Fibre stiffness

Constraint (1.1) does not dictate anything about the angle $\theta_k - \theta_m$ between two linked fibres k and m. In fact, this angle may vary depending on the interactions of the two fibres with the other agents in the system. However, since the collagen fibres this model seeks to replicate are not infinitely flexible but exhibit a certain level of stiffness, the authors assume that two linked fibre will tend to align with each other. This is translated into an alignment potential $W_{al}(\theta_k, \theta_m) = \widetilde{W}_{al} \sin^2(\theta_k - \theta_m)$ which is maximum if the fibres are perpendicular and vanishes if they are aligned.

Steric exclusion between adipocytes

Adipocytes are assumed to be incompressible and non-deformable objects which can not overlap. This means that, for all $i, j \in [\![1, N_{ad}]\!]$ such that $i \neq j$, the distance $\|X_i - X_j\|$ between the centres of adipocytes i and j must be greater than the sum $R_i + R_j$ of their radii. This steric exclusion constraint is expressed by the following condition :

$$\Phi_{steric}(\mathcal{A}) \coloneqq \left(\|X_i - X_j\|^2 - (R_i + R_j)^2\right)_{1\leqslant i<j\leqslant N_{ad}} \geqslant \mathbf{0}. \tag{1.2}$$

Fibre resistance to pressure

From a strict biological point of view, the ECM fibres should not overlap either between themselves or with adipocytes. However, strict steric exclusion constraints like the one defined above would make fibre crosslink as it was defined impossible. Therefore, these constraints are converted into a short-ranged repulsion potential W_{rep} emitted by each fibre. We denote by d_0 the maximal repulsion range (set to $2R_{fib}$ for a fibre-fibre interaction and to $R_{fib} + R_i$ for an interaction between a fibre and an adipocyte i) and by $\widetilde{W}_k$ the maximal intensity of the repulsion potential (which is achieved on the fibre segment and increases linearly with the local fibres alignment). The interaction between a fibre k and an adipocyte i generates a potential $W_{rep}(X_i, Y_k, \theta_k, R_{fib} + R_i, \widetilde{W}_k)$, while the interaction between two **non-linked** fibres k and m generates a total potential $W_{rep}(Y_m, Y_k, \theta_k, 2R_{fib}, \widetilde{W}_k) + W_{rep}(Y_k, Y_m, \theta_m, 2R_{fib}, \widetilde{W}_m)$.

It is assumed that linked fibres do not repel each other, as this would be in direct contradiction with their linking constraint. It may be noted that fibre-fibre repulsion was not included in the basal model described in [17]. The simulations presented in the article body were conducted with fibres both long and large (compared to the adipocytes) : smaller fibres would have been more biologically relevant, but this would have required a greater number of such fibres and thus lead to exceedingly time-consuming simulations. In appendix F of [17], the authors justified their approach by showing that their primary results were similar to those obtained with smaller fibres and a fibre-fibre repulsive interaction. It was observed that in both cases fibres tended to bundle together : without fibre-fibre repulsion, fibres were fully overlapping so the width of a cluster (a.k.a the width of the repulsive zone perceived by the adipocytes) was typically equal to $2R_{fib}$. On the other hand, when fibres were repelling each other the typical cluster width was larger than $2R_{fib}$. Thus, adipocytes would be affected in the same way by clusters of small autorepulsive fibres and clusters of large non-autorepulsive fibres.

Tissue growth through fat storage

The storage of additional fat is modelled by the volumic growth of existing adipocytes and the differentiation of immature cells into new adipocytes. Energy release and the ensuing adipocytes shrinking is not taken into account as the aim is to model AT morphogenesis (through fattening), and not homeostasis nor fasting.

Adipocyte growth is assumed to be essentially linear in volume with an average growth rate $K_{growth} \in \mathbb{R}^{+*}$ and to stop when adipocytes reach a given maximum radius R_{ad}^{max}. Stem cells being not represented individually, their differentiation is modelled by the creation (or "insemination") of a new adipocyte of minimal radius R_{ad}^{min} at a position randomly selected according to an uniform distribution in Ω. Successive insemination times follow a Poisson process of frequency ν_{ins}.

Synthesis and numerical implementation

The total mechanical energy of the system can be written as the sum of all elementary alignment and repulsion potentials :

$$\mathcal{W}(\mathcal{A}, \mathcal{F}, \mathcal{L}) = \sum_{(k,m)\in\mathcal{L}} W_{al}(\theta_k, \theta_m) + \sum_{\substack{1\leqslant i\leqslant N_{ad} \\ 1\leqslant k\leqslant N_{fib}}} W_{rep}(X_i, Y_k, \theta_k, R_{fib} + R_i, \widetilde{W}_k) + \sum_{(k,m)\notin\mathcal{L}} \Big(W_{rep}(Y_m, Y_k, \theta_k, 2R_{fib}, \widetilde{W}_k) + W_{rep}(Y_k, Y_m, \theta_m, 2R_{fib}, \widetilde{W}_m) \Big). \quad (1.3)$$

At each time-step, the system will tend towards its minimum energy state under the constraints (1.1)

and (1.2), that is :

$$(\mathcal{A}(t), \mathcal{F}(t)) = \underset{\widetilde{\mathcal{A}},\, \widetilde{\mathcal{F}}}{\operatorname{argmin}} \Big(\mathcal{W}(\widetilde{\mathcal{A}}, \widetilde{\mathcal{F}}, \mathcal{L}) \;\Big|\; \Phi_{link}(\widetilde{\mathcal{F}}, \mathcal{L}) = \mathbf{0},\ \Phi_{steric}(\widetilde{\mathcal{A}}) \geqslant \mathbf{0} \Big). \tag{1.4}$$

This constrained minimisation problem is solved using a type of Arrow-Hurwicz-Uzawa algorithm [120] whose implementation is detailed in appendix B4 and B5 of [17]. The four "biological" random processes (fibre crosslinking, fibre unlinking, adipocyte growth and adipocyte insemination) are updated between each "mechanical" time-step.

In order to reduce the computation time, the domain Ω is divided into sub-squares or "boxes" whose side length L_{box} scales with the maximum range of interactions. This allows interactions (both soft potentials and strict constraints) to be computed not between every pairs of agents but only between agents located in neighbouring boxes (since agents further apart than L_{box} do not interact). Periodic boundary conditions are implemented by adding on each border of the domain a layer of "ghost" boxes which replicate the content of the "real" boxes located on the opposite side.

4.3.2. Application to adipose tissue morphogenesis and reconstruction

In [17], the authors show that their model can produce three different types of tissue architecture : compact clusters of adipocytes wrapped either in a disordered fibre network (A) or in bundles of aligned fibres defining thick flexible threads (B), and elongated clusters not-quite separated by long, thin, rigid fibre threads (C). Comparison with biological data demonstrates that the various morphologies generated can be fitted to different regions of mice subcutaneous AT. Although (B), illustrated in Figure 1.8, is the most biologically relevant morphology, (A) and (C) can also be observed.

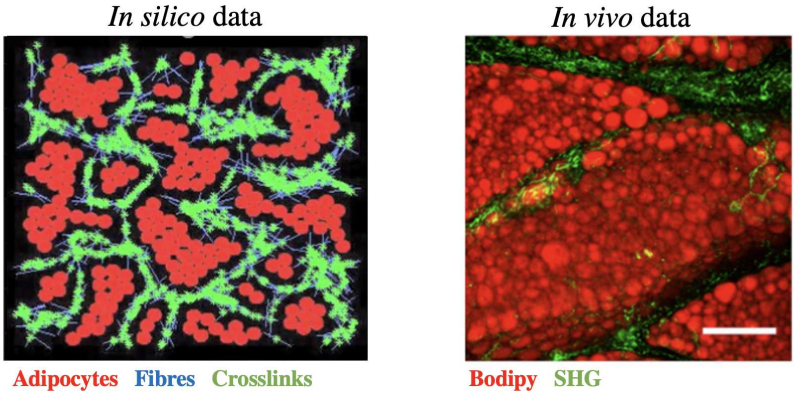


Figure 1.8: Illustration of an adipose tissue morphology of type (B), that is compact clusters of adipocytes wrapped in bundles of aligned fibres defining thick flexible threads. **Left :** *in silico* system generated by the model ATmorpho-2D. The adipocytes are shown in red, the fibres in blue and the crosslinks appear as green crosses. **Right :** Corresponding *in vivo* data, with adipocytes in red (fluorescence with bodipy staining) and collagen fibres in green (collagen second harmonic generation). The white scale bar at the bottom right correspond to 150 µm. **Data from Peurichard and al. [17].**

The authors also show that the morphology obtained in their simulations is mainly determined by the crosslinking dynamics of the ECM via parameters ν_{link} and ν_{unlink} (or equivalently χ_{link} and ν_{unlink}). For instance, morphology (B) is achieved with an equilibrium linked fibre fraction $\chi_{link} = 0.35$ and an unlinking frequency $\nu_{unlink} \in [0.01,\ 0.1]$. These results support the hypothesis that ECM properties (and the induced mechanical forces) are prominent in tissue structuring, a result that agrees well with recent biological findings [121].

Incidentally, these results suggest that the spatial fluctuations observed in the architecture of an adipose tissue depot could be modelled by making the making parameters χ_{link} and ν_{unlink} vary across space.

The authors also took a first step towards modelling the vascularisation network by biasing adipocyte insemination so that the new adipocytes had higher probability to be inseminated near pre-existing ones. This idea was that, since adipocytes need to be supplied by the vascularisation, the location already occupied by adipocytes could be assumed to be vascularised and thus more likely to see the differentiation of new adipocytes. However, the simulations showed that this spatially biased insemination process had no impact on the resulting architecture, leading to the surprising conclusion that accounting for the vascularisation was not necessary in a model aimed at reproducing tissue architecture. This in turn leads to the biological hypothesis that the vascularisation network could be secondary to tissue architecture.

In [37], the authors use a modified version of the model presented above to emulate tissue reconstruction after injury. The system is initialised with a type (B) morphology from which all agents in a central slice have been removed. The presence of a wound is characterised by a high gradient in the local density of both agent types (adipocytes and fibres). This generates inflammatory chemical signals prompting for the insemination of new fibres on the wound boundaries. This goes on until the wound closes (disappearance of the chemical signals). Concomitantly, new adipocytes are inseminated in areas already occupied by a not-too-much crosslinked ECM. The whole process is illustrated in Figure 1.9.

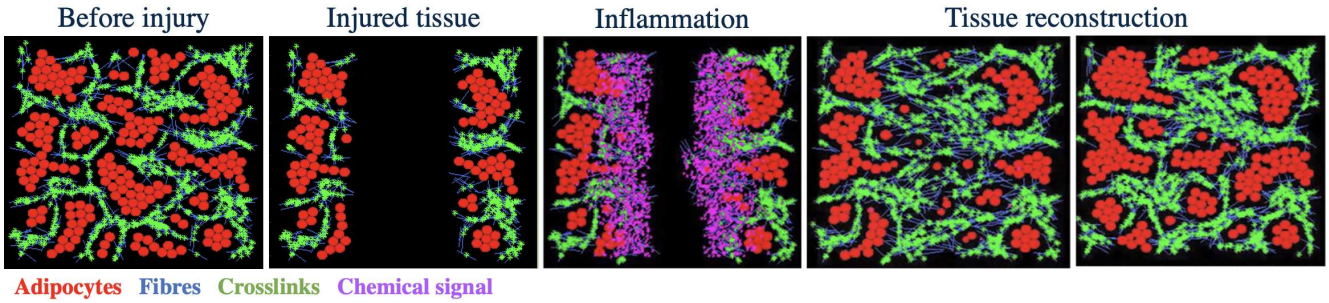


Figure 1.9: Modelling adipose tissue reconstruction after injury using a variant of the model ATmorpho-2D. The adipocytes are shown in red, the fibres in blue, the crosslinks in green and the inflammatory chemical signals in purple. **Data from Peurichard and al. [37].**

The authors show that the reconstructed section can either take the morphology of a scar (fibrosis with a highly crosslinked fibre network containing few adipocytes) or regain its original morphology (regeneration with similar features and number of adipocytes than before injury). Parametric analysis shows that the simulation outcome is mainly decided by the fibre insemination rate and the crosslinking probability during fibre insemination, leading to the hypothesis that the ECM reconstruction properties are the key driver of repair processes. This hypothesis opens up the prospect of a treatment window during which pharmacological action could steer tissue reconstruction towards regeneration.

Overall, these results constitute a convincing proof of concept regarding the possibility to reproduce and explain the emergence of biological tissues architecture with a mathematical model based on local mechanical interactions between a reduced set of components. Because *in vivo* tissues display complex 3D architectures, extending the model to 3D and confronting it to 3D biological data is highly necessary. However the transition to a 3D setting is not trivial : it involves many conceptual and technical challenges detailed in the previous sections of this chapter. This project constitutes the subject of the present thesis.

5. General view and plan of the manuscript

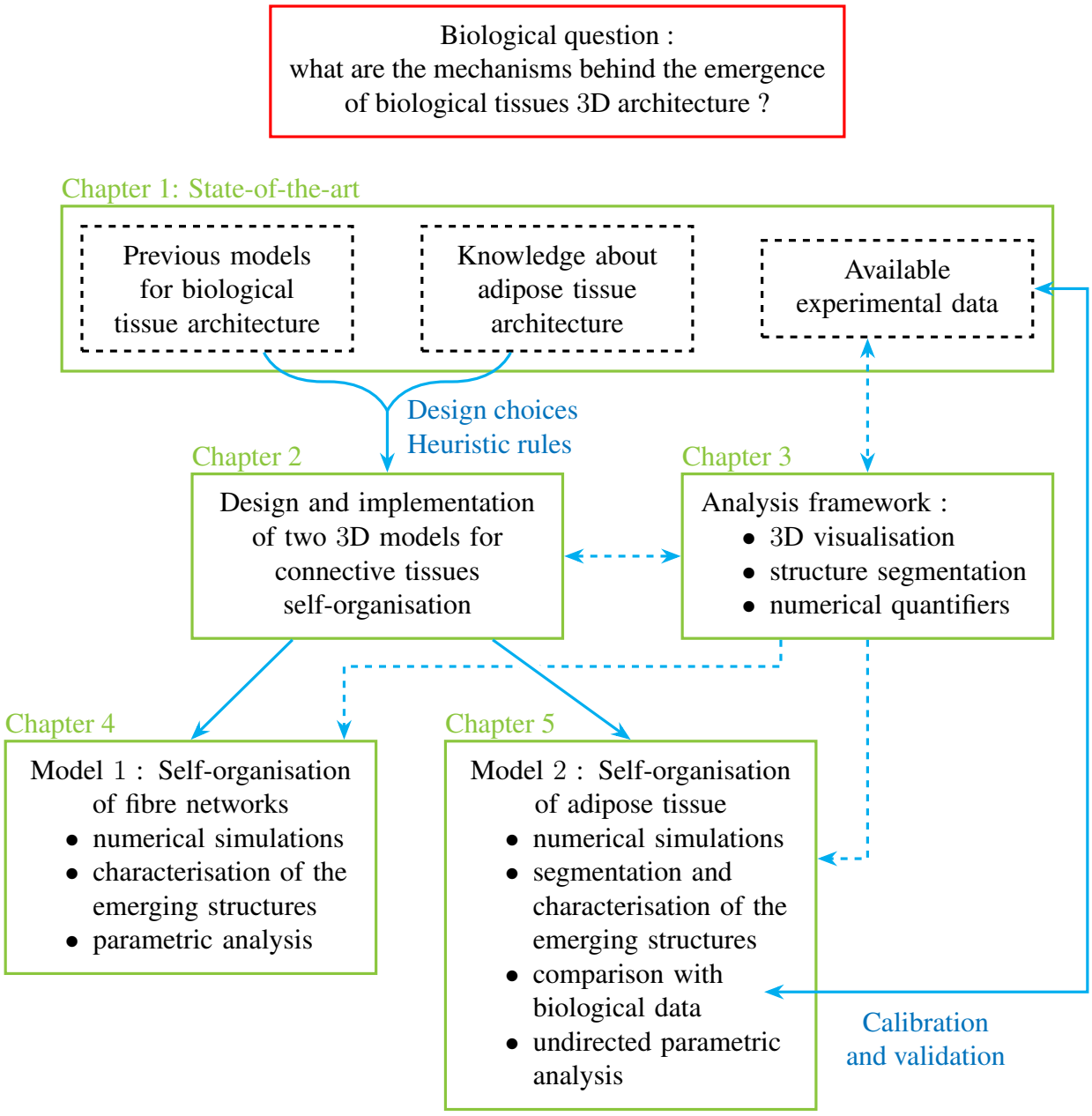


Figure 1.10: General view and plan of the manuscript chapter per chapter.

In this thesis, we study the mechanisms governing the emergence of the three-dimensional architecture of two types of connective tissues : a non-specialised connective tissue, which we reduce to its Extra-Cellular Matrix (ECM) component, and a specialised one, the Adipose Tissue (AT). To this end, we followed the methodology presented in Figure 1.10. This thesis was carried out in close collaboration with biologists and was co-hosted by a biological research unit, the Restore Institute (Toulouse), and a mathematical research unit, the Laboratoire Jacques Louis-Lions (Paris).

We first perform a thorough review of the literature on the various aspects of our subject : the mathematical modelling of biological tissues architecture, the challenges introduced by the transition from 2D to 3D and the current state of biological knowledge about the 3D architecture of connective tissues.

This review constitutes the present general introduction chapter 1. It shows that tissue architecture, and in particular the mechanisms governing its emergence, is largely understudied even though it is known to be a determining factor in tissue functionality. Recent improvements of 3D biological imaging techniques reveal that even the architecture of "simple" connective tissues like AT is much more complex than what biologists had believed until now. This makes us question the reliability of the conclusions achieved through otherwise validated 2D models such as [17] : would a similar 3D model be able to reproduce the intricate 3D architecture of AT ? Moreover, 2D models suffer from intrinsic limitations due to their dimensionality. In the case of [17] for example, it is not clear at all whether fibrous elements that are essentially 1D could envelop and delimitate groups of cells in 3D. Would they be able to organise themselves into planar structures ?

Based on the guidelines established above and on the heuristic rules deduced from biological observations, we design and implement two 3D Individual-Based Model for connective tissue self-organisation adapted respectively to a non-specialised connective tissue and to the Adipose Tissue. A comprehensive description of these models and of their optimised implementation is presented in chapter 2.

In chapter 3, we detail the unified analysis framework that we developed to investigate both the numerical simulations produced with our models and the biological data acquired at the Restore Institute, and compare them together. It contains :

- A common visualisation pipeline for *in silico* and *in vivo* data using the Paraview software. This pipeline was first developed in collaboration with three bachelor students, Charlotte Brunet, Juyeon Kim and Marion Saint-Pée, under the direction of Pr. Sinan Haliyo from the "Multiscale Interactions" team of the Institute of Intelligent Systems and Robotics (ISIR) at Sorbonne University.

- Two similar lobules segmentation protocols, both based on the watershed segmentation method but including different pre-processing steps depending on the type of data analysed (*in vivo* or *in silico*), and two quantifiers to characterise the shape of the lobules thus segmented. The *in vivo* segmentation was developed in collaboration with Laetitia Pierruccioni and Mathieu Vigneau from the CERT team of the Restore Institute at Toulouse III University.

- The description of two quantifiers enabling to characterise the local and global organisation of the fibre networks produced by our models.

In chapter 4, we conduct an in-depth study of the model ECMmorpho-3D by performing numerical simulations. The choice of the parameters tested is guided by the results of the 2D model ATmorpho-2D : we prioritise the parameters that were shown to have a predominant impact on the final architecture and, for each of them, we explore a wide range of values. We use the numerical quantifiers described in the previous chapter to identify and characterise the various architectures emerging in the modelled systems. We distinguish (i) aligned networks with a strong organisation of the fibres around one main direction, (ii) curved networks with a median, locally heterogeneous alignment of the fibres and a wide range of allowed directions living in a plane and (iii) unorganised networks with very low local alignment of the fibres and no preferential direction. A parametric analysis enables us to determine the influence of each parameter and to identify an intermediate emerging variable, namely the connectivity of the fibre network (expressed as the number of crosslinks per fibre), which explains and to some extent predicts the type of architecture produced. This gives us insight on the fundamental principles governing the emergence of biological tissues architecture and suggests a new therapeutic target to control tissue morphogenesis. Finally, we perform a temporal analysis which reveals that the characteristic time-scale of the organisation process is a function of the network remodelling speed and that there exists a unique evolutionary

pathway common to all systems.

In chapter 5, we study the structures produced by the model ATmorpho-3D. We first explore the influence of the cells over the organisation of the fibre structures, taking as reference the fibre-only systems characterised in the previous chapter. We show that the number of cells did not impact significantly the global organisation of the fibre network, though it could increase or decrease the local alignment of the fibres by at most 15%. We then characterise the cellular structures found in the simulations, which we classify into three morphological types : solid and elongated cell clusters possibly with multiple lobes (SE), solid and rounded cell clusters (SR) and branching cell clusters with many digitations (Br). Comparing these structures to those segmented from *in vivo* tissue samples, we show that they display very close morphological characteristics and that their distribution is similar, with many *in vivo* and *in silico* (SE) lobules, fewer (SR) lobules and even less (Br) lobules. Moreover, the proportion in which each morphology appears in a system is controlled mainly by the remodelling characteristic of the fibres, in agreement with the previous finding in 2D [17] and with several biological results demonstrating the impact of the ECM over AT architecture and function. This striking concordance suggests that the emergence of AT architecture is indeed controlled by local mechanical interactions between the adipocytes and the ECM fibres, and more generally that biological tissue architecture could follow much simpler organising principles than what was though until now.

Chapter 2
3D Individual-Based Models for connective tissues architecture emergence

In this chapter, we present two mathematical models that were developed during this thesis to study the emergence of the 3D architecture of two types of connective tissues : a non-specialised one which reduces to its Extra-Cellular Matrix (ECM) component and a specialised one, namely the Adipose Tissue (AT). These models will be thoroughly tested, and their results analysed and compared to biological data, in chapters 4 and 5 respectively.

The first section of the present chapter discusses some general modelling considerations that relate to both models. The next two sections are dedicated to the description of the models, highlighting the biological or numerical relevance of the various choices made during their conception. Finally, the fourth section addresses the details of the numerical implementation of these models, and in particular the optimisation strategies that were employed.

Summary of the chapter

1. Introduction : choice of the type of model

Because of the architecture of any biological tissue is fundamental to its functions, modelling the emergence of a tissue architecture will greatly improve our understanding of this tissue biology and plasticity in physiological or pathological conditions. In this thesis, we want to test the hypothesis that the architecture of a tissue could emerge spontaneously from simple local mechanical interactions between its components.

Numerous models for biological tissue can be found in the literature. Due to their simplicity and flexibility, the most widely used models are Individual-Based Models (IBMs), which describe the behaviour of each element (e.g. cell, fibre, protein) and its interactions with the surrounding elements over time [8, 9]. However, IBMs have a high numerical cost (in term of both memory space and computation time) which can become intractable when studying large scales, either spatial or temporal or for systems composed of too many elements. In such cases, continuous or mean-field kinetic models [21] may be preferred because they are less costly, at the expense of a loss of resolution at the microscopic level.

However, in our case this level of description can not be ignored for at least two reasons. The first is that our hypothesis focuses on local interactions. The second is that the macroscopic properties of the ECM, which are a key factor in the emergence of tissues architecture, have been shown to be modulated by microstructure configurations [122].

To ensure the relevance of our study, we must therefore face the computational cost of an IBM, while paying close attention to optimisation opportunities during the design and implementation stages (see section 4).

In view of the major impact of the ECM, we decided to first develop a model to study the self-organisation of this three-dimensional macrostructure inside a non-specialised connective tissue, so as to gain insights on the general principles driving the emergence of the various types of tissue architecture. In addition, such a model can serve as the basis for constructing more specific models, and indeed we use it to develop a second model aimed at mimicking the emergence of AT architecture. This tissue has major biomedical importance and is also a generalist object of study, with a relatively simple structure observed in other organs. Therefore, though we developed them with specific target tissues in mind, the two models presented in this chapter are in fact quite generic and can be adapted to other types of biological tissues.

Many models of fibre network account for the complex mechanical properties of the fibres by representing them as strings of beads connected either by elastic springs [16, 25, 27, 34, 35] or Euler-Bernoulli beams [13–15, 32, 123, 124]. Following a similar logic, Peurichard and al. [17, 37] discretised the fibres into small rigid segments linked at their intersection. In this approach, the mechanical properties did not stem from the segments themselves but from the fact that the links between them were dynamical and could spontaneously appear or break. Note that, among existing models, few of them feature dynamical crosslinking of ECM components. In [13, 19, 20], various models of 3D fibrous networks composed of permanent or transient crosslinks are proposed. However, these models feature ECM remodelling in reaction to external factors (applied load [13], migrating cells [19], contractile cells [20]), and the literature so far provides little cues on the mechanisms underlying fibre self-organisation. Considering that this is the central issue of our work, we choose to follow [17,37] approach and account for spontaneous, random remodelling of the ECM network.

The vascular system, which like the ECM is a common characteristic of all biological tissues, is vital to create and maintain tissue functions. Its complex layout emerges together with the global tissue architecture and is consubstantial to it. A lot of efforts have been devoted to study whether it is the pattern of the vascularisation network that drives the emergence of tissue architecture, or the tissue architecture that guides the development of the vascularisation network. This question remains open and could well be

akin to the chicken-and-egg dilemma. However, [17] showed that the 2D architecture of the subcutaneous white adipose tissue could be mimicked *in silico* without accounting for the vascularisation network. Based on their results, we choose to treat the development of the vascularisation network as secondary to the emergence of tissue architecture and did not represent it explicitly in our models.

Connective tissues host various types of stromal cells. Considering the small size of these cells, we choose to neglect their mechanical impact in a first approach and, like for the vasculature, we did not model them by individual agents (though we did account indirectly for some of their biological effects).

2. Modelling the Extra-Cellular Matrix

2.1. Model introduction and summary : discretisation of the fibre network

As described in chapter 1 (section 4.1), the ECM is a dynamical three-dimensional network of interacting fibres that provide a mechanical and biochemical support to surrounding cells as well as to the whole tissue, and is considered to be the key factor determining a tissue architecture.

In this thesis, we want to test the hypothesis that this macrostructure could spontaneously emerge from simple local mechanical interactions between dynamically interconnected fibres. To this aim, we construct an IBM model which accounts mainly for the local mechanical constraints and in which the complex chemical and biological processes are represented as simply as possible. We give here a brief description of this model, subsequently referred to as ECMmorpho-3D, before going into details in the following sections.

We choose to model the ECM fibres by discretising them into a large number N_{fib} of fibre units, consisting of non-stretching and non-flexible spherocylinders of uniform fixed length L_{fib} and radius R_{fib}. These fibre units have the ability to spontaneously link and unlink from their neighbours closer than a catching distance d_{link}^{max}, according to Poisson processes of respective frequencies ν_{link} and ν_{unlink}. As a result, the linked fibre ratio $\chi_{link} = \frac{\nu_{link}}{\nu_{link}+\nu_{unlink}}$ represents the equilibrium fraction of linked fibres among the pairs of neighbouring fibres. This dynamical crosslinking mechanism allows us to model both the overall temporal plasticity of the network and the complex physical properties of biological fibres such as elongation, bending, branching and growth, thus compensating our minimalistic description of the fibre units. Crosslinks are modelled as linear spring with constant stiffness κ_{rest} and equilibrium length d_{link}^{eq}, controlling the extension of long fibres made of a string of successively linked fibre units. They also have a constant rotational stiffness (or flexural modulus) α_{align} which controls the resistance to bending of long fibres. In addition, we suppose that overlapping fibre units repel each other, modelling volume-exclusion effects. This is achieved via a repulsive force based on Hertzian theory, whose maximum intensity E_{rep} controls the amount of overlap tolerated between fibres. Finally, as the Reynolds number in most biological tissues is very small, we suppose that inertial forces can be neglected and we consider an over-damped regime for fibre motion and rotation, associated with a dynamic viscosity of the fibres μ_{fib}.

All these features are summarised in Figure 2.1 and the corresponding parameters listed in Table 2.1. We now turn to the details of each of the model components.

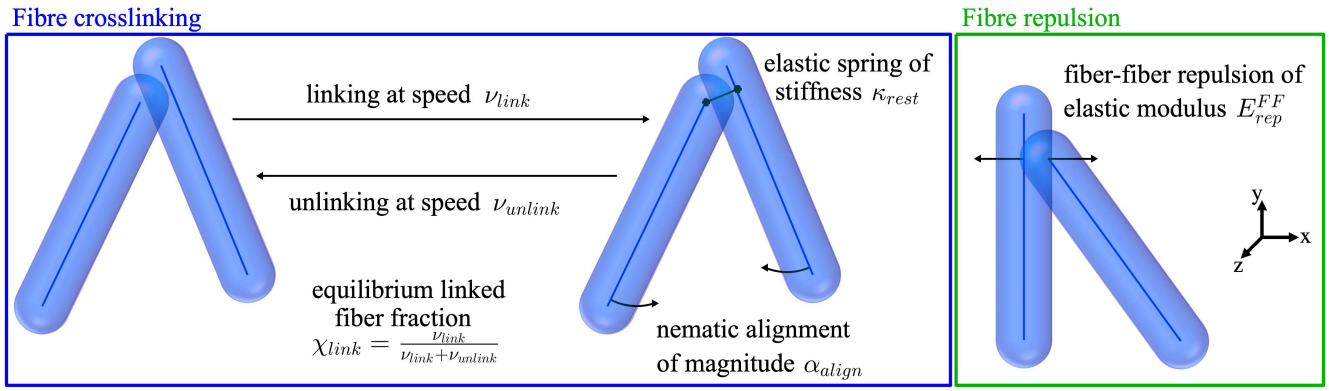


Figure 2.1: Scheme of all the biological processes and mechanical interactions included in the model ECMmorpho-3D.

Parameter	Domain	Dimension	Description
Fibre characteristics			
N_{fib}	$\mathbb{N}$	N/A	Number of fibres in the system.
L_{fib}	$\mathbb{R}^{+*}$	L	Fibre length.
R_{fib}	$\mathbb{R}^{+*}$	L	Fibre radius.
Crosslink processes			
ν_{link}	$\mathbb{R}^{+}$	T^{-1}	Fibre linking frequency.
ν_{unlink}	$\mathbb{R}^{+}$	T^{-1}	Fibre unlinking frequency.
d^{eq}_{link}	$[2R_{fib}, +\infty[$	L	Link equilibrium length.
d^{max}_{link}	$[d^{eq}_{link}, +\infty[$	L	Link maximum catching distance.
Mechanical interactions			
E_{rep}	$\mathbb{R}^{+}$	$M.L^{-1}.T^{-2}$	Elastic modulus of the fibre-fibre contact.
κ_{rest}	$\mathbb{R}^{+}$	$M.T^{-2}$	Link stiffness.
α_{align}	$\mathbb{R}^{+}$	$M.L^{2}.T^{-2}$	Maximum intensity of the alignment torque between linked fibres.
μ_{fib}	$\mathbb{R}^{+}$	$M.L^{-1}.T^{-1}$	Dynamic viscosity of fibres.

Table 2.1: Summary of all the parameters of the model ECMmorpho-3D.

2.2. Biological phenomena : fibre linking and unlinking

As previously mentioned, we discretise the ECM fibres into N_{fib} non-stretching and non-flexible spherocylindrical units of uniform fixed length L_{fib} and radius R_{fib} (see Figure 2.2 for a schematic representation). These fibre units are represented by the position of their centre $\boldsymbol{Y}_k(t) \in \Omega$ (with $\Omega \subset \mathbb{R}^3$ the finite spatial domain inside which the modelled tissue is developing) and by their unitary, non-oriented directional vector $\boldsymbol{\omega}_k(t) \in \mathbb{S}^+_2$ (with $\mathbb{S}^+_2$ the unit half-sphere). Figure 2.2 gives a schematic representation of one fibre unit and summarises the associated mathematical notations. Note that, for the sake of simplicity, we will tend to refer to fibre units merely as fibres.

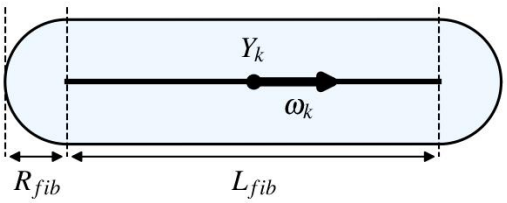


Figure 2.2: Schematic representation of a fibre unit, modelled by a spherocylinder of centre $\boldsymbol{Y}_k \in \Omega$ and directional vector $\boldsymbol{\omega}_k \in \mathbb{S}_2^+$, with fixed length $L_{fib} \in \mathbb{R}^{+*}$ and radius $R_{fib} \in \mathbb{R}^{+*}$.

The various types of fibres that make up the ECM are flexible materials with the ability to stretch and bend. They can grow and branch through the secretion of new baseline molecules by the surrounding cells, and also shrink due to degradation by specialised enzymes [67]. Moreover, ECM fibres are interconnected by molecular bonds that confer connectivity and elasticity throughout the whole network [68].

We modelled all these properties by means of a single crosslinking mechanism which randomly creates and destroys elastic springs between close-by fibres. Several consecutively crosslinked fibre units would model a long fibre with the ability to bend and/or stretch at the connection points. Spontaneous linking of an additional unit at the end of a string models fibre elongation [125], while spontaneous linking between two neighbouring strings models the natural biological crosslinking mechanism due to glycoproteins. On the other hand, spontaneous unlinking of a pair of crosslinked units allows for fibre breakage and, in a roundabout way, for fibre turnover (modelling at the same time degradation of linked fibres by metalloproteinases and secretion of new, still unlinked fibre material by fibroblast cells).

In practice, we decided that two fibre units whose central segments are closer than a catching distance d_{link}^{max} could develop a link according to a Poisson process of frequency ν_{link}. This link could then broke at anytime according to a Poisson process of frequency ν_{unlink}. As long as the link exists, the two fibres can not link together a second time (but they can link with other fibres). We denote by $p_{k,m}(t) \in \{0,1\}$ the linking state of fibres k and m at time t (equal to 1 if they are linked and to 0 otherwise) and by $d_{k,m}$ the distance between them (see later for details of its computation). With these notations, the probability of a link being created in the time-interval $[t, t+\Delta t]$ between two fibres which, at time t, are within catching distance and not already linked, is equal to :

$$\mathbb{P}\Big(p_{k,m}(t+\Delta t) = 1 \;\Big|\; p_{k,m}(t) = 0 \text{ and } d_{k,m}(t) \leqslant d_{link}^{max}\Big) = 1 - e^{-\nu_{link}\Delta t}. \tag{2.1}$$

Similarly, the probability that a link existing at time t breaks during the time-interval $[t, t+\Delta t]$ is equal to :

$$\mathbb{P}\Big(p_{k,m}(t+\Delta t) = 0 \;\Big|\; p_{k,m}(t) = 1\Big) = 1 - e^{-\nu_{unlink}\Delta t}. \tag{2.2}$$

The difference between d_{link}^{max} and the distance $2R_{fib}$ under which fibres overlap can be seen as a measure of the "reach" of the crosslinking molecules. We did not limit the number of crosslinks attached to the same unit so as to account for fibre branching. However, we limited the number of crosslinks between two given units to one, so as to account for the geometric and steric constraints on the attachment of the crosslinking molecules. In our model, two long fibres (made a several consecutively crosslinked fibre units) running parallel can link together at numerous points but not twice along the same pair of fibre units, meaning that the average distance between two successive crosslinks will be equal to the length L_{fib} of one unit.

Computation of two fibres closest points

We describe here the actual computation of the distance $d_{k,m}$ between two fibres k and m and of the

corresponding closest points $\boldsymbol{Y}_{k,m}$ and $\boldsymbol{Y}_{m,k}$ between these fibres. Readers not interested by technical details can go directly to the next section.

We define $d_{k,m}$ as the minimal distance between the two fibres k and m central segments. Parametrising the central segment of a fibre k by $F_k = \left\{\boldsymbol{Y}_k + u\boldsymbol{\omega}_k \;\middle|\; u \in [-L_{fib}/2, L_{fib}/2]\right\}$, we obtain :

$$d_{k,m} = \min_{u,v\in\left[-\frac{L_{fib}}{2},\frac{L_{fib}}{2}\right]} \|\boldsymbol{Y}_k + u\boldsymbol{\omega}_k - (\boldsymbol{Y}_m + v\boldsymbol{\omega}_m)\|. \tag{2.3}$$

We denote by $(l_{k,m}, l_{m,k})$ the argument pair minimising this equation, so that $\boldsymbol{Y}_{k,m} = \boldsymbol{Y}_k + l_{k,m}\boldsymbol{\omega}_k$ is the point of the central segment of fibre k that is closest to the central segment of fibre m and $\boldsymbol{Y}_{m,k} = \boldsymbol{Y}_m + l_{m,k}\boldsymbol{\omega}_m$ its counterpart on fibre m. Note that these two points may not be uniquely defined if the two fibres have parallel axes ($\boldsymbol{\omega}_k = \pm\boldsymbol{\omega}_m$). In that case, we arbitrarily chose the solution with the smallest $|l_{k,m}|$ value, i.e. where $\boldsymbol{Y}_{k,m}$ is closest to the fibre centre $\boldsymbol{Y}_k$.

The analytical expression of $(l_{k,m}, l_{m,k})$ depends on a number of conditions. We list below the various possible configurations and the corresponding expressions for $(l_{k,m}, l_{m,k})$. An illustration of each case is given in Figure 2.3.

1. General case $\boldsymbol{\omega}_k \neq \pm\boldsymbol{\omega}_m$: problem (2.3) admits a unique minimiser.

   We denote by (u^*, v^*) the positions (with respect to points $\boldsymbol{Y}_k$ and $\boldsymbol{Y}_m$ respectively) of the closest points of the two infinite lines passing through the fibre axes. Their analytical expression is :

$$\begin{cases} u^* = \dfrac{\boldsymbol{\omega}_k \cdot (\boldsymbol{Y}_m - \boldsymbol{Y}_k) - (\boldsymbol{\omega}_k \cdot \boldsymbol{\omega}_m)(\boldsymbol{\omega}_m \cdot (\boldsymbol{Y}_m - \boldsymbol{Y}_k))}{1 - (\boldsymbol{\omega}_k \cdot \boldsymbol{\omega}_m)^2}, \\ v^* = \dfrac{(\boldsymbol{\omega}_k \cdot \boldsymbol{\omega}_m)(\boldsymbol{\omega}_k \cdot (\boldsymbol{Y}_m - \boldsymbol{Y}_k)) - \boldsymbol{\omega}_m \cdot (\boldsymbol{Y}_m - \boldsymbol{Y}_k)}{1 - (\boldsymbol{\omega}_k \cdot \boldsymbol{\omega}_m)^2}. \end{cases} \tag{2.4}$$

   (a) If $(u^*, v^*) \in \left[-\dfrac{L_{fib}}{2}, \dfrac{L_{fib}}{2}\right]^2$ then the (infinite) lines optimum coincide with the (finite) segments optimum (see Figure 2.3(1a)). In other words, we have :

$$\begin{cases} l_{k,m} = u^* \in \left[-\dfrac{L_{fib}}{2}, \dfrac{L_{fib}}{2}\right] \\ l_{m,k} = v^* \in \left[-\dfrac{L_{fib}}{2}, \dfrac{L_{fib}}{2}\right] \\ (\boldsymbol{Y}_{k,m} - \boldsymbol{Y}_{m,k}) \perp \boldsymbol{\omega}_k, \boldsymbol{\omega}_m \end{cases}$$

   (b) One of the optimal points is located at the extremity of its fibre (see Figure 2.3(1b)). In that case we have either

$$\begin{cases} l_{k,m} = \operatorname{sign}(u^*)\dfrac{L_{fib}}{2} \\ l_{m,k} = (\boldsymbol{Y}_k + l_{k,m}\boldsymbol{\omega}_k - \boldsymbol{Y}_m) \cdot \boldsymbol{\omega}_m \in \left[-\dfrac{L_{fib}}{2}, \dfrac{L_{fib}}{2}\right] \\ (\boldsymbol{Y}_{k,m} - \boldsymbol{Y}_{m,k}) \perp \boldsymbol{\omega}_m \end{cases}$$

   or

$$\begin{cases} l_{m,k} = \operatorname{sign}(v^*)L_{fib}/2 \\ l_{k,m} = (\boldsymbol{Y}_m + l_{m,k}\boldsymbol{\omega}_m - \boldsymbol{Y}_k) \cdot \boldsymbol{\omega}_k \in \left[-\dfrac{L_{fib}}{2}, \dfrac{L_{fib}}{2}\right] \\ (\boldsymbol{Y}_{k,m} - \boldsymbol{Y}_{m,k}) \perp \boldsymbol{\omega}_k \end{cases}$$

(c) Both optimal points are located at the extremity of their fibres (see Figure 2.3(1c)). In that case we have $l_{k,m} = \text{sign}((\boldsymbol{Y}_m - \boldsymbol{Y}_k) \cdot \boldsymbol{\omega}_k)\frac{L_{fib}}{2}$ and $l_{m,k} = \text{sign}((\boldsymbol{Y}_k - \boldsymbol{Y}_m) \cdot \boldsymbol{\omega}_m)\frac{L_{fib}}{2}$.

2. Degenerate case $\boldsymbol{\omega}_k = \pm\boldsymbol{\omega}_m$: problem (2.3) may admit an infinite number of minimisers, among which we choose the one with the smallest $|l_{m,k}|$ value.

   (a) If $\left|(\boldsymbol{Y}_k - \boldsymbol{Y}_m) \cdot \boldsymbol{\omega}_m\right| \leqslant \frac{L_{fib}}{2}$ (see Figure 2.3(2a)), then according to our definition $l_{k,m} = 0$, $l_{m,k} = (\boldsymbol{Y}_k - \boldsymbol{Y}_m) \cdot \boldsymbol{\omega}_m$ and $(\boldsymbol{Y}_{k,m} - \boldsymbol{Y}_{m,k}) \perp \boldsymbol{\omega}_k, \boldsymbol{\omega}_m$.

   (b) If $\frac{L_{fib}}{2} < \left|(\boldsymbol{Y}_k - \boldsymbol{Y}_m) \cdot \boldsymbol{\omega}_m\right| \leqslant L_{fib}$ (see Figure 2.3(2b)), then according to our definition $l_{k,m} = (\boldsymbol{Y}_m - \boldsymbol{Y}_k) \cdot \boldsymbol{\omega}_k \pm \frac{L_{fib}}{2}$, $l_{m,k} = \text{sign}((\boldsymbol{Y}_k - \boldsymbol{Y}_m) \cdot \boldsymbol{\omega}_m)\frac{L_{fib}}{2}$ and $(\boldsymbol{Y}_{k,m} - \boldsymbol{Y}_{m,k}) \perp \boldsymbol{\omega}_k, \boldsymbol{\omega}_m$.

   (c) If $L_{fib} < \left|(\boldsymbol{Y}_k - \boldsymbol{Y}_m) \cdot \boldsymbol{\omega}_m\right|$ (see Figure 2.3(2c)) then problem (2.3) admits a unique argument of the minimum, whose analytical expression is $l_{k,m} = \text{sign}((\boldsymbol{Y}_m - \boldsymbol{Y}_k) \cdot \boldsymbol{\omega}_k)\frac{L_{fib}}{2}$ and $l_{m,k} = \text{sign}((\boldsymbol{Y}_k - \boldsymbol{Y}_m) \cdot \boldsymbol{\omega}_m)\frac{L_{fib}}{2}$.

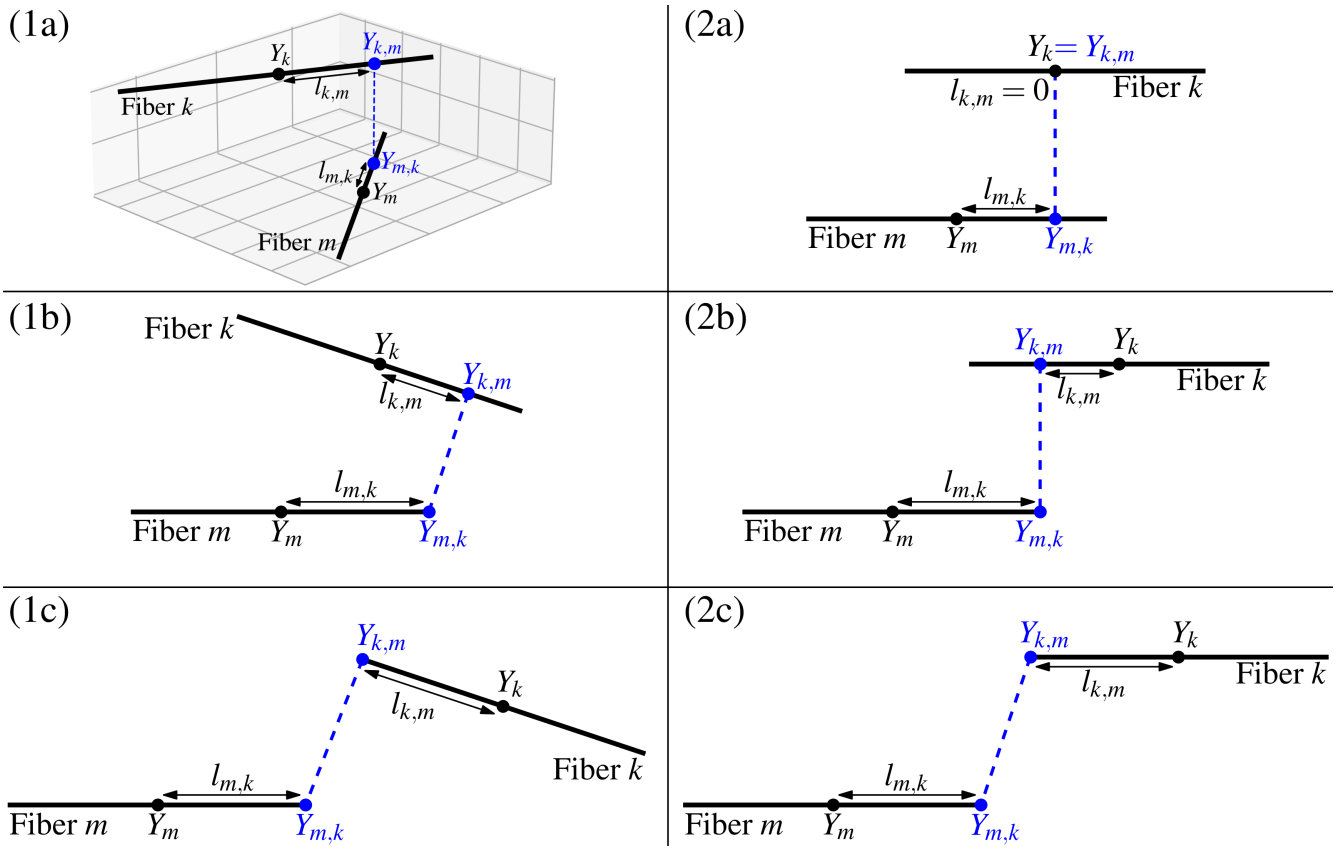


Figure 2.3: Illustration of all possible configurations for the closest points $\boldsymbol{Y}_{k,m}$ and $\boldsymbol{Y}_{m,k}$ of two finite segments of respective centre $\boldsymbol{Y}_k$ and $\boldsymbol{Y}_m$. We denote by $l_{k,m}$ (resp. $l_{m,k}$) the signed distance between $\boldsymbol{Y}_{k,m}$ and $\boldsymbol{Y}_k$ (resp. $\boldsymbol{Y}_{m,k}$ and $\boldsymbol{Y}_m$)

2.3. Mechanical interactions

In the following paragraphs, we will describe all the mechanical interactions between the fibres. These interactions generate both translational forces (that alter the position of the fibres) and rotational torques (that alter the direction of the fibres).

Fibres resistance to pressure and steric repulsion

From a physical point of view, the ECM fibres represented in our model are solid, possibly deformable objects : each of them occupies a certain volume from which other objects are excluded. This steric exclusion can have various effects upon contact between pairs of fibres (e.g. bouncing, deformation, contact stress, etc). We choose to model it using Hertz's theory for non-adhesive elastic contact between curved surfaces [126, 127].

According to this theory, two deformable solids coming into contact will deform and flatten against each other, creating a contact stress (or repulsive force) proportional to their contact radius r_c (that is the characteristic size of their contact area) and to their indentation depth δ (that is the depth of the overlap that would have happen in the absence of deformation). However, including agent deformation in our model would make it overly complicated as it would require to keep track of a complex 3D shape for each agent (instead of the quite simple mathematical representation given above). Therefore, we took a simplified approach where we do take into account the repulsive force generated by the contact, but do not modify the shape of the agents and instead allow them to overlap. Note that, because we do not register agent deformation upon contact, these deformations do not feed back on each other : a contact between two fibres will always be a contact between two spherocylinders, even if these fibres are already in contact with (and should be deformed by) many other agents.

A contact between two spherocylinders is not a standard case and is not described in the literature. We can divide it into three sub-cases :

- contact between the half-sphere extremities of the two spherocylinders;
- contact between the extremity of one spherocylinder and the cylindrical body of the other;
- contact between the cylindrical bodies of the two spherocylinders;

The first two sub-cases can be solved explicitly. The third one have an implicit solution which depends on the angle between the two fibres axes. Either way, except in the very marginal case of a contact between the cylindrical bodies of two perfectly parallel fibres (which is unlikely even with the tendency towards alignment of linked fibres, see below), the contact radius will be linearly proportional to $\sqrt{R_{fib}\delta}$. For the sake of simplicity, we will ignore the multiplicative factor and take the simplest possible expression, that is $r_c = \sqrt{R_{fib}\delta}$.

As above, we denote by $\boldsymbol{Y}_{k,m}$ and $\boldsymbol{Y}_{m,k}$ the closest points of the central segments of the two fibres k and m. If the two fibres touch each other then their indentation depth is $\delta = 2R_{fib} - \|\boldsymbol{Y}_{k,m} - \boldsymbol{Y}_{m,k}\| \geqslant 0$ and, according to the Hertzian theory, fibre m exerts on its neighbour k a repulsive force equal to :

$$\boldsymbol{F}^{rep}_{k,m} = \begin{cases} \frac{4}{3} E_{rep}\sqrt{R_{fib}}(2R_{fib} - \|\boldsymbol{Y}_{k,m} - \boldsymbol{Y}_{m,k}\|)^{3/2} \dfrac{\boldsymbol{Y}_{k,m} - \boldsymbol{Y}_{m,k}}{\|\boldsymbol{Y}_{k,m} - \boldsymbol{Y}_{m,k}\|} \text{ if } \|\boldsymbol{Y}_{k,m} - \boldsymbol{Y}_{m,k}\| \leqslant 2R_{fib}, \\ 0 \text{ otherwise,} \end{cases} \tag{2.5}$$

with E_{rep} the effective elastic modulus, whose dimension is a mass per length per time squared ($M.L^{-1}.T^{-2}$). Note that this physical factor depends on the intrinsic elastic properties of the agents involved in the interaction, properties which we assumed to be the same for all fibre units.

This force is applied at point $\boldsymbol{Y}_{k,m}$ and so induces on fibre k a rotational torque with respect to its centre-of-mass :

$$\boldsymbol{T}^{rep}_{k,m} = l_{k,m}\boldsymbol{\omega}_k \wedge \boldsymbol{F}^{rep}_{k,m}. \tag{2.6}$$

Crosslink elasticity

Following one of the two standard approaches to modelling fibre crosslinking [13, 16, 27, 34, 35], we choose to represent the link between two fibres by an elastic spring of stiffness κ_{rest} and equilibrium length d^{eq}_{link}, tied to the points of both fibre central segments that were closest at the time of the link creation. Note that the other classic approach is to model crosslinks by hinges [11, 12, 14, 20, 28, 32, 33, 40, 123, 124], either stiff or freely rotating.

Given two fibres k and m which develop a crosslink at time t_{link}, we denote by $(l^{link}_{k,m}, l^{link}_{m,k})$ the position (relative to their respective centre) of these two fibres closest points at time t_{link}. We assumed that the link does not slide along the fibres axes, or in other words that it remains tied to the same attachment sites $\boldsymbol{Y}^{link}_{k,m} = \boldsymbol{Y}_k(t) + l^{link}_{k,m}\boldsymbol{\omega}_k(t)$ and $\boldsymbol{Y}^{link}_{m,k} = \boldsymbol{Y}_m(t) + l^{link}_{m,k}\boldsymbol{\omega}_m(t)$. Its length under load is thus equal to $\left\|\boldsymbol{Y}^{link}_{k,m} - \boldsymbol{Y}^{link}_{m,k}\right\|$ and, according to Hooke's law, it exerts on fibre k an elastic restoring force of the form :

$$\boldsymbol{F}^{link}_{k,m} = \kappa_{rest}\big(d^{eq}_{link} - \left\|\boldsymbol{Y}^{link}_{k,m} - \boldsymbol{Y}^{link}_{m,k}\right\|\big)\frac{\boldsymbol{Y}^{link}_{k,m} - \boldsymbol{Y}^{link}_{m,k}}{\left\|\boldsymbol{Y}^{link}_{k,m} - \boldsymbol{Y}^{link}_{m,k}\right\|}. \tag{2.7}$$

This force is applied at point $\boldsymbol{Y}^{link}_{k,m}$ and so induces on fibre k a rotational torque :

$$\boldsymbol{T}^{link}_{k,m} = l^{link}_{k,m}\boldsymbol{\omega}_k \wedge \boldsymbol{F}^{link}_{k,m}. \tag{2.8}$$

Note that, to ensure coherence between the various elements of our model (namely steric repulsion, link elasticity and link creation), we prescribed that $2R_{fib} \leqslant d^{eq}_{link} \leqslant d^{max}_{link}$.

Fibre stiffness

The ECM fibres we seek to simulate are not infinitely flexible but exhibit a certain level of stiffness. We modelled this by applying to linked fibres a rotational torque that will tend to make them adopt and maintain the same direction. Considering that our fibres are non-oriented, we want this alignment process to be nematic, i.e. we want fibres to close their smallest relative angle (see Figure 2.1). Given two linked fibres k and m, we want $\boldsymbol{\omega}_k$ to rotate towards $\tilde{\boldsymbol{\omega}}_m = \boldsymbol{\omega}_m$ if $\arccos(\boldsymbol{\omega}_k \cdot \boldsymbol{\omega}_m) \leqslant \pi/2$ and towards $\tilde{\boldsymbol{\omega}}_m = -\boldsymbol{\omega}_m$ otherwise. Total alignment would be achieved by rotating the fibre k by an angle $\arcsin(\|\boldsymbol{\omega}_k \wedge \tilde{\boldsymbol{\omega}}_m\|)$ around the axis $\boldsymbol{\omega}_k \wedge \tilde{\boldsymbol{\omega}}_m$. Using Rodrigues's rotation formula [128], the corresponding rotation matrix is :

$$R_{k,m} = I + [\boldsymbol{\omega}_k \wedge \tilde{\boldsymbol{\omega}}_m]_\wedge + \frac{1 - \boldsymbol{\omega}_k \cdot \tilde{\boldsymbol{\omega}}_m}{\|\boldsymbol{\omega}_k \wedge \tilde{\boldsymbol{\omega}}_m\|^2}[\boldsymbol{\omega}_k \wedge \tilde{\boldsymbol{\omega}}_m]^2_\wedge, \tag{2.9}$$

where I is the 3×3 identity matrix and $[u]_\wedge = \begin{pmatrix} 0 & -u_z & u_y \\ u_z & 0 & -u_x \\ -u_y & u_x & 0 \end{pmatrix}$ the matrix form of the cross product by the vector u (i.e. $u \wedge v = [u]_\wedge \cdot v$ for all $v \in \mathbb{R}^3$).

If this rotation were to occur instantly, the variation of the vector $\boldsymbol{\omega}_k$ would be equal to $\Delta\boldsymbol{\omega}_k = \tilde{\boldsymbol{\omega}}_m - \boldsymbol{\omega}_k = R_{k,m} \cdot \boldsymbol{\omega}_k - \boldsymbol{\omega}_k = (R_{k,m} - I) \cdot \boldsymbol{\omega}_k$. However, such instantaneous alignment would amount to the combination of two linked fibre units displaying an infinite stiffness, thereby defeating one of our main purposes in introducing fibre crosslinking into our model (namely to account for fibre bending).

Thus, denoting α_{align} the maximum angular momentum per unit of time produced by our alignment torque, the alignment torque $\boldsymbol{T}_{k,m}^{align}$ sustained by fibre k due to the flexural stiffness of its link with fibre m is such that :

$$\boldsymbol{T}_{k,m}^{align} \wedge u = \alpha_{align}\left([\boldsymbol{\omega}_k \wedge \tilde{\boldsymbol{\omega}}_m]_\wedge + \frac{1 - \boldsymbol{\omega}_k \cdot \tilde{\boldsymbol{\omega}}_m}{\|\boldsymbol{\omega}_k \wedge \tilde{\boldsymbol{\omega}}_m\|^2} [\boldsymbol{\omega}_k \wedge \tilde{\boldsymbol{\omega}}_m]_\wedge^2 \right) \cdot u \quad \forall u \in \mathbb{R}^3, \tag{2.10}$$

where α_{align} is the maximum angular momentum per unit of time produced by the alignment interaction, whose dimension is $M.L^2.T^{-2}$.

Over-damped regime

From a biological point of view, the fibres we model move in a thick, viscous medium made of the interstitial fluid and of all the small cells that we did not represent by individual agents. Friction is therefore a major phenomenon and we can reasonably assume that the system is in an over-damped regime, i.e. that any movement will almost instantaneously reach terminal velocity. Under this assumption, the acceleration of any agent can be considered to be null.

We also assume that the agents move at relatively low speeds. Together with the high viscosity of the substrate, this hypothesis imply that the medium exhibits a Stokes flow (i.e. its Reynolds number is very small compared to 1). This allows us to determine the form of the friction force sustained by the fibres. We denote by μ_{fib} the friction coefficient (or dynamic viscosity) of the fibres with the substrate, expressed as a mass per length per time ($M.L^{-1}.T^{-1}$). The friction force experienced by a fibre k can be approximated by the expression :

$$\boldsymbol{F}_k^{fric}(t) = -\mu_{fib} L_{fib} \frac{d\boldsymbol{Y}_k}{dt}(t), \tag{2.11}$$

which is associated to the torque :

$$\boldsymbol{T}_k^{fric}(t) = -\mu_{fib} L_{fib}^3 \boldsymbol{\omega}_k(t) \wedge \frac{d\boldsymbol{\omega}_k}{dt}(t). \tag{2.12}$$

Note that, with this (very rough) approximation, we totally neglect the fibres anisotropy and treat fibres as spheres of radius L_{fib}. Since it is known that the friction phenomenon is highly dependent of the shape of the object considered, a more adequate modelling would be to replace the scalar coefficient μ_{fib} by a tensor $M_k = \mu_{fib}^{||} \boldsymbol{\omega}_k \otimes \boldsymbol{\omega}_k + \mu_{fib}^{\perp}(I - \boldsymbol{\omega}_k \otimes \boldsymbol{\omega}_k)$ depending on the direction of the fibre and on two friction coefficients $\mu_{fib}^{||}$ and $\mu_{fib}^{\perp}$ for motion either parallel or perpendicular to the fibre axis.

2.4. Equations of motion of the fibres

Under the assumption of an over-damped regime, Newton's second law of motion (a.k.a the fundamental principle of dynamics) states that the sum of the forces and the sum of the torques applied to any agent of the system are null :

$$\begin{cases} \boldsymbol{F}_k^{fric}(t) + \sum\limits_{\substack{m=1 \\ m \neq k}}^{N_{fib}} \left(\boldsymbol{F}_{k,m}^{rep}(t) + p_{k,m}(t) \boldsymbol{F}_{k,m}^{link}(t) \right) = 0 \\ \\ \boldsymbol{T}_k^{fric}(t) + \sum\limits_{\substack{m=1 \\ m \neq k}}^{N_{fib}} \left(\boldsymbol{T}_{k,m}^{rep}(t) + p_{k,m}(t) (\boldsymbol{T}_{k,m}^{link}(t) + \boldsymbol{T}_{k,m}^{align}(t)) \right) = 0 \end{cases} \quad \forall k \in [\![1, N_{fib}]\!]. \tag{2.13}$$

We denote by $\mathcal{F}(t)$ the coordinates of all fibre units in the modelled system at time t and by $\mathcal{L}(t)$ the coordinates of all crosslinks, that is :

$$\mathcal{F}(t) = \big\{(\boldsymbol{Y}_k(t), \boldsymbol{\omega}_k(t)) \;\big|\; 1 \leqslant k \leqslant N_{\mathit{fib}}\big\},$$
$$\mathcal{L}(t) = \big\{(k, m, l^{\mathit{link}}_{k,m}, l^{\mathit{link}}_{m,k}) \;\big|\; 1 \leqslant k < m \leqslant N_{\mathit{fib}} \text{ and } p_{k,m}(t) = 1\big\}.$$

Since the friction forces depend linearly on the derivatives $\dfrac{d\boldsymbol{Y}_k}{dt}(t)$ of the position of the agent centres while the other mechanical interactions can be directly computed from the sets $\mathcal{F}(t)$ and $\mathcal{L}(t)$, the first line of the above system can be easily reformulated as $\dfrac{d\boldsymbol{Y}_k}{dt}(t) = F_k(\mathcal{F}(t), \mathcal{L}(t))$.

The second line of system is a little trickier to reformulate. Given that, for all $k \in [\![1, N_{\mathit{fib}}]\!]$, the directional vector $\boldsymbol{\omega}_k(t)$ lives on the surface of the unit half-sphere $\mathbb{S}^+_2$, we know first that $\|\boldsymbol{\omega}_k(t)\| = 1$ and second that $\boldsymbol{\omega}_k(t) \perp \dfrac{d\boldsymbol{\omega}_k}{dt}(t)$. Thus, we have $\left(\boldsymbol{\omega}_k(t) \wedge \dfrac{d\boldsymbol{\omega}_k}{dt}(t)\right) \wedge \boldsymbol{\omega}_k(t) = \dfrac{d\boldsymbol{\omega}_k}{dt}(t)$ for all fibres k and at all times t.

Using all this, we can rewrite the system (2.13) as follows :

$$\begin{cases} \dfrac{d\boldsymbol{Y}_k}{dt}(t) = \dfrac{1}{\mu_{\mathit{fib}} L_{\mathit{fib}}} \displaystyle\sum_{\substack{m=1 \\ m \neq k}}^{N_{\mathit{fib}}} \big(\boldsymbol{F}^{\mathit{rep}}_{k,m}(t) + p_{k,m}(t) \boldsymbol{F}^{\mathit{link}}_{k,m}(t)\big) \\ \dfrac{d\boldsymbol{\omega}_k}{dt}(t) = \dfrac{1}{\mu_{\mathit{fib}} L^3_{\mathit{fib}}} \displaystyle\sum_{\substack{m=1 \\ m \neq k}}^{N_{\mathit{fib}}} \Big(\boldsymbol{T}^{\mathit{rep}}_{k,m}(t) + p_{k,m}(t)(\boldsymbol{T}^{\mathit{link}}_{k,m}(t) + \boldsymbol{T}^{\mathit{align}}_{k,m}(t))\Big) \wedge \boldsymbol{\omega}_k(t) \end{cases} \quad \forall k \in [\![1, N_{\mathit{fib}}]\!]. \quad (2.14)$$

It is important to note that this deterministic system apply only between the jumps of the various Poisson processes included in our model (fibres linking and fibres unlinking). In reality, our model is not described by a system of deterministic, Ordinary Differential Equations (ODEs) but by a system of Stochastic Differential Equations (SDEs). However, since there is no continuous components (e.g. a brownian motion) in the stochastic processes involved, but only jump components, the deterministic and stochastic part of the model can be solved separately (see section 4.1 below). It is thus simpler to write them separately.

3. Modelling the Adipose Tissue

3.1. Model introduction and summary : addition of adipocytes to the ECM

As described in chapter 1 (section 4.2), the 3D architecture of a mature Adipose Tissue is primarily made of clusters of fat-storing cells called adipocytes, surrounded but not quite separated by well-organised sheets of ECM. As in the case of the ECM, we want to explore whether this simple architecture can be achieved merely through local mechanical interactions between the adipocytes and the ECM. To test it, we constructed a three-dimensional IBM describing the insemination and growth of round cells in a network of dynamically interconnected fibres. This was done by adding adipocytes to the model ECMmorpho-3D described in the previous section, hereby creating a new model called ATmorpho-3D.

We give here a brief summary of the additional features associated to this model and refer the readers to the next sections for a full description. The adipocytes are represented by non-deformable spheres of variable radius. New adipocytes of minimum radius $R^{\mathit{min}}_{\mathit{ad}}$ are inseminated randomly according to a

Poisson process of frequency ν_{ins} until a maximum number N_{ad}^{max} is reached, and existing adipocytes undergo a linear volumic growth at constant speed K_{growth} until they reach the maximum radius R_{ad}^{max}. Both phenomena allow to model energy storage, respectively via hyperplasia and hypertrophy. All these adipocytes exert on each other and on the surrounding fibres a short-ranged repulsive force modelling volume-exclusion effects. This force is based on Hertzian theory and its maximum intensity E_{rep}^{AA} (resp. E_{rep}^{AF}) controls the amount of overlap tolerated between adipocytes (resp. between an adipocyte and a fibre). Finally, as for the fibres, we consider an over-damped regime for adipocyte motion associated with a dynamic viscosity μ_{ad}.

Apart from adipocytes, the adipose tissue hosts various cell types including adipocyte precursors, fibroblasts, vascular cells, nerves, multipotent mesenchymal stem-like cells and a large variety of immune cells. Two things must be noted. Firstly, mature adipocytes are much larger (between 50 and 100 µm in diameter) than any other type of cells (typically around 15 µm in diameter) and can thus reasonably be assumed to exert a far greater mechanical action on an individual level. Secondly, as a group adipocytes represent most of the volume of a mature adipose tissue. For these two reasons, we considered that the mechanical impact of individual cells other than adipocytes was negligible and did not represent these cells by agents. However, we did take their global impact into account by considering that they formed a viscous medium exerting a friction force over the modelled agents (see below). We also assumed that they were homogeneously distributed in space and time, implicitly fulfilling their biological roles as needed for any of the biological processes explicitly included in our model.

All the features of the model ATmorpho-3D are summarised in Figure 2.4 and the corresponding parameters listed in Table 2.2. We now turn to the details of the model components.

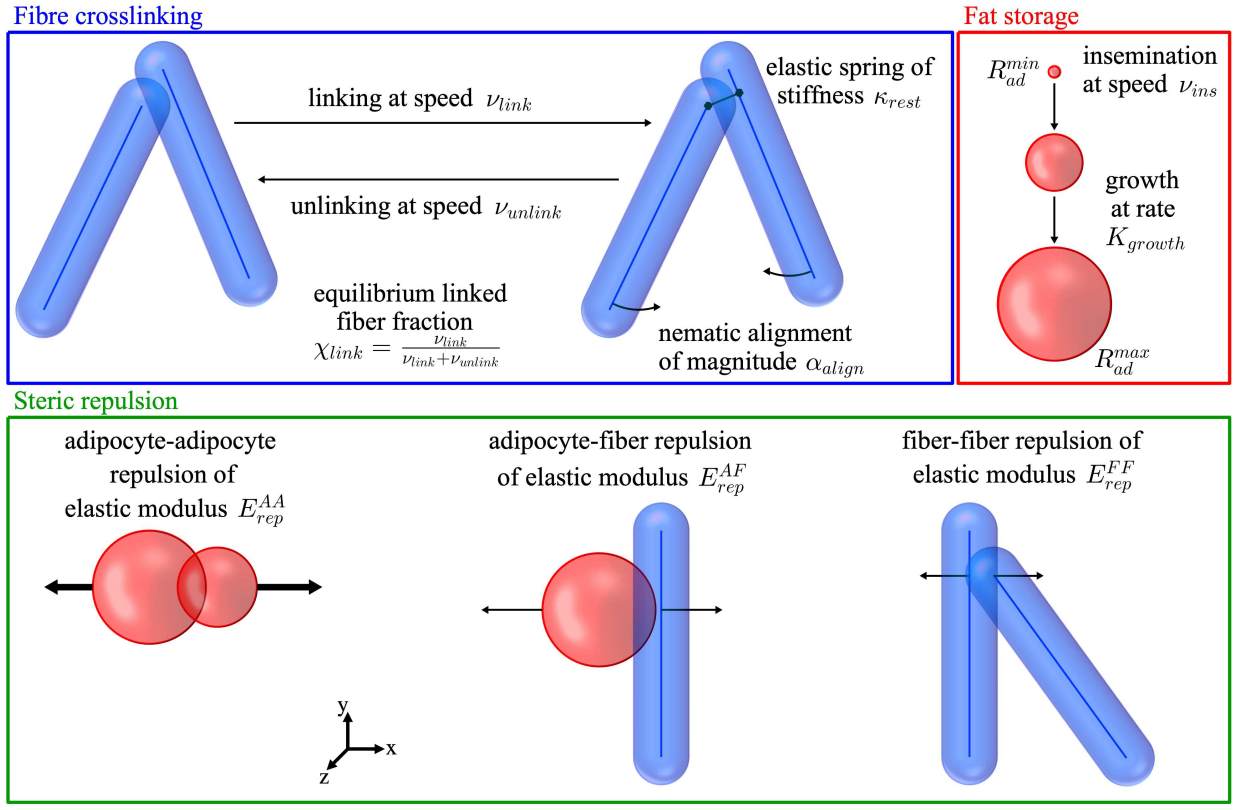


Figure 2.4: Scheme of all the biological processes and mechanical interactions included in the model ATmorpho-3D.

Parameter	Domain	Dimension	Description
Agent characteristics			
N_{ad}^{init}	$[\![0, N_{ad}^{max}]\!]$	N/A	Initial number of adipocytes in the system.
N_{ad}^{max}	$\mathbb{N}$	N/A	Maximum number of adipocytes in the system.
R_{ad}^{min}	$\mathbb{R}^{+*}$	L	Minimum adipocyte radius (= radius at insemination).
R_{ad}^{max}	$[R_{ad}^{min}, +\infty[$	L	Maximum adipocyte radius.
N_{fib}	$\mathbb{N}$	N/A	Number of fibres in the system.
L_{fib}	$\mathbb{R}^{+*}$	L	Fibre length.
R_{fib}	$\mathbb{R}^{+*}$	L	Fibre radius.
Adipocyte processes			
ν_{ins}	$\mathbb{R}^{+}$	T^{-1}	Adipocyte insemination frequency.
K_{growth}	$\mathbb{R}^{+}$	$L^3.T^{-1}$	Adipocyte volume growth per unit of time.
Crosslink processes			
ν_{link}	$\mathbb{R}^{+}$	T^{-1}	Fibre linking frequency.
ν_{unlink}	$\mathbb{R}^{+}$	T^{-1}	Fibre unlinking frequency.
d_{link}^{eq}	$[2R_{fib}, +\infty[$	L	Link equilibrium length.
d_{link}^{max}	$[d_{link}^{eq}, +\infty[$	L	Link maximum catching distance.
Mechanical interactions			
E_{rep}^{AA}	$\mathbb{R}^{+}$	$M.L^{-1}.T^{-2}$	Elastic modulus of the adipocyte-adipocyte contact.
E_{rep}^{AF}	$\mathbb{R}^{+}$	$M.L^{-1}.T^{-2}$	Elastic modulus of the adipocyte-fibre contact.
E_{rep}^{FF}	$\mathbb{R}^{+}$	$M.L^{-1}.T^{-2}$	Elastic modulus of the fibre-fibre contact.
κ_{rest}	$\mathbb{R}^{+}$	$M.T^{-2}$	Link stiffness.
α_{align}	$\mathbb{R}^{+}$	$M.L^{2}.T^{-2}$	Maximum intensity of the alignment torque between linked fibres.
μ_{ad}	$\mathbb{R}^{+}$	$M.L^{-1}.T^{-1}$	Dynamic viscosity of adipocytes.
μ_{fib}	$\mathbb{R}^{+}$	$M.L^{-1}.T^{-1}$	Dynamic viscosity of fibres.

Table 2.2: Summary of all the parameters of the model ATmorpho-3D.

3.2. Biological phenomena : adipocyte differentiation and growth

Adipocytes being relatively spherical cells, we choose to represent them by spheres of centre $\boldsymbol{X}_i(t) \in \Omega$ and radius $R_i(t) \in \mathbb{R}^{+*}$ (see Figure 2.5.A). Figure 2.5 gives a schematic representation of the two type of agents included in the model ATmorpho-3D and summarises the associated mathematical notations. We denote by :

$$\mathcal{A}(t) = \left\{ (\boldsymbol{X}_i(t), R_i(t)) \;\middle|\; 1 \leqslant i \leqslant N_{ad}(t) \right\}$$

the coordinates of the $N_{ad}(t)$ adipocytes present in the modelled system at time t.

The main role of AT is to maintain the body energy balance by storing excess energy (carbohydrates

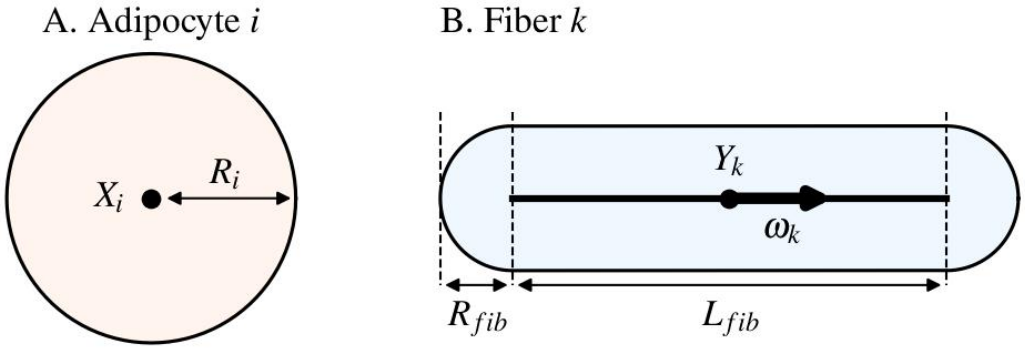


Figure 2.5: Schematic representation of the two types of agents found in the model ATmorpho-3D. **A :** adipocytes are modelled by spheres of centre $\boldsymbol{X}_i \in \Omega$ and radius $R_i \in \mathbb{R}^{+*}$. **B :** Fibres are modelled by spherocylinders of centre $\boldsymbol{Y}_k \in \Omega$ and directional vector $\boldsymbol{\omega}_k \in \mathbb{S}_2^+$, with fixed length $L_{fib} \in \mathbb{R}^{+*}$ and radius $R_{fib} \in \mathbb{R}^{+*}$.

and lipids) as triglyceride droplets inside adipocyte cells and releasing it according to metabolic needs. Since we want to model the emergence of AT architecture, we are specifically interested in AT morphogenesis, that is the creation of a new layer of adipose tissue to store a continued (not occasional) excessive influx of energy. For this reason, we will only include in our model energy storage and not energy release.

This storage has two direct consequences : adipocytes increase in volume as they store more fat (hypertrophy), and pre-adipocyte cells differentiate into full-fledged adipocytes (hyperplasia). We account for the latter phenomenon by inseminating a new adipocyte with a small radius R_{ad}^{min} at a position randomly selected according to an uniform distribution in the domain Ω (thus assuming an uniform repartition of the progenitor cells). We generate the successive insemination times using a Poisson process of frequency ν_{ins}.

We model adipocyte growth, in a simplified way, by a linear volume growth of constant rate K_{growth} capped by a maximal adipocyte radius R_{ad}^{max} :

$$R_i^3(t+\Delta t) = \max\big(R_i^3(t) + K_{growth}\Delta t,\ (R_{ad}^{max})^3\big). \tag{2.15}$$

To avoid overcrowding the modelled tissue, we set a maximal domain occupancy rate for adipocytes p_{ad}, from which we determine the maximal number of adipocytes allowed in the system :

$$N_{ad}^{max} = p_{\text{ad}} \times \frac{3|\Omega|}{4\pi(R_{ad}^{max})^3}. \tag{2.16}$$

We stop the insemination process when $N_{ad}(t)$ reaches N_{ad}^{max}.

Note that this is equivalent to assuming that a vascularisation network was implicitly developing "behind the scene" in the modelled system, in a way that supported all the biological phenomena we explicitly modelled. For instance, we assumed that all adipocytes always received all the oxygen and nutrients necessary to their survival. Note that this is equivalent to assuming an infinite influx of nutrients everywhere in the tissue : it is an extremely rough and unrealistic approximation. However, modelling a realistic vascularisation network and metabolism is a complex project, outside the scope of the current work. Moreover, as stated in introduction, [17] showed that spatially biasing the biological processes to account for an heterogeneous distribution of the vasculature did not impact the architectures generated by their model.

3.3. Mechanical interactions

Similarly to the fibres, the adipocytes modelled here are physical objects occupying a given volume from which other objects are excluded. Adipocytes undoubtedly classify as deformable objects because their cell walls, cytoplasm and triglyceride droplets are elastic (with an average overall cell stiffness lesser than 1 kPa [129]), so that two adipocytes coming into contact will deform and flatten against each other. We thus once again use Hertz's theory for non-adhesive elastic contact between curved surfaces [126, 127] to model the interaction between two adipocytes or an adipocyte and a fibre.

According to this theory, the deformation of two elastic spheres pushing on each other creates a contact area in the shape of a disk and a contact stress proportional to the radius of that disk and the indentation depth. Given two adipocytes i and j that touch each other, their indentation depth is $\delta = R_i + R_j - \|\boldsymbol{X}_i - \boldsymbol{X}_j\| \geqslant 0$ and the radius of their contact area is $\sqrt{R_{\mathit{eff}}\delta}$, with $R_{\mathit{eff}} = \dfrac{R_i R_j}{R_i + R_j}$ the effective radius. Therefore, according to the Hertzian theory, adipocyte j exerts on its neighbour i a repulsive force equal to :

$$\boldsymbol{F}_{i,j}^{AA} = \begin{cases} \dfrac{4}{3} E_{\mathit{rep}}^{AA} \sqrt{\dfrac{R_i R_j}{R_i + R_j}} \left(R_i + R_j - \|\boldsymbol{X}_i - \boldsymbol{X}_j\|\right)^{3/2} \dfrac{\boldsymbol{X}_i - \boldsymbol{X}_j}{\|\boldsymbol{X}_i - \boldsymbol{X}_j\|} \text{ if } \|\boldsymbol{X}_i - \boldsymbol{X}_j\| \leqslant R_i + R_j, \\ 0 \text{ otherwise,} \end{cases} \tag{2.17}$$

with E_{rep}^{AA} the adipocyte-adipocyte effective elastic modulus. As for the fibres, we assume that all adipocytes have the same elastic properties, so that this physical factor does not depend on the agent involved.

To model the repulsion between an adipocyte and a fibre, we assumed that the geometry of a contact between a sphere and a spherocylinder was similar to that of a contact between two spheres. This is obviously exact if the contact happens at one of the hemispherical end of the fibre, and is a reasonable approximation if the contact happens along the cylindrical body of the fibre.

Given an adipocyte i and a fibre k, we denote by $\boldsymbol{Y}_{k,i}$ the point on the central segment of the fibre which is closest to the centre of the adipocyte, i.e. the orthogonal projection of $\boldsymbol{X}_i$ on segment F_k :

$$\boldsymbol{Y}_{k,i} = \boldsymbol{Y}_k + l_{k,i}\boldsymbol{\omega}_k = \boldsymbol{Y}_k + \mathrm{Sat}((\boldsymbol{X}_i - \boldsymbol{Y}_k) \cdot \boldsymbol{\omega}_k)\boldsymbol{\omega}_k, \tag{2.18}$$

with Sat the saturation function over $[-L_{\mathit{fib}}/2, L_{\mathit{fib}}/2]$:

$$\mathrm{Sat}(x) = \begin{cases} -\dfrac{L_{\mathit{fib}}}{2} \text{ if } x < -\dfrac{L_{\mathit{fib}}}{2}, \\ x \text{ if } x \in \left[-\dfrac{L_{\mathit{fib}}}{2}, \dfrac{L_{\mathit{fib}}}{2}\right], \\ \dfrac{L_{\mathit{fib}}}{2} \text{ if } x > \dfrac{L_{\mathit{fib}}}{2}. \end{cases}$$

If the two agents are in contact then their indentation depth is $\delta = R_i + R_{\mathit{fib}} - \|\boldsymbol{X}_i - \boldsymbol{Y}_{k,i}\| \geqslant 0$ and the radius of their contact area is $\sqrt{R_{\mathit{eff}}\delta}$, with $R_{\mathit{eff}} = \dfrac{R_i R_{\mathit{fib}}}{R_i + R_{\mathit{fib}}}$ the effective radius. Following the Hertzian theory, fibre k exerts on adipocyte i a repulsive force equal to :

$$\boldsymbol{F}_{i,k}^{AF} = \begin{cases} \dfrac{4}{3} E_{\mathit{rep}}^{AF} \sqrt{\dfrac{R_i R_{\mathit{fib}}}{R_i + R_{\mathit{fib}}}} \left(R_i + R_{\mathit{fib}} - \|\boldsymbol{X}_i - \boldsymbol{Y}_{k,i}\|\right)^{3/2} \dfrac{\boldsymbol{X}_i - \boldsymbol{Y}_{k,i}}{\|\boldsymbol{X}_i - \boldsymbol{Y}_{k,i}\|} \text{ if } \|\boldsymbol{X}_i - \boldsymbol{Y}_{k,i}\| \leqslant R_i + R_{\mathit{fib}}, \\ 0 \text{ otherwise,} \end{cases} \tag{2.19}$$

with E_{rep}^{AF} the adipocyte-fibre effective elastic modulus. Conversely, adipocyte i exerts on fibre k a repulsive force equal to $\boldsymbol{F}_{k,i}^{FA} = -\boldsymbol{F}_{i,k}^{AF}$ according to Newton's third law. This force is applied at point $\boldsymbol{Y}_{k,i}$, inducing a rotational torque with respect to the fibre centre-of-mass $\boldsymbol{Y}_k$ of the form :

$$\boldsymbol{T}_{k,i}^{FA} = l_{k,i}\boldsymbol{\omega}_k \wedge \boldsymbol{F}_{k,i}^{FA}, \tag{2.20}$$

with $\wedge$ the cross product.

For the sake of standardisation, we rename the fibre-fibre effective elastic modulus introduced in the model ECMmorpho-3D from E_{rep} to E_{rep}^{FF}. Similarly, in this model the repulsive interaction between two fibres k and m will be denoted $\boldsymbol{F}_{k,m}^{FF}$ and the associated torque $\boldsymbol{T}_{k,m}^{FF}$ (defined by equations (2.5) and (2.6) respectively).

3.4. Equations of motion of the agents

We denote by μ_{ad} the friction coefficient (or dynamic viscosity) of the adipocytes with the substrate. Using the same hypotheses as in the model ECMmorpho-3D, we assume that the acceleration of all adipocytes is null and that the friction force experienced by an adipocyte i is equal to :

$$\boldsymbol{F}_i^{fric,A}(t) = -\mu_{ad} R_i(t) \frac{d\boldsymbol{X}_i}{dt}(t). \tag{2.21}$$

For the sake of standardisation, we rename the friction force sustained by a fibre k as $\boldsymbol{F}_k^{fric,F}$ and the corresponding torque as $\boldsymbol{T}_k^{fric,F}$ (see equations (2.11) and (2.12)). Applying the fundamental principle of dynamics, we obtain a system of equations :

$$\begin{cases} \displaystyle\sum_{\substack{j=1 \\ j\neq i}}^{N_{ad}} \boldsymbol{F}_{i,j}^{AA}(t) + \sum_{k=1}^{N_{fib}} \boldsymbol{F}_{i,k}^{AF}(t) + \boldsymbol{F}_i^{fric,A}(t) = 0 & \forall i \in [\![1, N_{ad}]\!], \\ \displaystyle\sum_{j=1}^{N_{ad}} \boldsymbol{F}_{k,j}^{FA}(t) + \sum_{\substack{m=1 \\ m\neq k}}^{N_{fib}} \left(\boldsymbol{F}_{k,m}^{FF}(t) + p_{k,m}(t)\boldsymbol{F}_{k,m}^{link}(t)\right) + \boldsymbol{F}_k^{fric,F}(t) = 0 & \forall k \in [\![1, N_{fib}]\!], \\ \displaystyle\sum_{j=1}^{N_{ad}} \boldsymbol{T}_{k,j}^{FA}(t) + \sum_{\substack{m=1 \\ m\neq k}}^{N_{fib}} \Big(\boldsymbol{T}_{k,m}^{FF}(t) + p_{k,m}(t)(\boldsymbol{T}_{k,m}^{link}(t) + \boldsymbol{T}_{k,m}^{align}(t))\Big) & \\ \qquad + \boldsymbol{T}_k^{fric,F}(t) = 0 & \forall k \in [\![1, N_{fib}]\!], \end{cases} \tag{2.22}$$

which, as in the case of the model ECMmorpho-3D, can be rewritten as follows :

$$\begin{cases} \dfrac{d\boldsymbol{X}_i}{dt}(t) = \dfrac{1}{\mu_{ad} R_i(t)} \left[\displaystyle\sum_{\substack{j=1 \\ j\neq i}}^{N_{ad}} \boldsymbol{F}_{i,j}^{AA}(t) + \sum_{k=1}^{N_{fib}} \boldsymbol{F}_{i,k}^{AF}(t) \right] & \forall i \in [\![1, N_{ad}]\!], \\ \dfrac{d\boldsymbol{Y}_k}{dt}(t) = \dfrac{1}{\mu_{fib} L_{fib}} \left[\displaystyle\sum_{j=1}^{N_{ad}} \boldsymbol{F}_{k,j}^{FA}(t) + \sum_{\substack{m=1 \\ m\neq k}}^{N_{fib}} \left(\boldsymbol{F}_{k,m}^{FF}(t) + p_{k,m}(t) \boldsymbol{F}_{k,m}^{link}(t) \right) \right] & \forall k \in [\![1, N_{fib}]\!], \\ \dfrac{d\boldsymbol{\omega}_k}{dt}(t) = \dfrac{1}{\mu_{fib} L_{fib}^3} \left[\displaystyle\sum_{j=1}^{N_{ad}} \boldsymbol{T}_{k,j}^{FA}(t) \right. & \\ \left. + \displaystyle\sum_{\substack{m=1 \\ m\neq k}}^{N_{fib}} \left(\boldsymbol{T}_{k,m}^{FF}(t) + p_{k,m}(t) (\boldsymbol{T}_{k,m}^{link}(t) + \boldsymbol{T}_{k,m}^{align}(t)) \right) \right] \wedge \boldsymbol{\omega}_k(t) & \forall k \in [\![1, N_{fib}]\!]. \end{cases} \tag{2.23}$$

Again, it is important to note that this deterministic system apply only between the jumps of the various Poisson processes included in our model (adipocyte insemination, fibres linking and fibres unlinking). In practice, the deterministic and stochastic part of the model are solved separately (see section 4.1 below).

4. Numerical implementation

In this section, we will describe the main features (including cost-optimisation ones) of the algorithm we wrote to conduct numerical simulations of the model ATmorpho-3D. Note that the model ECMmorpho-3D can be simulated using the same algorithm by merely setting the number of adipocytes to 0. The overall structure of the algorithm is outlined in Figure 2.6. The implementation was done in Fortran 90, a low level programming language offering two major advantages :

- its extremely good performances which enable us to keep numerical costs at a manageable level;
- its optimised mathematical libraries, some of which have no equivalent in other languages.

The resulting program is available on GitHub.

4.1. Euler-Maruyama scheme

The solution of the system (2.23) can be well approximated numerically using an explicit Euler scheme with adaptive time-step. The hypothesis of an over-damped regime therefore has the side benefit of saving computational resources at the level of the discretisation scheme. But, as mentioned above, our model does not reduce to system (2.23) : it also includes time-dependent stochastic processes.

Models involving both deterministic and stochastic components, the former usually modelling mechanical interactions and the latter chemical or biological processes, are now common in the field of mathematical biology. They are usually formulated as SDEs and there exists various numerical schemes to approximate their solutions, which are typically based on Taylor-type approximations [130]. The simplest

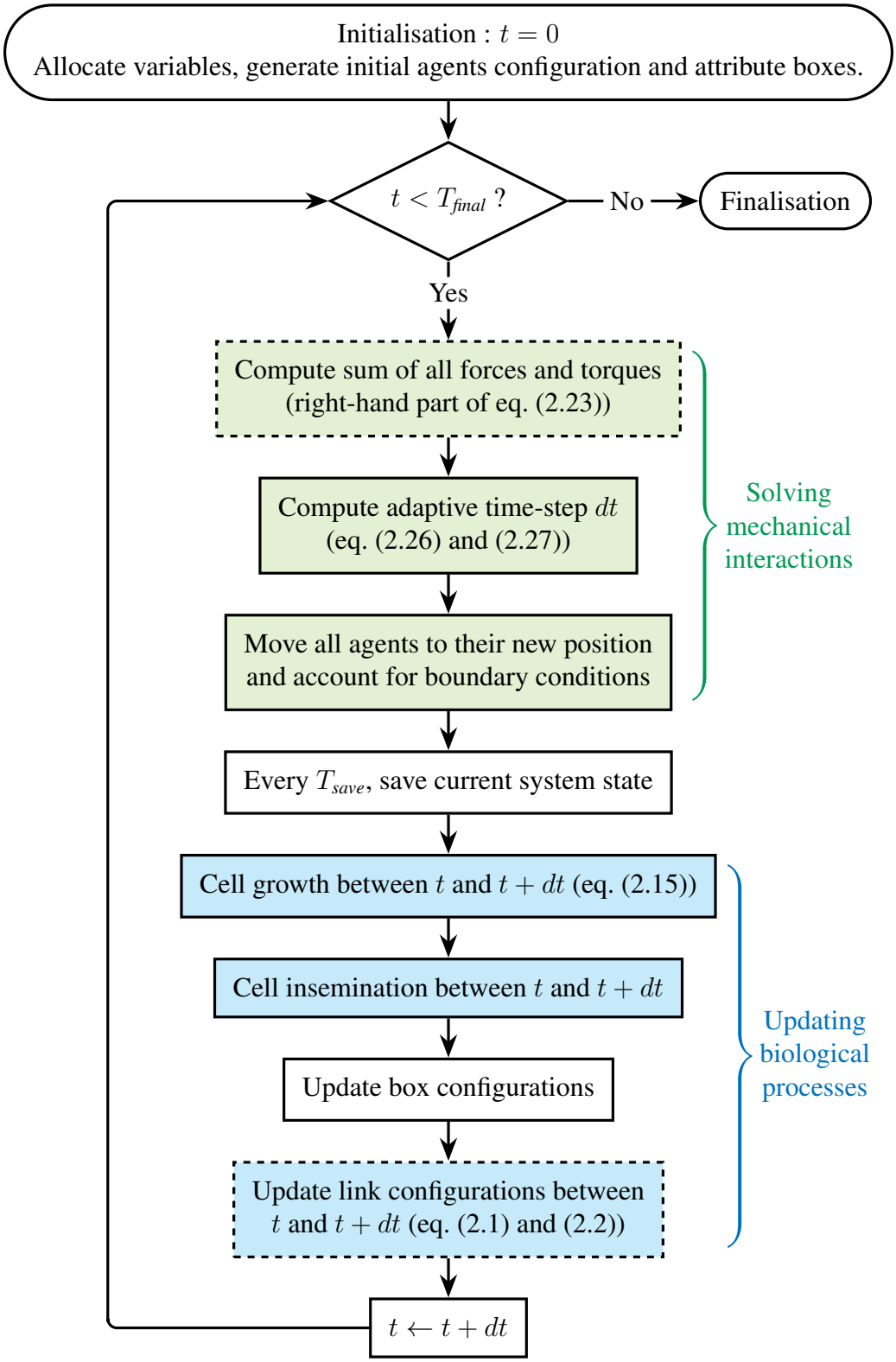


Figure 2.6: Flowchart of the algorithm for the model ATmorpho-3D. Parallelised sections are enclosed in dashed lines.

of these is the Euler-Maruyama schemes, which is an extension to SDEs of the classic Euler method for ODEs. Consider the following SDE :

$$\frac{dX}{dt}(t) = a(t, X(t)) + b(t, X(t))\frac{dS}{dt}(t) \quad \forall t \in [t_0, T], \tag{2.24}$$

where a, b are continuous functions and S is a stochastic process independent of the configuration of the system. Given a discretisation $t_0 < t_1 < ... < t_N = T$ of the time-interval $[t_0, T]$ and a sample function or realisation $\widetilde{S}$ of S, the Euler-Maruyama iterative scheme is defined by :

$$\begin{cases} X_0 = X(t_0), \\ X_{n+1} = X_n + a(t_n, X_n)(t_{n+1} - t_n) + b(t_n, X_n)(\widetilde{S}(t_{n+1}) - \widetilde{S}(t_n)) \quad \forall n \in [\![0, N-1]\!]. \end{cases} \tag{2.25}$$

The resulting sequence $(X_n)_{0 \leqslant n \leqslant N}$ is a time discrete approximation of the continuous time stochastic process which is solution of (2.24).

The time-steps $dt_n = t_{n+1} - t_n$ must be adapted to the quickest phenomenon of the model, either deterministic or stochastic. A special case occurs when the stochastic part of the SDE is not a continuous process (e.g. a brownian motion) but a jump process (e.g. a Poisson process) evolving on a much longer time-scale than the deterministic part. The two parts can then be dissociated in the following way :

- At a given time t_n, randomly draw the time t_{n+1} of the next jump.
- Discretise the interval $[t_n, t_{n+1}]$ in $N_n + 1$ steps $t_n = t_{n,0} < t_{n,1} < ... < t_{n,N_n} = t_{n+1}$ using a time-step adapted to the deterministic components of the model.
- Compute the approximated solution at time t_{n+1}^- using a classic, deterministic iterative scheme.
- Add the contribution due to the jump of the stochastic process to obtain the approximated solution at time t_{n+1}.

In our model however, the linking and unlinking processes are *not* independent of the system configuration. It is especially true for the linking processes, which explicitly depend on the relative position of the fibres. The unlinking processes, on the other hand, only depend on it insofar as a new stochastic process spawns each time a link is created.

Thus, the sample function $\widetilde{S}$ can not be generated in advance and we must be very careful as to when the jumps are applied, to ensure that the approximated solution $(X_n)_{0 \leqslant n \leqslant N}$ is a succession of (relative) mechanical equilibria. Indeed, one must see that, in our model, the stochastic processes are essentially disrupting the mechanical equilibrium that would otherwise be reached through the deterministic system (2.23). For instance, suppose that at time t_n none of the agents overlap and all links are at their equilibrium length : the system is thus in a perfect mechanical equilibrium and the contribution of the mechanical interactions between t_n and t_{n+1} will reduce to 0. Suppose now that during the interval $[t_n, t_{n+1}]$ a number of links are created or destructed and that adipocytes are inseminated at positions already occupied by other agents. Then, at time t_{n+1}, the position of all agents will be exactly the same as that of time t_n (plus the new adipocytes), but the system will be very far from mechanical equilibrium. This is not a realistic configuration and it should be quickly reorganised by the newly generated interactions, but meanwhile it is not correct to store it as a point X_{n+1} on the trajectory of the approximated solution.

A more correct approach would be to take the jumps into account before computing the mechanical interactions. If the time-step is constant this is easily done, but an Euler scheme with constant time-step

usually produces poor results (due to the build-up of truncation errors) and switching to an adaptive time-step is the very first step to take to improve accuracy. In that case, since the time-step $dt_n = t_{n+1} - t_n$ is generally not know before computing the mechanical interactions, the jumps must be generated with respect to the previous time-step $dt_{n-1} = t_n - t_{n-1}$. The resulting modified Euler-Maruyama scheme is :

(i) Considering an equilibrated configuration X_n at time t_n, randomly draw the stochastic jumps that happened during the interval $[t_{n-1}, t_n]$ based on X_n.

(ii) Apply them to obtain the non-equilibrated configuration X_n^+ at time t_n^+.

(iii) Compute the mechanical interactions in configuration X_n^+.

(iv) Evaluate the appropriate time-step dt_n.

(v) Compute the equilibrated configuration X_{n+1} at time $t_{n+1} = t_n + dt_n$ by applying the classic Euler scheme to configuration X_n^+.

Since the initial randomly generated configuration can be considered a non-equilibrated state, in practice our algorithm start at step (iii), as can be seen on the flowchart in Figure 2.6.

In the next section, we will discuss the choice of the adaptive time-step.

4.2. Adaptive time-step

The adaptive time-step is chosen according to two constraints : two agents must not be allowed to swap positions or "teleport" behind each other without ever coming into contact, and the probability of any given Poisson process (adipocyte insemination, pair-wise fibre linking or pair-wise fibre unlinking) to trigger more than once during a single time-step must be negligible.

The first condition is met if the mobility of the agents is restricted to half their smallest dimension per time-step. We also restrict the rotation of a fibre to $\arctan(0.1) \approx 5.71°$. This leads to the following upper limits for the computational time-step :

$$\begin{cases} dt_{ad}(t) = \min\limits_{1 \leqslant i \leqslant N_{ad}} \left(\dfrac{R_i}{2} \times \dfrac{\mu_{ad} R_i}{\|d\boldsymbol{X}_i/dt\|} \right), \\ dt_{fib}(t) = \min\limits_{1 \leqslant k \leqslant N_{fib}} \left(\dfrac{R_i}{2} \times \dfrac{\mu_{fib} L_{fib}}{\|d\boldsymbol{Y}_k/dt\|} \right), \\ dt_{rot}(t) = \min\limits_{1 \leqslant k \leqslant N_{fib}} \left(0.1 \times \dfrac{\mu_{fib} L_{fib}^3}{\|d\boldsymbol{\omega}_k/dt\|} \right). \end{cases} \tag{2.26}$$

Regarding the second condition, we arbitrarily considered as negligible a probability of less than 0.1. Given a Poisson process of frequency ν, the probability to have more than two occurrences in a time-span dt is equal to $p_{\geqslant 2}(dt) = 1 - e^{-\nu dt} - \nu dt e^{-\nu dt}$. The function $p_{\geqslant 2}$ is increasing over the interval $[0, +\infty[$ and $p_{\geqslant 2}(0.5/\nu) \approx 0.09$, so $p_{\geqslant 2}(dt) < 0.1$ for all $dt \leqslant 0.5/\nu$. Applying this to the Poisson processes of our model gives :

$$dt_{Poisson}(t) = \begin{cases} \min\left(\dfrac{0.5}{\nu_{ins}}, \dfrac{0.5}{\nu_{link}}, \dfrac{0.5}{\nu_{unlink}} \right) \text{ while } N_{ad}(t) < N_{ad}^{max}, \\ \min\left(\dfrac{0.5}{\nu_{link}}, \dfrac{0.5}{\nu_{unlink}} \right) \text{ after completion of the insemination process.} \end{cases} \tag{2.27}$$

Finally, the computational time-step is chosen as the minimum of all these upper-limits (with a reasonable floor value dt_{min} to prevent the program from being overly slow) :

$$dt_n = \max(\min(dt_{ad}(t_n),\ dt_{fib}(t_n),\ dt_{rot}(t_n),\ dt_{Poisson}(t_n)), dt_{min}). \tag{2.28}$$

Note that the limits defined in (2.26) depend on the actual configuration of the system and must be recomputed at each step, which is why the time-step is said to be "adaptive" and not constant. The limit $dt_{Poisson}$ on the other hand is a piecewise constant function that changes value only once (at the end of the adipocyte insemination process) throughout the simulation.

4.3. Data recording and disk access management

In the absence of a reliable and objective criterion to assert that a system has reached its equilibrium state, we chose to run the numerical simulations for an arbitrary, user-defined period of time. This total time of evolution is denoted T_{final} and is given in U_t. It differs from the computational time of the simulation, which is the real-life run time of the simulation.

To follow the temporal evolution of the system, its state (namely the value of each variable corresponding to the datasets $\mathcal{A}(t)$, $\mathcal{F}(t)$ and $\mathcal{L}(t)$) is recorded at regular intervals of duration T_{save}. The computational time-step is adjusted so that the algorithm saves data at exactly the times pT_{save} ($p \in \mathbb{N}$) : this makes it much easier to post-process the data and to display them as a dynamical sequence (see chapter 3). In other words, if at a given iteration n of the Euler scheme we have :

$$\begin{cases} pT_{save} \leqslant t_n < (p+1)T_{save}, \\ \\ t_n + dt_n \geqslant (p+1)T_{save}, \end{cases}$$

then the actual value of the time-step dt_n is reduced to $(p+1)T_{save} - t_n$ so that the next iteration occurs at time $t_{n+1} = (p+1)T_{save}$.

Let $N_{save} = \left\lceil \dfrac{T_{final}}{T_{save}} \right\rceil$ be the number of recording intervals in the total time of evolution (with $\lceil \cdot \rceil$ the ceil function). For each $0 \leqslant p \leqslant N_{save}$, the algorithm will produce :

- a file `cells`p`.dat` containing $N_{ad}(pT_{save})$ rows with the position and radius of all adipocytes (only if the system includes adipocytes, i.e. $N_{ad}^{max} > 0$);
- a file `fibres`p`.dat` containing N_{fib} rows with the position and direction of all fibres (only if the system includes fibres, i.e. $N_{fib} > 0$);
- a file `links`p`.dat` containing $N_{links} = |\mathcal{L}(t)|$ rows with the indexes k and m of the linked fibres and the positions $l_{k,m}$ and $l_{m,k}$ of their attachment sites (only if the links are dynamical, i.e. ν_{link} or $\nu_{unlink} \neq 0$).

All these files will be placed in a folder created at the beginning of the simulation, whose name is a user-defined string followed by the current date and time in the format yy.MM.dd_HH.mm.

Producing these files in real time when reaching $t = pT_{save}$ is the easiest and most straightforward method, but requires to make frequent disk access requests. This will not be a problem on most machines. However, due to their high computational cost, simulations of IBM are usually ran on computation servers with two peculiarities :

- As their name indicates, they are designed for high performance calculations and not for data management, so they are not able to perform a large number of simultaneous disk accesses.

- For safety reasons (e.g. in case of power outage or network connection loss), their operating system disables cache write. In other words, it ensures that the data is actually written to the hard disk, and not simply in a cache, before moving on to the next operation or giving hand back to the user.

Therefore, because the standard practice is to take full advantage of a server by having multiple processes running simultaneously (whether they be launched by the same person or by several users), a bottleneck may appear at the data saving step. This is called a data contention : the various processes have to pause while waiting to be granted access to the disk, sometimes resulting in a significant increase of the run time.

A way of bypassing this problem is to manually create a cache buffer, storing the data into dedicated variables and writing them in the disk only at the end of the simulation or when the RAM allocated to this buffer is full. Grouping the writing operations reduce the probability of interfering with the requests of another process, and thus the risk of bottleneck. The gain from this operation varies greatly depending on the server load and is thus difficult to quantify.

4.4. Cell linked-list for neighbour detection

To further reduce the computational time of our algorithm, we used a "Cell linked-list" approach to locate all agents on a numerical grid and decrease the number of pair-wise interactions to be computed at each time-step. The method is classic and widely used when implementing IBMs [131–134]. We describe it here in full for the sake of completeness.

The general idea is as follows. Consider a system of N agents interacting through mechanical forces of maximum range p_{max}. The straight-forward way to compute all pair-wise interactions has a $\mathcal{O}(N^2) = \mathcal{O}(\rho_{agent}^2|\Omega|^2)$ complexity (with ρ_{agent} the average agent density and $|\Omega|$ the volume of the domain). Increasing either the agent density (within a fixed domain) or the domain size (at fixed agent density) leads to a quadratic increase of the computational cost. While nothing can be done regarding the first case, the second case is not consistent with the mathematical number of interactions for short-ranged forces : since there is no interaction between agents further apart than p_{max}, a large number of the N^2 calculations are unnecessary.

Instead, let us divide the domain into identical rectangular cuboid boxes of side lengths greater than p_{max}. The number of boxes N_{box} will scale with the size of the whole domain, whereas the volume of a single box V_{box} will be fixed by the model parameters. It can be seen that, for a given agent, interactions need to be computed only with agents located in the same box or one of its 26 neighbours. The number of agents in one box being on average $N/N_{box} = \rho_{agent}V_{box}$, the algorithmic complexity is reduced to $\mathcal{O}(N^2/N_{box}) = \mathcal{O}(\rho_{agent}^2|\Omega|V_{box})$. As stated above the complexity is still a quadratic function of the agent density, but it has become linear with respect to the domain size. The study of large-sized systems being one of the major issues with IBMs, this gain is particularly appreciable.

We now turn to the technical implementation of this concept. Let $\Omega = [-L_x, L_x] \times [-L_y, L_y] \times [-L_z, L_z]$ be a domain centred on the origin and divided into N_{box} boxes of side length d_x, d_y and d_z. We arbitrarily chose to preserve the symmetry of the domain with respect to the origin of the coordinate system by imposing an even number of boxes in each direction. We denote by N_x (resp. N_y and N_z) the half number of boxes in the direction x (resp. y and z), so that $N_{box} = 8N_xN_yN_z$. Taking into account the fact that these numbers must be integers, that the side lengths must be greater than p_{max} and that reducing

V_{box} (i.e. increasing N_{box}) reduce the algorithmic complexity, we have :

$$N_a = \left\lfloor \frac{L_a}{p_{max}} \right\rfloor \quad \text{and} \quad d_a = \frac{L_a}{N_a} \quad \forall a \in \{x, y, z\}, \tag{2.29}$$

where $\lfloor \cdot \rfloor$ is the floor function. The procedure is illustrated in Figure 2.7. Note that, in our case, the maximum range of the mechanical interactions is equal to :

$$p_{max} = \max \left(\underbrace{2R_{ad}^{max}}_{\substack{\text{ad-ad} \\ \text{steric} \\ \text{repulsion}}}, \underbrace{R_{ad}^{max} + \frac{L_{fib}}{2} + R_{fib}}_{\text{ad-fib steric repulsion}}, \underbrace{L_{fib} + 2R_{fib}}_{\substack{\text{fib-fib steric} \\ \text{repulsion}}}, \underbrace{L_{fib} + d_{link}^{max}}_{\substack{\text{fibre} \\ \text{crosslinking}}} \right). \tag{2.30}$$

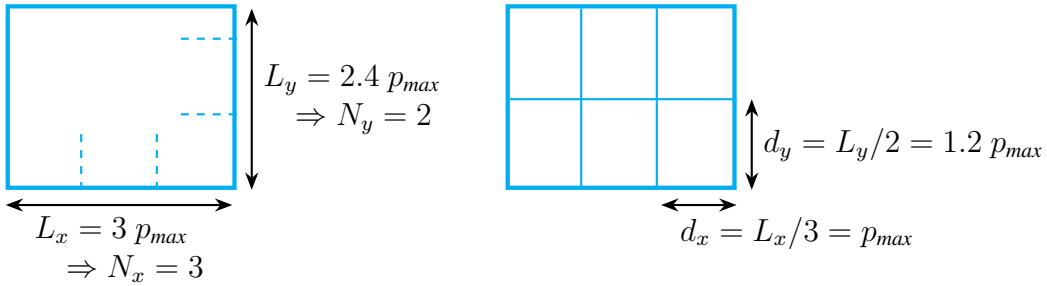


Figure 2.7: Scheme (in 2D) of the division of a domain into rectangular boxes. Only the subset $[0, L_x] \times [0, L_y]$, which represents a quarter of the domain, is drawn.

The boxes are indexed starting from 1 in the corner $(-L_x, -L_y, -L_z)$. Each box is therefore assigned a unique triplet (i_x, i_y, i_z) of coordinates along the three axes and a unique index $l \in [1, N_{box}]$ given by :

$$l = 1 + i_x + 2N_x i_y + 4N_x N_y i_z. \tag{2.31}$$

An agent located at point $X = (x, y, z)$ is thus inside the box of coordinates (i_x, i_y, i_z) defined by :

$$i_x = \left\lfloor \frac{x + L_x}{d_x} \right\rfloor, \quad i_y = \left\lfloor \frac{y + L_y}{d_y} \right\rfloor \quad \text{et} \quad i_z = \left\lfloor \frac{z + L_z}{d_z} \right\rfloor. \tag{2.32}$$

We will now see how to organise the agents in the various boxes. To do that, we typically use four lists :

- $BoxFirstAgent$ (of size N_{box}) which contains the index of the first agent in each box;
- $BoxLastAgent$ (of size N_{box}) which contains the index of the last agent in each box;
- $LinkedAgentList$ (of size N), where each cell j contains the index of the next agent in the same box as agent j;
- $AgentBox$ (of size N), which contains the index of the box each agent is in.

The agents are indexed starting from 1, so that a 0 in $BoxFirstAgent$ indicates an empty box and a 0 in $LinkedAgentList$ means there is no more agent in the box. For instance, the system depicted on the left of Figure 2.8 leads to the lists on the right, constructed using algorithm 1 (which has a $\mathcal{O}(N) = \mathcal{O}(\rho_{agent}|\Omega|)$ complexity). They can be read as follows :

- box No. 1 contains agents 3 and 4;
- box No. 9 contains agents 1, 2, 5 and 7;
- box No. 11 contains agents 6, 9 and 10;
- box No. 15 contains agent 8;
- all other boxes are empty.

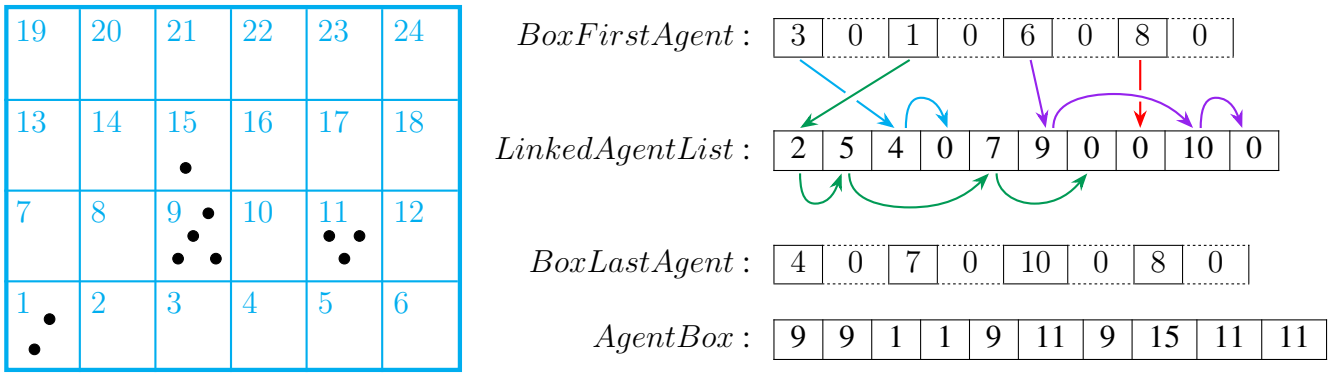


Figure 2.8: Example of agents localisation using the Cell linked-list method. Left : domain Ω divided in 24 boxes (numbered in blue) and containing a total of 11 agents (marked with black dots). Right : linked-lists constructed with algorithm 1. The dashed parts indicate multiple successive zero-valued cells.

Algorithm 1 Attribute each agent to a box.

$BoxFirstAgent$: [0]*N_{box}
$BoxLastAgent$: [0]*N_{box}
$LinkedAgentList$: [0]*N
$AgentBox$: [0]*N
for $i = 1, \ldots, N$ **do**
  $l =$ index of the box containing agent i ▷ *See eq.* (2.31) *and* (2.32).
  if $BoxFirstAgent(l) = 0$ **then**
    $BoxFirstAgent(l) = i$ ▷ *Store i as the first agent in box l.*
  else
    $LinkedAgentList(LastAgent(l)) = i$ ▷ *Add i at the end of the linked-list of box l.*
  end if
  $BoxLastAgent(l) = i$
  $AgentBox(i) = l$ ▷ *Store value of l to avoid future recalculations.*
end for

The sum of all mechanical interactions sustained by any agent can now be computed and stored in the list $Grad$ using algorithm 2. Note that the list $AgentBox$ has been added to save time by limiting recalculations in algorithm 2.

Algorithm 2 Compute all pair-wise interactions.

$Grad$: [0]*(N,3)
for $i = 1, \ldots, N$ **do**
  for all $(\delta_x, \delta_y, \delta_z) \in \{-1, 0, 1\}^3$ **do**
    $l_{neigh} = AgentBox(i) + \delta_x + N_x\delta_y + N_xN_y\delta_z$ ▷ *l_{neigh} is a neighbour of box $AgentBox(i)$. See section 4.5 for the handling of boundary conditions.*
    $j = BoxFirstAgent(l_{neigh})$ ▷ *Get first agent in box l_{neigh}.*
    while $j \neq 0$ **do** ▷ *There are still unvisited agents in the box.*
      if $j > i$ **then** ▷ *Pair-wise interactions only need to be computed once per pair.*
        Compute force F exerted by agent j over agent i
        $Grad(i,:) = Grad(i,:) + F$
        $Grad(j,:) = Grad(j,:) - F$ ▷ *Reciprocal interaction.*
      end if
      $j = LinkedAgentList(j)$ ▷ *Get next agent in the box.*
    end while
  end for
end for

4.5. Periodic boundary conditions

When modelling a system, an important question is the choice of the domain of modelling : it can be either finite or infinite and, in the former case, must be associated with boundary conditions.

For IBMs (which by definition contains a finite number of agents), an infinite domain is appropriate only if the system modelled can reasonably be assumed to be isolated from other entities by a large amount of empty space : a flock of birds in the sky, a cluster of stellar bodies in the void, etc. Non isolated systems may be modelled in a finite domain with various types of boundary conditions representing their interaction with the outside world. When the modelled system is too big (in term of the number of agents) to be numerically simulated as a whole, the usual method is to resort to a finite domain with periodic boundary conditions : this is equivalent to assuming that the exact same system is duplicated ad infinitum in every directions, as if the small system was immersed in a larger one.

Our model belongs to the last category. To implement the periodic boundary conditions, we used a combination of the "ghost boxes" and "periodic wrapping" approaches described below.

The ghost boxes approach, illustrated on Figure 2.9, consists in copying the external layer of boxes (hereafter referred to as the "frontier layer") of the domain, as well as all the agents within, and to paste them on the opposite border. The extended domain is denoted $\Omega^{BC} = [-L_x - d_x, L_x + d_x] \times [-L_y - d_y, L_y + d_y] \times [-L_z - d_z, L_z + d_z]$ and contains $8(N_x + 1)(N_y + 1)(N_z + 1)$ boxes numbered starting from corner $(-L_x - d_x, -L_y - d_y, -L_z - d_z)$ with the same convention as before. An agent located in the frontier layer is copied :

- 1 times if in a face-box of Ω (an edge-box in 2D),
- 3 times if in an edge-box of Ω (a vertex-box in 2D),
- 7 times if in a vertex-box of Ω.

The copies, called "ghost agents", are referenced in the data structure in the same way as their parent agent and inherit all their properties except their position (which must be translated to account for the copy/paste process).

Hence, the data structure must allow for the existence of as much as eight times the real number N of agents in the system (in the worst case where all real agents are located in vertex-boxes). This high memory cost is the main drawback of the ghost boxes approach. Its advantage, on the other hand, is that it makes it extremely easy to account for boundaries in algorithm 2 : the `for` loop iterates only over the N real agents, located in real boxes which always have 27 neighbouring boxes (including themselves). Agents in the frontier layer will have ghost boxes and so ghost agents among their neighbours (see Figure 2.9 for examples), accessible through the extended $BoxFirstAgent$ and $LinkedAgentList$ variables. The only modification to make is to test if j is a real agent (i.e. $j \in [\![1, N]\!]$) before storing the reciprocal interaction in the variable $Grad$.

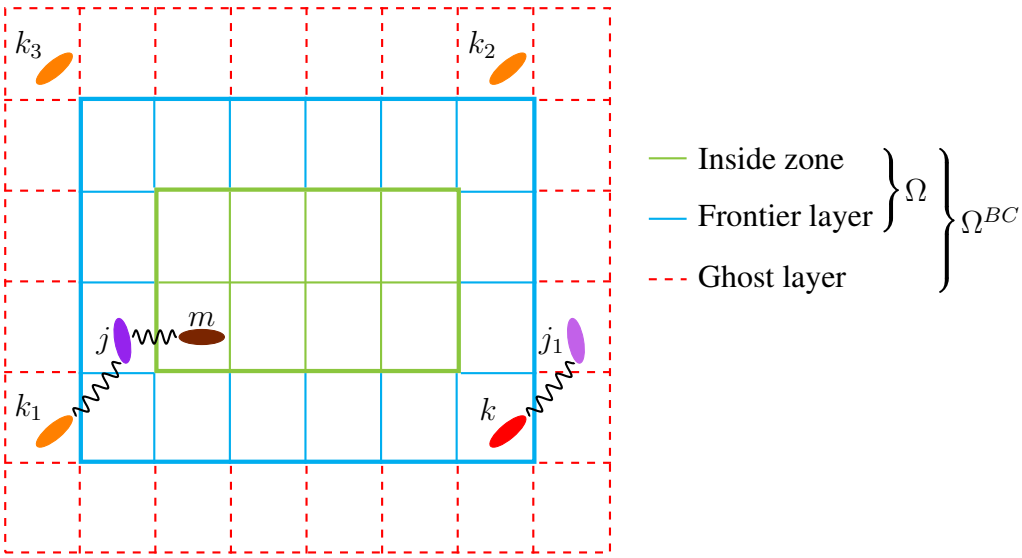


Figure 2.9: Scheme of the implementation of periodic boundary conditions on all sides of a 2D domain Ω. Inside zone : subset of the primary domain Ω that does not interact with the periodic boundaries. Frontier layer : subset of the primary domain Ω, composed of the outer layer of boxes, that does interact with the periodic boundaries. Agents located in this zone generate copies in the ghost layer. Ghost layer : extraneous layer of boxes added to the primary domain Ω to produce the extended domain Ω^{BC} (i.e. $\Omega^{BC} \backslash \Omega$).
For instance : agent k generates three ghosts k_1 (horizontal), k_2 (vertical) and k_3 (diagonal), while agent j only generates one ghost j_1 (horizontal). Agent m does not generate any ghost. Agents j an m interact naturally, while agents j and k interact through the periodic boundaries via the pairs (j, k_1) and (k, j_1).

The periodic wrapping approach, on the other hand, is more efficient and does not create any additional memory cost, but is also more complex to implement. It consists of identifying all pairs (B_i, B_j) of boxes of the frontier layer that interact through the periodic boundaries. We denote by $q_{j \to i}$ the wrapping vector from B_j to B_i, that is the correction to be applied to the coordinates of box B_j to wrap it around the domain to the neighbourhood of B_i. The pair-wise distance between two agents $\boldsymbol{X}_i \in B_i$ and $\boldsymbol{X}_j \in B_j$ can thus be computed as $\|\boldsymbol{X}_j + q_{j \to i} - \boldsymbol{X}_i\|$.

This method allows to compute interactions through the periodic boundaries without having to actually duplicate agents. The trouble is how to store the pairs (B_i, B_j) and the corresponding wrapping

vectors in a way that is both algorithmically efficient and easily readable.

In our code, we used the periodic wrapping method and solved the aforementioned problem by picking up the idea of the extended domain Ω^{BC} from the ghost boxes method. We did increase the number of boxes to include ghost boxes in the lists $BoxFirstAgent$ and $BoxLastAgent$, but we did not duplicate the data by creating ghost agents : instead, ghost boxes point to the same agents as the real boxes they mimic and we only store the translation vector that should have been applied according to the ghost boxes method, which is really the wrapping vector of the periodic wrapping method.

Since each ghost box l_{ghost} only mimics a single real box l_{real} (but a real box can be mimicked by multiple ghost boxes), the first part is done by setting $BoxFirstAgent(l_{ghost}) = BoxFirstAgent(l_{real})$ at the end of algorithm 1 (alternatively, we could have stored a pointer from l_{ghost} to l_{real}). The second part is done by creating a list $BoxType$ which stores the type (0 for real and 1 for ghost) of each box and a list $BoxWrap$ which stores their wrapping vectors (**0** for real boxes).

The correspondence between ghost and real boxes is as follows : a box has the type ghost if any of its coordinates i_a (with $a \in \{x, y, z\}$) is equal to its maximum $(2N_a + 1)$ or minimum (0) allowed value. The corresponding coordinate $\widetilde{i_a}$ of the associated real box is then the penultimate element on the other end of the definition interval (i.e. 1 or $2N_a$) and the wrapping vector along the a-axis is equal to $(i_a - \widetilde{i_a} + 1)d_a = \pm(L_a + d_a)$. Note that the lists $BoxType$ and $BoxWrap$ are constructed at initialisation and do not need to be updated during the simulation.

With this approach, algorithm 2 only needs to be slightly modified so that, if $BoxType(l_{neigh}) = 1$, the distance between agents i and j is not $\|\boldsymbol{X}_j - \boldsymbol{X}_i\|$ but $\|\boldsymbol{X}_j + BoxWrap(l_{neigh}) - \boldsymbol{X}_i\|$.

4.6. Parallelisation

Despite the optimisation of the two most time-consuming features (namely calculating the pair-wise interactions and recording the data, see sections 4.3 and 4.4 above), our algorithm takes a long time to run.

For instance, in chapter 5 we explore the behaviour of our model by conducting numerical simulations on a comparatively small domain ($L_x = L_y = L_z \simeq 2p_{max}$), up to a final time $T_{final} = 2000\ U_t$ allowing us to reasonably expect that the system will have reached equilibrium. The maximal adipocyte density and fibre density tested are both equal to 50%, corresponding respectively to 955 and 2578 agents. Simulations with such densities and intermediate linking dynamics ($\nu_{link} = 0.01$ and $\chi_{link} = 0.4$) take an average of 18 hours to complete. Parallelising the two main `for` loops of our algorithm (the computation of all forces and torques and the update of the linking configuration, see Figure 2.6) reduces this computational time to 1.85 hours !

Such parallelisation can easily be implemented in a Fortran 90 program using the library OpenMP (Open Multi-Processing) : the programmer only needs to add banners (respectively `!$OMP PARALLEL DO` and `!$OMP END PARALLEL DO`) at the beginning and end of the `for` loop they want to parallelise. However, this simplicity is misleading and hides two main traps : data dependency and concurrent memory access, both leading to what is called a race condition.

Let us start with a brief explanation of how parallelisation works. Each instance (or ongoing execution) of a program is called a "process". Each process is independent from other processes running on the same machine and can not communicate with them. A "thread" on the other hand is the unit of execution inside a process. Sequential processes are single-threaded, while parallel processes can have any number of threads. Contrary to multiple processes on the same machine, threads within the same process share memory and resources.

In a parallel process, the number of threads is usually not constant. Sequential regions of the program will be single-threaded. Upon reaching the start banner of a parallel `for` loop, the "master thread" will spawn multiple sub-threads and distribute the various iterations of the loop between them. This is called "scheduling" and can be done in three ways : static, dynamic and guided. Static scheduling splits the iterations evenly between the threads and is best suited for regular tasks (i.e. when each iteration of the loop takes approximately the same time). Dynamic scheduling splits the iterations into chunks (i.e. contiguous non-empty subsets) of fixed size and distributes them along the way to threads that have finished their previous task. It is better suited for irregular tasks because it allows a thread to complete a large number of quick tasks while another performs a small number of slow tasks. Guided scheduling operates in the same way except that the chunk size is not constant : it decreases with each successive allocation in order to reduce the end time difference between the various threads. The default behaviour is compiler-dependent. When all iterations of the loop are done, sub-threads merge into the master thread. Note that scheduling requires the total number of iterations, or loop trip count, to be known : `while` loops are thus not eligible to parallelisation.

Parallel loops can be nested : each sub-thread will then become the master thread of its own batch of sub-threads. The maximal number of threads created on each occasion can be set globally with the function `omp_set_num_threads` (with the possibility to define different values depending on the nesting level) or specified in the start banner of each loop. The default value is environment-dependent : it is usually either equal to 1 (i.e. no parallelisation) or to the number of Central Processing Unit (CPU) cores of the machine the process is running on.

Obviously, all these are technical considerations that should only affect the program run time, not its final outcome. When it is not the case, that is when the program outcome does depend on the order in which the iterations are carried out, it is referred to as a race condition.

Because the order in which the various iterations of the loop are accomplished depends on many factors, such as the choice of the scheduling method, the number of allocated threads and the run time of each iteration, it is very hard to predict. Hence all the loop iterations must be independent of each other, otherwise the behaviour of the program can not be guaranteed. This is why time-loops, where the state of a given iteration depends on the results of the previous ones (in sequential ordering), are not suitable for parallelisation. Such cases of data dependency are usually easily identifiable and may or may not be bypassed, depending on the circumstances.

The second main catch to parallelisation is the problem of concurrent memory access. This happens when two (or more) sub-threads perform a read-and-write operation on the same memory address at nearly the same time. A typical example would be a simple counter $var = var + 1$: if a thread tries to perform this instruction while another one is already in the process of doing so, then the value initialising both operations will be the same because the stored value of var will not yet have been modified by the first thread when the second one reads it. The variable var will thus be updated twice with the value $var + 1$ instead of reaching $var + 2$.

An example of concurrent memory access can be found in algorithm 2, where the same cell of the array $Grad$ may be accessed simultaneously by different sub-threads. For instance, if one thread is computing the interaction between agents $i = 1$ and $j = 8$ while another is computing the interaction between agents $i = 8$ and $j = 12$, both threads will update the value of the variable $Grad(8)$ with the result of their own computation, resulting in a race condition.

There are various ways to solve this issue :

1. If possible, modify the initial algorithm to circumvent the problem. In our case, it can be done by removing the "`if` $j > i$" condition and not computing the reciprocal interaction (see algorithm 3).

This roughly doubles the number of operations in the loop, but allows for a safe parallelisation without other additional cost. The final result (when parallelising over 10 threads) is a 5-fold gain.

Algorithm 3 Dissociation of the computation of the two reciprocal interactions of the agent pair (i, j).

if $j \neq i$ **then** ▷ *No self-interaction*
  Compute force F exerted by agent j over agent i
  $Grad(i) = Grad(i) + F$
end if

2. Put the banners `!$OMP CRITICAL` and `!$OMP END CRITICAL` around the lines $Grad(i) = Grad(i) + F$ and $Grad(j) = Grad(j) - F$, so that only one sub-thread at a time can execute these instructions. This option would be interesting if the threads did not encounter the critical region very often (for example, if we were counting the occurrences of a rare phenomenon), but in our case it create such a bottleneck that we lose nearly all gains from the parallelisation.

3. Put a banner `!$OMP ATOMIC UPDATE` before each of the two problematic lines. The instructions will then be performed atomically, i.e. allowing only one thread at a time to access a particular memory location. Compared to a critical construct, an atomic construct reduces the bottleneck issue because different threads can access different cells of the array $Grad$ simultaneously, but it can only be used for simple instructions of the form $x = x$ `op` exp, with `op` being either $+, -, *, /$ or a binary operator and exp an expression which does not involve x. However, the strong memory flush performed by atomic operations make them more costly than normal ones, a problem which increases with parallelisation. In our case, when parallelising over 10 or 20 threads, algorithm 2 with atomic operations has the same performances (in term of run time) than algorithm 3.

For the sake of simplicity, we settled on solution (1) (which has the best performances, ex-aequo with solution (3)).

Another important thing to note is that the variables used in a parallelised loop are either "shared" (between all threads) or "private" (a separate copy being created for each thread). For instance, in algorithm 2 the variables i, δ_x, δ_y, δ_z, l_{neigh}, j and F should be private because they are used to store values specific to each iteration of the loop. On the other hand, the variables $AgentBox$, N_x, N_y, $BoxFirstAgent$, $Grad$ and $LinkedAgentList$ should be shared since they contain values common to all iterations. Apart from $Grad$, whose case is discussed above, all these variables are never modified inside the loop so there is no risk of concurrent memory access.

Unintentional sharing of variables is a common source of race condition. Hence, it is crucial that the programmer declares as private all the variables that should be so (by default, only the loop index is private). In the same way that one should avoid implicit typing of variables, a good practice is to explicitly indicate the shared or private status of all variables (either at the beginning of the program or in the start banner of the parallel region) instead of relying on the implicit sharing behaviour.

Finally, an often overlooked point is the way pseudorandom generators behave in a parallel setting. All pseudorandom generators work by using a seed and a function to produce a random-looking number and the next seed. Because the seed value must be retained between two calls of the generator, the variable storing it is not local but global. For non thread-safe generators, different threads read from and write to the same memory address, compromising the performance of the generator and creating a race condition. In our case, we used the function `random_number` which, together with the `gfortran` compiler, ensures that the seed variable is private to each thread and that each copy is initialised with a

different value in a way that minimise the probability of aliasing. This generator has a period of $2^{256} - 1$, so each thread (up to a maximal number of $2^{128} \approx 3.10^{38}$) can generate $2^{128} \approx 3.10^{38}$ random numbers before any aliasing occurs.

Chapter 3

Analysis framework

In this chapter, we present the data processing framework that we developed to analyse the results of our two mathematical models ECMmorpho-3D and ATmorpho-3D presented in chapter 2, as well as the *in vivo* data acquired at the Restore Institute.

We start with the challenges and tools around 3D data visualisation, and present a new Paraview visualisation pipeline that we developed in collaboration with Charlotte Brunet, Juyeon Kim, Marion Saint-Pée and Sinan Haliyo from the Institute of Intelligent Systems and Robotics (ISIR) at Sorbonne University. We then present the segmentation methods we developed for the automatic detection of cell clusters from either *in silico* or *in vivo* 3D images, in collaboration with Laetitia Pieruccioni and Mathieu Vigneau from the Centre for Expertise and Technological Resources (CERT) at the Restore Institute, Toulouse. These two sections gave rise to a computational tool in the form of two Paraview macros and a Python module, which will be submitted as a methodological article (currently in preparation) and is freely available on GitHub.

Finally, the third section is devoted to the two numerical quantifiers that we developed to characterise the *in silico* fibre structures produced by our models.

Throughout this chapter, we will illustrate the discussion with both *in silico* and *in vivo* data :

- Simulation 1 : numerical simulation of the model ECMmorpho-3D.
- Simulation 2 : numerical simulation of the model ATmorpho-3D.
- Biological sample : light-sheet imaging of a mouse subcutaneous adipose tissue depot.

Summary of the chapter

1. 3D data visualisation

As discussed in chapter 1, the visualisation of tridimensional data, whether produced by mathematical models (typically through numerical simulations) or acquired from biological samples using tridimensional imaging techniques, is a complex issue. A number of softwares include specific features to plot 3D data on screen and manipulate the view. In this section, we will present our successive attempts at data visualisation with a computational software (Matlab®), a biological data processing software (Imaris) and a polyvalent visualisation engine (Paraview). The elements to visualise are the spherical cells and the spherocylindrical fibres described in chapter 2.

1.1. Using a computational software : Matlab®

When they need to visualise data, modellers tend to turn first towards computational softwares such as Python or Matlab, with which they are familiar and which provides specific libraries for data visualisation, including 3D data. However, as we will see from the illustrations below, these libraries are mainly intended for mathematical data such as parametric curves, surface plots, point clouds, etc., potentially 2D images in .png or .jpeg format. They can display polygonal 3D objects and allow interactive manipulation of the view angle, but they will quickly begin to lag as the number of displayed objects increases. Moreover, they do not have tools for performing operations such as slicing, clipping or thresholding : these results may be obtained by manual implementation, but this is cumbersome and can be very time-consuming if not well optimised.

Figure 3.1.A shows a visual of a simulation of our model ATmorpho-3D using Matlab, where we represented the adipocytes with red spheres (using `sphere`, a native Matlab function) and the fibres with blue spherocylinders (using the `spherocylinder` function by Sathish Sanjeevi [135]). As one can see, the result is too dense to distinguish anything beyond the outer layer, even when making the face of the displayed objects slightly transparent. This is due to the fact that the parameters chosen for this simulation (and for most of our simulations) produce a very dense system with closely packed agents and non-negligible amounts of overlap between agents. Moreover, although Matlab allows to interactively manipulate the view of a 3D visual by zooming, panning and rotating, these operations are made extremely slow by the large number of objects to alter and cause typical laptops to lag.

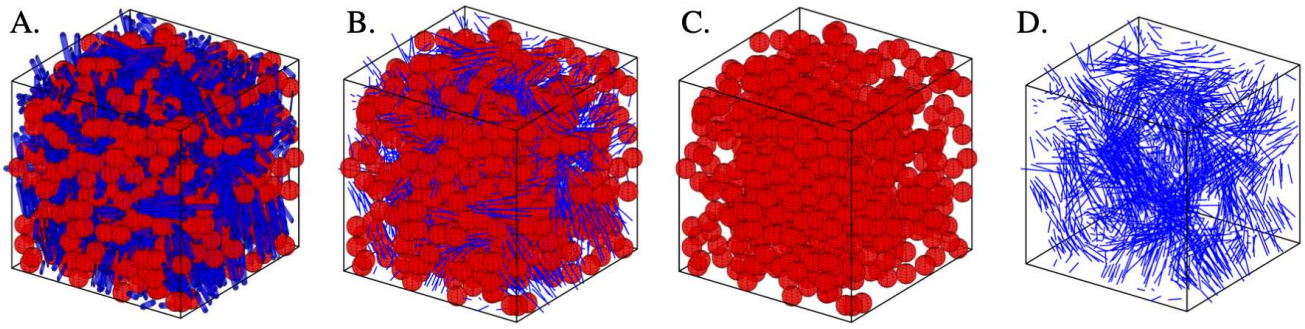


Figure 3.1: Visualisation of simulation 2 with Matlab. **A :** 3D image of the whole system, with fibres represented by spherocylinders. **B :** 3D image of the whole system, with fibres represented by segments. **C :** 3D image of the adipocytes. **D :** 3D image of the fibres, represented by segments.

In Figure 3.1.B, we show another visual where the fibres are only represented as line segments : the rendering is better but still hard to manipulate. In Figure 3.1.C and D, we display only one type of agents

at a time (adipocytes in panel C and fibres in panel D). With this last strategy the visualisation is much less obstructed, but the sense of perspective is still bad, the view manipulation very slow and we lack tools such as slicing.

1.2. Using a microscopic data processing software : Imaris

Another software tool to visualise data is the Imaris (Interactive Microscopy Image Analysis) software commonly used by the biological community. It is a specialised software especially adapted to data produced by biological imaging techniques, with a strong focus on 3D and 4D (temporal) rendering. It also includes modules for segmentation and morphological operations, making it both a visualisation and an analysis tool. This software is used at the Restore Institute to visualise the data obtained through light-sheet imaging of biological samples (see Figure 3.2.A). Since one of our goals is to compare the results of our numerical simulations with these *in vivo* data, it would be a real advantage to be able to visualise the two types of data (*in vivo* and *in silico*) using the same software. However, as many other biological data processing softwares (which are often paired with the equipment acquiring the data), Imaris is very restrictive concerning the format of the input data : it can only read pixelated images (2D), stacks of pixelated images (3D) or several stacks of pixelated images (4D), whereas the natural way to register the state of an IBM is to list the features of every components (see chapter 2, section 4.3). To use Imaris we thus had to manually create, from our raw simulation data, stacks of pixelated images representing successive slices of the system.

We display in Figure 3.2.B one such slice (created using Matlab) of a system simulated by our model ATmorpho-3D, with adipocytes represented in red and fibres in blue. For illustration purpose, here the two types of agents (cells and fibres) are represented on the same image with semi-transparent cores and solid edges. In practice, we created two stacks of black-and-white images, one for the adipocytes and one for the fibres (with objects in white and background in black), which we supplied to Imaris as two separate colour channels. This allowed us to interactively modify the colour and the level of transparency of each type of agent via Imaris colour setting.

After overcoming the initial difficulty of the input format, we could explore the various 3D visualisation tools provided by Imaris. Figure 3.2.C displays an example of 3D view of a simulation of our model ATmorpho-3D which, like in our previous attempt with Matlab, is rather obstructed. Figure 3.2.D displays the same view with application of Imaris interactive slicing tool, allowing to observe the inside of the system. Another possibility is to use the section view illustrated in Figure 3.2.D, that is synchronised slices along the three orthogonal directions.

Finally, it is possible to supply to Imaris several files at the same time if these files follow a nomenclature of the type "*CommonName_Cxxx_Tyyy*" where the *C* prefix references colour channels and the *T* prefix references time series. The time series allow us to visualise the temporal evolution of a system, while the colour channels can be used to manage separately the display characteristics (colour, transparency, etc.) of different components. As mentioned above, we used the latter feature to manage separately the visualisation of the adipocytes and the fibres.

Like Matlab, Imaris has the two disadvantages of being a commercial software and of being quite slow on laptops. Moreover, the pre-processing step needed to create the input image stack(s) is a huge drawback :

- It is computationally costly.
- It requires the use of a computational software in addition to Imaris, making the pipeline more complex.

- Achieving a satisfying resolution requires a high number of pixels in the images, leading to the production of rather large files (which must be stored and possibly transferred to another computer, see previous point).

- Including any supplementary information, such as which adipocyte pertains to which lobule, can only be done via a colour code, thus inducing the creation of one image stack per supplementary information, with no easy way of switching between the representations.

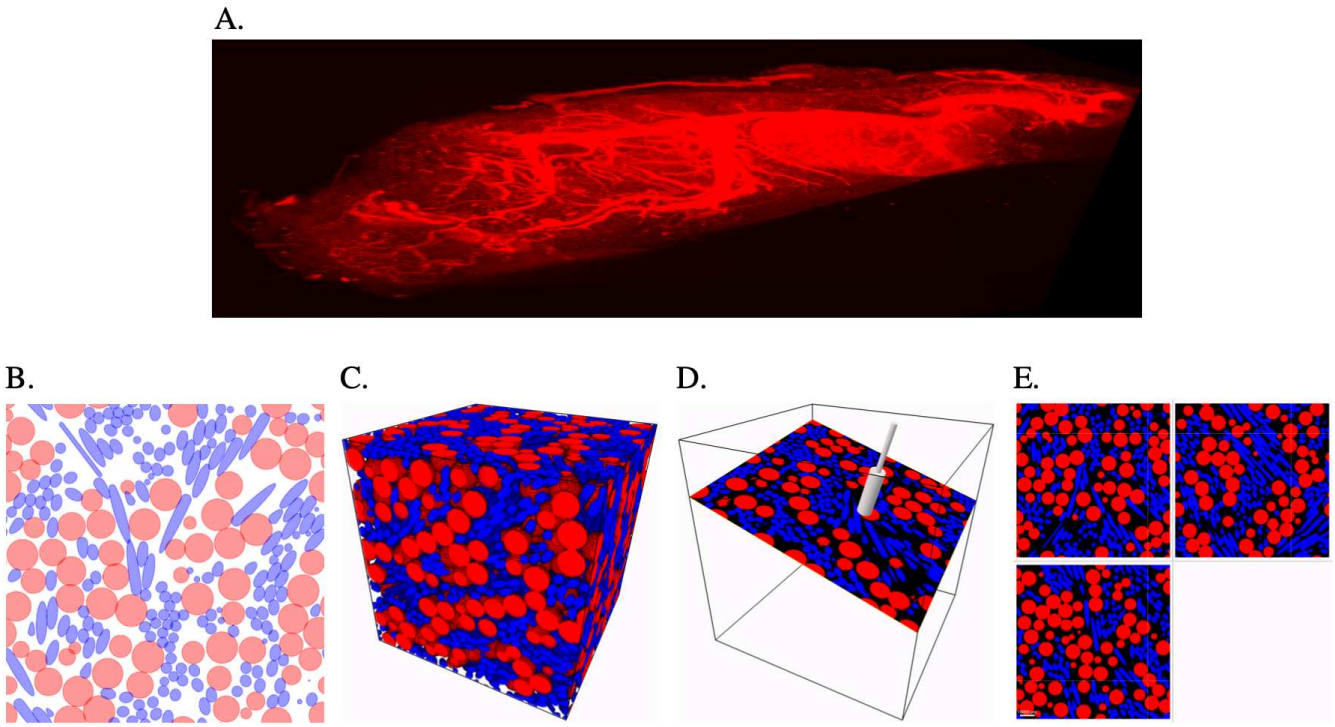


Figure 3.2: Visualisation with Imaris. **A :** 3D image of the biological sample. **B :** 2D slice of the system for simulation 1 (image created with Matlab). **C :** 3D image of simulation 1. **D :** Application of the slicer tool to simulation 1. **E :** Section view of simulation 1.

1.3. Using a polyvalent visualisation software : Paraview

Our aim is to develop a visualisation pipeline adapted to both *in silico* data and *in vivo* images, with seamless interactive view manipulation and as user-friendly as possible. All these objectives can be met using the software Paraview. It is a powerful, open-source software designed for the interactive visualisation of large datasets. The user-interface is quite intuitive and offers a large number of predesigned filters to construct and customise 2D or 3D representation from data files, as well as an easy way to create macros for the automation of established protocols. Paraview is a front-end application of the Visualisation Toolkit (VTK), a well-maintained software platform containing an expansive set of native functionalities and providing a robust foundation for scientific visualisation.

In collaboration with Pr. Sinan Haliyo (head of the "Multiscale Interactions" team of the Institute of Intelligent Systems and Robotics (ISIR) at Sorbonne University), Charlotte Brunet, Juyeon Kim and Marion Saint-Pée (students of the L3 CMI Mécanique of Sorbonne University), we developed a fully

automated 3D visualisation pipeline for data in the form of text files listing the properties of either spherical or rod-like objects.

This protocol uses Paraview `Sphere` filter to represent spherical objects (in our case, the adipocytes) and the `Arrow` filter to represent rod-like objects (in our case, the fibres). The latter choice was made after careful consideration of the two following points :

- In most of our simulations, the agent density is such that depicting each fibre by a full-fledged spherocylinder produces a very crowded image in which it is difficult to actually see anything past the outer layer, as was seen during our first visualisation with Matlab.
- In practice, spherocylinders as well as line segments convey a poor sense of depth.

The two issues can be solved by using thin double-headed arrows. Note that double-headed arrows are not among the predesigned glyphs proposed by Paraview (as of version 5.10), but can be obtained using two single-headed arrows per object.

The whole process implies several steps, but Paraview has an integrated tool to automatise such sequences by recording the user's actions through one manual iteration and saving them as a macro. This macro takes the form of a .py file (written in pvpython, the underlying language of Paraview and VTK) which can be integrated into any Paraview setup. In our case, we created two macros, one for displaying spherical objects and the other for rod-like objects (i.e. the adipocytes and fibres), because the data for these two types of agents were originally stored in separate files (see chapter 2, section 4.3). These macros, edited with ParaView-5.10, are available online at this address.

The result is displayed in Figure 3.3.A. Notice how the shading on the sphere glyphs enhances the sense of perspective. The user can easily switch to a fibre-only or adipocyte-only view (Figure 3.3.B and C) by simply ticking a box. The view can be manipulated smoothly through actions such as zooming, panning and rotating but also slicing and thresholding. The latter takes as input any feature registered in the data file provided to Paraview : thus, if we add to the data files produced by our simulations numerical quantifiers such a those described in section 3, we will be able to selectively display agents with a quantifier value above or below a given threshold. In the same way, each glyph (sphere or arrow) can be coloured depending on the value of any registered feature.

In addition Paraview can, like Imaris, display the temporal evolution of system by loading a sequence of files (with the same name except for a number) as a time-series. Any processing done to the series is in fact applied to each file separately : this can be a little slow if the number of files is high, but once the computation is done the visual rendering will be seamless. As an illustration, movies displaying the temporal evolution of a few simulations using Paraview can be found online at this address.

Finally, Paraview can display our biological sample images nicely : the data only need to be converted from their native proprietary format (czi) to a format readable by Paraview (e.g. tiff). The result is illustrated in Figure 3.3.E. As with our *in silico* data, the 3D view as well as the colour intensity and transparency can be controlled at will. Depending on the number of pixels in the image the operations may be slow or overflow a laptop memory. This can be circumvented either by using a workstation or by downsizing the image before visualisation.

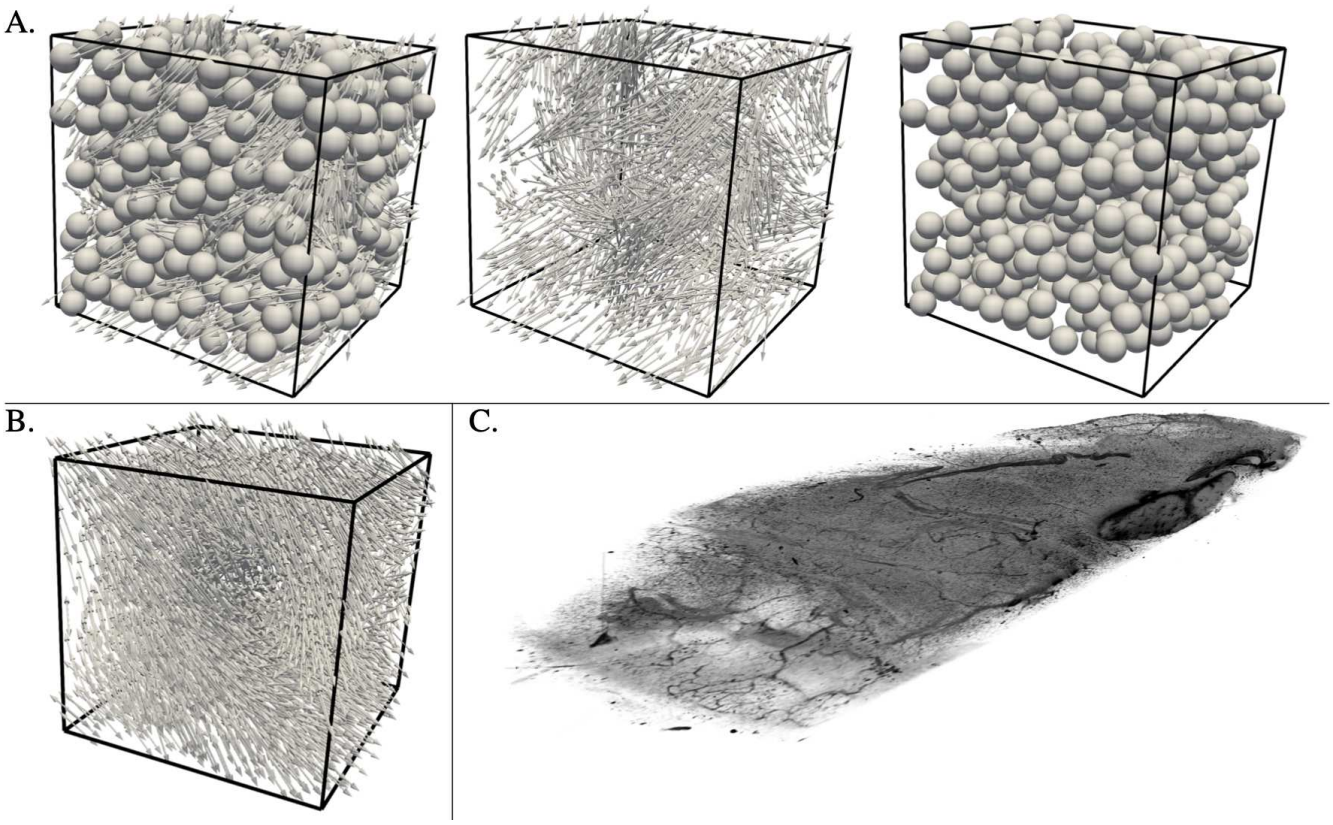


Figure 3.3: Visualisation of our three datasets with Paraview. **A :** 3D images of simulation 2. Adipocytes are represented by spheres, fibres by double-headed arrows and the edges of the spatial domain Ω are drawn in black. The left-hand panel displays the whole system, the central panel only the adipocytes and the right-hand panel only the fibres. **B :** 3D image of simulation 1. Fibres are represented by double-headed arrows and the edges of the spatial domain Ω are drawn in black. **C :** 3D image of the biological sample.

2. Characterisation of the cellular structures : lobules segmentation in 3D

As described in chapter 1 (section 4.2), the main structural block of AT is the lobule, i.e. a complex-shaped cluster of adipocytes more or less enveloped in a sheet of ECM, with a poly-lobed core surrounded by multiple string-like appendages called digitations which connect it to its neighbours.

The first step towards validating our model is thus to check if the architectures it produces do contain lobule-like structures. The second is to compare these *in silico* lobules, if they exist, to real *in vivo* lobules. To achieve these two objectives, we need a protocol to identify clusters of adipocytes in our *in silico* data and another, as similar as possible, to do the same with *in vivo* data. Note that this issue, and the solution we came up with, are not specific to our problem but can be generalised to any system or model where the agents organise in clusters (e.g. organoid, tumour, collective motion, etc).

There was a number of technical and theoretical obstacles :

1. We actually lack a clear and precise definition of what is a lobule : while 2D images suggested that they were well separated, ellipsoid-shaped clusters, 3D images revealed that they are in fact complex, interconnected poly-lobed structures. This makes the evaluation of a segmentation difficult

and at any rate subjective.

2. At first (and for nearly a year), we lacked of a clear, practical way to visualise our data in 3D. We were thus unable to check the relevance of the segmentations performed by our algorithms.

3. Concerning our *in silico* data, the fact that we choose to apply periodic boundary conditions in our model (see chapter 2) means that, most probably, some clusters will straddle the border of the simulation domain. Our *in silico* segmentation protocol thus has to be able to deal with periodic boundary conditions.

We gave in chapter 1 (section 3.2) a brief overview of the various existing methods to solve the widespread problem of objects segmentation and clustering. We tried a few of them, which proved to be inappropriate for various reasons :

- As the lobules have complex poly-lobed shapes, we can not define a "model" of their shape to supply to a model-based segmentation algorithm;
- Because they are all interconnected, connectivity-based segmentation methods give poor results;
- Because they are densely nested together, the separation between them are better characterised by the presence of fibres than a high distance between the cells, which makes density-based segmentation methods fail.

The solution we finally constructed rely on the watershed transformation which, as will be explained below, can divide a continuous shape into several regions based on the presence of cores or bulks : this property makes it ideally suited to our lobule segmentation issue. In this section, we will first outline the general principle of the watershed transformation, then explain how we adapted it to the two cases at hand (namely segmenting lobules in our numerical simulations and in biological sample images) and finally introduce our numerical quantifiers for the lobules thus segmented.

2.1. Image segmentation using the watershed transformation

The watershed transformation is based on the geographical concept of watershed ridges (or drainage divide), that is the lines separating neighbouring catchment (or drainage) basins. There are multiple definitions (and correspondingly multiple algorithms) [59] for how the catchment basins should be delimited :

- Watershed by flooding : water sources are placed at each local minimum of the map and gradually flood the entire map. Watershed lines are drawn where water from different sources meet.
- Watershed by water runoff : watershed lines are drawn at points from which a drop of water could flow down towards distinct minima.
- Watershed by topographic distance : a point of the map is in the catchment basin of its nearest local minimum in terms of topographic distance, that is the point located at the end of the path of steepest descent.

We choose to use Matlab `watershed` function, which is based on Meyer's flooding algorithm [136]. An example of the classic use of this function is displayed in Figure 3.4, with two equivalent representation of the same data : in panel A, a matrix of dimension 2 is seen as an altitude map for a 3D landscape,

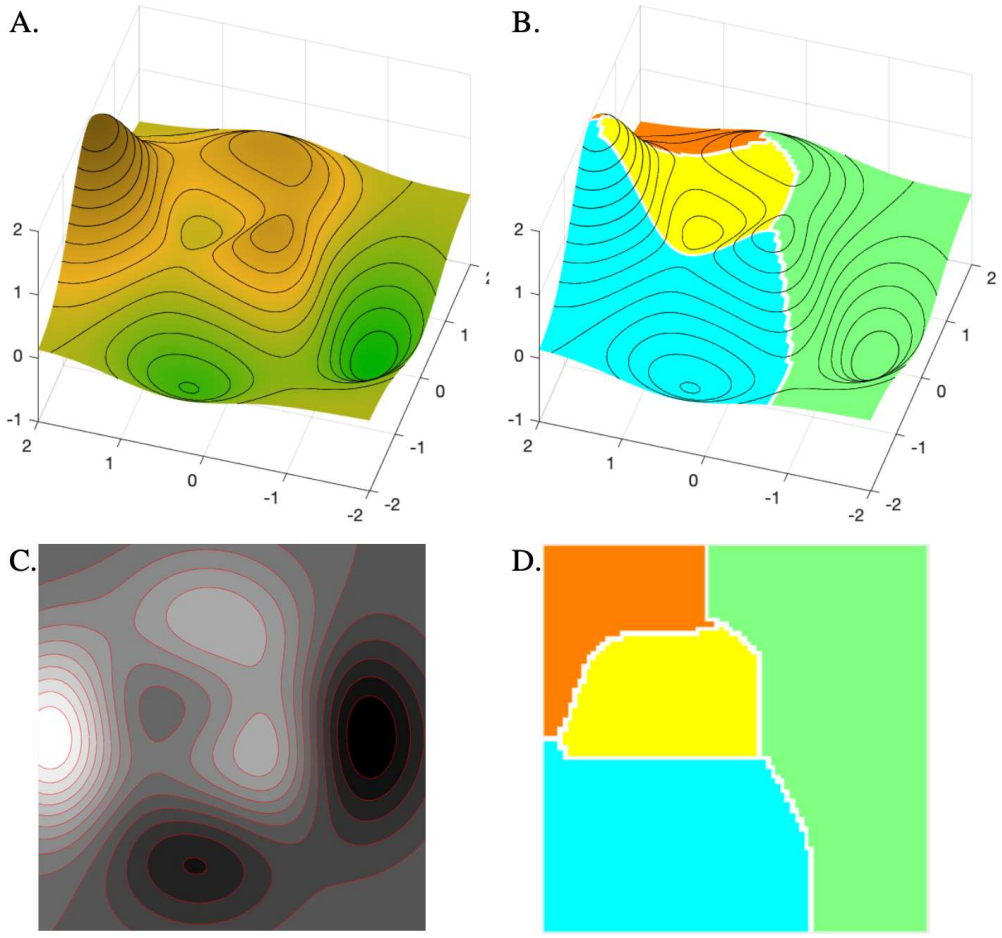


Figure 3.4: Illustration of the classic use of the watershed transformation. **A :** 3D landscape with colour indicating the altitude and isolines drawn in black. One can note the presence of three holes constituting local minima. There is actually a fourth local minima in the top left corner of the landscape. **B :** Landscape segmented into four catchment basins using Matlab `watershed` function, with colour indicating the basins and watershed lines drawn in white. **C :** Landscape viewed as a 2D grayscale image, with "isocolour" lines drawn in red for the sake of comparison. **D :** segmentation into four catchment basins viewed as a 2D RGB image, with watershed lines drawn in white.

whereas in panel C the same matrix is seen as a 2D grayscale image. Panels B and D display the result of Matlab `watershed` function applied to this matrix using both representations.

The watershed transformation is often used in the field of image analysis to divide a contiguous region of a black-and-white image into several touching objects or structures. The process, illustrated in Figure 3.5, can be described as follows :

- Consider white shapes over a black background (see Figure 3.5.A).
- For each white pixel, compute its Euclidean distance to the nearest black pixel. This creates a distance map, that is a grayscale image with a black background and grey shapes growing whiter away from their edges.
- Invert this distance map so that the shapes are now "valleys" (lower elevation than the baseline) instead of "mountains" (higher elevation than the baseline), with one or more local minima inside each shape (see Figure 3.5.B).

- Apply the `watershed` function to this grayscale image to obtain a labeled region map (see Figure 3.5.C). Note that the background is also divided between the different regions. It can be removed by setting the value of all original background pixels back to 0 (see Figure 3.5.D).

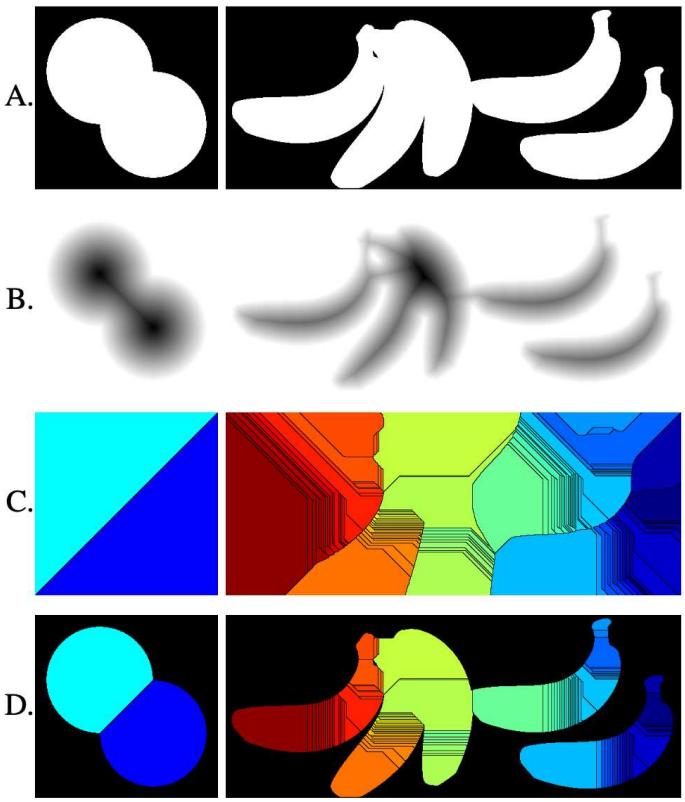


Figure 3.5: Illustration of the use of the watershed transformation to segment objects in an image. **A :** Two 2D black-and-white images with objects in white over a black background. **B :** Euclidean distance transform of these images computed using Matlab `bwdist` function. **C :** Region maps obtained by applying Matlab `watershed` function to the previous distance maps. **D :** Same region maps with original background pixels set to 0.

Watershed can be applied in exactly the same way to 3D images, using the 3D Euclidean distance and flooding in the three dimension simultaneously. This amounts to considering 3D images as 4D landscapes.

As one can see with the second example in Figure 3.5, the main drawback of the watershed transformation is its marked tendency to over-segmentation compared to what a human would typically expect. This comes from the fact that small irregularities and mere non-convexity on the edges of the shapes can create additional local minima. This is especially troublesome when the data includes noise. The issue can be solved by removing shallow minima from the list of water sources. Of course the definition of a shallow minima depends on the context and on prior knowledge of the data (e.g. estimation of the noise intensity and of the size of the objects to segment). It is usually recommended to only remove minima that are a few pixels deep, otherwise there is a non-negligible risk of sub-segmentation.

When using a pre-implemented watershed algorithm (especially if the function is proprietary, as is the case with Matlab), one of the easiest ways to remove shallow minima is to smooth the distance map beforehand. In Matlab this can be done using the `imhmin` function, which operates a preliminary flooding of the map up to a user-defined level. This fill up all local minima of depth inferior to that level, without modifying the rest of the image.

2.2. Application to our *in silico* data

To apply the watershed transformation to the systems produced by our numerical simulations of the model ATmorpho-3D, three steps were required : transform our data into a 3D black-and-white image, chose an appropriate smoothing value to avoid both over- and sub-segmentation, and above all periodise the transformation over the image edges to account for the periodic boundary conditions in the simulations.

For the first step, we defined a 3D pixelated image where all pixels were initially set to 0 (black) and in which we inserted the adipocytes by putting to 1 (white) all pixels whose centre was located at a distance less than R_i from the adipocyte centre $\boldsymbol{X}_i$. To restrict the numerical cost of this insertion operation, it was not performed on the whole 3D image but only on the portion contained in the bounding box of the spherical adipocyte.

We noted that our *in silico* adipocytes, being non-deformable objects which repulsed each other, tended to leave small holes between themselves even in the core of densely packed structures. Since these holes would have caused the apparition of multiple local minima in the core of a lobule, resulting in over-segmentation by the watershed transformation, we had to remove them. Hole-filling image operations performing poorly on this specific case, we instead chose to "dilate" the radius of the adipocytes by a factor *coeff_dil* before inserting them in the images. After testing, the value of this factor was fixed to 1.5 so as to ensure a full overlap inside dense structures while not allowing the largest adipocytes to expand right to the other side of a neighbouring fibre (i.e. $0.5R_{ad}^{max} < 2R_{fib}$ in our simulations).

On the other hand, because the fibres act as separator between the different lobules, we chose to put all pixels pertaining to a fibre back to 0 even if they were also inside an adipocyte (due to adipocyte-fibre overlap). This ensured that our previous dilatation did not accidentally merge two neighbouring clusters across the fibres separating them and also smoothed the bumpy edges of the clusters, further reducing the risk of over-segmentation.

We initially expected the watershed transform to be likely to over-segment our data precisely because of the bumpy aspect of the cluster edges : many adipocytes located at the periphery of their lobule would be sufficiently protruding for the distance map to have a local minima at their centre. This would lead to the segmentation of individual adipocytes all around the periphery of the lobules. However, using the fibres to trim the border of the lobules solved this issue almost entirely, leaving only slightly jagged edges. We merged the multiple neighbouring local minima produced by these irregularities by applying a preliminary flooding whose level *flood* was determined empirically from our data : we tested several flooding levels on several simulations and retained one which was both small (compared to the characteristic size of the agents) and allowed to greatly reduce the number of local minima (dividing it by 10 on average). This retained value is equal to 0.5 length unit.

We turn now to the issue of the periodic boundary conditions. This setting is commonly used in mathematical modelling because of its numerous conceptual and computational advantages, but rarely occur in the physical world – and, when it does, it is not photographable (e.g. a world map has periodic edges but is not a photography). In consequences, pre-implemented image operations never include an option for periodic boundary conditions. This in turn entails that it is not trivial to implement a proper segmentation protocol for systems produced by a mathematical model with periodic boundary conditions.

In our case, the proper approach would have been to use a periodised version of both the Euclidean distance transform and the watershed transform, so that water reaching the edge of the image would flood the pixels located on the opposite side. However, because Matlab pre-implemented `watershed` function is a black box, we have no way to modify it. We thus took an indirect approach which consists in rep-

licating our system in all three directions before applying watershed, and then automatically merging the various replica of each region (i.e. regions representing the same data replicated through the periodic boundaries). This approach is illustrated in Figure 3.6. It is computationally costly since it involves conducting all the segmentation steps (Euclidean distance transform, pre-flooding and watershed transform) over an image $3^3 = 27$ times larger, in number of pixels, than the initial one. To keep the computational costs at a manageable level, we selected the largest pixel size (and accordingly the smallest number of pixels) guaranteeing a reasonable representation of all the agents.

The complete segmentation pipeline can be summarised as follows :

1. Create a 3D binary image with a user-chosen pixel size (parameter *pixelsize*), containing the system replicated in all three directions so that each agent is represented 27 times (step B of Figure 3.6).

   - Considering a user-chosen *pixelsize* and a simulation domain $\Omega = [-L_x, L_x] \times [-L_y, L_y] \times [-L_z, L_z]$, create an 3D matrix of size $\frac{6L_x}{pixelsize} \times \frac{6L_y}{pixelsize} \times \frac{6L_z}{pixelsize}$ filled with 0.
   - Set to 1 all pixels located at a distance less than *coeff_dil* $\times R_i$ from the point $\boldsymbol{X}_i + 2k_xL_x\boldsymbol{e_x} + 2k_yL_y\boldsymbol{e_y} + 2k_zL_z\boldsymbol{e_z}$, with $k_x, k_y, k_z \in \{-1, 0, 1\}$ (each adipocyte is thus represented 27 times).
   - Set back to 0 all pixels located both inside an adipocyte and a fibre.

2. Compute the Euclidean distance transform using Matlab `bwdist` function.

3. Flood the resulting map up to a level *flood* using Matlab `imhmin` function with threshold $\frac{flood}{pixelsize}$.

4. Apply the watershed transformation using Matlab `watershed` function and number all the resulting regions (step C of Figure 3.6, illustrated on one of our simulations in Figure 3.7.A).

5. Attribute to each replica of an adipocyte the number of the region to which it pertains.

6. Considering that all the replica of an adipocyte ultimately pertain to the same cluster, merge the clusters containing (replica of) the same adipocytes.

   (a) List all the clusters containing at least one "real" adipocyte (not a replica) (step D of Figure 3.6).

   (b) Consider as their size the number of real adipocytes they contain.

   (c) Select the largest cluster and reattribute to it all real adipocytes having a replica inside it. For each reattributed adipocyte, save the position of said replica as its "cluster position". Remove the cluster treated from the list.

   (d) Recompute the number of real adipocytes inside each cluster after the reattribution step.

   (e) Iterate the last two steps until the list is empty.

7. Consider as "isolated cells" all adipocytes pertaining to clusters of less than *MinNumberOfObjects*$= 10$ objects, and consequently attribute them the number 0.

8. List all the (merged) clusters containing at least one real adipocyte and renumber them continuously starting from 1 (step E of Figure 3.6, illustrated on one of our simulations in Figure 3.7.B).

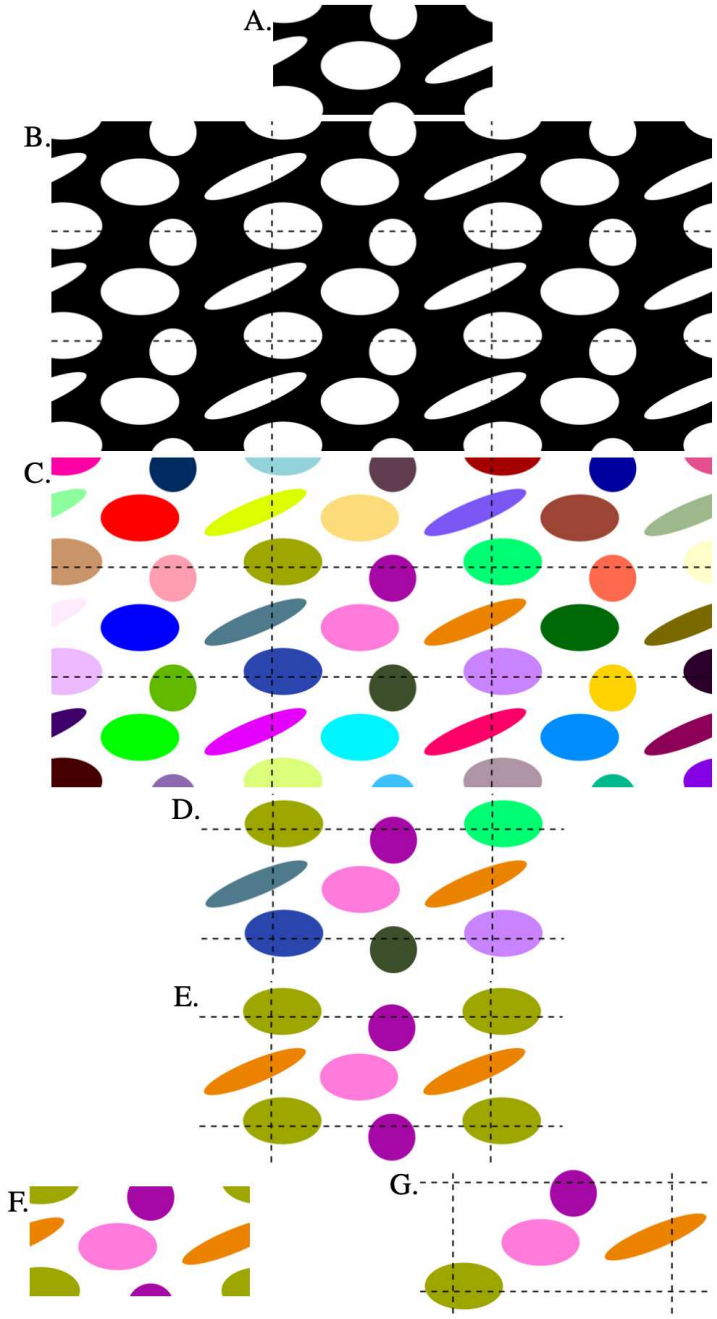


Figure 3.6: Scheme (in 2D) of the protocol developed to segment lobules in a periodic space. **A :** Initial binary image, with stylised shapes in white over a black background. **B :** Replication of this image in the $\pm x$ and $\pm y$ directions. **C :** Regions segmented by watershed transformation. **D :** Cropping to keep only the regions that pass through the central area. **E :** Merging of the various replica of each shape, i.e. regions representing the same data replicated through the periodic boundaries. **F :** Cropping to the central area, resulting in a periodic segmentation map of the initial image. **G :** Alternative representation of the system where each segmented region is represented by its replica having the largest intersection with the central area, instead of being split in several components strictly contained in the initial image.

The segmented system can then be visualised by displaying all adipocytes, coloured according to their cluster number, either at their real position (step G of Figure 3.6, illustrated on a simulation in Figure 3.7.C) or at their "cluster position" identified at step 6(c) of the above algorithm (step H of Figure 3.6,

illustrated on one of our simulations in Figure 3.7.E).

This pipeline, which was originally implemented in Matlab, has been rewritten as an open-source Python module and is freely available on GitHub.

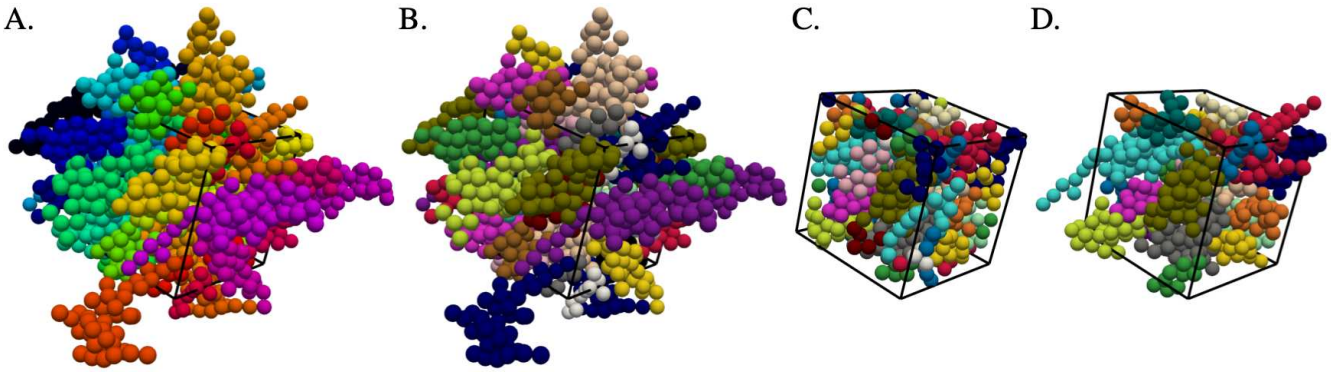


Figure 3.7: Illustration on simulation 2 of our periodised watershed segmentation protocol. **A :** Representation of the system at step D. Each cluster passing through the central area is coloured in a specific shade and the border of that central area are drawn in black lines. Because of the large number of clusters at that step (53), the different shades can be hard to distinguish. **B :** Representation of the system at step E. The various replica of the same cluster are coloured in the same shade and the border of the central area are drawn in black lines. After merging, only 23 distinct clusters are left. **C :** Representation of the system at step F. Each adipocyte is displayed only once, at its real position and coloured according to the cluster to which it pertains. The border of the central area are drawn in black lines. With this representation the simulation domain Ω is preserved, but the shapes of the clusters are hard to picture. **D :** Representation of the system at step G. Each adipocyte is displayed only once, at the position of one of its replica and coloured according to the cluster to which it pertains. The border of the central area drawn in black lines. With this representation the clusters extend over the borders of the simulation domain Ω, but their shapes are clearly visible.

2.3. Application to our *in vivo* data

Compared to the *in silico* case, the technical challenges posed by the *in vivo* data were very different : there was obviously no question of periodic boundary conditions but, on the other hand, the native images could not be directly segmented and required several pre-processing steps, illustrated in Figure 3.8. We described here the complete pipeline applied to these *in vivo* data, starting from their acquisition.

The *in vivo* data are acquired through light-sheet imaging of a mouse subcutaneous adipose tissue depot. The tissue samples were collected by different members of the Restore Institute, then prepared and imaged by Laetitia Pieruccioni (CERT, Restore Institute). The tissue is first embedded in Agarose Low Melting 1% (to prevent it from oscillating during the image acquisition), then transparized by immersion in a mixture of methanol, benzyl alcohol and benzyl benzoate (to homogenise the refractive indices to a final value of 1.55), and finally imaged using a light-sheet microscope "Light sheet Z7" from Zeiss. The sample is placed in a tank containing transparizer medium. The illumination (5x) and detection (5x) lenses are set to correct to a refractive index of 1.55. To image the whole depot with a sufficient resolution, the area to acquire is split into a mosaic of small square tiles (with a theoretical overlap of 10% between neighbouring tiles). Each tile is then imaged at several different depths. Due to technical constraints, the distance between two successive "slices" is greater than the pixel size in the image of one

slice. The resulting 3D image will thus have cuboid, not cubic, pixels. The same values were used for all samples, namely 0.92 µm in the x and y directions and 10 µm in the z direction. The total size of the image depends on the sample. Figure 3.8.A displays an example of a tissue thus imaged and recombined over the whole mosaic, both in 2D (single slice) and 3D view.

Once the images are acquired, several digital processing operations are performed by Mathieu Vigneau (CERT, Restore Institute) to transform them into black-and-white images on which we can apply our watershed segmentation algorithm.

The first step aims to improve image quality and is carried out using Fiji software [137]. It consists in applying, on each tile of the mosaic, a 3D median filter with ellipsoidal kernel (to remove noises) followed by a background subtraction using the rolling ball algorithm (to enhance contrast). The second step is the detection of adipocytes in the image using the `cyto2` machine learning model of the Cellpose software [138, 139]. The model only works on 2D images and have a high computational cost which can be reduced by working on each tile separately, instead of trying to treat the whole mosaic at once. The model is trained on several slices from several tiles, with manual annotation of false positives (detection of inexistent adipocytes) and false negatives (undetected adipocytes). Once trained, it is applied to each slice of each tile of the image to generate a prediction map. The average diameter of the objects to detect is set to 60 pixels, that is ≈ 55 µm.

The third step consists in separating the inside of the lobules from the interlobular spaces and is performed using Fiji software. For each tile, the adipocyte prediction map is binarised to create a black-and-white 3D image representing the adipocytes in white and all the rest, that is the data background, the interlobular spaces and the adipocyte membranes, in black. This image is treated with a hole-filling operation to turn white all black regions smaller than 5000 pixels ($\approx 42\,000$ µm^3), considered to correspond to adipocyte membranes.

Finally, the various tiles are fused to reconstruct the whole mosaic. In theory, this last step should be straightforward. However, because the total acquisition time is very long (approximately 10 hours) and the sample very large, it gradually slips from the clamp holding it : as a result, the tiles are slightly (or badly) misaligned and do not all have the same overlap rate. Accurate automatic reconstruction is achieved using the Stitching plugin [140] of the Fiji software. An example of the resulting image is displayed in Figure 3.8.B in both 2D and 3D view.

Having reached this stage, we can apply the classic watershed segmentation protocol for black-and-white images. Note that all the above operations, and especially adipocyte detection, require images with the highest possible resolution. The watershed algorithm on the other hand performs well on lower-resolution images. Since its computational cost scales with the number of pixels, we choose to downsize the data by a factor 20 in the plane (x, y), so that the 3D pixels are now 18.4 µm $\times 18.4$ µm $\times 10$ µm. We also clean the image to remove all objects smaller than the maximum adipocyte volume ($\approx 1.4 \cdot 10^6$ µm^3), which correspond to isolated cells on the border of the tissue.

We then compute the Euclidean distance map in 3D, using a modified distance function `bwdist` [141] which accounts for the pixels anisotropy (see above), and apply to that map a pre-flooding of 100 µm (that is approximately the diameter of one adipocyte) using the `imhmin` function. Finally, we use the `watershed` function to obtain a segmented map of the lobules (see Figure 3.8.C for an illustration).

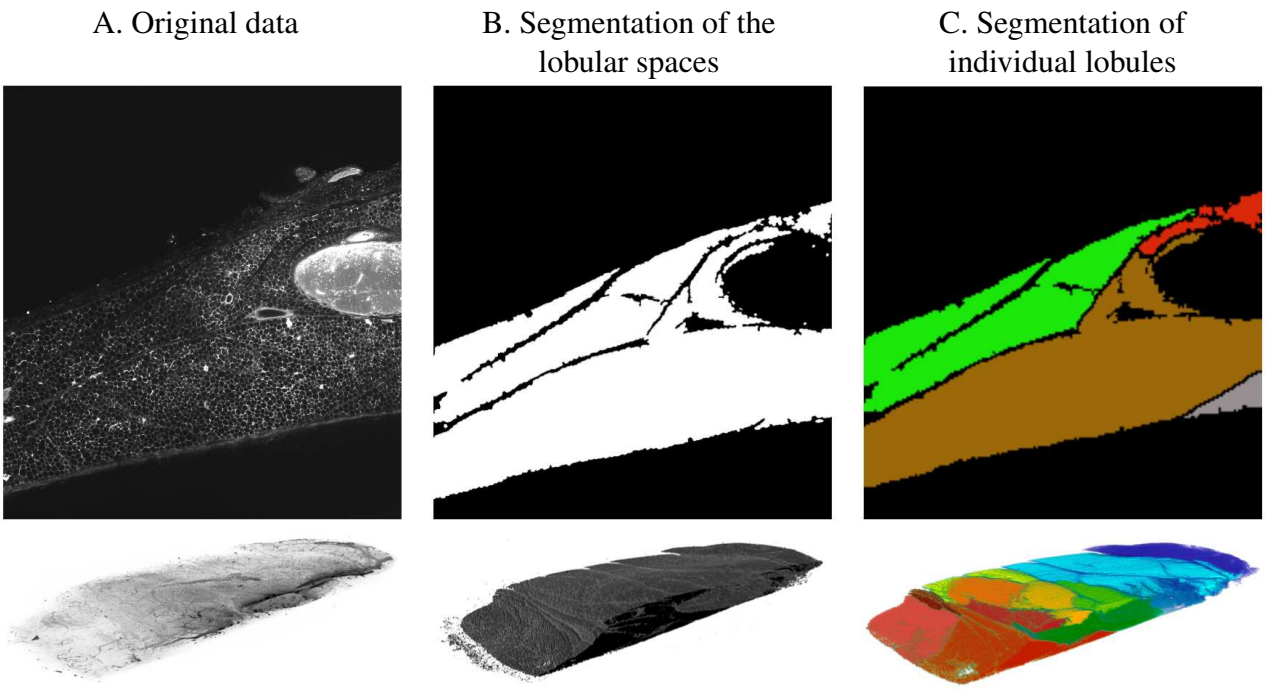


Figure 3.8: Illustration of the complete lobules segmentation pipeline for *in vivo* samples. **A :** Image acquisition using light sheet imaging, by Laetitia Pieruccioni (CERT, Restore Institute). **B :** Identification of the lobule components (adipocytes and their membranes) using several image processing algorithms, by Mathieu Vigneau (CERT, Restore Institute). **C :** Segmentation of the lobules using a watershed algorithm with pre-flooding up to 100 µm, by Pauline Chassonnery.

2.4. Ellipsoidal fit on the lobules

Finally, we will characterise the size and shape of the lobules segmented both from *in vivo* and *in silico* data. To enable a systematic classification and comparison of the lobules, we want this characterisation to be as simply as possible.

To that end, we will not take into account the possible digitations and focus on the core of each lobules. Moreover, though *in vivo* lobules are polylobed, it has been observed that all the lobes of one lobule tend to be arranged in the same direction. The general shape of its core is therefore a good first approximation for describing the shape of a lobule.

A criterion commonly used by biologists is the elongation of the lobules, which can be computed in several manners. Considering that our *in silico* lobules are composed of a relatively small number of cells, most of these manners would be very sensitive to the presence of peripheral digitations or surface irregularities : this is not what we are interested in. In consequence, we choose to fit each lobule with an ellipsoid and compute the elongation (or eccentricity) of that ellipsoid. This is done using Matlab `regionprops3` function, which performs an ellipsoidal fit on each region of the segmented lobule map (see section 2.2 above) and compute many characteristics of the regions and ellipsoids such as their volume, centre-of-mass, orientation, elongation, etc.

One of these characteristics is the region solidity, which measures its deviation from convexity. It is expressed as the ratio between the volume of the region and the volume of its convex hull (i.e. the smallest convex polygon containing the region). This quantifier is useful to distinguish between lobules with similar elongation (see first three panels of Figure 3.9). It especially allows to identify, in our *in silico* data, lobules with loose, non-compact and convoluted shapes which are quite unrealistic (see first panel of Figure 3.9).

We will thus characterise our lobules using three scalar quantifiers : their elongation $E_{lob} \in [0, 1]$ (equal to 0 is the lobule as no preferred direction and to 1 if it is strictly unidirectional), their solidity $S_{lob} \in [0, 1]$ (equal to 1 is the lobule is convex) and their volume V_{lob} (expressed as a percentage of the total number of cells in the system). They are illustrated on a few *in silico* lobules in Figure 3.9.

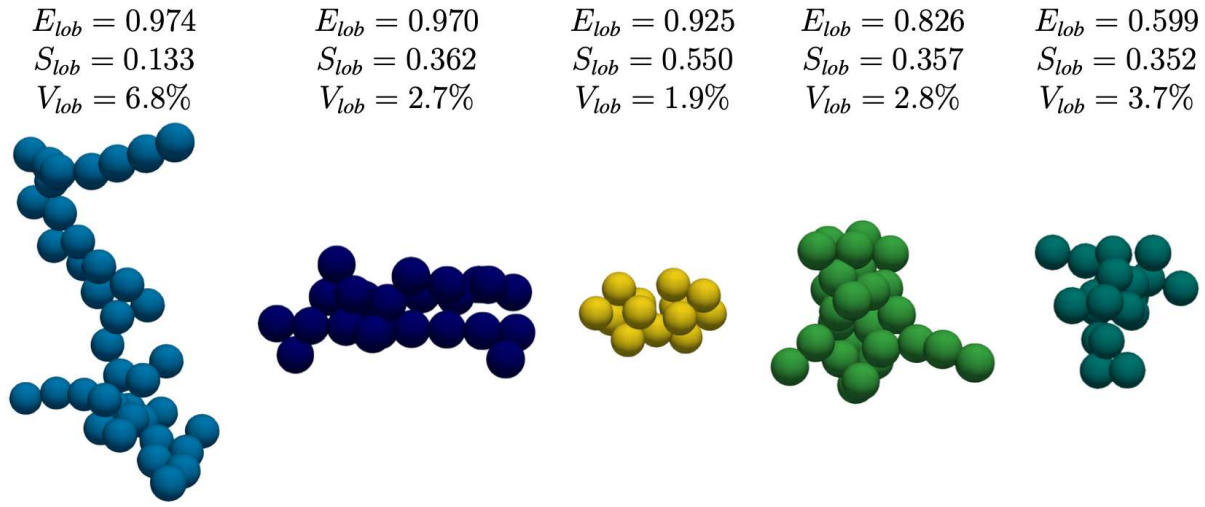


Figure 3.9: Illustration of the quantifiers E_{lob}, S_{lob} and V_{lob} over a few *in silico* lobules generated by the model ATmorpho-3D. Note that these lobules come from different simulations.

3. Characterisation of the fibrous structures

The goal of this section is to define quantifiers allowing to quantitatively describe the local and global organisation of the fibre network observed in the numerical simulations of our models ECMmorpho-3D and ATmorpho-3D. We will first describe how we quantify the local alignment of the fibres by computing the fraction anisotropy Al_k of the fibre directional vectors in a neighbourhood of fibre k, then present a quantifier A_{max} for the fibres organisation on a global scale, obtained from the stereographic projection of all fibres directional vectors.

3.1. Local alignment indicator of the fibres

In healthy adipose tissue, the fibres of the ECM form long, aligned bundles that surround the lobules. To check if and how much this kind of organisation is present in our results, we define a numerical quantifier for the local alignment of the fibres, which we call the alignment indicator. It is computed as follows.

Let $R_{interact}$ denote the maximum interaction distance between two fibres : in our models, it is equal to the maximum crosslinking distance d_{link}^{max}, which is always greater or equal to the repulsion range $2R_{fib}$ (see Table 2.1 in chapter 2). That is, two fibres whose closest points are more than $R_{interact}$ apart can not interact. For any fibre k, we define its interaction neighbourhood $\mathcal{B}_k$ as the set of all fibres with which it can interact, plus itself. The projection matrix on the directional vector of a fibre m is $p_m = \boldsymbol{\omega}_m \otimes \boldsymbol{\omega}_m$, with $\otimes$ the outer product. The mean of the projection matrices of the fibres in $\mathcal{B}_k$ is thus given by :

$$P_k = \frac{1}{N_k} \sum_{m \in \mathcal{B}_k} p_m, \tag{3.1}$$

where N_k denotes the number of fibres in $\mathcal{B}_k$ and is at least equal to 1. The eigenvalues λ_1, λ_2, λ_3 of P_k can be used to quantify the level of alignment of fibre k with its neighbours via the fractional anisotropy formula [142] :

$$Al_k = \sqrt{\frac{3}{2} \times \frac{\lambda_1^2 + \lambda_2^2 + \lambda_3^2 - 3\bar{\lambda}^2}{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}}, \tag{3.2}$$

with $\bar{\lambda} = (\lambda_1 + \lambda_2 + \lambda_3)/3$. Al_k is called the local alignment indicator of fibre k.

In order to demonstrate a few properties of this quantifier, let us first list the following points :

- As a sum of symmetric positive-definite matrices, P_k is itself symmetric positive-definite and its eigenvalues are in $\mathbb{R}^+$.
- As projectors over 1-subspaces, the matrices p_m have traces equal to 1. By linearity of the trace function P_k has itself a trace of 1, so the sum of its eigenvalues is equal to 1 (and $3\bar{\lambda}^2 = 1/3$).
- Combined, the two above properties imply that $\lambda_1, \lambda_2, \lambda_3 \in [0, 1]$.
- Studying the extrema of the function $f : (x, y, z) \mapsto x^2+y^2+z^2$ under the constraints $x, y, z \in [0, 1]$ and $x + y + z = 1$ demonstrates that $\lambda_1^2 + \lambda_2^2 + \lambda_3^2 \in [1/3,\ 1]$.
- P_k has eigenvalues $\{0,\ 0,\ 1\}$ (i.e. it is a projector over a 1-subspace) if and only if all the fibres in $\mathcal{B}_k$ have the same directional vector.
- P_k has eigenvalues $\{1/3,\ 1/3,\ 1/3\}$ if and only if the directional vectors of the fibres in $\mathcal{B}_k$ are equally distributed in the three directions.

In consequence, the quantifier Al_k is in the range $[0, 1]$, is equal to 1 if all the neighbours of fibre k have the same directional vector, and is equal to 0 if on the contrary the fibres in $\mathcal{B}_k$ display no preferred direction.

Figure 3.10 shows the same simulations as Figures 3.3.A and B, but here the fibres have been coloured as a function of their local alignment indicator. This enables us to distinguish areas with higher or lower level of alignment and can highlight the presence of interesting local structures such as the rotating node in the centre of panel A. In the next chapters, we will use this quantifier as a visualisation tool to support the qualitative observation of locally organised states in our numerical simulations.

Note that, if the directional vectors of the fibres in $\mathcal{B}_k$ are randomly distributed according to an uniform law in $\mathbb{S}_2^+$, then $Al_k = 0$ theoretically, but in practice it is rarely the case.

This is because the actual sampling of an uniform law is usually not fully isotropic, especially if the number of elements in the sample is small. Figure 3.11 displays the value of the alignment indicator obtained for various distribution of fibres and various sample sizes : it can be seen that an uniform distribution produces alignment indicators ranging from 0.1 (when the sample size is large) to as much as 0.55 (when the sample size is small), and that there is a large discrepancy between different samples.

These biases are much smaller for non-isotropic distributions : for mainly two- or one-directional distributions, the values computed are nearly the same regardless of the sample size and the discrepancy between different samples is small. For a two-directional distribution (i.e. when the fibre directional vectors describe a disk), the eigenvalues on the mean projection matrix are theoretically $\{0,\ 1/2,\ 1/2\}$, leading to a theoretical alignment indicator of $1/\sqrt{2} \approx 0.707$. This is very close to the value observed in our calibration tests (see the yellow curve on Figure 3.11). Nearly two-directional distributions, where the fibre directional vectors describe a "band" or thick disk, give lower and lower alignment indicator as

the prominence of the third direction (i.e. the band width) increases (see the green curves on Figure 3.11). Likewise, conical distributions, which are mainly one-directional, give an alignment indicator close to 1 which becomes lower and lower as the aperture angle of the cone increases (see the red curves on Figure 3.11).

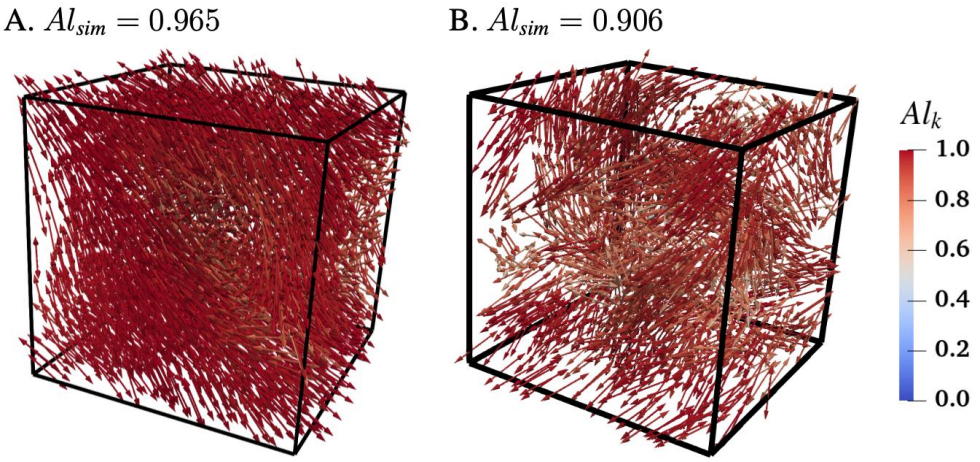


Figure 3.10: Illustration of the local alignment indicator Al_k on two simulations. **A :** 3D view of the fibre network at the end of simulation 1, with edges of the spacial domain Ω drawn in black and each fibre represented by a double-headed arrow coloured according to its local alignment indicator (from blue : $Al_k = 0$, to red : $Al_k = 1$). This system has an mean alignment indicator $Al_{sim} = 0.965$. **B :** Same view for simulation 2 (adipocytes are not represented for a better visibility). This system has an mean alignment indicator $Al_{sim} = 0.906$.

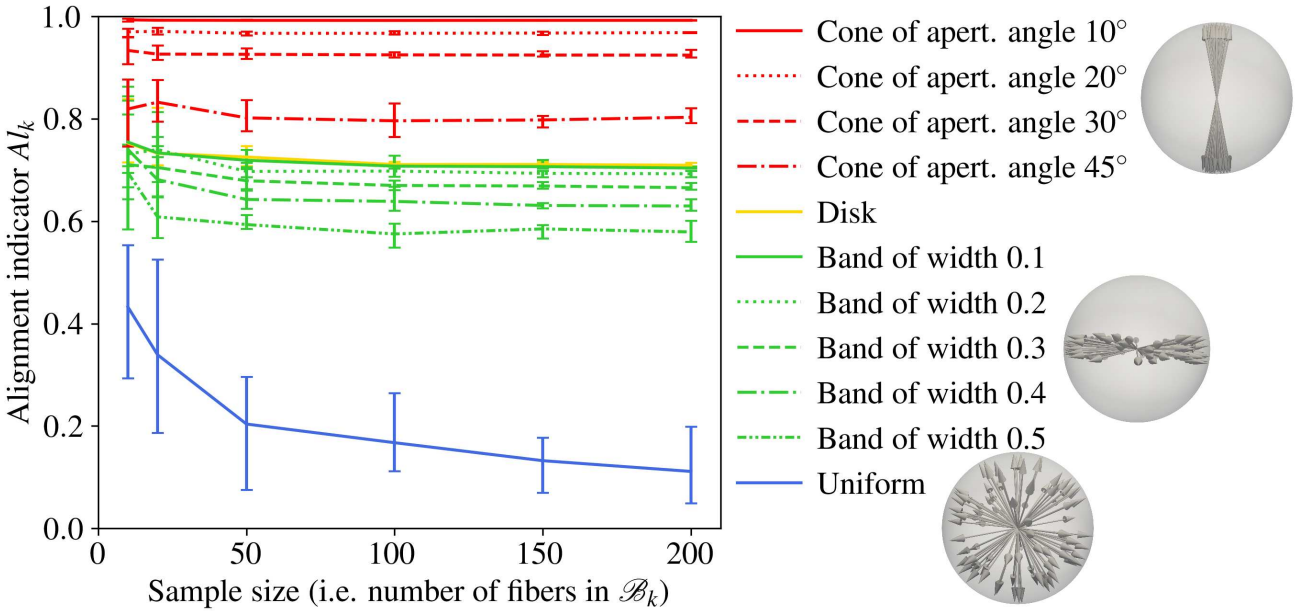


Figure 3.11: Calibration of the local alignment indicator on random sets of directional vectors, for various distribution laws and sample sizes. The displayed values correspond to the average and standard deviation over 10 random draws with the same characteristics.

In the next chapters, we will use the mean alignment indicator of all the fibres in the system, denoted $Al_{sim} = \sum_{k=1}^{N_{fib}} Al_k$, to characterise the state of local organisation of the fibre networks produced by our models : Al_{sim} is high if, in general, the fibres are locally aligned, and low if they are locally disorganised. When needing to account for the stochastic variability of our models, we will use the mean and standard deviation of Al_{sim} over 10 simulations conducted with the same set of parameters, denoted by Al_{mean} and Al_{STD}.

3.2. Stereographic projection of the fibres

The alignment indicator is a good quantifier to describe the local organisation of the fibres, but it can not be used to distinguish large-scale structures. Indeed, in the simulation displayed in Figure 3.10.B the fibres are organised in structures that are very aligned locally (with high values of Al_k) but rotate on the length scale of the whole system. To achieve a characterisation of this large-scale organisation, we designed a global numerical quantifier A_{max} which can be roughly described as the "size" of the point cloud defined by the fibres directional vectors. We describe it in details below.

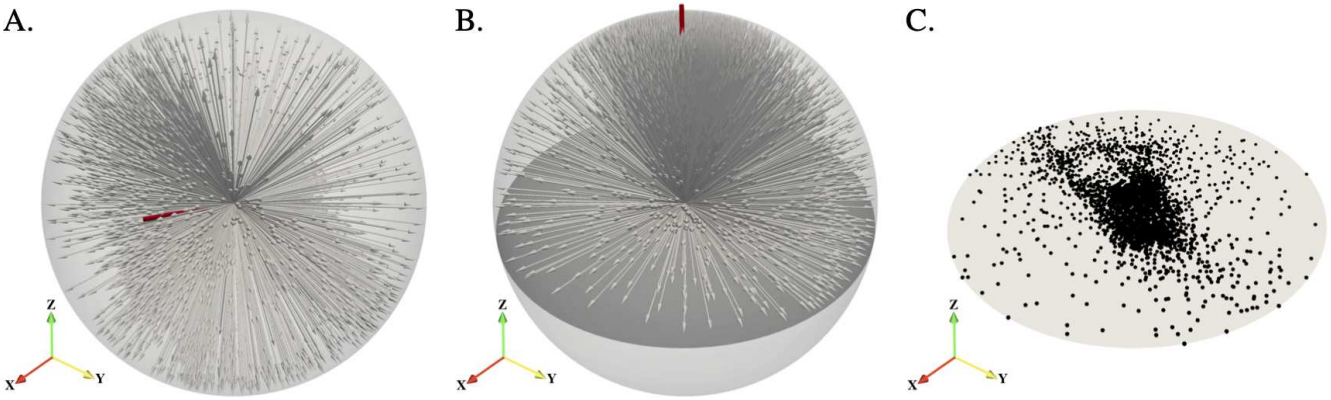


Figure 3.12: Illustration of the stereographic projection. The orientation axis are shown for reference. **A :** Natural distribution of the fibres directional vectors on the unit sphere $\mathbb{S}_2$, with main direction indicated by a red line. **B :** Rotation of the vectors set so that its main direction (in red) now lies along the z-axis. The definition-space of the vectors have been reduced by central symmetry to the "north hemisphere", that is to the subset $\mathbb{S}_2^+$ in the new rotated coordinates system. The equatorial plane is shown in dark-grey. **C :** Projection of the vectors onto the equatorial plane, shown in 3D perspective.

Disregarding the spatial position of the fibres (that is, the position of their centres), we can observe the distribution of their directional vectors in the unit sphere $\mathbb{S}_2$ as shown in Figure 3.12.A. However, this is not very practical as it implies a 3D visualisation which, in the present case, can be avoided. Indeed, there is multiple ways to transform a sphere (which is ultimately a 2D surface) into a plane. We choose the stereographic projection, a transformation which locally preserves shapes (though not distances nor areas). We also use the fact that the fibres are non-oriented to restrict the definition-space of the directional vectors to a half-sphere (by applying a central symmetry to vectors located in the unretained half), so that the image of the projection will be a disk instead of the whole plane. The half in question is not chosen at random : to reduce the risk of interesting structures being located on the edge of the half-sphere, and so cut into two parts when symmetrising, we retain the half-sphere whose pole is the main direction of the set of vectors.

More precisely, we define the main direction of the fibres as the eigenvector associated to the largest eigenvalue of their total projection matrix :

$$P_{tot} = \frac{1}{N_{fib}} \sum_{k=1}^{N_{fib}} \boldsymbol{\omega}_k \otimes \boldsymbol{\omega}_k. \tag{3.3}$$

In the unlikely case that there is two or three equally represented directions (associated to equal eigenvalues), one of them is randomly selected. This main direction is represented on Figure 3.12.A by a red line.

We then rotate the set of directional vectors so that its main direction lies on the z-axis and apply to the (rotated) vectors located in the "south hemisphere" a central symmetry, so that the definition-space is now restricted to the "north hemisphere" of the unit sphere. Finally, we apply to our set of (rotated and symmetrized) directional vectors the projection :

$$p_{stereo} \; : \; \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} \dfrac{x}{1+z} \\ \dfrac{y}{1+z} \end{pmatrix}, \tag{3.4}$$

that is a stereographic projection from the south pole onto the equatorial plane (see Figure 3.12.C).

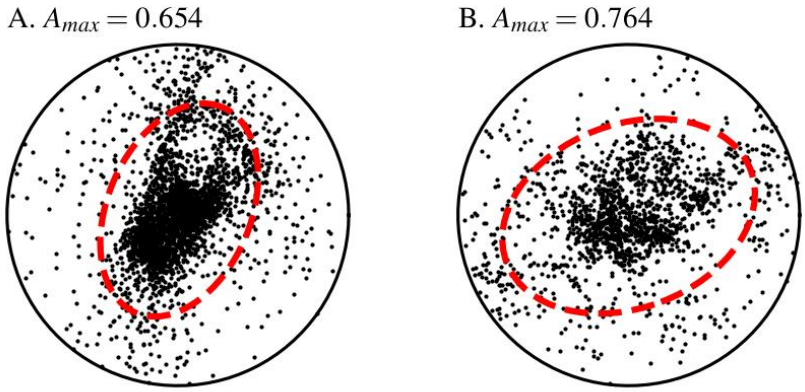


Figure 3.13: Illustration of the global quantifier A_{max} on two simulations. **A :** Stereographic projection of the fibres directional vectors at the end of simulation 1, with the unit circle drawn in black solid line and the covariance ellipse of the point cloud in red dashed line. This covariance ellipse has a semi-major axis length $A_{max} = 0.654$. **B :** Stereographic projection of the fibres directional vectors at the end of simulation 2, with the unit circle drawn in black solid line and the covariance ellipse of the point cloud in red dashed line. This covariance ellipse has a semi-major axis length $A_{max} = 0.764$.

Figure 3.13 shows the stereographic projection of the two simulations displayed in Figures 3.3.A and B. As one can observe, in both cases the dots are not uniformly distributed but densely packed at the centre of the unit circle, indicating the existence of a main preferential direction in the two networks. However, not all fibres have a directional vector close to this main direction : a non negligible number of dots are distributed all around the circle, meaning that all possible directions are represented in these networks. Furthermore, a closer look at Figure 3.13.A reveals the presence of a "circular branch" in the top-right part of the point cloud and allows to identify the locally twisting structure that can be observed at the centre of the network in Figure 3.3.A : in this part of the system, nearby fibres have similar but

gradually shifting directional vectors such that, on the scale of the whole structure, the fibres directional vectors describe a circle (in the domain $\mathbb{S}_2^+$).

Thus, this representation enables us to quickly grasp the distribution of the fibres directional vectors around one or more poles. It must be noted that proximity on the stereographic projection indicates similar directional vectors, but not necessarily spatial proximity. Nonetheless, we can gain insights into the overall architecture of the network by characterising the distribution of the fibre directional vectors via this stereographic point cloud. We found that, whether or not adipocytes were present in the system, the stereographic point cloud was usually quite concentrated around the centre of the unit circle with limited sub-clustering. It is thus relevant to fit it with a single ellipse. We used the 3σ-confidence ellipse, that is the ellipse which would enclose 98.9% of the points if the data were randomly distributed according to a gaussian law. This ellipse can be characterised by a number of variables : its semi-major and semi-minor axis lengths, its area, its eccentricity, etc. In practice, we will mostly use the semi-major axis length A_{max}, which gives a good idea of the dispersion of the point cloud and so of the global organisation of the fibre network : if A_{max} is small the fibre directional vectors are concentrated around a main direction and the network exhibits a strong global organisation, whereas if A_{max} is large the directional vectors display at least a two-directional, and possibly a three-directional distribution with no global organisation of the network.

Chapter 4

Study of dynamically cross-linked fibre networks : how matrix connectivity drives the emergence of ordered structures

In this chapter, we investigate the behaviour of the dynamically connected fibres networks produced by the model ECMmorpho-3D presented in chapter 2. We observe the emergence of various types of 3D structures that we characterise using the tools developed in chapter 3. We show that these structures can be explained by the network connectivity, which is an emerging variable of the system, and also partly by the link life-expectancy, which is a parameter of the model. We also show that the structures emerge homogeneous in space and on a timescale controlled by the network remodelling characteristics.

This work gave rise to a scientific paper titled "Fiber crosslinking drives the emergence of order in a 3D dynamical network model", written with Jenny Paupert, Anne Lorsignol, Childérick Sévérac, Marielle Ousset, Pierre Degond, Louis Casteilla and Diane Peurichard [143], which has been accepted under minor revision at the Royal Society Open-Science journal.

The present text is closely based on the accepted version.

Summary of the chapter

1. Introduction and motivation

As explained in chapter 1, the adequate architecture of any organ is mandatory for their efficient physiological function and any change is associated with function impairment and putative developing dysfunctions and diseases [69,144,145]. The tissue architecture depends mainly on the mechanical forces exerted by the Extra-Cellular Matrix (ECM) which provides spatial information for cells and largely participate to mechanical constraints [63] (see section 4.1 in chapter 1). For example, alignment as well as accumulation of ECM observed in fibrosis lead to a loss of function [69, 144]. Because the global architecture of fibre networks seems to be fundamental for controlling tissue functions, modelling the process of ECM structure emergence will greatly improve our understanding of tissue biology and plasticity in physiological or pathological conditions.

Despite the great variability of proteins that make up the ECM (macromolecules such as collagen, glycoproteins etc), it can be seen as a dynamic physical network of fibres interconnected by molecular bonds, i.e. crosslinks, generating a connected and elastic environment for the surrounding cells [69].

In the present work, we test the hypothesis that this macrostructures could spontaneously emerge without appealing to contact guidance or external mechanical challenges, as a result of simple mechanical interactions between the fibre elements composing the ECM network.

We assess this hypothesis using the three-dimensional Individual-Based Model (IBM) for the self-organisation of ECM that we presented in chapter 2 (section 2). ECM fibres are discretised into non-stretching and non-flexible units with the ability to spontaneously link to and unlink from their close neighbours. This dynamical crosslinking mechanism allows us to model both the overall temporal plasticity of the network and the complex physical properties of biological fibres such as elongation, bending, branching and growth, thus compensating our minimalistic description of the fibre units. The relevance of such discretisation was previously validated in the frame of adipose tissue morphogenesis and regeneration in 2D [17, 37].

Through computational simulations and exhaustive parametric analysis, we demonstrate that organised macrostructures can spontaneously emerge without external guidance. Overall, this study provides a comprehensive view on the role of ECM connectivity on tissue architecture emergence.

- The model first reveals that dynamic remodelling is essential for the generation of ordered ECM structures.

- We surprisingly find that, for dynamical networks, tissue architecture at equilibrium is simply controlled by the number of crosslinks per fibre in the network, an emerging variable not directly linked to the model parameters. This simple emerging variable therefore becomes an important putative target to control and predict the development of the architecture of biological tissues. Because of its simplicity, this variable is amenable to experimental measurements and could represent a major target for the development of therapeutic drugs to induce tissue recovery after injury, prevent tissue degradation during ageing, or help in the design of engineering collagen scaffolds for tissue regeneration.

- A deep exploration of the model parameters reveals that this emerging variable, and therefore the global organisation abilities of tissues, depend on a complex interplay between the model parameters related to the crosslinks, i.e their remodelling speed and their linked fibre fraction. These results rationalise how even subtle changes in ECM dynamical crosslinking can drive tissue reorganisation and suggest that the development of biological crosslinkers to control ECM connectivity as a target for tissue reconstruction must carefully account for different parameters such as tissue remodelling activities.

2. Description of the experiments

To test our hypothesis, we perform numerical simulations of our model ECMmorpho-3D, that we briefly recap here (see chapter 2, section 2 for a full description of the model components).

The ECM fibres are discretised into N_{fib} rigid spherocylindrical fibre units of fixed uniform length L_{fib} and radius R_{fib}. Overlapping fibres (linked or not) exert on each other a steric repulsion force of maximal intensity E_{rep}. Moreover, fibres closer than d_{link}^{max} can crosslink with each other according to a Poisson process of frequency ν_{link}. Existing crosslinks can break spontaneously according to a Poisson process of frequency ν_{unlink}. As a result, the linked fibre fraction $\chi_{link} = \dfrac{\nu_{link}}{\nu_{link} + \nu_{unlink}}$ represents the equilibrium fraction of linked fibres among the pairs of neighbouring fibres. Each crosslink is represented by an elastic spring of stiffness κ_{rest} and unloaded length d_{link}^{eq}. Linked fibres are also subjected a nematic alignment torque proportional to the flexural modulus α_{align}. Finally, all fibres are subjected to a friction force (with dynamic viscosity μ_{fib}) large enough to nullify their acceleration.

All this leads to the differential system (2.14), which we repeat below :

$$\begin{cases} \dfrac{d\boldsymbol{Y}_k}{dt}(t) = \dfrac{1}{\mu_{fib} L_{fib}} \displaystyle\sum_{\substack{m=1 \\ m \neq k}}^{N_{fib}} \left(\boldsymbol{F}_{k,m}^{rep}(t) + p_{k,m}(t) \boldsymbol{F}_{k,m}^{link}(t) \right) \\ \dfrac{d\boldsymbol{\omega}_k}{dt}(t) = \dfrac{1}{\mu_{fib} L_{fib}^3} \displaystyle\sum_{\substack{m=1 \\ m \neq k}}^{N_{fib}} \left(\boldsymbol{T}_{k,m}^{rep}(t) + p_{k,m}(t) (\boldsymbol{T}_{k,m}^{link}(t) + \boldsymbol{T}_{k,m}^{align}(t)) \right) \wedge \boldsymbol{\omega}_k(t) \end{cases} \quad \forall k \in [\![1, N_{fib}]\!]. \quad (2.14)$$

This system and all the stochastic processes included in our model are numerically solved using the algorithm described in chapter 2, section 4. The spatial domain of the simulations is $\Omega = [-L_x, L_x] \times [-L_y, L_y] \times [-L_z, L_z]$ with periodic boundary conditions. At initialisation, all N_{fib} fibres are randomly inseminated according to a uniform law for both their position in Ω and their directional vector in $\mathbb{S}_2^+$. The latter is obtained by generating the polar and azimuth angles (θ, ϕ) of each vector according to the following law :

$$\mathbb{P}(\theta = x, \varphi = y) = \frac{\sin(y)}{2\pi} \quad \forall x, y \in [0, \pi]. \quad (4.1)$$

The physical scaling of all the parameters of the simulations are described in Table 4.1. A few points may be noted :

- the maximum distance for link creation d_{link}^{max} and the link unloaded length d_{link}^{eq} are both equal to their minimum acceptable value $2R_{fib}$;
- the domain is relatively small compared to the agents (its side length is approximately 4 times the size of a fibre along its main axis);
- the fibre aspect-ratio $\dfrac{L_{fib}}{2R_{fib}} = 6$ is quite small compared to the values used in other models of the ECM, which usually varies between 250 and 10^4 [25, 34, 35]. This compensates the fact that these models directly account for fibre bending and/or fibre elongation, as one elastic fibre from such models amounts to a large number of end-to-end crosslinked rigid fibre units from our model.

We denote by ϕ_{fib} the fibre density of the network, that is the ratio between the total volume of fibres (without overlapping) and the volume of the spatial domain :

$$\phi_{fib} = \frac{N_{fib} V_{fib}}{|\Omega|} = \frac{N_{fib} \left(\pi R_{fib}^2 L_{fib} + (4/3) \pi R_{fib}^3 \right)}{8 L_x L_y L_z}, \quad (4.2)$$

Parameter	Value	Dimension	Description
Agents characteristics			
N_{fib}	$[1500, 3000]$	N/A	Number of fibres
L_{fib}	6	L	Fibre length
R_{fib}	0.5	L	Fibre radius
Mechanical interactions			
E_{rep}	13.26	$M.L^{-1}.T^{-2}$	Magnitude of the repulsion force
κ_{rest}	5.0	$M.L^{-2}$	Magnitude of the elastic restoring force
α_{align}	2.0	$M.L^{2}.T^{-2}$	Magnitude of the alignment torque
d_{link}^{max}	1.0	L	Perception distance for link creation
d_{link}^{eq}	1.0	L	Link equilibrium length
μ_{fib}	1.0	$M.L^{-1}.T^{-1}$	Dynamic viscosity of the fibres
Link processes			
ν_{link}	$[0, 10]$	T^{-1}	Network remodelling speed
χ_{link}	$[0.1, 0.9]$	N/A	Equilibrium linked fibre fraction
Numerical parameters			
$L_x = L_y = L_z$	15	L	Side half-length of the cuboid domain
T_{final}	5.10^4	T	Total time of simulation

Table 4.1: Model and numerical parameters used for the simulations.

where V_{fib} is the volume of one fibre and $|\Omega|$ the volume of the spatial domain.

The quantity ϕ_{fib} can be compared to the packing density, that is the maximal fraction of the domain that can be occupied by densely packed, non-overlapping fibres. When considering an ordered packing, the optimal configuration for spherocylinders is to have parallel cylindrical stems and alternated (quincunx) half-spheres, as represented in Figure 4.1.A in a 2D perspective. The space is thus divided into alternating slices that are either :

- occupied by cylinders, with a packing density equal to $\dfrac{\pi}{2\sqrt{3}}$ (same as circles in 2D) and a height L_{fib} (in green on the scheme);
- occupied by half-spheres, with a packing density equal to $\dfrac{\pi}{3\sqrt{2}}$ (same as whole spheres) and a height $\sqrt{3}\,R_{fib}$ (in blue on the scheme).

The result is an overall average packing density equal to :

$$\phi_{order} = \left(\frac{\pi}{2\sqrt{3}} L_{fib} + \frac{\pi}{3\sqrt{2}} \sqrt{3}\, R_{fib} \right) \times \frac{1}{L_{fib} + \sqrt{3}\, R_{fib}} \tag{4.3}$$

which in our case gives $\phi_{order} = 0.89$.

However, since in our model fibres are randomly inseminated, the situation is closer to what is called a random or amorphous packing : particles are generated randomly with a volume exclusion constraint until it is no longer possible to inseminate another one. Examples of such configurations can be found in

Figure 4.1.B. The average density reached at the end of the insemination process is called the maximal random packing density ϕ_{random}. For spherocylinders with an aspect-ratio of 6, the literature gives us $\phi_{random} \approx 0.4$ [146].

Thus, we may say that a system is "sparse" if its fibre density is below ϕ_{random}, "dense" if it is between ϕ_{random} and ϕ_{order}, and "hyperdense" if it is above ϕ_{order}. For this set of experiments, we fix the size of the domain and test two types of systems : dense systems containing $N_{fib} = 3000$ fibres ($\phi_{fib} = 0.58$) and sparse systems with $N_{fib} = 1500$ fibres ($\phi_{fib} = 0.29$).

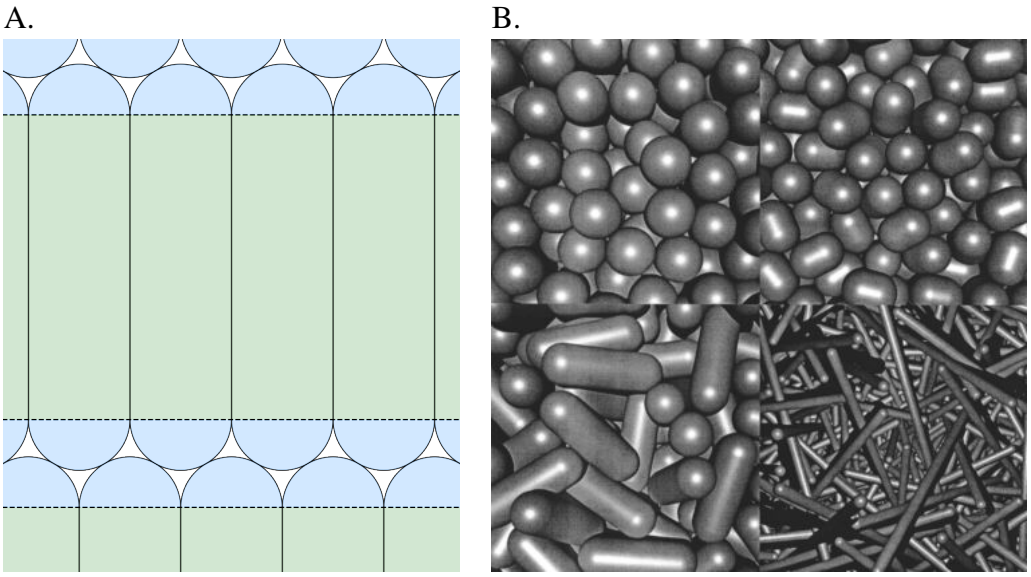


Figure 4.1: **A :** Scheme in 2D of the optimal packing for spherocylinders. **B (from [146], Fig. 1) :** Images (ray tracings) of tightly packed isotropic spherocylinders for several aspect ratios. Aspect ratios of (clockwise from top left) $\alpha = 0$ (spheres), $\alpha = 0.4$, $\alpha = 40.0$ and $\alpha = 2.0$. For $\alpha = 0.4$, the highest packing density $\phi = 0.70$ is achieved (see [146], Fig. 2). The packing for $\alpha = 2.0$ is already to the right of the density maximum, and has a density $\phi = 0.616$, which is close to that of the random sphere packings.

For each of the three types of mechanical interactions in the system, we define the "characteristic interaction time" as the time needed, for two isolated fibres submitted only to this interaction and initially positioned in the most unfavourable configuration, to reach 99% of the (asymptotic) equilibrium configuration.

- T_{rep} is the time needed for two fully overlapped fibres ($\boldsymbol{Y}_1 = \boldsymbol{Y}_2$ and $\boldsymbol{\omega}_1 \ /\!/ \ \boldsymbol{\omega}_2$) to move apart by 99% of their equilibrium distance $2R_{fib}$ (i.e. $||\boldsymbol{Y}_1 - \boldsymbol{Y}_2|| = 0.99 \times 2R_{fib}$) due to the repulsion force alone.

- T_{rest} is the time needed for two fibres that are initially fully overlapping and crosslinked at their centre to move apart by 99% of their equilibrium distance d^{eq}_{link} due to the elastic restoring force alone.

- T_{align} is the time needed for two perpendicularly intersecting fibres ($\boldsymbol{Y}_1 = \boldsymbol{Y}_2$ and $\boldsymbol{\omega}_1 \perp \boldsymbol{\omega}_2$) crosslinked at their centre to reach a relative angle $\arccos(\boldsymbol{\omega}_1 \cdot \boldsymbol{\omega}_2) = 0.9°$ due to the alignment torque alone.

Explicit computation leads to the following formula (numerical values correspond to the parameters given in Table 4.1) :

$$\begin{cases} T_{rep} = \dfrac{27\mu_{fib}L_{fib}}{4\sqrt{2}\ R_{fib}\ E_{rep}} = 4.32\,U_t, \\ \\ T_{rest} = \ln(100)\dfrac{\mu_{fib}L_{fib}}{\kappa_{rest}} = 5.53\,U_t, \\ \\ T_{align} = 4.8\dfrac{\mu_{fib}L_{fib}^3}{\alpha_{align}} = 523\,U_t. \end{cases} \tag{4.4}$$

It may be noted that the alignment interaction is much slower than the repulsive and elastic restoring forces.

3. Results

3.1. Characterisation and quantitative assessment of various 3D architectures

In Figure 4.2, we present an overview of the various structures that can be obtained with the model ECMmorpho-3D by playing on the parameters in the ranges indicated in Table 4.1 (see later sections for the influence of the parameters). We use the data processing framework described in chapter 3 : 3D visualisation using Paraview (see chapter 3, section 1) with fibres coloured according to their local alignment indicator Al_k (see chapter 3, section 3.1) and stereographic projection of the fibres directional vectors (see chapter 3, section 3.2). We recall that Al_{sim} denotes the average alignment indicator of all the fibres and A_{max} the semi-major axis length of the stereographic projection covariance ellipse.

The structures obtained at equilibrium range from highly organised systems with high alignment indicator and one clear main direction, referred to as "aligned states" (Figure 4.2.A), to disordered systems with a low alignment indicator and no preferential direction, referred to as "unorganised states" (Figure 4.2.C). An intermediate state that emerges consists in curved patterns with a median, locally heterogeneous alignment indicator and a wide range of directional vectors describing a plane (see Figure 4.2.B). We will refer to these last structures as "curved states". These three different types of stable structures display very different distribution of fibres directional vectors, as observed in the stereographic projections shown below their respective 3D visualisations. As expected, the fibres directional vectors are very concentrated around the centre of the disk for "aligned states" (Figure 4.2.D), homogeneously distributed on the disk for "unorganised states" (Figure 4.2.F), and distributed along a preferential axis (corresponding to the plane in which the curve-patterns unfold) for "curved states", with complete depletion in the direction perpendicular to this axis (Figure 4.2.E).

The conclusion is that the different states of our networks can be fully characterised by combining a quantifier for local structuring such as Al_{sim} and a quantifier for global organisation such as A_{max}. We considered a system to be locally aligned if the local distribution of its fibres directional vectors was mainly unidirectional, that is Al_{sim} above 0.7 (see calibration of the quantifier Al_k in chapter 3, section 3). At the same time, we considered that a system was globally aligned, in the sense that it displayed a global main direction, if its stereographic projection covariance ellipse had a semi-major axis smaller than 0.45 (implying that the point cloud covers less than 20% of the whole projection disk). We therefore classified the simulations outcomes into three different states (unorganised, curved and aligned) using Table 4.2. We ran a total of 1080 numerical simulations, exploring various values of the parameters ν_{link}, χ_{link} and N_{fib} in the broad ranges indicated in Table 4.1, and counted among their outcomes : 180 unorganised

states (all occurring in non dynamical systems, i.e. $\nu_{link} = 0$), 661 curved states and 239 aligned states (among which only 12 occurred in sparse systems).

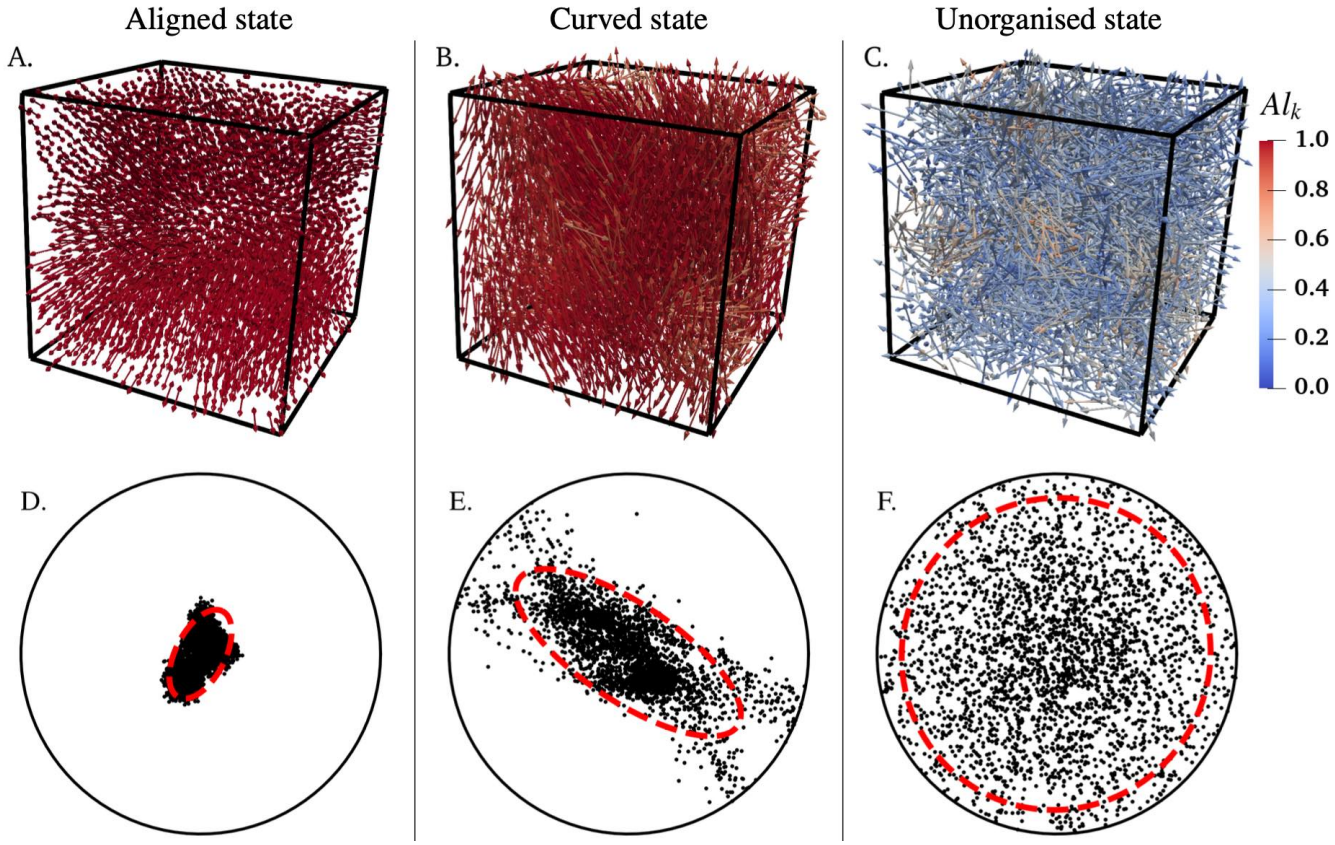


Figure 4.2: Illustration of the three types of structure obtained at final time. **Top row :** 3D visualisation with each fibre coloured according to its local alignment indicator, from blue ($Al_k = 0$) to red ($Al_k = 1$). **Bottom row :** Stereographic projection of the fibres directional vectors (black dots) and its covariance ellipse (red dashed line).

		A_{max}	
		$\leqslant 0.45$	> 0.45
Al_{sim}	$\geqslant 0.7$	**Aligned state :** alignment both local and global.	**Curved state :** alignment local but not global.
	< 0.7	(alignment global but not local)	**Unorganised state :** no alignment, either local or global.

Table 4.2: Classification of the simulations outcomes into different states based on the local quantifier Al_{sim} and the global quantifier A_{max}. The case $\{Al_{sim} < 0.7 \ \& \ A_{max} \leqslant 0.45\}$ never occurs in our simulations and is thus unnamed.

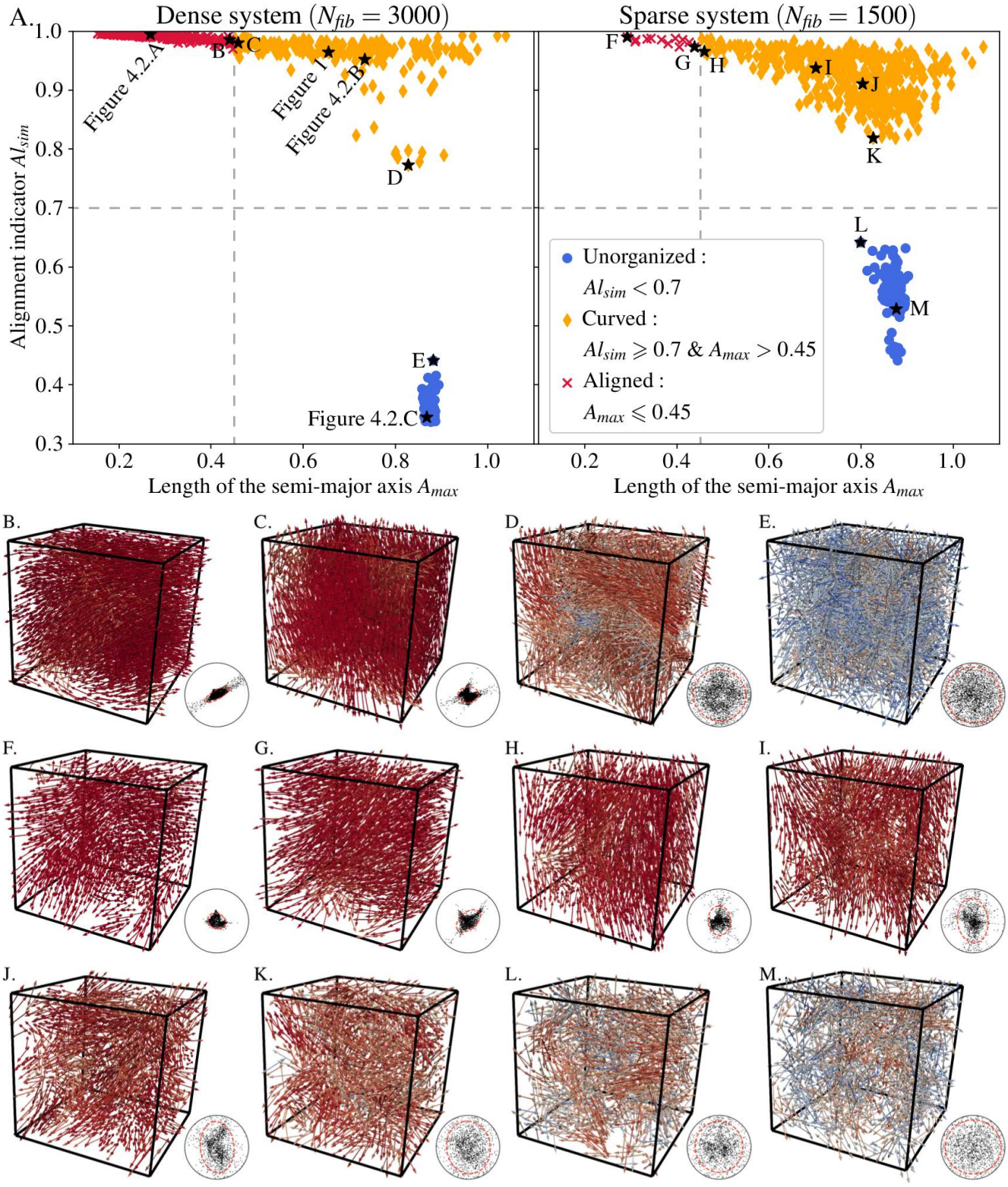


Figure 4.3: **A :** Alignment indicator Al_{sim} versus semi-major axis length of the covariance ellipse of the stereographic projection A_{max}, for each simulation of either dense systems ($N_{fib} = 3000$, on the left) or sparse systems ($N_{fib} = 1500$, on the right). The limits between each class of structures are drawn in dashed lines. **B–M :** Equilibrium state (3D view and stereographic projection) of a few simulations illustrating either typical or borderline cases. Their position on the corresponding diagram, as well as that of the simulations previously displayed in Figure 4.2, are indicated with a black star.

The top panels of Figure 4.3 compare the values of quantifiers Al_{sim} and A_{max} when the simulation has reached equilibrium. The simulations already displayed in Figure 4.2 are indicated with a black star : it can be seen that they are quite typical of their respective states in terms of the quantifiers Al_{sim} and A_{max}. Twelve other simulation outcomes are singled out with black stars and illustrated with a 3D view and stereographic projection in the panels B to M below : the first four (panels B–E) correspond to borderline cases of dense systems, the other eight (panels F–M) to both typical and borderline cases of sparse systems.

We first observe that the unorganised states form a small, compact group of points with large semi-major axis length while the aligned states make a long thin group with very high alignment indicator. On the other hand, the curved states form a scattered cloud of points with a broad range of values for both the semi-major axis length and the alignment indicator.

It can be seen that the transition between unorganised and curved states is very sharp : notice the gap between the two groups of points in both panels. Indeed, no simulation displays an alignment indicator at equilibrium between 0.65 and 0.77, and there is a marked difference between the least aligned of the curved states (illustrated in Figure 4.3.D for dense systems and Figure 4.3.K for sparse systems) and the most aligned of the unorganised states (illustrated in Figure 4.3.E and Figure 4.3.L for dense and sparse systems respectively). This strengthens our choice of 0.7 for the threshold value between unorganised and curved states and makes the partition objective since any value in the range $[0.65, 0.77]$ would give the same result. On the contrary, the transition from curved to aligned states is not a clear switch but a continuum of structures that can be illustrated by the two borderline cases in Figure 4.3.B and C for dense systems and Figure 4.3.G and H for sparse systems. Thus, one must be aware that the partition between curved and aligned states is partly arbitrary and depends on the choice for the threshold.

Finally, we notice that aligned states are quite common among dense systems (representing 42% of the simulations outcomes) but are very rare among sparse systems (only 2.2% of the simulations outcomes), which tend to favour curved states.

The following sections are devoted to the parametric analysis of our model, in order to identify the role of the parameters in the shape of the emerging structures.

3.2. ECM architecture emergence is driven by a complex interplay between the model parameters

Following the results of [17], which served as a basis for our own model, we studied the impact of the network remodelling dynamics on the different tissue architectures (aligned/curved/unorganised) obtained with our model. In Figure 4.4, we show the distribution of the simulations outcomes at equilibrium, depending on the values of the network remodelling speed ν_{link}, the linked fibre fraction χ_{link} and the number of fibres N_{fib}.

The first observation is that systems with non-dynamical crosslinks ($\nu_{link} = \nu_{unlink} = 0$, far left column of panels A and B) are systematically unorganised, independently of the equilibrium linked fibre fraction or the fibre density. This major result shows that non-dynamical networks are locked in mechanically constrained configurations, preventing the system from reorganising on a local or global scale. On the contrary, dynamical networks ($\nu_{link} > 0$) never equilibrate in an unorganised state : their plasticity (i.e. their ability to rearrange their connections) favours the formation of more organised states than non-dynamical networks. This shows that the discontinuous phase transition between unorganised and curved equilibrium states, revealed in Figure 4.3, is controlled by ν_{link}.

Note that the number of neighbours of a fibre is very stable throughout all our simulations. It ranges from 20 to 25 in dense systems and from 10 to 15 in sparse systems. Non-dynamical networks display mean alignment indicators between 0.3 and 0.45 for dense systems and between 0.4 and 0.65 for sparse

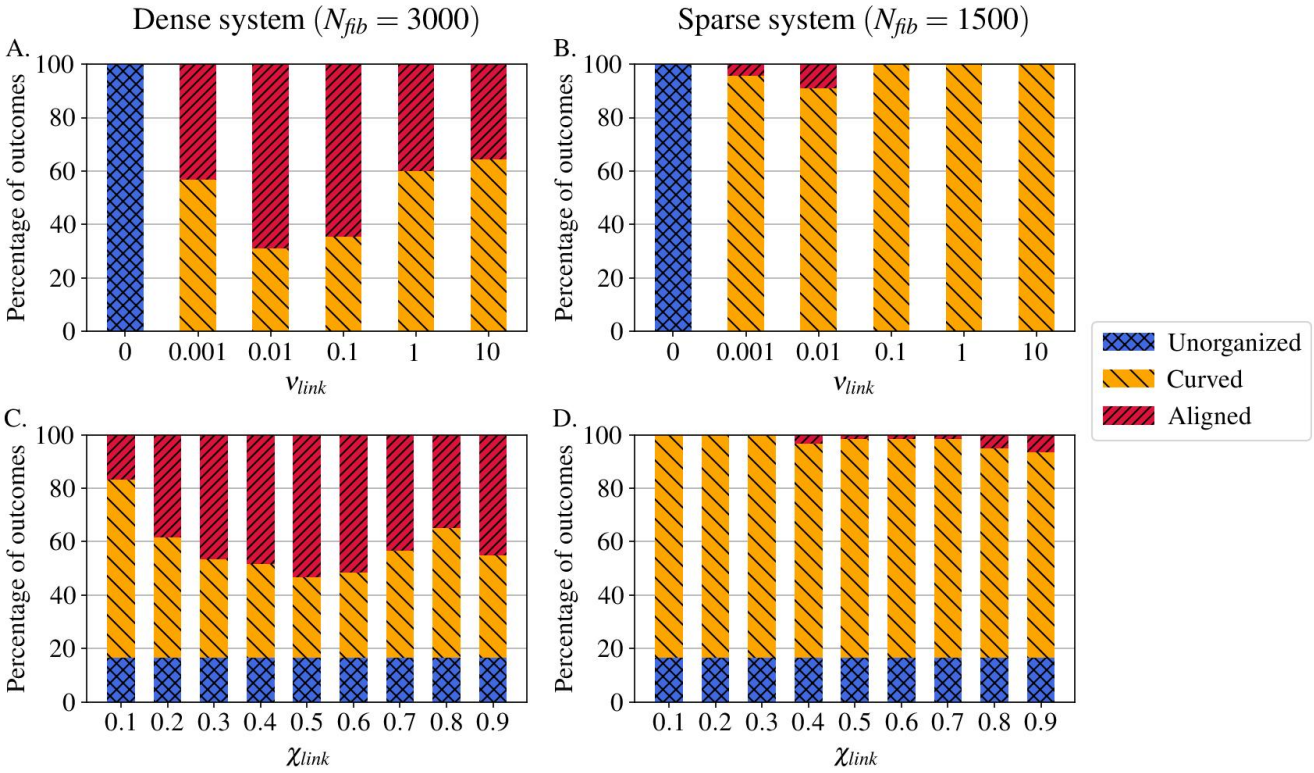


Figure 4.4: Distribution of the outcomes of all the simulations between the different categories (unorganised, curved and aligned). To account for the stochastic components of our model, we ran 10 simulations for each set of parameters. We tested 2 values of N_{fib}, 6 values of ν_{link} and 9 values of χ_{link}, for a total of 1080 simulations. **A and B :** Each bar gives the percentage of each category among the outcomes of the 90 simulations conducted with a given value of ν_{link} (on the x-axis) and N_{fib} (dense for panel **A** and sparse for panel **B**). **C and D :** Each bar gives the percentage of each category among the outcomes of the 60 simulations conducted with a given value of χ_{link} (on the x-axis) and N_{fib} (dense for panel **C** and sparse for panel **D**).

systems : these values are comparable to those observed in our calibration tests for an uniform distribution with similar sample size.

In contrast, the transition between curved and aligned state is not controlled by a unique model parameter but is the interplay between several parameters. Indeed, we first observe in Figure 4.4 that dense dynamical networks seem to have a greater ability to create aligned states than sparse networks, which tend to favour curved states (compare the red zones between the left and right panels). Moreover, we also observe that, for both fibre densities, networks with a moderate remodelling speed ($\nu_{link} \approx 0.01$) seem to have a greater ability to reorganise into aligned states than low dynamical networks ($\nu_{link} \approx 0.001$) or fast remodelling networks ($\nu_{link} \geqslant 0.1$) (compare the red zones between each bar within panels A and B). These results suggest that there exists a remodelling speed maximising the network alignment.

Looking at the impact of the equilibrium linked fibre fraction χ_{link}, we observe different behaviours depending on the fibre density of the system. For sparse networks (Figure 4.4.D), increasing the equilibrium linked fibre fraction tends to favour a higher level of organisation by increasing slightly the number of aligned states. On the contrary, dense networks (Figure 4.4.C) exhibit a more complex behaviour where intermediate fibre fraction $\chi_{link} \in [0.4, 0.6]$ generate more aligned states, implying that there exists an equilibrium linked fibre fraction maximising the alignment of the system.

These results show that the overall ability of our model to generate different types of tissue architectures is mainly controlled by the network remodelling speed ν_{link} and the linked fibre fraction χ_{link}, but the relationship between these two parameters is not trivial and also depends on the fibre density.

3.3. ECM architecture emergence can be explained by the network connectivity

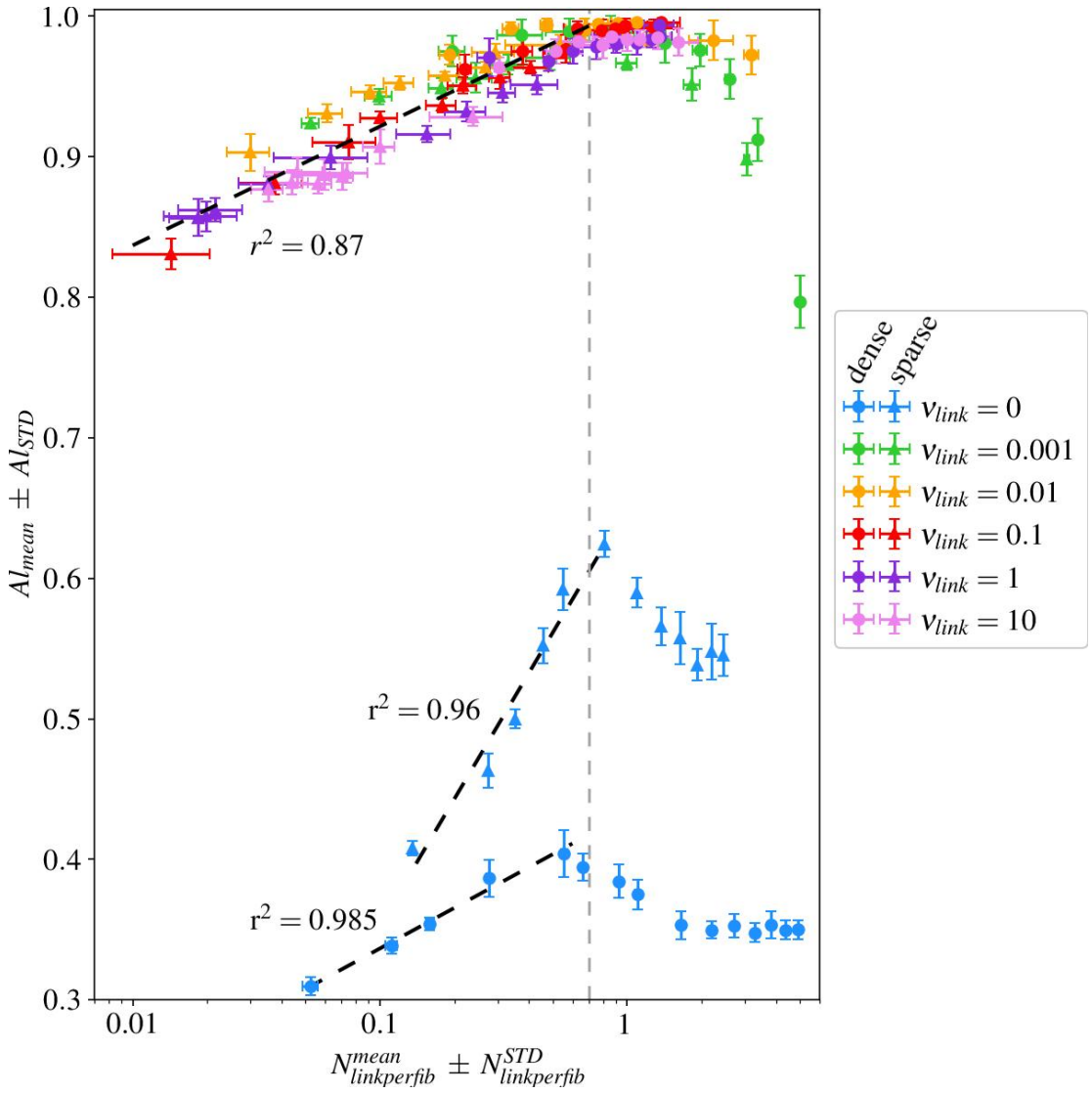


Figure 4.5: Value of Al_{mean} according to $N^{mean}_{linkperfib}$ at equilibrium, with vertical and horizontal error-bars indicating the standard deviations Al_{STD} and $N^{STD}_{linkperfib}$ respectively. The value of the fibre density is indicated with marker, that of ν_{link} with colour and, inside each colour series, χ_{link} is increasing with $N^{mean}_{linkperfib}$. The grey dashed-line indicates the critical threshold N_{critic} and the black dashed-lines the three logarithmic fits obtained for $N^{mean}_{linkperfib} < N_{critic}$.

Since the equilibrium state of a system depends both on the crosslinks characteristics and the system density, the network connectivity or mean number of links per fibre $N_{linkperfib}(t) = N_{links}(t)/N_{fib}$ could play a significant role in the simulation outcome. If the system is in a state of dynamical equilibrium, this connectivity is theoretically equal to $N_{catch}(t) \times \chi_{link}/N_{fib}$, with $N_{catch}(t)$ the current number of pair of fibres that could be (and potentially are) crosslinked because they are closer than the link catching

distance d_{link}^{max}. Note that $N_{catch}(t)$ directly depends on the spatial configurations of the system components at the current time and is thus not easily predictable, which means that the connectivity of a network can not be known in advance (except for non-dynamical networks, which keep the same number of links throughout the simulation). Moreover, when the system is not at equilibrium, the number of links also depends on the spatial configurations of the system at earlier times (that is, the number of links depends on the system history) and is not equal to its theoretical value. Since we could not predict the value of $N_{linkperfib}$ directly with the parameters, we resorted to measuring it.

To account for stochastic variability (due to the random initial condition and the stochastic linking and unlinking processes), we computed the mean and standard deviation of Al_{sim} over 10 simulations conducted with the same set of parameters, denoted by Al_{mean} and Al_{STD}. Similarly, we denote by $N_{linkperfib}^{mean}$ and $N_{linkperfib}^{STD}$ the average and standard deviation over 10 simulations of the mean number of links per fibre $N_{linkperfib}$.

Figure 4.5 presents Al_{mean} as a function of $N_{linkperfib}^{mean}$ at final time : we can see that there is a striking correlation between these two quantities. First, when $N_{linkperfib}^{mean}$ is inferior to a critical threshold $N_{critic} \approx 0.7$ links per fibre, there is a logarithmic relationship between the network connectivity and its alignment indicator :

$$Al_{mean} \approx \alpha \log(N_{linkperfib}^{mean}) + \beta, \tag{4.5}$$

with

- $\alpha = 0.037$, $\beta = 1.006$ and coefficient of determination $r^2 = 0.87$ for dynamical systems (non-blue markers);
- $\alpha = 0.129$, $\beta = 0.651$ and coefficient of determination $r^2 = 0.96$ for sparse non-dynamical networks (blue triangles);
- $\alpha = 0.042$, $\beta = 0.433$ and coefficient of determination $r^2 = 0.985$ for dense non-dynamical networks (blue dots).

All these correlations are shown on Figure 4.5 with black dashed lines. Then, when $N_{linkperfib}^{mean} > N_{critic}$, we observe an abrupt drop of the equilibrium value of Al_{mean}.

Our interpretation is that, when the number of links per fibre is small, the network is weakly constrained. In this configuration, an increase in the number of links per fibre improves the transmission of information in the network and thus enhances the alignment process. The logarithmic scaling indicates that the higher the number of links per fibre, the less prominent this feature becomes, until the gain (in terms of the equilibrium alignment indicator) becomes null. The system then shifts into a constricted regime where each new link adds to the constriction of the network and impedes its reorganisation, thus decreasing its final alignment indicator.

Surprisingly and very interestingly, for dynamical systems ($\nu_{link} > 0$) there is no difference in alignment induced by the fibre density or the link characteristics ν_{link} and χ_{link} : the correlation observed is the same for all sets of points. This means that, as long as a network is slightly dynamical, it can reach an equilibrium state with an extremely high level of local alignment if its connectivity is around the threshold value N_{critic}.

The second major observation from Figure 4.5 is the difference between non-dynamical and dynamical networks at equilibrium. In agreement with the results of section 3.2, we see that non-dynamical networks, which have a fixed connectivity, are systematically less aligned than dynamical ones, independently of the other model parameters. Moreover, although we do find the same type of relationship between the fibre local alignment and the network connectivity, for non-dynamical networks this correlation significantly depends on the fibre density.

Altogether, these results show that the emergence of local network architecture is mainly controlled by the network connectivity, playing the role of an ordering parameter.

For completion, in the next section we studied the impact of the life-expectancy of the links.

3.4. ECM architecture emergence is partly driven by the link life-expectancy

Here, we explore whether the network organisation abilities could be partly controlled by the link life-expectancy, which depends of both ν_{link} and χ_{link} (but not N_{fib}) via the following relation :

$$T_{link\text{-}life} = \frac{1}{\nu_{unlink}} = \frac{\chi_{link}}{(1-\chi_{link})\nu_{link}}. \tag{4.6}$$

Because $T_{link\text{-}life}$ does not take into account the fibre density, we can not expect it to explain the emerging architecture as well as $N^{mean}_{linkperfib}$. However, it would have the great advantage to be predictable (i.e. entirely determined by the simulation parameters) instead of being an emerging variable.

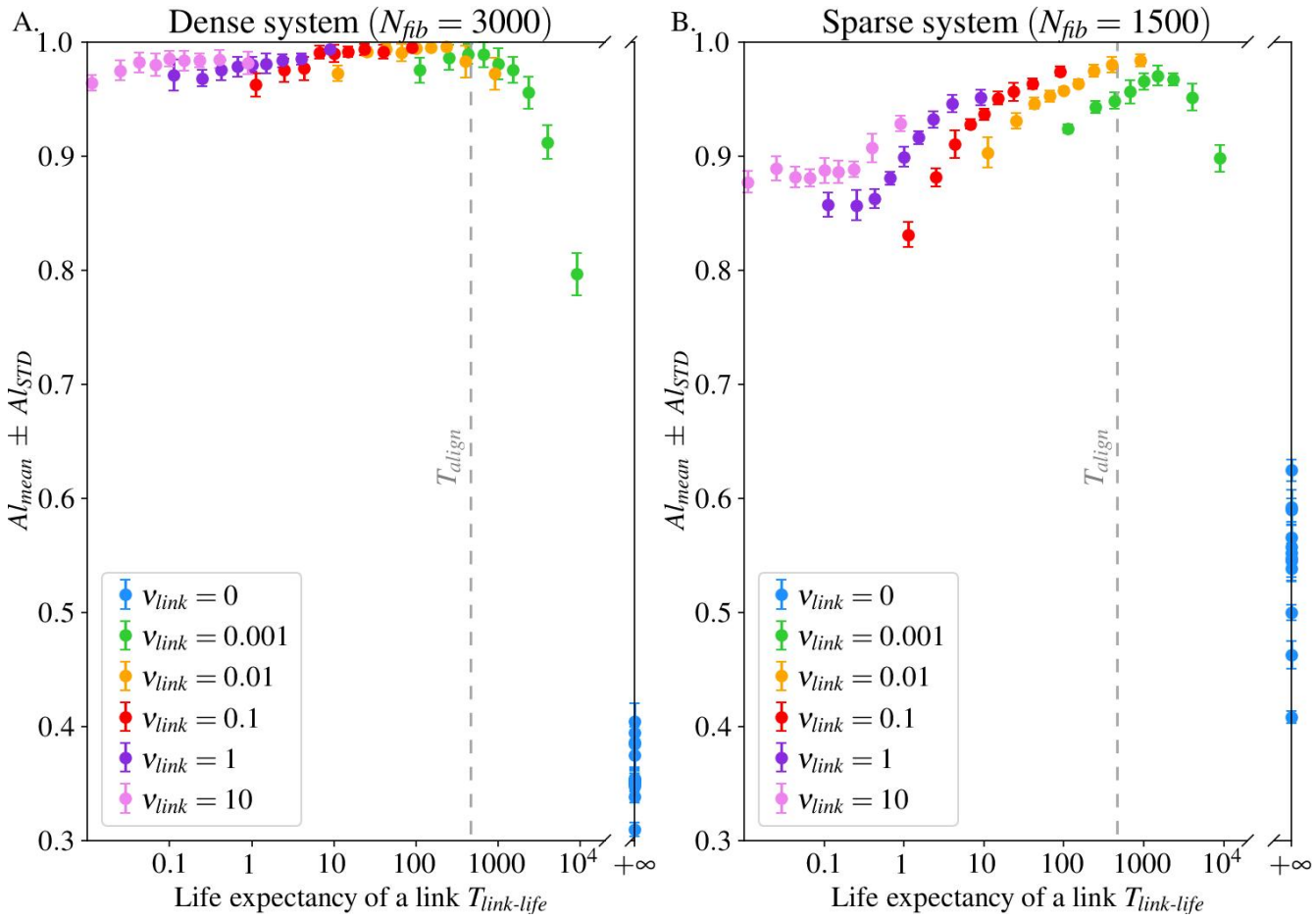


Figure 4.6: Value of Al_{mean} according to $T_{link\text{-}life}$ at equilibrium for either dense ($N_{fib} = 3000$, panel **A**) or sparse systems ($N_{fib} = 1500$, panel **B**), with vertical error-bars indicating the inter-simulation standard deviation Al_{STD}. The value of ν_{link} is indicated with colour and, inside each colour series, χ_{link} is increasing with $T_{link\text{-}life}$. The characteristic time of the alignment interaction $T_{align} = 523\,U_t$ is indicated with a grey dashed-line for the sake of comparison.

Figure 4.6 displays the value of Al_{mean} at equilibrium as a function of $T_{link\text{-}life}$. We observe that, in the case of dense systems (panel A), the alignment indicator displays a flat maximum for $T_{link\text{-}life} \in [10, 500]\ U_t$, whereas for sparse system (panel B) it reaches its highest value at $T_{link\text{-}life} \approx 500\ U_t$. This can be explained by the fact that, when the average life-expectancy of a link $T_{link\text{-}life}$ is very small compared to the characteristic time of the alignment interaction $T_{align} = 523\ U_t$, the links do not persist long enough to fully exert their aligning influence and the equilibrium alignment indicator is lesser. This is especially true for sparse systems, which display a clear drop for $T_{link\text{-}life} < 500\ U_t$. For dense systems, the drop is slower and less pronounced.

On the other hand, when $T_{link\text{-}life}$ is large compared to T_{align}, the links last on average longer than necessary to wield their full effect and lock the system in non-optimal configurations by obstructing the action of other links. Though these locally locked structures will disappear over time, others will appear - or, to put it another way, the transmission of information (i.e. the fibre direction) in the network is too slow for all the agents to synchronise and the system will not be able to reach an extremely aligned equilibrium state. Going to the extreme, when $T_{link\text{-}life} = +\infty$ the links are non-dynamical and the communication lines are totally frozen, making the system unable to align even locally.

The link life-expectancy $T_{link\text{-}life}$ is thus a relatively good predictor of the simulation equilibrium state, even if it does not explain the discrepancy between the behaviour of dense and sparse systems.

It must be noted that all these considerations only relate to the equilibrium state reached by the systems on a long time-scale. Thus, although some parameters have little or no influence on the final state of the system, they may still have a major impact on its time dynamics. This is discussed in the next section.

3.5. ECM local alignment emerges on a timescale controlled by its remodelling characteristics

In this section, we study the evolution of the quantifier Al_{mean} over time. Figure 4.7 displays this evolution for dense systems with various values of ν_{link} and χ_{link}.

Our main observation is that, for all parameters, the alignment indicator follows an inverted exponential growth, that is a quick initial growth followed by a slow convergence towards an asymptotic value. We computed the time-constant τ_{Al} of this growth, whose classical definition is the time needed for the quantifier to reach 63% of its asymptotic value, and plotted it on the corresponding curve with a black circle. It can be seen that, in our case, it corresponds approximately to the time at which Al_{mean} crosses the 0.7 threshold between unorganised and curved states.

We observe that, for a given value of ν_{link}, the shorter the time-constant, the higher the equilibrium value of the alignment indicator (compare the curves inside each panel). By comparing the panels from left to right, we see that the faster the remodelling of the network, the faster the convergence of the system towards its equilibrium value. Moreover, by comparing the extreme cases $\chi_{link} = 0.1$ (blue curve) with $\chi_{link} = 0.9$ (pink curves) of panels A and C, we see that the dependence of the reorganisation time τ_{Al} on the equilibrium linked fibre fraction is not trivial. Indeed, fast remodelling networks seem to reorganise faster when the equilibrium linked fibre fraction is large than low, while the reverse is observed for slow remodelling networks. Altogether, these results suggest that for each network dynamics, there exists a most efficient range of equilibrium linked fibre fraction allowing for quicker convergence to equilibrium.

To explore in more details the dependence between the convergence speed and the parameters of the networks, in Figure 4.8 we plot τ_{Al} as a function of χ_{link}, for different values of ν_{link}. The left panel

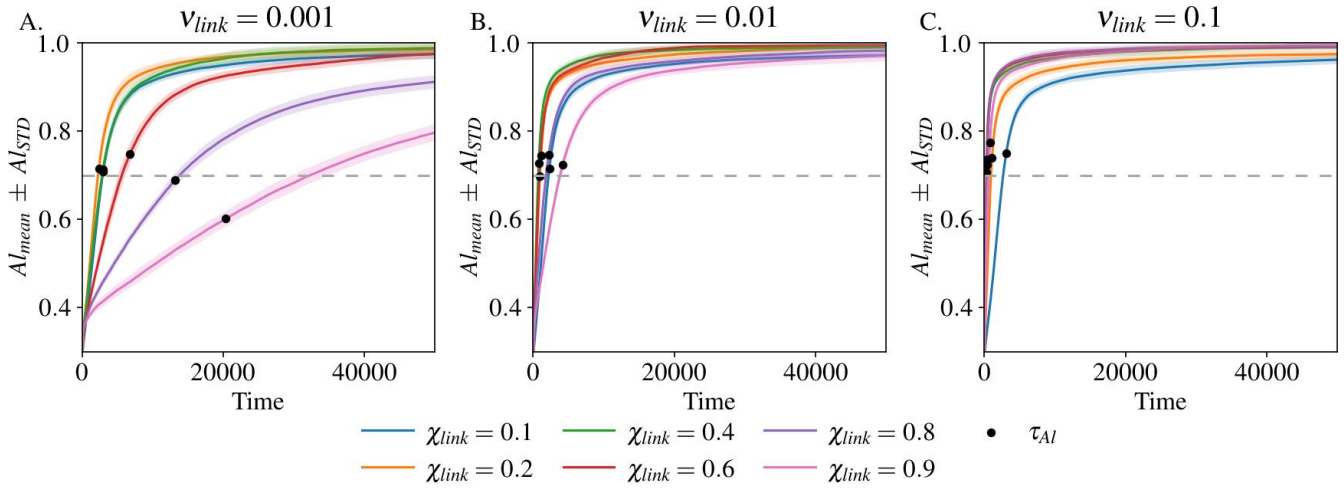


Figure 4.7: Temporal evolution the quantifier Al_{mean} for dense systems ($N_{fib} = 3000$) and various linking dynamics ($\nu_{link} = 0.001$ for panel **A**, $\nu_{link} = 0.01$ for panel **B** and $\nu_{link} = 0.1$ for panel **C**), with shading indicating the inter-simulation standard deviation Al_{STD}. The time-constant τ_{Al} of this inverted exponential growth is indicated with a black circle and the limit between unorganised and curved or aligned states is drawn with a dashed line.

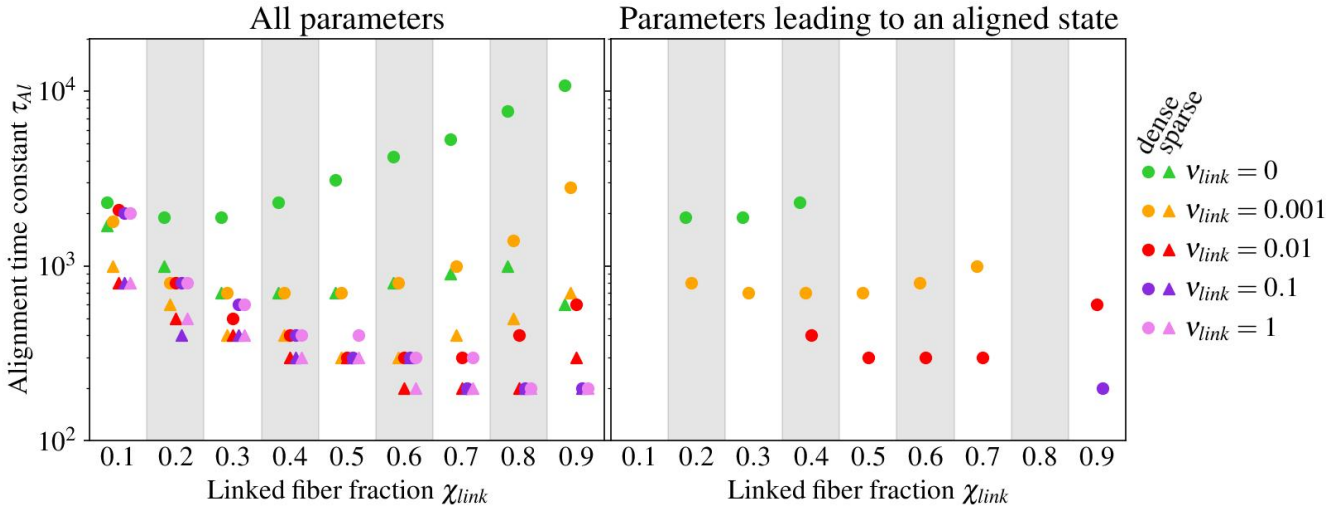


Figure 4.8: Time-constant τ_{Al} of the average alignment indicator (Al_{mean}) according to the equilibrium linked fibre fraction χ_{link}, with colour depending on the remodelling speed ν_{link} and marker depending on the fibre density. **Left :** Results for all tested sets of parameters, whatever the outcomes of the simulations. **Right :** Results for all tested sets of parameters which lead, on average, to an aligned state.

displays the value of τ_{Al} for each set of parameters while the right panel only shows τ_{Al} for the sets of parameters leading, on average over 10 simulations, to an aligned equilibrium state.

We can first see on the left panel that τ_{Al} decreases when ν_{link} increases according to a non-linear relationship which saturates for $\nu_{link} \geqslant 0.1$ (compare the different colour points). These results show that fast remodelling networks relax faster to their steady-states than slow-dynamical networks. Moreover, sparse systems organise quicker than dense systems at low linking dynamics ($\nu_{link} \leqslant 0.01$, compare the dot and triangle markers for the green and yellow populations), whereas there is no difference between

dense and sparse systems for fast remodelling networks ($\nu_{link} \geqslant 0.1$ where dot and triangle markers are superimposed).

For each value of ν_{link}, there is a most efficient range of equilibrium linked fibre fraction χ_{link} allowing for a lower value of τ_{Al} and so a quicker convergence to equilibrium. For slow remodelling networks ($\nu_{link} = 0.001$, green markers) this range lays between $\chi_{link} = 0.2$ and $\chi_{link} = 0.3$, because systems with too much crosslinks will undergo stiffening and take longer to relax, but systems with too few crosslinks will have difficulty to align themselves. As one can observe, the range of χ_{link} allowing the fastest convergence to equilibrium shifts towards 1 as the network remodelling speed ν_{link} increases. As the network remodelling increases, a greater number of crosslinks will then promote a quicker alignment.

When looking only at parameter sets which, on average, lead to aligned equilibrium states (right panel of Figure 4.8), we can see that these parameter sets cover all remodelling dynamics and correspond to the range of equilibrium linked fibre fraction leading to fastest convergence for each remodelling speed. We conclude that the most efficient systems (which organise the fastest) are also those that align most.

3.6. ECM architecture emergence follows a unique evolutionary path

In this section, we study the temporal evolution of the spatial structures : this gives us more insights into the role of the parameters in tissue structuring. We use the time-constant τ_{Al}, that emerged from the analysis in the previous section, as a time-scale for the temporal evolution of the system in the following discussion.

Movies displaying the full temporal evolution of a few simulations are available online at this address. In Figure 4.9.A–A"' and B–B"', we show the 3D view and stereographic projection of two of these simulations (respectively from *Movie3.mp4* and *Movie4.mp4*) at a few well chosen time frames (namely $0.5\tau_{Al}$, τ_{Al}, $3\tau_{Al}$ and T_{final}). They correspond to dense systems with $\chi_{link} = 0.8$ and two different crosslink dynamics : fast remodelling network $\nu_{link} = 0.1$ (panels A–A"', *Movie3.mp4*) and slow remodelling network $\nu_{link} = 0.001$ (panels B–B"', *Movie4.mp4*). These screenshots enable us to answer the important question of how the network global structure emerges. It is not by accretion around a few structured areas that gradually merge together, but by an overall homogeneous structuring. Indeed, one can observe that the directional vectors gradually concentrate around a main direction without creating clustered points that merge together. This behaviour can be observed both for very aligned networks (panels A-A"') or curved states (panels B–B"'), and in fact in all our simulations, independently on the network density. Therefore, our model suggests that the emergence of tissue architecture occurs on a global scale.

We turn now towards the analysis of the time trajectories of the spatial structures observed within our different networks. Panels C–C" of Figure 4.9 display the trajectory in the phase plane A_{max} vs Al_{sim} of individual simulations conducted with various set of parameters. We chose this one-run representation instead of the usual 10-runs average because the two quantifiers exhibit a non-negligible inter-simulations variability, so that plotting the standard deviation Al_{STD} as shading would blur the graphic but plotting only the average value Al_{mean} would give a limited and partial view of the situation.

It can be seen that all the trajectories follow a common pattern. It begins with a sharp increase of the alignment indicator (from 0.15 to between 0.4 and 0.5) while maintaining a quasi-constant semi-major axis length : this corresponds to the partial depletion of one direction (denoted d_1) in the family of the fibres directional vector, thus shifting from the initial uniform distribution to a mainly two-directional distribution (see chapter 3, section 3 for more details on this interpretation). Non-dynamical networks, which are not represented on these graphics, do not go past that first stage.

The trajectories then diversify : the alignment indicator keeps increasing while the semi-major axis length either decreases, stays constant or slightly increases. The first case is the most common and

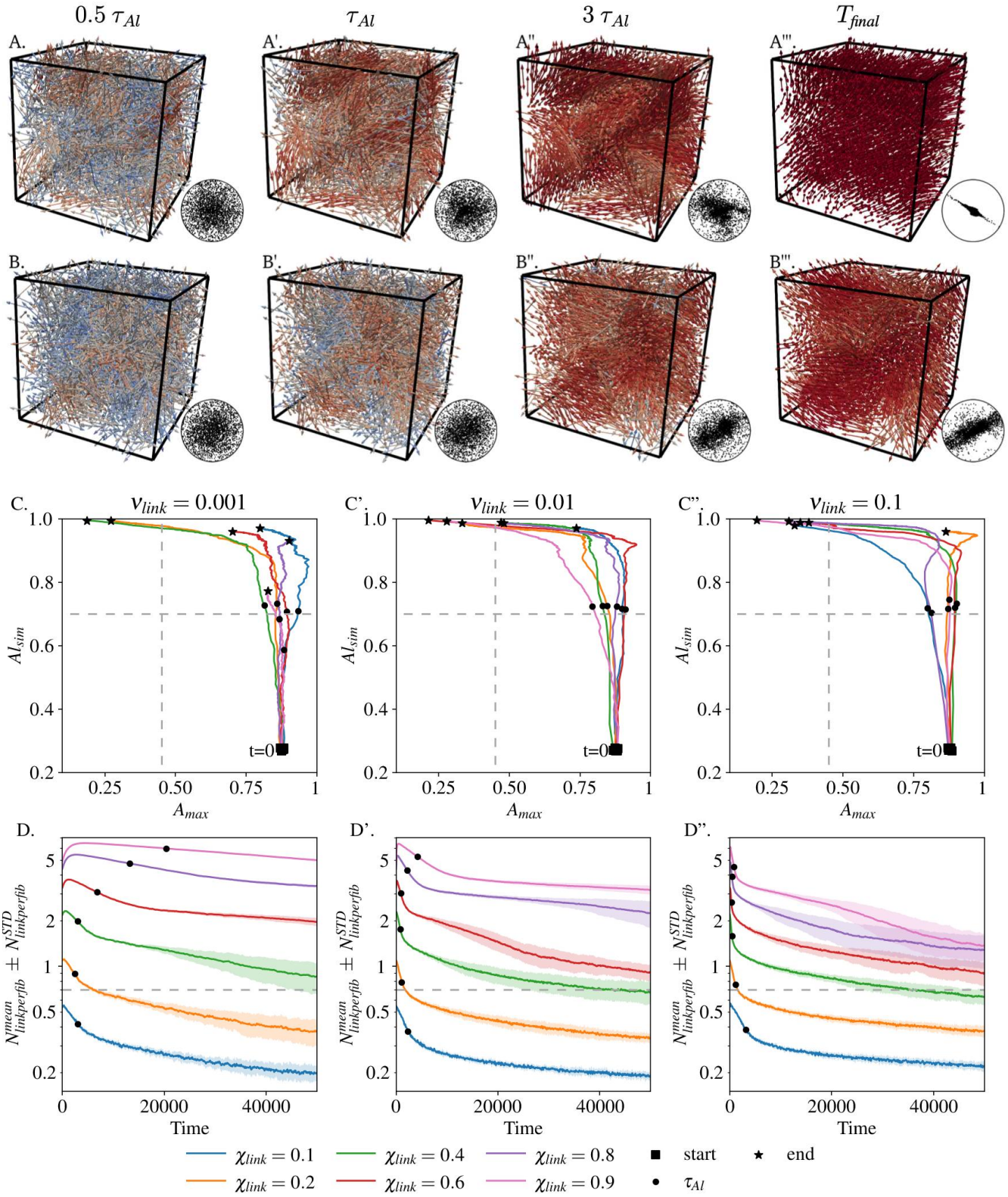


Figure 4.9: Temporal evolution of dense systems. **A–A"' and B–B"' :** 3D view and stereographic projection of the system at various times, for one simulation with $\chi_{link} = 0.8$ and either $\nu_{link} = 0.1$ (**A–A"'**) or $\nu_{link} = 0.001$ (**B–B"'**). **C–C" :** Trajectory in the plane A_{max}–Al_{sim} of a few individual simulations. The limits between each class of structures are drawn in dashed lines. **D–D" :** Evolution of $N^{mean}_{linkperfib}$, with shading indicating the inter-simulation standard deviation $N^{STD}_{linkperfib}$. The critical value N_{critic} is marked with a dashed line.

indicates that, while direction d_1 keeps depleting until near extinction, one of the two remaining directions starts to deplete as well. This diversification happens on the scale of the time-constant τ_{Al} of the alignment indicator (marked on the trajectories of Figure 4.9.C–C" with a black circle).

Lastly, simulations ending in an aligned state and part of those ending in a curved state display a stage of condensation of the fibres directional vectors around a main direction. This is marked by a shrinking of the covariance ellipse and a slow increase of the alignment indicator, which has already nearly reached its steady state (compare with the stabilisation of Al_{mean} in Figure 4.7). This last point comes from the local quality of the quantifier Al_{sim} (and by extension Al_{mean}) : a system can be very aligned locally, but not globally, if the main direction of the local structures varies smoothly across space. Thus, the transition between a curved and an aligned state is mostly characterised by a gradual shifting of multiple local structures towards the same direction, a phenomenon better registered by the quantifier A_{max} than Al_{sim}.

Finally, we observe that the number of links per fibre (displayed in Figure 4.9.D–D") undergoes a transient increase followed by a two-stage exponential decay (appearing as a piece-wise linear decrease on the semi-logarithmic scale). The initial accumulation of crosslinks is more pronounced, in the sense that the peak is higher and the subsequent decrease slower, when χ_{link} is high and ν_{link} is low. For the extreme case of slow remodelling networks with a large linked fibre fraction ($\nu_{link} = 0.001$ and $\chi_{link} = 0.9$), the phenomenon is so strong that only the first stage of exponential decay is observed during the time of the simulation. On the other hand, for fast remodelling networks ($\nu_{link} = 0.1$, panel D") and/or small equilibrium linked fibre fraction ($\chi_{link} = 0.1$, blue curve on each panel), we do not observe any crosslinks accumulation.

This behaviour can be explained by comparing the linking dynamics to the characteristic time of the repulsive interaction $T_{rep} = 4.32\ U_t$. Parameter χ_{link} describes the proportion of linked fibres among all linkable fibres at equilibrium, but this equilibrium takes time to establish (inversely proportional to ν_{link}). If the repulsion interaction operates faster than the links remodelling (i.e. $T_{rep} \ll 1/\nu_{link}$), then the linkable configurations will change before the linking/unlinking processes could equilibrate on the current configuration : new links will appear between newly overlapping fibres while former overlapping fibres will still be linked even if not overlapping anymore, leading to an accumulation of links in the system. This happens all the more if the disparity between the frequencies ν_{link} and ν_{unlink} is more favourable to linking than unlinking ($\nu_{link} > \nu_{unlink}$, i.e. if $\chi_{link} > 0.5$).

The system thus exhibits a global, macroscopic relaxation phenomenon which emerges from its various local, microscopic properties. It can be seen that the characteristic time-scale of this relaxation is comparable to the time-constant of the alignment indicator τ_{Al} (see position of the black circles on the curves in Figure 4.9.D–D", which indicates the value of τ_{Al} for the corresponding set of parameters).

We conclude that slow remodelling networks with a high equilibrium linked fibre fraction χ_{link} first build up increasing stress and stiffen before slowly relaxing, whereas networks with low χ_{link} or fast remodelling networks exhibit stress relaxation and do not undergo high stiffening. As a result, these last types of networks reach higher local alignment at equilibrium.

These results demonstrate a nonlinear dependence between the network properties and the type and proportion of its crosslinks. A high number of long lasting crosslinks promotes crosslink accumulation resulting in medium/low alignment, while fast remodelling reduces the mechanical action of the individual links on the overall network, resulting in lowly connected networks being unable to align. The network alignment ability therefore requires a number of links adapted to its remodelling speed : fast remodelling networks need a high equilibrium linked fibre fraction to quickly reach a high alignment indicator, whereas slow remodelling networks need a low equilibrium linked fibre fraction to prevent crosslink accumulation and the increase of matrix stiffness.

4. Discussion

In this work, we have implemented a 3D model for fibre networks composed of fibre elements capable to dynamically crosslink or unlink each others, to align with each others at the crosslinks and to repel their nearest neighbours to prevent fibres from cluttering. We showed that this model can spontaneously generate various types of macrostructures whose emergence can be finely described. The model reveals that the different macrostructures (i) can be easily explained by a single emerging intermediate variable, namely the connectivity of the ECM network, (ii) emerge homogeneously in space and not in a fragmented way, and (iii) follow the same unique evolutionary path for all structures and not multiple paths.

To our knowledge, this work is the first exhaustive study questioning the mechanisms of tissue architecture emergence via a simple mechanical model of dynamical fibre networks in 3D. This framework reveals that the different tissue architectures at equilibrium is directly controlled by a simple intermediary variable, the network connectivity (see section 3.3). Our interpretation is that, when the number of links per fibre is inferior to the critical threshold N_{critic}, the network is weakly constrained. In this configuration, an increase in the number of links per fibre improves the transmission of information in the network and thus enhances the alignment process. The logarithmic scaling indicates that the higher the number of links per fibre, the less prominent this feature becomes, until the gain (in terms of the equilibrium alignment indicator) becomes null. The system then shifts into a constricted regime where each new link adds to the constriction of the network and impedes its reorganisation, leading to a decrease of the local alignment.

The fact that we observe the same correlation for all dynamical networks means that, as long as a network is slightly dynamical, its final alignment is mostly controlled by its connectivity rather than by its remodelling dynamics or its density. On the other hand, non-dynamical networks are locked in mechanically constrained configurations, preventing the system from reorganising efficiently compared to dynamical ones and leading to a much lower level of alignment. However, we showed that non-dynamical networks still contain some degrees of freedom allowing for spatial matrix reorganisation, and that this organisation is controlled again by the network connectivity but also by the matrix density, which becomes an important factor. Indeed, denser networks are even less organised than sparse networks : this is due to the fact that denser networks are overcrowded, preventing any reorganisation of their fibres.

The existence of a simple emerging variable such as the network connectivity to control tissue structuring can have major therapeutic implications in systems where the architecture of the ECM is impacted (scarring, fibrosis, ageing), but can also prove very useful in the field of tissue engineering. It is noteworthy that this variable is not prescribed by model parameters but emerges from the initial simple rules as a combination of ECM remodelling dynamics, linked fibre fraction and fibre spatial organisation, independently of supplementary complex interactions involving external factors such as migrating cells, contractile forces etc. However, the correlation between network connectivity and fibre alignment only gives local information on the long-time structures (mean local alignment of the fibres at equilibrium).

The second major contribution lies in the analysis of the fine time evolution of the spatial structures. This documents the different temporal evolution of the structures as function of the ECM remodelling speed and reveals an unique trajectory for all architecture combined with internal and transient temporal windows during which they self-organise.

The equilibrium structures obtained with our model can be classified into three types : (a) aligned states with a strong organisation around one main direction, (b) curved states with a median, locally heterogeneous alignment indicator and a wide range of directional vectors living in a plane, named curved patterns and (c) unorganised states with very low alignment indicator and no preferential direction. Unorganised states were exclusively obtained for non-dynamical networks composed of permanent crosslinks ($\nu_{link} = 0$), whose plasticity was very low due to their inability to rearrange their crosslinks. In contrast,

dynamical networks exhibited a mixture of aligned and curved states. These results point to the essential role of matrix remodelling in ECM structuring, consistent with several results in the literature (see [147] and references therein).

In emerging systems, the characteristics of the final outcome cannot be predicted from the initial rules of the system and the paths from the initial interactions to the final equilibrium can be numerous and complex corresponding to a stochastic evolution. This is not completely the case in our model because, if indeed the emerging macrostructures cannot be predicted from the initial rules and the emergence must be understood as a whole, the path is simple and unique and can be strongly predicted by an intermediate emerging variable (the connectivity of the ECM network). Our study suggests that the very aligned structures observed in fibrotic tissues could be mainly due to excess accumulation of crosslinks, consistent with the alterations of ECM structure observed as a consequence of increased crosslinking in lung fibrosis [148] or cancer [71], or again with previous studies on tissue-induced alignment of fibrous ECM [144,149]. Such deciphering of the emergence would open numerous perspectives for future investigations. Indeed, because of its simplicity, this emerging variable (the connectivity of the ECM network) is amenable for experimental measurements and represents a new putative target for the development of therapeutic drugs one could develop to restore the architecture of various biological tissues after external or internal alterations. In vivo experiments must be conducted to definitively validate this hypothesis and are out of the scope of this manuscript.

Finally, the temporal evolution of the structures revealed that dynamical networks composed of long-lasting links exhibited a phase of crosslink accumulation followed by a "relaxation" phase (reduction of the network connectivity) associated with a spatial reorganisation of its fibres, whereas fast remodelling networks exhibited only the "relaxation" phase. These results suggest possible mechanisms for crosslink accumulation observed for instance in ageing tissues [76]. Moreover, the new insights into the temporal evolution of the structures as function of the ECM remodelling speed could prove useful in the field of tissue engineering, where there is a need to design efficient biological crosslinkers [150, 151].

5. Perspectives

An in-depth exploration of the structures produced by our model revealed the existence of a rare, secondary type of curved state which can be described as a twist or an helicoid : at equilibrium, the directional vectors of the fibres display a fully planar (two-directional) distribution, and the direction of one fibre depends on its position along the axis perpendicular to that plane. Figure 4.10 displays one such case where the structure is especially striking because it happens to be nearly aligned with the axes of the simulation domain : the fibres are all perpendicular to the y-axis, but their directional vector varies $\pm x$ to $\pm z$ according to their position along this same y-axis. In order to highlight this continuous twisting, the fibres in the image are coloured not according to their local alignment indicator (which is uniformly high) but according to the magnitude of the z-coordinate of their directional vector. It can be seen that this quantity varies smoothly along the y-axis (including across the periodic border), being equal to 0 near the plane $\{y = 0\}$ (the fibres are then perpendicular to the z-axis) and to 1 near the planes $\{y = \pm L_y\}$ (the fibres are then oriented along the z-axis).

The planar distribution of the fibres directional vectors is evidenced by their stereographic projection, which describe a thin band of nearly uniform width spanning the whole length of the unit circle (see Figure 4.10). Indeed, helicoidal states are characterised by their very high semi-major axis length A_{max} (equal to 0.9675 for the simulation displayed Figure 4.10) which are among the twenty highest registered in all dense simulations (including number one, two and four). They also display very high local alignment indicator Al_{sim} (equal to 0.9924 for the system displayed in Figure 4.10). Since all the fibres are

almost perfectly aligned with their neighbours, these structures are very stable once they have emerged.

However, their emergence itself is very sensitive to the stochastic variability of our model. Among the six parameters sets that produced an helicoidal state, only one did it twice out of ten independent simulations. Figure 4.11 displays the trajectories in the plane A_{max} vs Al_{sim} of all the simulations ending in an helicoidal states. It can be seen that they all exhibit a sharp bifurcation during their third phase of evolution, which should have lead them towards an aligned state : as the covariance ellipse had begun to condense, it suddenly expand again while maintaining a quasi-constant alignment indicator. This shows that helicoidal states occur when a curved structure evolving towards an aligned state randomly happens to "close" on itself before the fibres directional vectors can condensate around a single main direction. From this point on, the helicoid will only strengthen itself and will never break down.

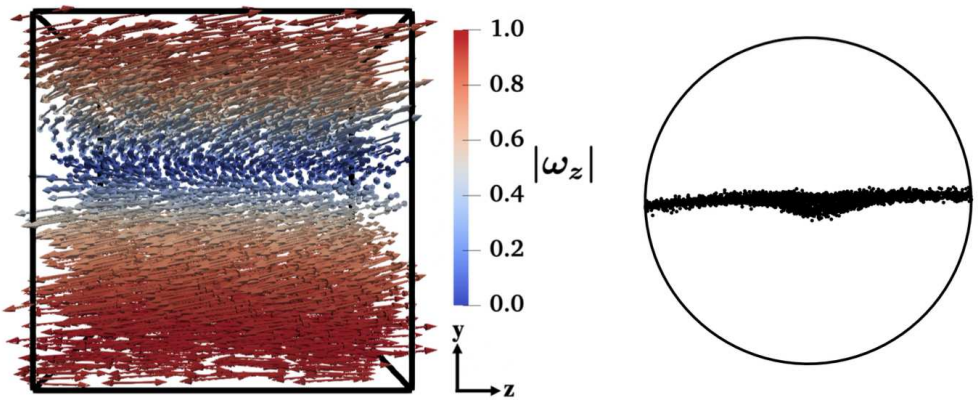


Figure 4.10: Illustration of an helicoidal state. **Left :** 3D view of the system with fibres coloured according to the magnitude of the z-coordinate $|\omega_z|$ of their directional vector. The axis are indicated at the bottom right corner of the domain. **Right :** Stereographic projection of the fibres directional vectors.

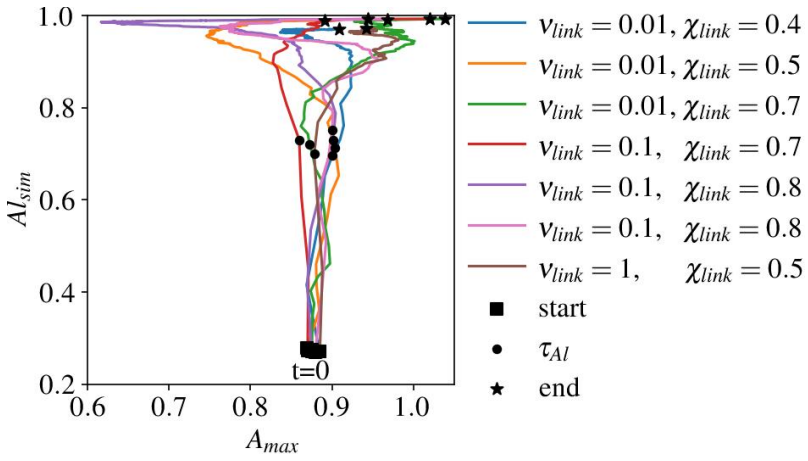


Figure 4.11: Trajectories in the plane A_{max}–Al_{sim} of the seven simulations ending in an helicoidal state. The simulation illustrated in Figure 4.11 correspond to the purple curve. Note that helicoidal states only occur in dense systems ($N_{fib} = 3000$) and that two of them resulted from the same parameters set.

It would be very interesting to study these structures in more details. One thing to do would be to run a large number of simulations around the range of parameters which produced them, to better delineate the range in which they can be formed and accurately compute their probability of occurrence.

However, we first need to make sure that helicoidal states are a genuine, intrinsic outcome of our model and not an artefact produced by the periodic boundary conditions in a small-sized domain. Indeed,

because the side-length of the simulation domain is only about 4 times the length of a fibre, the boundary conditions enforces strong constraints on the network and can reasonably be expected to facilitate the occurrence of locked rotating structures. The question is whether helicoidal states could emerge in a larger, less constraining spatial domain.

More generally, the small size of the simulation domain is a clear limitation of our study. It was chosen as a reasonable balance between the risk of interferences from the boundary conditions and the computational cost of our model, which allowed use to explore a large range of parameters. Now that we have clearly outlined the structures produced by our model, an important validation step would be to reproduce these results in a larger spatial domain.

In this study, we demonstrated the ability of fibre networks to spontaneously self-organise as function of the kinetics of their crosslinks. Our model features active crosslinks, i.e crosslinks that generate an alignment of the fibres they are attached to. As a result, our fibre networks self-organise without needing to be subjected to any external mechanical stimuli. However, it would be interesting to study how they react to mechanical stimuli such as tensile/compressive stress, shear, etc. sustained before or after having reached their autonomous equilibrium state. For example, it is natural to expect that an aligned and a curved network will not displays the same overall elastic moduli when subjected to tensile stress. Moreover, the dynamic aspect of our model obviously raise the issue of its plasticity : will a network regain its former state once the stimuli stops ? It could be presumed that quickly remodelling networks would oppose low resistance to any type of stimuli, resulting in an irreversible (plastic) deformation even at low stress, whereas slow remodelling networks would exhibit a more elastic (rather than plastic) response. On the other hand, the plastic or elastic nature of the deformation may be more related to the duration of the stimuli compared to the remodelling speed of the network. For instance, all dynamical networks sustaining a compressive stress should exhibit stress-stiffening due to crosslink accumulation when their density increase. Quickly remodelling networks would stiffen very quickly, resisting the compression and deforming little, but also relaxing quickly after the end of the stimuli. At the same time, slow remodelling networks would undergo crosslink accumulation only under prolonged compression, but the phenomenon may then be quasi-permanent, preventing them from returning to their initial size.

Future works will be devoted to studying the mechanical properties of the dynamical networks produced by our model. Another interesting perspective would be to add cells having the ability to generate locally biophysical cues such as tension, stiffness and fibre production/degradation and study these effects on the structure and mechanical properties of the ECM networks. As first step in this direction has already been taken with the introduction, in the model ATmorpho-3D, of spherical cells exerting a steric pressure on the neighbouring fibres. The consequences over the network organisation of the addition of this new type of agent are studied in chapter 5.

Finally, it is noteworthy that our model features networks composed of only one type of crosslinks (permanent or transient with a given life expectancy). A natural perspective of our works would be to study the self-organisation abilities of networks composed of heterogeneous crosslinks, following the works of [20]. One could represent the various crosslinking molecules of the ECM by different type of crosslinks, each with its own linking and unlinking dynamics but also mechanical properties.

For example, one could imagine a blend of very elastic, long reaching crosslinks (with low κ_{rest} and large d_{link}^{max}) and stiff, short reaching crosslinks (with high κ_{rest} and small d_{link}^{max}).

Chapter 5

Study of adipose tissue architecture emergence

In this chapter, we investigate the behaviour of the systems produced by the model ATmorpho-3D, which are composed of a dynamically connected fibres network inside which spherical cells appear and grow. This model is aimed at reproducing the emergence of Adipose Tissue (AT) architecture and is fully described in chapter 2.

We observe the emergence of lobule-like clusters of cells that we segment and characterise using the tools developed in chapter 3. We show that the cells have a marginal impact on the global organisation of the fibre network, whereas the intrinsic properties of the fibre network exert a major influence on both this network organisation and the shape of the cell clusters, in agreement with previous 2D findings [17]. Lastly, we show that our *in silico* lobules compare quantitatively well to *in vivo* lobules segmented in AT.

We end the chapter with a discussion on the major findings and the many perspectives of this work.

Summary of the chapter

1. Introduction and motivation

In this chapter, we aim to investigate numerically the 3D Individual-Based Model for Adipose Tissue architecture emergence (ATmorpho-3D). This model is presented in full in chapter 2 (section 3) and briefly summarised in the next section of the present chapter.

As in our previous study (chapter 4), we aim to explore whether tissue architecture emergence could be mainly driven by simple mechanical interactions between the tissue core components and a few well-chosen biological phenomena (Extra-Cellular Matrix (ECM) remodelling, cell differentiation and growth). As mentioned in the introduction of this manuscript (chapter 1, section 4.2), studies on AT have been mostly focusing on its cellular composition and on the molecular pathways driving cell differentiation rather than on its architecture. As a result, studies on tissular organisation of AT depots are scarce and descriptive, and the basic mechanisms driving the emergence of its architecture remain poorly understood. The model ATmorpho-3D is built on a different viewpoint and aims to identify the main mechanisms at play in AT architecture emergence. It is closely based on the model for fibre networks ECMmorpho-3D analysed in the chapter 4, which has been extended to account for the appearance and growth of a population of spherical cells in a 3D dynamical ECM network.

As mentioned in chapter 1 (section 4.2), the seminal works of Wassermann [107] revealed that, during development, the AT architecture emerges from primitive structures constituted of a disorganised fibre network containing endothelial cells and fibroblast-like cells (believed to be pre-adipocytes). In adult AT, mature adipocytes (cells responsible for fat storage and release) are organised in 3D clusters called lobules, separated from each other by well-structured sheets of ECM called septa [107]. These structures are robust throughout adult life and large perturbations of this tissue architecture, for instance the increased fibrosis observed in excessive development of AT occurring during obesity, is associated with adipocyte dysfunctions [117–119]. Therefore, modelling the process of lobule emergence could greatly improve our understanding of AT biology and plasticity in physiological or pathological conditions and give insights into the link between tissue architecture and function.

In this chapter, we demonstrate that the simple 3D model we propose is able to produce realistic lobule-like structures of cells surrounded by fibres organised in sheet-like structures. Quantitative comparison between *in silico* and *in vivo* lobules is achieved via the definition of appropriate morphological quantifiers (solidity and elongation) for the segmented cell structures, using the tools presented in chapter 3. These major results show that 3D tissue architecture emergence could indeed be mainly driven by mechanical interactions between the ECM network and its functional cells (here adipocytes). In addition, via an in-depth sensibility analysis of the model parameters, we find that the overall architecture of the system (shape of the lobules and alignment of the ECM), is mostly controlled by the parameters related to the ECM network (namely its remodelling rate and its connectivity), and that cells have a secondary impact on the global architecture of the tissue. These results point to the essential role of matrix remodelling in tissue structuring. We end this chapter by discussing these results in the light of recent findings in the literature.

2. Description of the experiments

To test our hypothesis, we perform numerical simulations of our model ATmorpho-3D presented in chapter 2, that we briefly recap here.

The ECM fibres are discretised into N_{fib} rigid spherocylindrical fibre units of fixed uniform length L_{fib} and radius R_{fib}, while the adipocytes are represented by rigid spheres of variable radius. We include the following biological phenomena : *(i) Fibre linking/unlinking :* as with the model ECMmorpho-3D, fibres

closer than d_{link}^{max} can crosslink with each other according to a Poisson process of frequency ν_{link}. Existing crosslinks can break spontaneously according to a Poisson process of frequency ν_{unlink}. As a result, the linked fibre fraction $\chi_{link} = \dfrac{\nu_{link}}{\nu_{link} + \nu_{unlink}}$ represents the equilibrium fraction of linked fibres among the pairs of neighbouring fibres. *(ii) Adipocyte differentiation and growth :* the cells undergo a linear volumic growth of constant rate K_{growth} (modelling adipocyte hypertrophy) until they reach a maximum radius R_{ad}^{max}. In addition, new adipocytes of minimum radius R_{ad}^{min} are inseminated randomly inside the spacial domain (modelling adipocyte hyperplasia) according to a uniform law in space and a Poisson process of frequency ν_{ins} in time, until the number $N_{ad}(t)$ of adipocytes in the modelled system reaches its maximum value N_{ad}^{max}. Together, these two processes allow to model the creation and expansion of a new AT depot due to excessive energy intake.

Concerning the mechanical interactions, we consider that overlapping agents exert on each other a steric repulsion force of maximal intensity E_{rep}^{FF} for a fibre-fibre contact, E_{rep}^{AF} for an adipocyte-fibre contact and E_{rep}^{AA} for a contact between two adipocytes. Each crosslink is represented by an elastic spring of stiffness κ_{rest} and unloaded length d_{link}^{eq}. Linked fibres also sustain a nematic alignment torque proportional to the flexural modulus α_{align}. Finally, all agents are subjected to a friction force (with dynamic viscosity μ_{fib} for fibres and μ_{ad} for adipocytes) large enough to nullify their acceleration.

Using Newton's second law of motion in an over-damped regime, cell motion, fibre motion and fibre rotation follow the differential system (2.23), which reads as :

$$\begin{cases} \dfrac{d\boldsymbol{X}_i}{dt}(t) = \dfrac{1}{\mu_{ad} R_i(t)} \left[\displaystyle\sum_{\substack{j=1 \\ j\neq i}}^{N_{ad}} \boldsymbol{F}_{i,j}^{AA}(t) + \sum_{k=1}^{N_{fib}} \boldsymbol{F}_{i,k}^{AF}(t) \right] & \forall i \in [\![1, N_{ad}(t)]\!], \\ \dfrac{d\boldsymbol{Y}_k}{dt}(t) = \dfrac{1}{\mu_{fib} L_{fib}} \left[\displaystyle\sum_{j=1}^{N_{ad}} \boldsymbol{F}_{k,j}^{FA}(t) + \sum_{\substack{m=1 \\ m\neq k}}^{N_{fib}} \left(\boldsymbol{F}_{k,m}^{FF}(t) + p_{k,m}(t) \boldsymbol{F}_{k,m}^{link}(t) \right) \right] & \forall k \in [\![1, N_{fib}]\!], \\ \dfrac{d\boldsymbol{\omega}_k}{dt}(t) = \dfrac{1}{\mu_{fib} L_{fib}^3} \left[\displaystyle\sum_{j=1}^{N_{ad}} \boldsymbol{T}_{k,j}^{FA}(t) \right. & \\ \qquad \left. + \displaystyle\sum_{\substack{m=1 \\ m\neq k}}^{N_{fib}} \left(\boldsymbol{T}_{k,m}^{FF}(t) + p_{k,m}(t) (\boldsymbol{T}_{k,m}^{link}(t) + \boldsymbol{T}_{k,m}^{align}(t)) \right) \right] \wedge \boldsymbol{\omega}_k(t) & \forall k \in [\![1, N_{fib}]\!]. \end{cases} \tag{2.23}$$

This system and all the stochastic processes included in our model are numerically solved using the algorithm described in chapter 2, section 4. As in the previous chapter, which was concerned with the model ECMmorpho-3D, we consider a spatial domain $\Omega = [-L_x, L_x] \times [-L_y, L_y] \times [-L_z, L_z]$ with periodic boundary conditions. At initialisation, the number of adipocytes is null ($N_{ad}(0) = 0$) and all N_{fib} fibres are randomly inseminated according to a uniform law for both their position in Ω and their directional vector in $\mathbb{S}_2^+$ (see chapter 4).

Parameter	Value	Dimension	Description
Agents characteristics			
N_{ad}^{init}	0	N/A	Initial number of adipocytes in the system
ϕ_{ad}	$[0.3, 0.5]$	N/A	Maximum adipocytes density
R_{ad}^{min}	0.1	L	Minimum adipocyte radius (= radius at insemination)
R_{ad}^{max}	1.5	L	Maximum adipocyte radius.
ϕ_{fib}	$[0.3, 0.5]$	N/A	Fibres density
L_{fib}	6	L	Fibre length
R_{fib}	0.5	L	Fibre radius
Mechanical interactions			
E_{rep}^{AA}	5.0	$M.L^{-1}.T^{-2}$	Magnitude of the adipocyte-adipocyte repulsion force
E_{rep}^{AF}	8.9	$M.L^{-1}.T^{-2}$	Magnitude of the adipocyte-fibre repulsion force
E_{rep}^{FF}	12.5	$M.L^{-1}.T^{-2}$	Magnitude of the fibre-fibre repulsion force
κ_{rest}	5.0	$M.L^{-2}$	Link stiffness
α_{align}	2.0	$M.L^{2}.T^{-2}$	Magnitude of the alignment torque between linked fibres
d_{link}^{max}	1.2	L	Perception distance for link creation
d_{link}^{eq}	1.0	L	Link equilibrium length
μ_{ad}	1.0	$M.L^{-1}.T^{-1}$	Dynamic viscosity of the adipocytes
μ_{fib}	1.0	$M.L^{-1}.T^{-1}$	Dynamic viscosity of the fibres
Adipocyte processes			
ν_{ins}	1	T^{-1}	Adipocyte insemination frequency.
K_{growth}	0.05	$L^{3}.T^{-1}$	Adipocyte volume growth per unit of time.
Link processes			
ν_{link}	$[0, 0.1]$	T^{-1}	Network remodelling speed
χ_{link}	$[0.1, 0.8]$	–	Equilibrium linked fibre fraction
Numerical parameters			
$L_x = L_y = L_z$	15	L	Side half-length of the cuboid domain
T_{final}	2000	T	Total time of simulation

Table 5.1: Model and numerical parameters used for the simulations.

The physical scaling of all the parameters of the simulations are described in Table 5.1. A few points may be noted :

- A new adipocyte is inseminated every $t_{ins} = 1/\nu_{ins} = 1\ U_t$ on average, and subsequently takes $t_{growth} = \dfrac{(R_{ad}^{max})^3 - (R_{ad}^{min})^3}{K_{growth}} = 67.5\ U_t$ to grow up to its maximum radius. These numbers are difficult to relate to actual biological time scales because adipocyte growth rate and differentiation rate are highly variable from one individual to another, depending on many factors such as age, sex, diet and body mass index but also on an innate tendency to favour either adipocyte hypertrophy

(growth) or hyperplasia (insemination). However, it as been shown that adipocyte turnover (i.e. the proportion of the total number of adipocytes renewed each year) is relatively homogeneous between individuals independently of all the above-mentioned factors, and is approximately equal to 10% per year [152, 153]. Given that the total number of adipocytes in our simulations varies between 573 ($\phi_{ad} = 0.3$) and 955 ($\phi_{ad} = 0.5$), a turnover of 10% per year would be roughly equivalent to the production of 100 adipocytes per year, that is $t_{ins} = 1\ U_t \approx 3$ days. This value will serve as a reference time scale for the present study.

- The fibre network can be regarded as having undergone significant remodelling when 10% of its linkable configurations (i.e. existing links and non-linked neighbouring fibre pairs) have been renewed through link creation and/or destruction. Given that each configuration is remodelling independently of the others, and considering that ν_{link} and ν_{unlink} are of the same order of magnitude, a tenth of the crosslinking processes will activate in $t_{remodelling} = 0.1/\nu_{link}$ on average. Excluding the case of non-dynamical networks, this remodelling time ranges from 3 to 300 days, which is around the estimated ECM remodelling time of 15 days [154].

- The total time of simulation T_{final} is chosen to be at least twice as long as the total adipocyte insemination time $N_{ad}^{max} t_{ins}$. In contrast, it is too short for the fibre networks, except the fastest remodelling ones, to fully equilibrate. This choice was made because long time-scales (compared to the total adipocyte insemination time) are more related to tissue homeostasis than tissue morphogenesis, which is the subject of the present study.

- Recalling that the average diameter of a fully grown white adipocyte in the subcutaneous depot is 100 µm (see chapter 1, section 4.2), we can estimate the length scale of our simulations to be $U_l = 33$ µm.

- From this length scale, we see that the fibre units considered here are $2R_{fib} = 33$ µm wide, in agreement with the measures of the width of collagen fibril bundles in loose connective tissues such as adipose tissue [155].

- On the other hand, the length $L_{fib} = 200$ µm of the fibre units is twice the size of fully grown adipocytes. In our model, the flexibility of the fibre network is entirely due to the crosslinks since the fibre units themselves are treated as rigid. Thus, in order for our network to be able to bend around the adipocytes like actual collagen fibres, the length of the fibre units should be about the diameter of an adipocyte. This means that, in practice, the fibre networks modelled here are stiffer (on the length scale the adipocytes) than they should. However, using shorter fibre units while maintaining the same fibre density would imply a higher number of agents, leading to intractable computational costs (see chapter 1, section 2.3).

- Contrarily to the previous experiments with the model ECMmorpho-3D, here the maximum distance for link creation d_{link}^{max} is slightly larger than its minimum acceptable value of $2R_{fib}$. This allows fibres to link together even if the pressure exerted by the adipocytes tend to move them apart.

- In the absence of empirical data concerning the dynamic viscosity of adipocytes and ECM fibres, we make the simplifying assumption that $\mu_{ad} = \mu_{fib}$.

We denote by ϕ_{ad} the maximum adipocyte density of the modelled system, that is the ratio between the maximum volume of adipocytes (after full insemination and growth, without overlapping) and the

volume of the spatial domain :

$$\phi_{ad} = \frac{N_{ad}^{max} \times \frac{4}{3}\pi (R_{ad}^{max})^3}{8 L_x L_y L_z}. \tag{5.1}$$

Similarly, we denote by ϕ_{fib} the fibre density of the network (see chapter 4). The number of agents are chosen so that both of these densities are in the range $[0.3, 0.5]$. These values can be compared to that of the fibre networks presented in chapter 4, which varied between $\phi_{fib} = 0.29$ and 0.58. The fibre networks studied in the present chapter have similar densities, but are immersed in an environment which is at least dense (with a total agent density $\phi_{tot} = 0.6$) and at most hyperdense ($\phi_{tot} = 1$), guaranteeing a high level of mechanical stress.

For each type of mechanical interactions in the system, we define the "characteristic interaction time" as the time needed for two isolated agents submitted only to this interaction and initially positioned in the most unfavourable configuration to reach 99% of the (asymptotic) equilibrium configuration. Explicit computation leads to the following formula (numerical values are given for the parameters presented in Table 5.1).

- T_{rep}^{AA} is the time needed for two fully overlapped adipocytes of equal radius to move apart by 99% of their equilibrium distance (i.e. from $\boldsymbol{X}_i = \boldsymbol{X}_j$ to $||\boldsymbol{X}_i - \boldsymbol{X}_j|| = 0.99(R_i + R_j)$, with $R_i = R_j$) due to the repulsion force alone.

$$T_{rep}^{AA} = \frac{27 \mu_{ad}}{4 E_{rep}^{AA}} = 1.35\, U_t. \tag{5.2}$$

- T_{rep}^{AF} is the time needed for a fibre and an adipocyte fully overlapped to move apart by 99% of their equilibrium distance (i.e. from $\boldsymbol{X}_i = \boldsymbol{Y}_k$ to $||\boldsymbol{X}_i - \boldsymbol{Y}_k|| = 0.99(R_{ad}^{max} + R_{fib})$) due to the repulsion force alone.

$$T_{rep}^{AF} = \frac{27}{2\sqrt{R_{ad}^{max} R_{fib}}\, E_{rep}^{AF}} \times \frac{\mu_{ad} R_{ad}^{max} \mu_{fib} L_{fib}}{\mu_{ad} R_{ad}^{max} + \mu_{fib} L_{fib}} = 2.10\, U_t. \tag{5.3}$$

- T_{rep}^{FF} is the time needed for two fully overlapped fibres to move apart by 99% of their equilibrium distance (i.e. from $\boldsymbol{Y}_k = \boldsymbol{Y}_m$ to $||\boldsymbol{Y}_k - \boldsymbol{Y}_m|| = 0.99 \times 2R_{fib}$, with $\boldsymbol{\omega}_k \parallel \boldsymbol{\omega}_m$) due to the repulsion force alone.

$$T_{rep}^{FF} = \frac{27 \mu_{fib} L_{fib}}{4\sqrt{2} R_{fib} E_{rep}^{FF}} = 4.58\, U_t. \tag{5.4}$$

- T_{rest} is the time needed for two fibres that are initially fully overlapping and crosslinked at their centre to relax their link to 99% of its equilibrium length d_{link}^{eq} due to the elastic restoring force alone.

$$T_{rest} = \ln(10) \frac{\mu_{fib} L_{fib}}{\kappa_{rest}} = 2.76\, U_t. \tag{5.5}$$

- T_{align} is the time needed for two perpendicularly intersecting fibres ($\boldsymbol{Y}_k = \boldsymbol{Y}_m$ and $\boldsymbol{\omega}_k \perp \boldsymbol{\omega}_m$) crosslinked at their centre to reach a relative angle $\arccos(\boldsymbol{\omega}_k \cdot \boldsymbol{\omega}_m) = 0.9^\circ$ due to the alignment torque alone.

$$T_{align} = -\ln\left(\sqrt{2}\sin\left(\frac{0.9^\circ}{2}\right)\right) \frac{\mu_{fib} L_{fib}^3}{2\alpha_{align}} = 462\, U_t. \tag{5.6}$$

It may be noted that $T_{rep}^{AA} < T_{rep}^{AF} < T_{rep}^{FF}$: although their elastic modulus is lesser (indicating a lesser resistance to pressure), adipocytes are technically less prone to overlap than fibres. As before, the alignment interaction is much slower than the repulsive and elastic restoring forces.

3. Overview of the various possible structures

In Figure 5.1, we present an overview of the various structures that can be obtained with the model ATmorpho-3D by playing on the parameters in the ranges indicated in Table 5.1 (see later sections for the influence of the parameters). We use the data processing framework described in chapter 3 : 3D visualisation using Paraview (see chapter 3, section 1) with fibres coloured according to their local alignment indicator Al_k (see chapter 3, section 3.1) and adipocytes coloured according to the lobule number and located at their "lobule coordinates" (see chapter 3, section 2.2). For more visibility, we represented the same system three times, displaying either both type of agents (in the first column) or only one of them (adipocytes in the second column and fibres in the third column).

The first thing to note is that lobules, here defined as cell clusters containing at least 10 cells, appear in all the simulations. Though most of them are relatively elongated, they display various shapes ranging from branching structures with many digitations (see Figure 5.1.A) to non-branching ellipsoidal (see Figure 5.1.B) or poly-lobed (see Figure 5.1.C) structures. Moreover, some of them are compact while others exhibit holes between their constituent cells (see for example the khaki lobule on the right-hand side of Figure 5.1.A).

The fibres structures on the other hand range from locally well organised networks exhibiting a global preferred direction (see Figure 5.1.A) or no preferred direction at all (see Figure 5.1.B) to disordered networks with low local alignment (see Figure 5.1.C).

As with the model ECMmorpho-3D (see chapter 4), the local alignment rate of the fibres is quite homogeneous inside one system : contrarily to our expectation, the presence of cells does not cause any significant heterogeneity. However, the spacial distribution of the fibres is not homogeneous, especially in locally aligned networks. In these networks (see Figure 5.1.A and B), fibres seem to bundle and form locally planar structures bending around lobules, which could be likened to biological fascia (see chapter 1, section 4.2).

Visually it seems that the lobules are relatively homogeneous inside one simulation, so that studying the average of the quantifiers E_{lob} and S_{lob} introduced in chapter 3 (which characterise respectively the elongation and the solidity of an individual lobule) over all the lobules of a system could inform us on the overall cellular architecture of this system. However, it is not so : all systems display similar average value of E_{lob} and S_{lob}, and the variation observed between different systems are of the same order of magnitude then the typical standard deviation inside one system. To be exact, if we denote by E_{sim} and S_{sim} the average eccentricity and solidity of all the lobules of a system, we find that :

- E_{sim} varies between 0.80 and 0.94 (with an average value over all simulations equal to 0.88), but the typical variation of E_{lob} between lobules pertaining to the same system is 0.06;
- S_{sim} varies between 0.21 and 0.47 (with an average value over all simulations equal to 0.37), but the typical variation of S_{lob} between lobules pertaining to the same system is 0.08.

This means that we will not be able to characterise the organisation of a system based on the average characteristics of its lobules. Instead, we will study the lobules individually and try to classify them into different morphological categories that could be related to the biological observations given in introduction of this manuscript (see chapter 1, section 4.2). Indeed, from a biological viewpoint it is more significant to look at how lobule morphologies are distributed inside a system than to look at the average morphology.

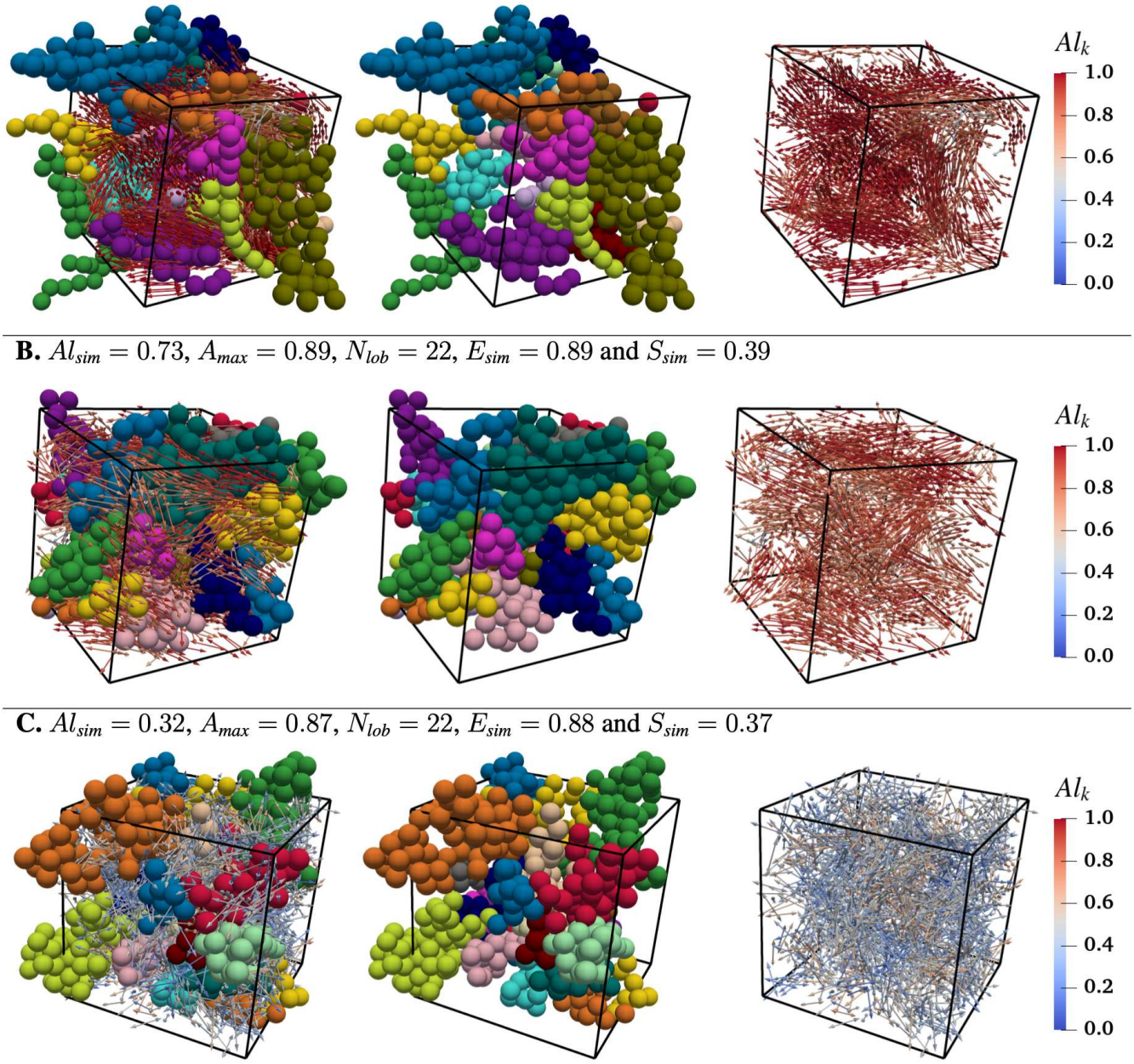


Figure 5.1: Illustration of the various types of structures generated by the model ATmorpho-3D. Each row displays the 3D view of the system at the end of one simulation, with fibres coloured according to their local alignment rate (blue: $Al_k = 0$, red: $Al_k = 1$) and adipocytes coloured according to their lobule number and translated according to their lobule coordinates. The left panels display both types of agents, the middle panels only the fibres and the right panels only the adipocytes. For each simulation, the value of a few quantifiers are given. Note that we denote by E_{sim} (resp. S_{sim}) the average value of the quantifier E_{lob} (resp. S_{lob}) over all the lobules of a system.

In the following sections, we will first explore qualitatively the influence of the presence of cells over the organisation of the fibre network, taking as a basis our study of fibre-only systems in the previous chapter. We will then turn to the quantitative characterisation of the lobule-like structures produced by

the model ATmorpho-3D, studying both their distribution according to the model parameters and their heterogeneity inside one simulation. Finally, we will compare these structures to the lobules segmented in a few *in vivo* samples.

4. Characterisation of the fibre structures : comparison with the model ECMmorpho-3D

Each row of Figures 5.2 and 5.3 illustrates the fibre structures produced by systems with the same set of parameters except for the adipocyte density, which increases from left to right (with $\phi_{ad} = 0$, $\phi_{ad} = 0.3$ and $\phi_{ad} = 0.5$ for each column respectively). Figure 5.2 focus on systems with a high fibre density ($\phi_{fib} = 0.5$) and Figure 5.3 is the equivalent for systems with a low fibre density ($\phi_{fib} = 0.3$). The first row of these two figures displays non-dynamical, moderately crosslinked networks ($\nu_{link} = 0$ and $\chi_{link} = 0.3$), the second and third rows display slow remodelling networks ($\nu_{link} = 0.001$) with either moderate or high linked fibre fraction ($\chi_{link} = 0.3$ and 0.8 respectively), and the last two rows display quickly remodelling networks ($\nu_{link} = 0.1$) again with either moderate or high linked fibre fraction ($\chi_{link} = 0.3$ and 0.8 respectively).

As can be seen by comparing the three columns of Figure 5.2, the presence of adipocytes does not have much impact on the organisation of dense fibre networks : all that can be noted is that it may decrease slightly the local alignment rate of the fibres (compare the various panels in the second, fourth and fifth rows of Figure 5.2). The influence of the number of adipocytes is more pronounced over sparse fibre networks (Figure 5.3), where it can be noted that a medium adipocyte density may increase the level of local and global organisation of networks with small crosslinked fibre fraction (compare the left and middle panels of the second and fourth rows of Figure 5.3) whereas a high adipocyte density decreases the level of organisation especially in non-dynamical (first row) or highly crosslinked (third and fifth row) networks.

We recall that the local organisation of a fibre network can be characterised by its mean alignment rate Al_{sim} (defined in chapter 3, section 3.1) and its global organisation by the semi-major axis length of its stereographic projection covariance ellipse A_{max}, (defined in chapter 3, section 3.2). To take account of the stochastic variability of our model, parametric analyses are done on the average Al_{mean} and A_{max}^{mean} of these two quantifiers over 10 simulations conducted with the same set of parameters.

Hence, to investigate the impact of the adipocytes over the fibres organisation in a systematical manner, we plotted in Figure 5.4 the variations of the quantifiers Al_{mean} and A_{max}^{mean} according to the value of the adipocyte density ϕ_{ad}, all other parameters being fixed (values given in the legend). To facilitate the comparison of the various curves, we took as reference the value of these quantifiers for $\phi_{ad} = 0$ (no adipocyte in the modelled system). That is, for a given set of parameters (excluding ϕ_{ad}), we plotted $Al_{mean}(\phi_{ad}) - Al_{mean}(0)$ (with errorbars corresponding to $Al_{STD}(\phi_{ad})$) and $A_{max}^{mean}(\phi_{ad}) - A_{max}^{mean}(0)$ (with errorbars corresponding to $A_{max}^{STD}(\phi_{ad})$) as functions of ϕ_{ad}. These curves are displayed in Figure 5.4.

The first observation is that the adipocyte density does not have a significant impact on the degree of global organisation of the fibre network : the variations observed for the quantifier A_{max}^{mean} (top row of Figure 5.4) are largely covered by the standard deviations between simulations with the same parameters, which measure the stochastic variability due to the random initialisation and the random processes included the model ATmorpho-3D. In other word, the influence of the number of adipocytes on the global organisation of the fibre networks is not distinguishable from the stochastic variations due to the random components of the models.

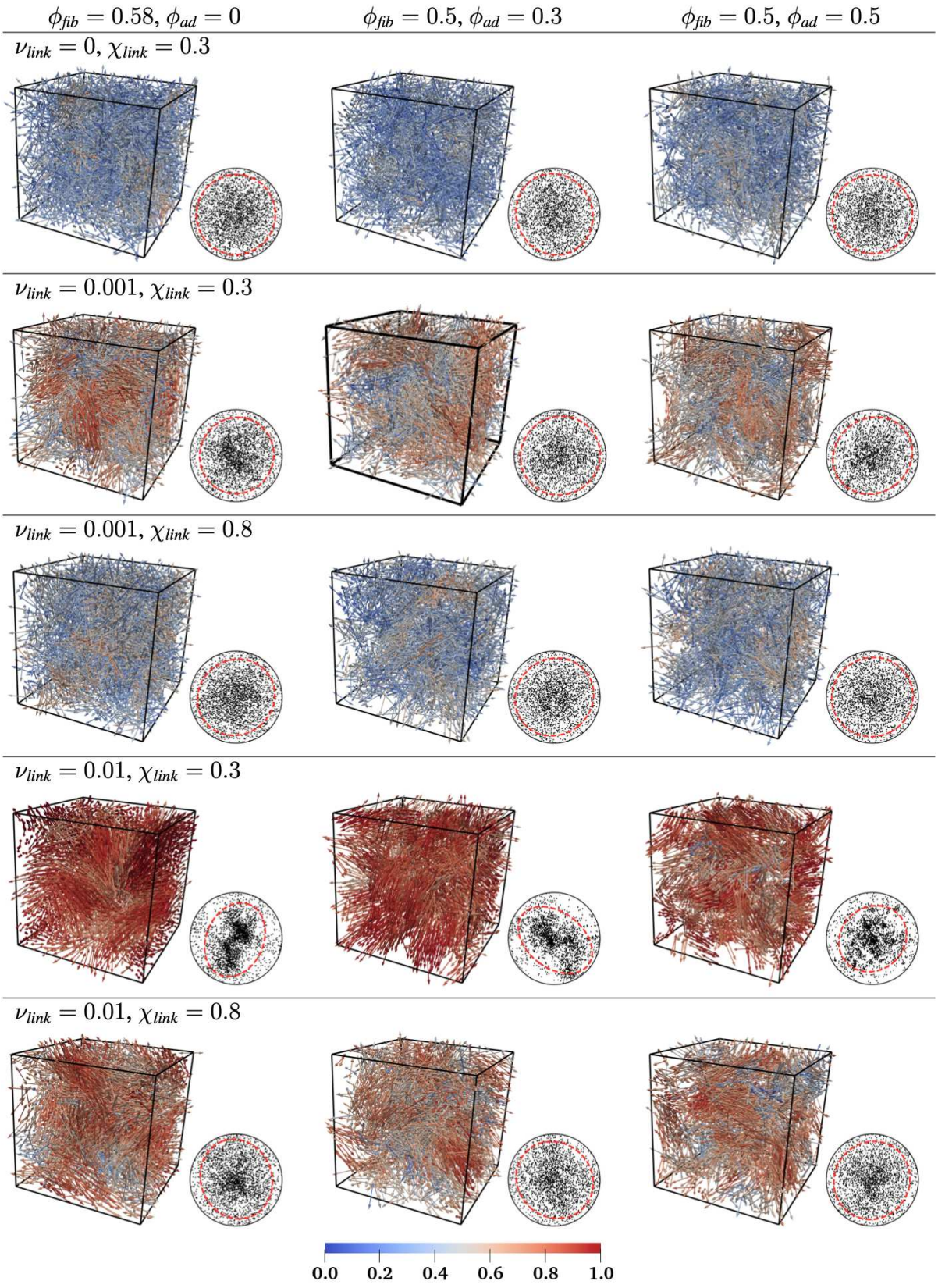


Figure 5.2: Illustration of the impact of the presence of cells on the structure of the fibre network, for systems with high fibre density ($\phi_{fib} = 0.5$) and various linking dynamics. Each panel display a 3D view of the fibre network at the end of the simulation ($T_{final} = 2000\ U_t$), with fibres coloured according to their alignment rate (blue: $Al_k = 0$, red: $Al_k = 1$), and the corresponding stereographic projection with covariance ellipse drawn in red dashed line. **From left to right :** no cell ($\phi_{ad} = 0$), low cell density ($\phi_{ad} = 0.3$) and high cell density ($\phi_{ad} = 0.5$).

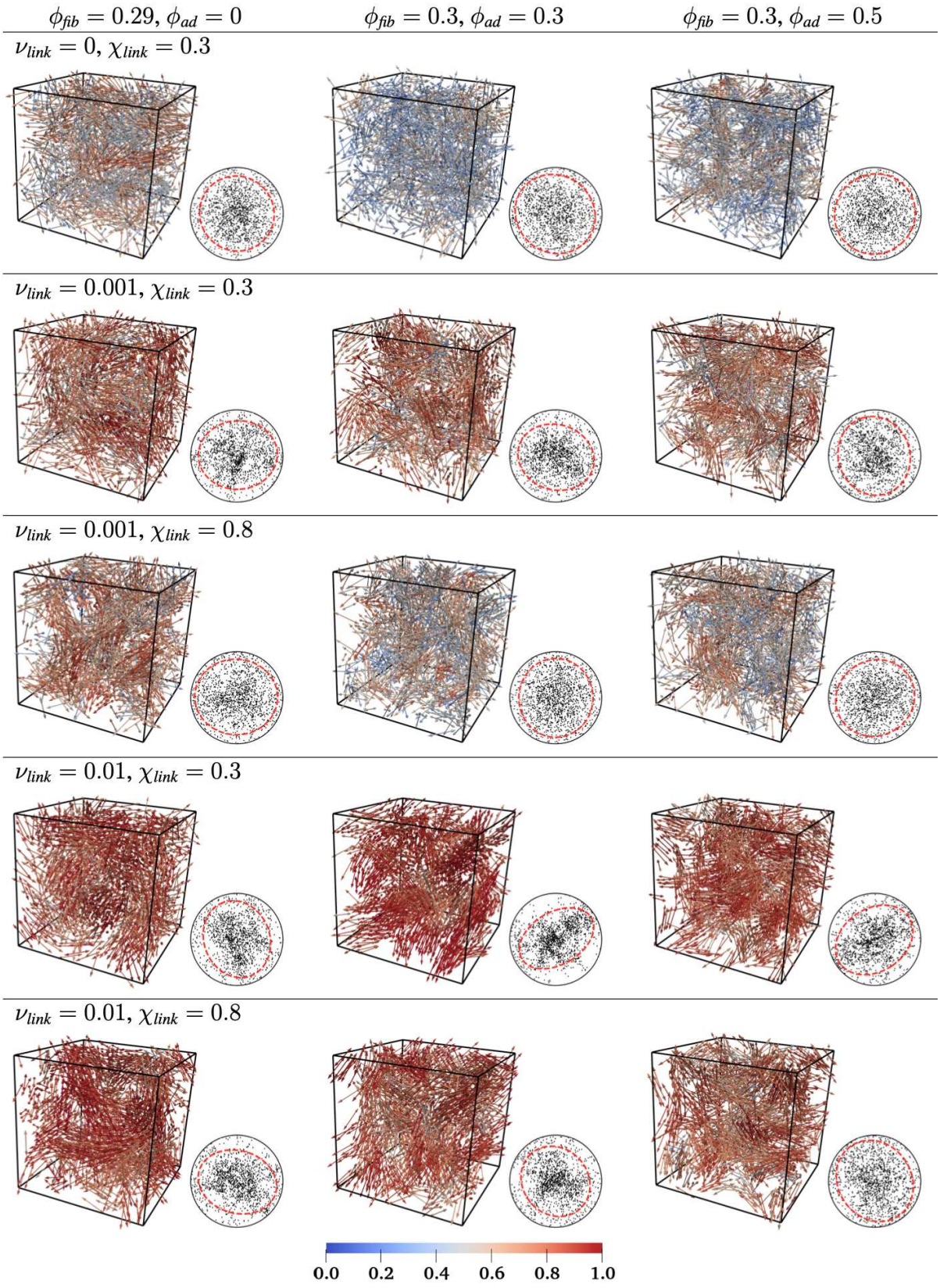


Figure 5.3: Illustration of the impact of the presence of cells on the structure of the fibre network, for systems with low fibre density ($\phi_{fib} = 0.3$) and various linking dynamics. Each panel display a 3D view of the fibre network at the end of the simulation ($T_{final} = 2000\ U_t$), with fibres coloured according to their alignment rate (blue: $Al_k = 0$, red: $Al_k = 1$), and the corresponding stereographic projection with covariance ellipse drawn in red dashed line. **From left to right :** no cell ($\phi_{ad} = 0$), low cell density ($\phi_{ad} = 0.3$) and high cell density ($\phi_{ad} = 0.5$).

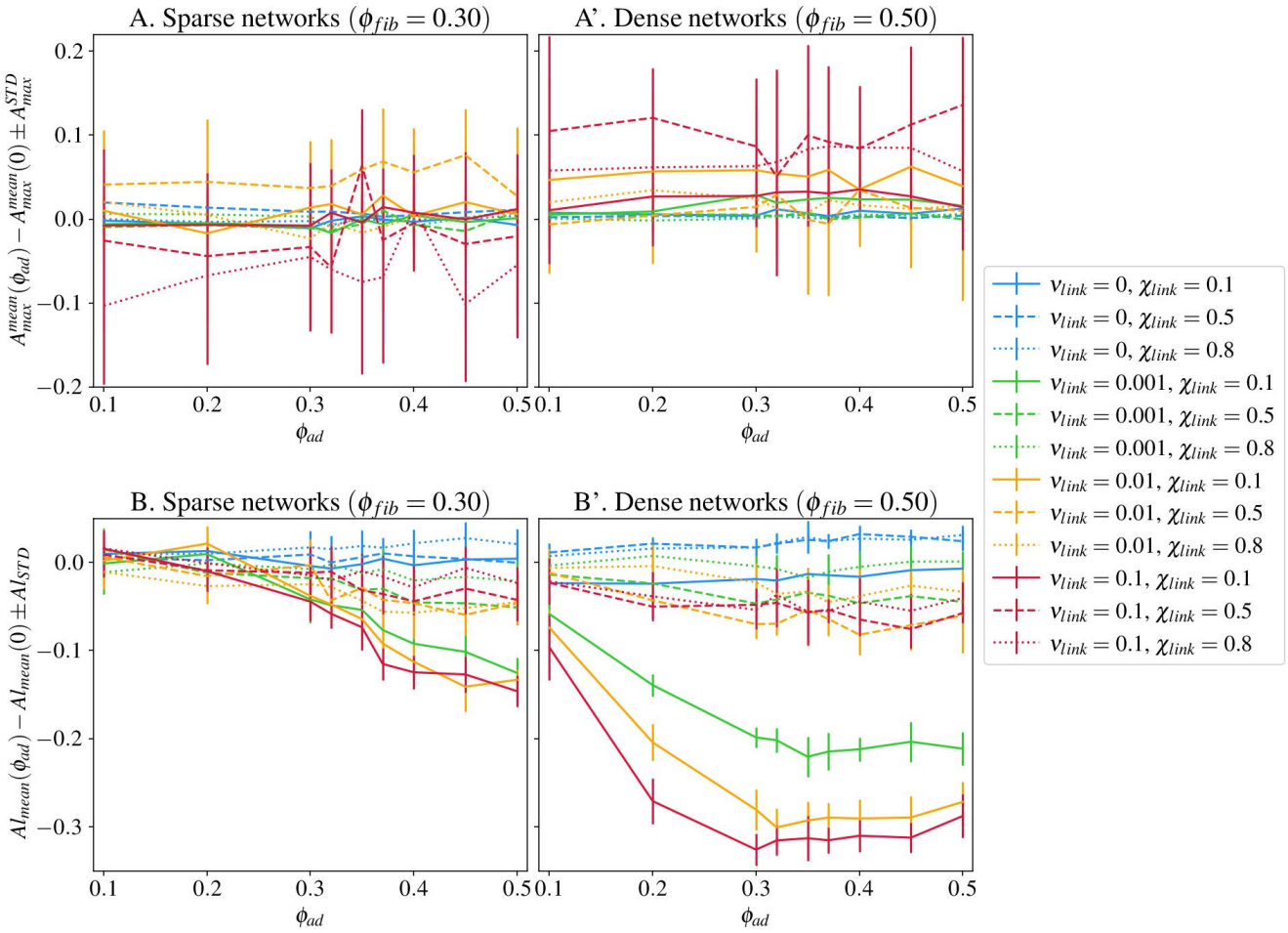


Figure 5.4: Variations of the quantifiers A_{max}^{mean} (top row) and Al_{mean} (bottom row) as a function of the adipocyte density ϕ_{ad}, compared to their value for $\phi_{ad} = 0$. Each curve represents the variation for a fixed set of parameters (excluding parameter ϕ_{ad}), with colour indicating the network remodelling rate ν_{link} and line type indicating the linked fibre fraction χ_{link}. Vertical errorbars indicate the value of A_{max}^{STD} and Al_{STD} respectively. **Left column :** Variation for sparse fibre networks ($\phi_{fib} = 0.3$). **Right column :** Variation for dense fibre networks ($\phi_{fib} = 0.5$).

On the contrary, the adipocyte density does have a significant, negative impact on the local alignment of dynamical fibre networks, measured by the quantifier Al_{mean} (see panels B and B' of Figure 5.4). Non-dynamical fibre networks, on the other hand, are unaffected by the presence of adipocytes (see blue curves in Figure 5.4.B and B'). The amplitude of the variations is stronger for dense networks (up to -0.34, see Figure 5.4.B') than sparse networks (up to -0.17, see Figure 5.4.B), and for networks with low linked fibre fraction (compare the curves with the same colour inside Figure 5.4.B and B'). We interpret this behaviour in terms of steric hindrance : the addition of adipocytes exerting a high repulsion force (see discussion on the parameters values in section 2) locally increases the fibre density by pushing the fibres together into bundles. Since dense fibre networks are, as their name indicates, relatively dense structures by themselves, the addition of adipocytes tend to make them overly constricted, quickly hindering their ability to align locally. Meanwhile, sparse fibre networks are less constricted to begin with and thus less sensitive to the congestion created by adipocytes. Alternatively, networks with a high linked fiber fraction are rigid enough to resist pressure from adipocytes and undergo a lesser increase in local fiber density

than networks with low linked fiber fraction.

Lastly, there is no discernible correlation between the quantifiers Al_{mean} and the values of $\phi_{ad} + \phi_{fib}$, $\frac{\phi_{ad}}{\phi_{fib}}$ or $\frac{\phi_{ad}}{\phi_{fib} + \phi_{ad}}$ (data not shown). This means that the increase in local fibre density can not be easily related to the value of ϕ_{ad} due to the influence of other structuring parameters such as ν_{link} and χ_{link}.

From all this, we conclude that the addition of round, growing cells to a dynamical fibre network (i) causes the fibres to bundle together in clusters with higher local fibre density than the overall density ϕ_{fib}, (ii) can modify the local alignment of the fibres by up to ± 0.15 and (iii) does not influence the global organisation of the fibres. This is a more limited impact than what could have been expected. However, this analysis remains mostly qualitative. The next step would be to measure the effective local fibre density of the systems, in order to quantitatively assess firstly the influence of the adipocyte density ϕ_{ad} over the spacial distribution of the fibres and secondly the relationship between the local fibre density and the quantifier Al_{mean}. An in-depth study of the ability (or not) of these potential fibre bundles to form planar structures is also required.

It should be noted that we did not recover the striking correlation between the number of links per fibre $N^{mean}_{linkperfib}$ and the average alignment rate Al_{mean} at equilibrium that was found in the previous chapter for fibre-only systems (data not shown). This may be because the networks considered here are not at equilibrium (see discussion on the parameters values in section 2). On the other hand, the transient crosslink accumulation phenomenon that we identified previously for networks with small unlinking frequency ν_{unlink} can be expected to play a predominant role in the network organisation on short time scale and to be reinforced by an increased local fibre density. This phenomenon thus have to be investigated to see if, for example, it can prevent the fibre bundles to organise into planar structures.

We will now turn to the study of the lobule-like structures produced by our model and show that, whereas the cells have little impact on the organisation of the fibres, the intrinsic properties of the fibre network exert a major influence on the cell clusters.

5. Characterisation of the lobules : comparison with *in vivo* data

Figure 5.5.A displays the repartition of the $266\ 420$ lobules produced in our 12600 simulations (that is an average of 21 lobules per modelled system) in the plane S_{lob} versus E_{lob}, with colour indicating the lobules volume V_{lob} (expressed in number of cells). For a better readability of both the density of lobule and average lobule volume at a given point of the phase plane, the diagram is presented in the form of a pixelated map (where each pixel is coloured according to the average volume of the lobules that fall within it) overlaid with density isolines indicating the number of lobules in the pixels. A few individual lobules are displayed below the diagram for illustration.

The first point to note is that the volume of the lobules is strongly correlated with their solidity : the larger the lobules, the less solid they are. This is counter-intuitive because discretisation bias would tend to have the opposite effect, considering that groups made of many spherical elements should have less difficulty describing a regular shape than groups made of only a few elements. It means that the observed correlation reflects a real geometrical difference between large lobules with irregular shapes (holes, digitations, etc.) and small lobules with regular, compact shapes. This can be seen by comparing the lobules n°6, 10 and 11 to the n°9, 14 and 15 in Figure 5.5.B.

On the other hand, there is no correlation between the volume of the lobules and their elongation : both small and large lobules range from round shapes ($E_{lob} \approx 0.4$) to unidirectional shapes ($E_{lob} \approx 1$).

The second observation is that the distribution of the lobules in the plane S_{lob}–E_{lob} shows a single,

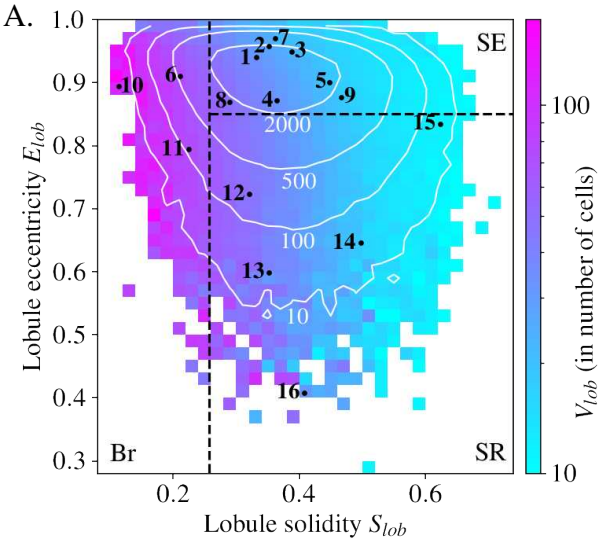
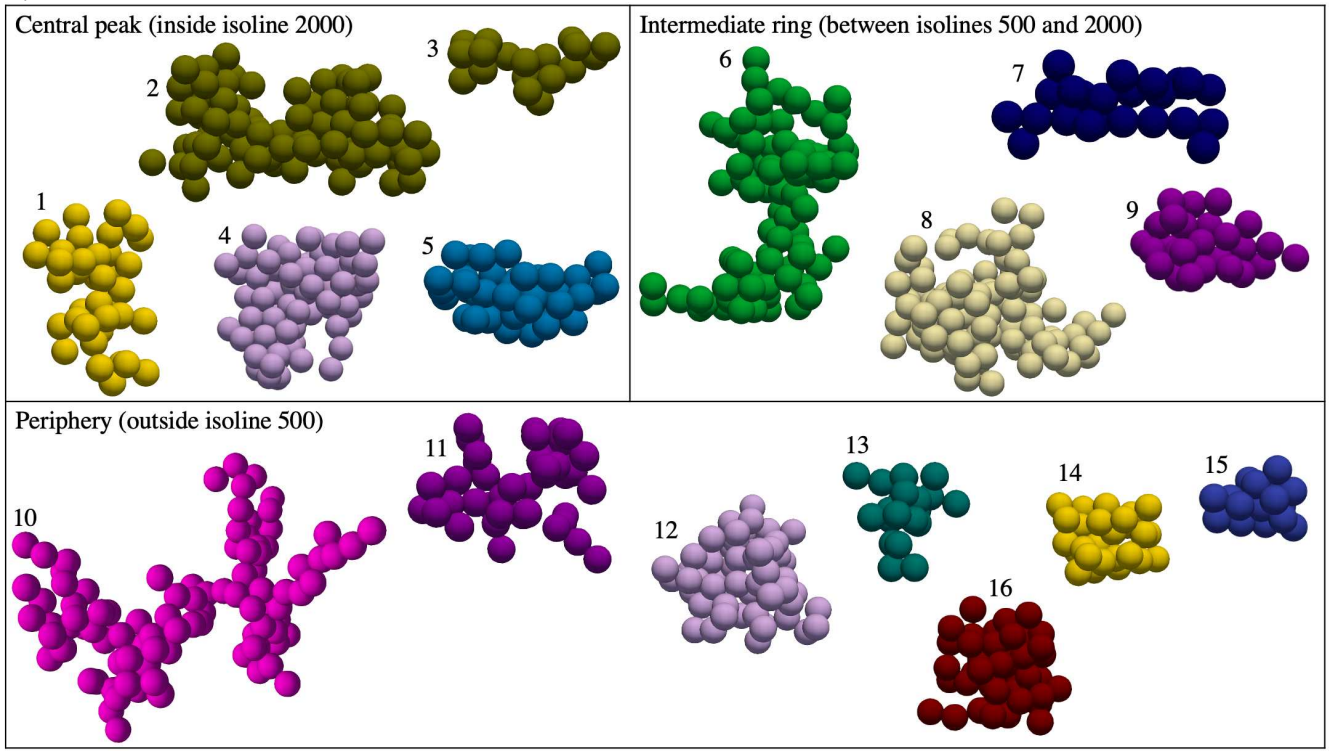


Figure 5.5: Characterisation of the various lobule morphologies produced by the model ATmorpho-3D. **A :** Distribution of all *in silico* lobules in the plane S_{lob}–E_{lob}. Each pixel is coloured according to the average volume (expressed in number of cells) of the lobules that fall within it, with a logarithmic scaling. Density isolines, indicating the number of lobules in the pixels, are drawn in white solid lines and the limits between each class of structures (branching (Br), solid elongated (SE) and solid rounded (SR)) in black dashed lines. The numbered black dots correspond to the lobules displayed in the table below. **B :** 3D view of a few isolated lobules from various simulations.

B.



very large peak centred on $S_{lob} = 0.34$ and $E_{lob} = 0.9$, meaning that the vast majority of our *in silico* lobules are moderately solid and very elongated. Based on this repartition, we choose to classify the lobules into three categories : (i) branching lobules (denoted as "Br") located on the left of the peak delineated by the 2000-isoline, that is $S_{lob} < S_{critic} = 0.257$; (ii) rounded solid lobules (denoted as "SR") located below the peak, that is $S_{lob} \geqslant S_{critic}$ and $E_{lob} < E_{critic} = 0.851$; and (iii) elongated solid lobules (denoted as "SE") corresponding to the peak itself ($S_{lob} \geqslant S_{critic}$ and $E_{lob} \geqslant E_{critic}$). The limits between these three categories are drawn on Figure 5.5.A in black dashed lines.

Note that branching lobules correspond almost exclusively to very large lobules, which as explained above have irregular shapes (see lobules n°6, 10 and 11 in Figure 5.5.B). Non-branching lobules, on the other hand, have fairly regular appearances with possibly a few short digitations. Rounded solid lobules do look round (see lobules n°12 to 16 in Figure 5.5.B) while elongated solid lobules have more complex

shapes sometimes evoking several lobes (see lobules n°1 to 5 and 7 to 9 in Figure 5.5.B). In fact, visually this category corresponds surprisingly well to the biological observations [109].

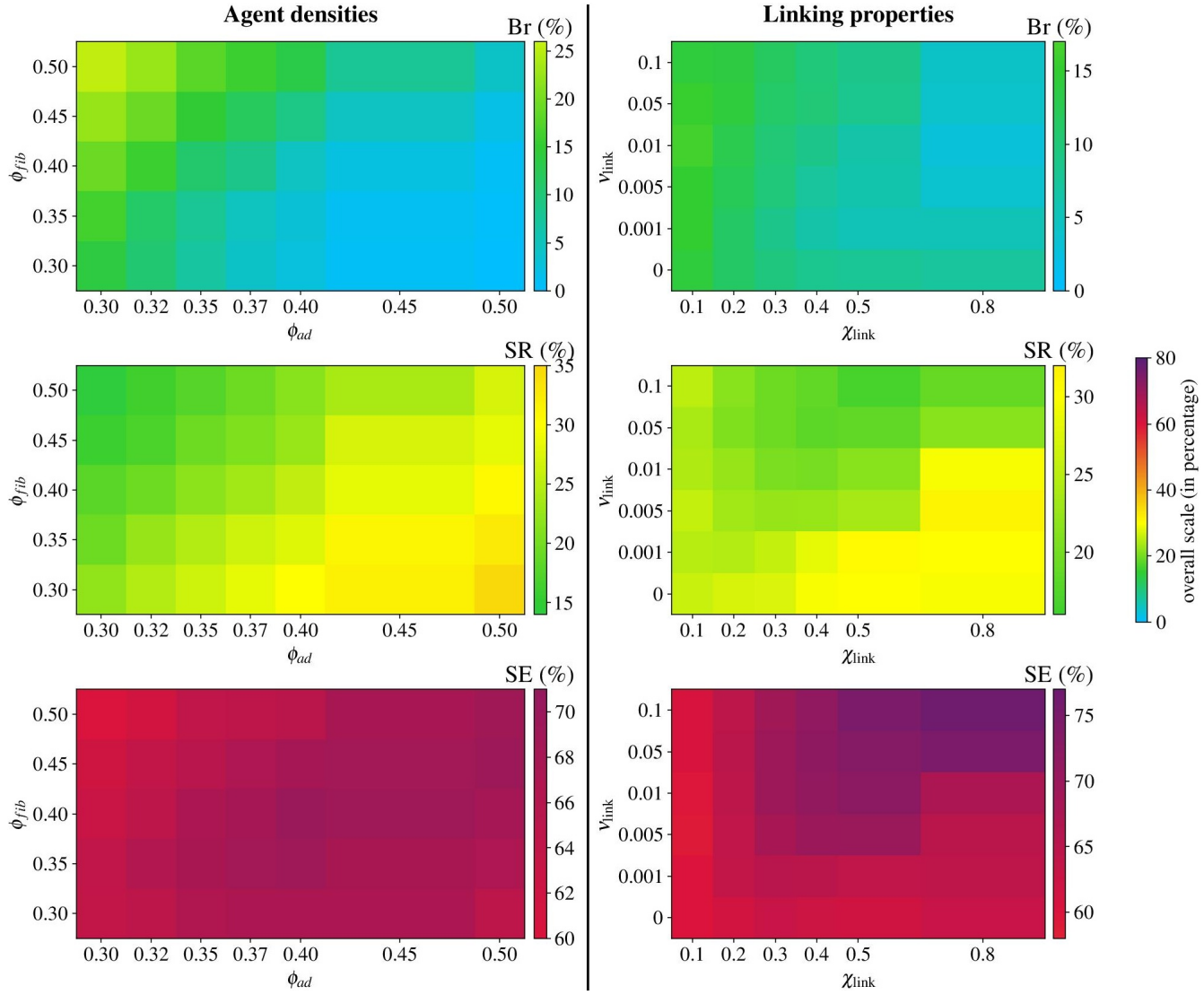


Figure 5.6: Distribution of the three categories of lobule morphology according to the model parameters. Each pixel of a map represents the percentage of lobules from a given morphological category (from top to bottom : branching (Br), solid rounded (SR) and solid elongated (SE)) found in all the simulations conducted with the indicated parameters. The overall colour bar, common to all panels, is displayed on the right-hand side of the figure. **Left column :** Distribution according to the agent densities ϕ_{ad} (in abscissa) and ϕ_{fib} (in ordinate). Each pixel represents the average value over 360 simulations. **Right column :** Distribution according to the linking characteristics χ_{link} (in abscissa) and ν_{link} (in ordinate). Each pixel represents the average value over 350 simulations.

Figure 5.6 displays heat-maps of the percentage of each type of lobules (from top to bottom : branching, solid rounded and solid elongated) in the system according to model parameters (agent densities ϕ_{ad} and ϕ_{fib} on the left and fibre linking properties ν_{link} and χ_{link} on the right). All heat-maps use the same colormap displayed on the right-hand side of the figure, with the appropriate range magnified on the right of each panel.

First, we note that solid elongated lobules are rather homogeneously distributed across the range of parameters tested, representing in average between 58% and 77% of all the lobules of a system.

Exploring the influence of the agent densities ϕ_{ad} and ϕ_{fib} (left column of Figure 5.6), we see that branching lobules occur quasi-exclusively in systems containing more fibres than cells ($\phi_{fib} > \phi_{ad}$), whereas solid rounded lobules occur more often in systems containing more cells than fibres and solid elongated lobules slightly favour dense, balanced systems ($\phi_{ad} \approx \phi_{fib} \geqslant 0.4$). Our interpretation is that the presence of a majority of fibres creates strong mechanical constraints on the cells which have to fit between them, promoting the formation of loose, dislocated branching cluster of cells not really spatially segregated from the fibres. On the other hand, a majority of cells will push back the fibres and form round clusters, which is the optimal geometric configuration for round agents. When the two types of agents are present in balanced quantities, mechanical constraints tend to separate them into distinct areas, creating bundles of fibres and clusters of cells, and the tendency towards alignment of the fibre bundles will force the cell clusters to take on elongated shapes.

Turning now to the analysis of the influence of the fibre linking properties ν_{link} and χ_{link} (right column of Figure 5.6), we observe a similar transition between the three lobule categories. Branching lobules occur either in non-dynamical fibre networks or in quickly remodelling, lowly crosslinked fibres networks. The former have very limited remodelling abilities which prevent them from making room for the cells, so that there will be no clear segregation between the two types of agents. The latter combine a low rigidity with an ability to self-organise rather quickly (see chapter 4, section 3.5), so that they will already have reached a high local alignment rate by the time a significant number of cells have appeared in the system. Their low rigidity will then prevent them from forcing the cells to organise into compact clusters, but they will also tend to maintain their local alignment which constitutes a local mechanical equilibrium. In consequence, there will again be no clear segregation between the two types of agents and the lobules will take tortuous geometries, branching around the fibre threads. On the other hand, solid rounded lobules occur preferentially in highly crosslinked, slowly remodelling fibre networks which we know (see chapter 4, section 3.5) are slow to self-organise, meaning that the cells appear in a rigid, nearly disordered environment exerting a strong, isotropic pressure. This leads to the formation of solid lobules with isotropic (i.e. round) shapes. Lastly, "balanced" networks which organise at moderate speed and have a moderate rigidity tend to favour the formation of solid lobules with elongated shapes dictated by fibre alignment.

Having characterised the different lobule morphologies produced by the model ATmorpho-3D, we can compare them to the morphologies of actual *in vivo* lobules. Figure 5.7.A displays the distribution in the plane S_{lob}–E_{lob} of the 55 *in vivo* lobules we segmented in four biological samples, using the protocol presented in chapter 3 (section 2.3). The limits between the three morphological classes defined above (branching, solid elongated and solid rounded) are drawn in black for comparison.

It can be seen that the distribution these *in vivo* lobules corresponds strikingly well to that of our *in silico* lobules, with a large majority of solid elongated lobules (74.5%), some solid rounded lobules (20%) and only a few branching lobules (5.5%). The existence of extremely elongated lobules, with $E_{lob} \approx 1$, may be noted. These lobules all pertain to the same biological sample, which was extracted from an older mouse than the other three samples (5 months instead of 2 months). Moreover, as the tissue sample was larger, it was not technically possible to image it entirely and most of the segmented lobules are truncated (they touch the border of the image). Hence, the specific lobule morphology observed in this sample may reflect a real, age-induced morphological difference or be due merely to the lowest quality of the segmentation. This calls for further investigation. In any case, collecting more biological samples would enable a more quantitative comparison between *in silico* and *in vivo* lobule morphologies.

On the other hand, the *in vivo* lobules do not exhibit any correlation between their shape and their

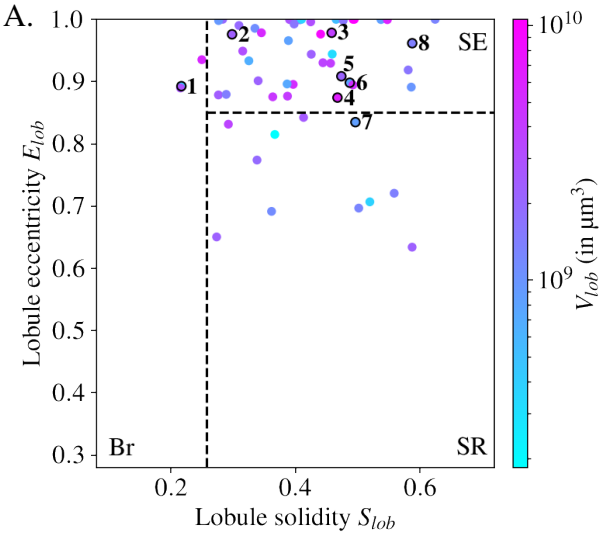
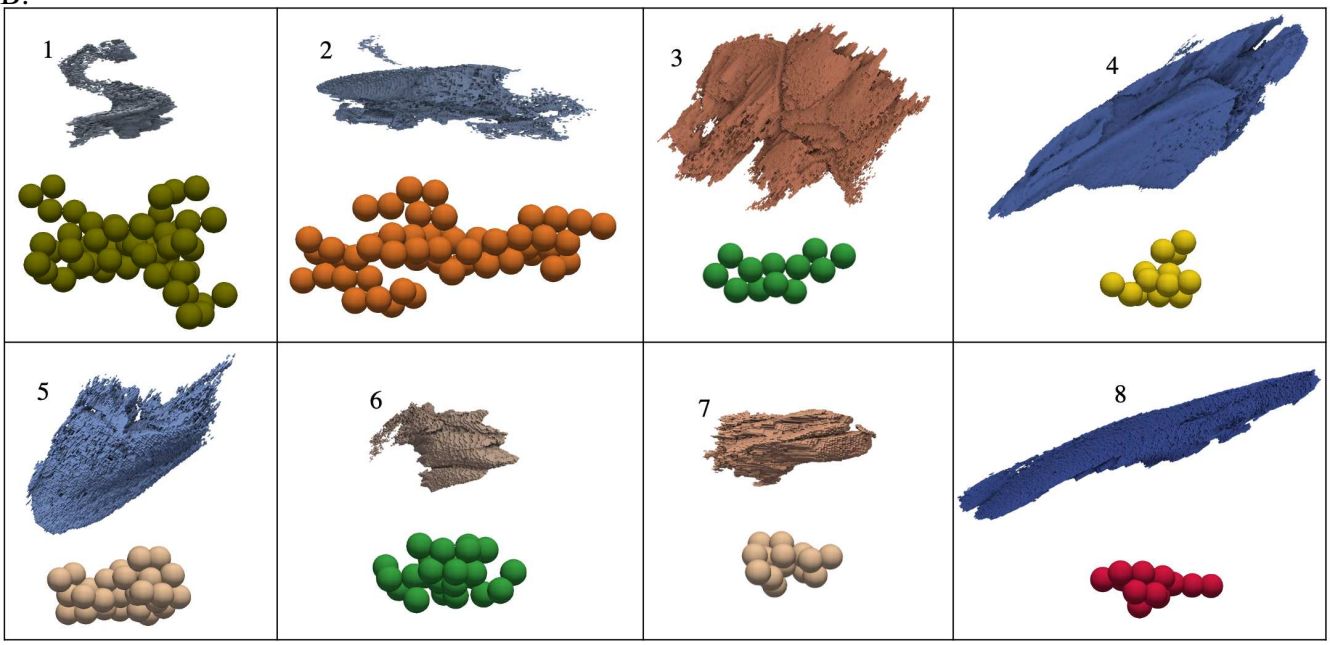


Figure 5.7: Characterisation of the various lobule morphologies observed in our *in vivo* samples. **A :** Distribution of the 55 lobules segmented in our *in vivo* samples in the plane S_{lob}–E_{lob}, colored according to their volume V_{lob} (in µm^3). The limits between each class of structures (branching (Br), solid elongated (SE) and solid rounded (SR)) are drawn in black dashed lines. The numbered black circles correspond to the lobules displayed in the table below. **B :** 3D view of a few isolated lobules from one sample, compared to the *in silico* lobule with the closest morphological characteristics.

B.



volume, contrarily to what was found in our *in silico* data. This may mean that the correlation produced by the model ATmorpho-3D does not match any biological reality, but it may also be an effect of the small amount of biological data used for this comparison and/or of the small size of the simulation domain. Indeed, we did not expect to be able to compare the volume of the *in silico* and *in vivo* structures given that the spatial domain used for the simulations have a side length equivalent to 1 mm whereas the biological samples are approximately 8 times larger. In consequences, the largest *in silico* lobule recorded have a volume $V_{lob} = 356$ cells $\approx 2.10^8$ µm^3 comparable to that of the smallest *in vivo* lobules.

As in the previous chapter, the small size of the simulation domain is a clear limitation to our study but a necessary choice considering the computational cost of our model. It allowed use to explore a large range of parameters, observing and characterising the various types of structures that can be generated by the model and their dependance to the parameters. An important validation step would be to reproduce these results in a larger spatial domain and to compare them to more numerous biological samples. It is already noteworthy that the difference between the smallest and the largest lobules is of the order of a $1 : 50$ ratio for both the *in vivo* and *in silico* data, indicating that the size variability of our model is

comparable to that found in biological tissues.

Figure 5.7.B illustrates the various lobule morphologies that can be found in a single biological sample and compares them to the *in silico* lobule with the closest values of E_{lob} and S_{lob}. It can be seen that these structures compare qualitatively well, except in the case n°3 where the difference in volume is too high to allow the *in silico* lobule to display a real poly-lobed shape. This highlights the relevance of the morphological quantifiers E_{lob} and S_{lob}, which enable a quantitative comparison between *in vivo* and *in silico* lobules.

6. Discussion

In this work, we have implemented a 3D model for a growing population of cells evolving in an interconnected and dynamical ECM fibre network. Our mathematical model is able to reproduce, in 3D, a tissue architecture composed of clustered adipocytes in lobular units, surrounded by sheet-like structures of ECM fibres, with morphological characteristics similar to the ones computed on experimental images of mature AT tissues. The originality of the model lies in its apparent simplicity. It demonstrates that the emergence of the 3D architecture of a complex biological tissue such as the AT could be the result of a self-organisation process driven by simple mechanical interactions between few components, here the adipocyte cells and the ECM fibres, with very few biological phenomena – namely fibre remodelling (chemical crosslinking/unlinking), cell insemination (pre-adipocyte differentiation) and cell growth (lipid storage). Furthermore, the emergent architecture appears as mainly driven by fibre interactions, suggesting a secondary role for cells in this process.

To our knowledge, our model is the first 3D individual-based model interfaced with experimental data for cells interacting within a dynamical fibrous network. This model is the natural extension of the 2D model of AT morphogenesis studied in [17], which focused on the cross-sectional organisation of the lobules. Although restricted to the 2D case, this previous model represented a proof-of-concept that lobular organisation in tissue morphogenesis could be mainly driven by mechanical interactions between the cells and the ECM. It was later extended to study the mechanisms of tissue reconstruction after injury [37], again pointing to the crucial role of ECM connectivity in the fate of injury outcome. This led to the development of a combined study involving *in silico* and *in vivo* experiments, which suggested a new putative target for controlling ECM connectivity to induce regeneration in otherwise scarring tissues [156], conferring to the 2D model a "digital twin" aspect. The extension to 3D that is presented in this thesis represents an important and necessary step forward in the conception of a physiologically relevant model of tissue architecture. Indeed, growing amount of evidence highlight great differences in cell function and behaviour between 2D and 3D cultures (see [157] and references therein). Therefore, 3D models for tissue structuring could prove to be invaluable tools for studying real biological systems, with applications in particular to cell migration [158], cancer invasion [159] and more generally in the field of tissue engineering [160]. It is noteworthy that our 3D model is able to generate realistic tissue architectures without the need for a complex geometric representation of the individual fibres. Indeed, the model produces septa (fibrous sheet-like structures surrounding the lobular cell clusters) emerging from mechanical interactions between spherocylindrical crosslinked units (fibre elements) and spherical objects (adipocytes). The fact that this simple modelling of the fibre units is able to produce complex two-dimensional surfaces surrounding cell clusters constitutes another surprising and major result of this study.

Our model can spontaneously generate various types of macrostructures. We classified the different cell clusters morphologies into three types : solid and elongated clusters (SE), solid and round clusters (SR), and branching structures with many digitations (Br). By coupling image segmentation techniques

and morphological quantifiers, we found that the lobules segmented from experimental images belonged mostly to the first category (SE), a few to the second (SR) and even less to the third (Br). Surprisingly, the different structures emerging from our model by playing on the different parameters followed the same repartition. The model was therefore able to reproduce different lobule geometries, with the same characteristics and distribution as the ones observed in real tissues. We note however that some experimental samples showed a higher number of very elongated lobules compared to our simulations.

A profound parametric analysis of the model enabled us to disentangle the respective role of the different parameters on the tissue architecture. We first found that, surprisingly, the presence of cells in the ECM did not impact significantly the ECM global organisation, although it could decrease the local alignment of the fibres according to the type of network they evolve into. Indeed, we found that cells could negatively impact the local alignment quantifier in dense and sparse ECM by 35% and 17% respectively. The segregation of the spherical objects (cells) and spherocylindrical units (fibre elements) and the emergence of a nematic order (fibre alignment) revealed in our model is consistent with several theoretical and experimental studies on binary mixtures of isotropic and anisotropic agents in 3D [161–163]. In particular, several simulations featuring hard spherocylinders and spheres [164, 165] showed that the addition of the spherical crowders have a negative impact upon the stability of rod-like particles in a nematic phase (alignment). In all these works, the phase behaviour is regarded as entropically driven via excluded volume effects, known as "depletion interaction" or "macromolecular crowding", between differently shaped particles. The different arrangements of fibres and cells in our model bear evident analogies with these observations, although the interactions in our model stem from soft-repulsive potentials instead of volume exclusion effects. As a result, our model is able to generate segregation in the absence of noise because the interaction remains active even without noise, contrary to the case of sole volume exclusion effects. Another fundamental difference compared to previous studies is the active nematic alignment of the rod-like elements due to the alignment torque at the (dynamical) crosslinks. This enables to study the impact of spherical particles in actively remodelling networks and points to the essential role of ECM rigidity on tissue architecture, results not accessible in binary mixture of passive particles interacting only via volume exclusion effects. Finally, the previous studies focused on the nematic order of the rod-like particles with little interest in the morphology of the clusters composed by the spherical particles. Here, we also provide a deep analysis of the cell structures, showing that the nematic ordering of the ECM network reflects in the elongation of the spherical particles clusters. This constitutes an interesting result and, to our knowledge, a novel quantitative exploration of this effect.

The parametric analysis on the cell structures showed that the different lobule morphologies were mostly controlled by the ECM network remodelling characteristics. Indeed, (Br) structures, although rarer in general, were favoured by networks with a low proportion of crosslinks (lowly constrained networks), while (SR) structures were found more often in slow remodelling networks with a large proportion of crosslinks (very constrained networks, rigid). Finally, the more common lobule structures (SE) were found more often in highly dynamical networks with a large proportion of crosslinks (moderately constrained networks with large plasticity). Altogether, these results showed that ECM rigidity was determinant in the shape of the lobule structures, by transmitting mechanical constraints to the growing cell population. Indeed, lowly constrained networks did not seem to exert enough mechanical action on the cell population, favouring the appearance of less organised cell clusters with tortuous geometries (Br). On the contrary, the fibre structures in rigid networks were less able to spontaneously self-organise, leading to more disordered networks exerting high mechanical pressure isotopically on the cell population. As a result, the cells were forced to regroup into compact and round clusters (SR). In dynamical and highly crosslinked networks however, fibres could arrange more easily into organised structures thanks to the torque acting on crosslinked fibres, imposing local directional constraints to the cell clusters, and favouring cell cluster elongation (SE). All these results point again to the major role of ECM mechanical

constraints in tissue architecture, and are consistent with several results demonstrating the impact of the ECM and its dysregulation (fibrosis) on AT functions [166].

Further improvements of the model could be made. If the choice of modelling cells as non-deformable 3D spheres proved to be sufficient for reproducing realistic AT architectures, considering more realistic deformable cell models such as [167] would enable to generalise the model to different systems where cell deformation plays an important role [168–170]. On the biological viewpoint, designing new experiments allowing for *in vivo* tracking of ECM organisation and cell positions during AT development would provide invaluable data on the time dynamics of the morphogenesis process, and enable to quantitatively compare the model predictions with the fibrous structures in normal or pathological conditions.

A key issue about the relevance of our model is the observation that, during development, newly differentiated adipocytes appear close to the existing differentiated adipocytes, suggesting key local cues for lobule emergence potentially associated with energy partitioning due to vasculature that generates heterogenous and gradient distribution. In our 3D model, the energy required for cell growth and ECM remodelling is prescribed and equally partitioned across the whole volume of the simulation. However, in the 2D model we previously demonstrated that random or biased insemination led to the same structuring, suggesting that the most efficient vasculature architecture compatible with lobule emergence has been selected during evolution. From an evolutionary point of view, the vasculature structure should be a consequence of mechanical constraints. Such hypothesis should also be tested in the 3D model.

In the same vein, the model could be extended to account for energy exchanges and metabolism to model the long-term evolution of tissue architecture (ageing, changes due to diet, obesity, etc). These aspects will be the subject of future works and promise exciting challenges such as the determination of complex feedback loops between energy intakes and local growth laws. Moreover, the model showed that cells had a relatively small mechanical impact on the ECM global architecture and that rather the main mechanical driver of tissue structuring lied in ECM rigidity. However, the model did not account for chemical interactions between the cell and the ECM. A natural perspective of this work would be to consider cells having chemical interactions with the ECM. For instance, one could consider the linking/unlinking of the fibres to be dependent on the local density of cells, to model matrix degradation by metalloproteinase generated by the adipocytes as evidenced in several biological systems [171]. Size or pressure-dependent growth laws for the cells and for matrix remodelling could also be implemented, with the aim to explore the coupling between energy exchanges and mechanics, in light of several results in the literature [172–174].

On the mathematical viewpoint, the derivation of macroscopic models for the cell and/or fibrous phases of the present model would produce a computationally efficient continuous model containing as much as possible the mechanisms of the microscopic scale, enabling us to address the question of the architecture at the level of the whole tissue. A step in this direction has been made in [42] where a continuum model for crosslinked fibre was derived from an individual-based model and further analysed in [175]. Several continuum models for the cell phase have also been derived from interacting particle systems [176–180], but major works still remain to explore how specific microscopic effects translate at the macroscopic scale for binary mixtures of isotropic and anisotropic particles.

Whatever the case and whatever the outlook, our model appears as relevant to mimic the emergence of tissue architecture and to decipher its basic rules. Its simplicity questions about its links with molecular cues usually investigated in biology but represents a key advantage for its manipulation and its use to test numerous hypotheses.

Conclusion

In this thesis we studied the emergence of the 3D architecture of biological tissues, and especially connective tissues which are the most abundant in animals. We developed two Individual-Based Models (IBMs), one aimed at reproducing a non-specialised connective tissue and reduced to the Extra-Cellular Matrix (ECM) component, and another designed for a type of specialised connective tissue : the Adipose Tissue (AT). Both models are based on the hypothesis that a tissue architecture could emerge from local mechanical interactions between its components, with a minimalistic representation of the biological processes involved.

To visualise the structures produced by our models, we developed an automated visualisation pipeline using the Paraview software, presented in chapter 3. In order to compare the cell clusters observed in our numerical simulations with those found in biological tissue samples, we developed a Python segmentation protocol to automatically detect cell clusters from both *in vivo* and *in silico* 3D images. These two tools are highly generic and can be used to study any system made of spherical and/or rod-like objects, in and outside the field of biology. During their development, we have taken great care to make them as user-friendly as possible and to compel with open-source format. They are freely available online along with detailed user manuals, and will be submitted as a methodological article. We also constructed various numerical quantifiers to characterise both the fibrous and cellular structures produced by our two models.

The first model, ECMmorpho-3D, is studied in detail in chapter 4 of the present manuscript. It enabled us to identify the key mechanisms governing the self-organisation of 3D dynamical fibre networks and to suggest an emerging variable, namely the number of crosslinks per fibre (or network connectivity), as decisive for the fate of the simulation and a potential therapeutic target for regulating the architecture of biological fibre networks.

The second model, ATmorpho-3D, is studied in chapter 5 and confronted to biological data acquired through 3D imaging of *in vivo* tissue samples. We showed that our model generates 3D cellular structures whose morphological characteristics are very close to those of the structures segmented from *in vivo* samples. The fact that this very simple mathematical model yields results close to what has hitherto appeared to be a complex biological reality calls for two considerations. Firstly, the biological complexity may only be apparent and cover simple general organising principles. Secondly, the founding hypotheses of the model seem to be correct concerning these underlying principles : biological tissue architecture could stem simply from local mechanical interactions.

In this respect, it should be noted that the only modification we carried out concerned mechanical repulsion interactions, which were initially modelled by linear forces. This emphasises the key role of mechanical forces in ours models. In the initial version, the effect of the repulsion forces depended on the relative size of the agents in a non-physical manner. This prompted us to change for a more complex, but physically more relevant, modelling based on the Hertzian theory of contact. This need for accuracy regarding the mechanical interactions again underlines their predominant role in the emergence of tissue architecture.

For the future, several perspectives can be proposed. First, it would be very interesting to try to apply our models to other types of biological tissues. The structures produced by the model ATmorpho-3D could be compared to tissues displaying a lobular architecture similar to that of subcutaneous AT : visceral AT depots, liver, pancreas, etc. Minimalistic modifications, bearing not on the core of the models but on the boundary conditions used during numerical simulations, could enable us to mimic tissues subjected by their environment to major mechanical constraints. For example, coupling the model ECMmorpho-3D with Neumann boundary conditions (i.e. elements located on the border can not move) on two opposite

sides of the simulation domain could allow us to mimic muscles and sinews. Applying Dirichlet boundary conditions (i.e. no element can pass through the border) on one side of the domain could model the presence of a physical frontier for the tissue, such as the skin or a nearby gland. Additionally, the implementation of these new types of boundary conditions would enable us to conduct tension/compression and shearing experiments to characterise the mechanical properties of our *in silico* systems, such as their elasticity, plasticity and potential breakdown limit.

Another perspective is to confer other properties to the cellular agents to induce local heterogeneity in the mechanical interactions, such as the ability to modify the linking and unlinking probabilities of nearby fibres according to the cell size. Another change would be to enable physical links between cells and fibres to mimic the adhesion capacities of certain cells.

Lastly, another possible direction for future works would be to adapt the models (and their potential variations) to other contexts beyond that of tissue morphogenesis. For instance, they could be used to study tissue reconstruction after injury, following the work of [37] in 2D.

All this discussion highlights the generic aspect of our models, which can potentially be adapted to many types of tissues and biological contexts. This opens exciting perspectives concerning our understanding of how biological tissue architecture emerges and how alteration of basic mechanical properties in various diseases can lead to a loss of tissue architecture and function.

Bibliography

[1] Leah Edelstein-Keshet. *Mathematical Models in Biology*. Society for Industrial and Applied Mathematics, Philadelphia (USA), 2005. doi:10.1137/1.9780898719147.

[2] Ricardo López-Ruiz. "Mathematical Biology: Modeling, Analysis, and Simulations". *Mathematics*, 10(20):3892, Oct 2022. doi:10.3390/math10203892.

[3] Ernst Mayr. *The Growth of Biological Thought: Diversity, Evolution and Inheritance.* The Belknap Press of Harvard University Press, Cambridge (MA) and London (England), 1982. ISBN 0674364465.

[4] Alexey Kolodkin. "Systems Biology Through the Concept of Emergence". In *Philosophy of Systems Biology: Perspectives from Scientists and Philosophers*, edited by Sara Green, pp. 181–191. Springer International Publishing, Cham, 2017. ISBN 978-3-319-47000-9. doi:10.1007/978-3-319-47000-9_17.

[5] Emily Herring and Gregory Radick. "Emergence in Biology". In *The Routledge Handbook of Emergence*, edited by Sophie Gibb, Robin Findlay Hendry and Tom Lancaster, chapter 28, pp. 352–362. Routledge, Abingdon (Oxon, UK), 2019. ISBN 9781138925083. doi:10.4324/9781315675213-29.

[6] Richard N. Mitchell and Frederick J. Schoen. "Functional Tissue Architecture, Homeostasis, and Responses to Injury". In *Biomaterials Science*, edited by William R. Wagner, Shelly E. Sakiyama-Elbert, Guigen Zhang and Michael J. Yaszemski, pp. 683–700. Academic Press, 4th edition, 2020. ISBN 978-0-12-816137-1. doi:10.1016/B978-0-12-816137-1.00044-1.

[7] Jiaming Zhong, Wei Huang and Huamin Zhou. "Multifunctionality in Nature: Structure–Function Relationships in Biological Materials". *Biomimetics*, 8(3):284, Jul 2023. doi:10.3390/biomimetics8030284.

[8] Dirk Drasdo. "On Selected Individual-based Approaches to the Dynamics in Multicellular Systems". In *Polymer and Cell Dynamics: Multiscale Modeling and Numerical Simulations*, edited by W. Alt, M. Chaplain, M. Griebel and J. Lenz, pp. 169–203. Birkhäuser, Basel, 2003. ISBN 978-3-0348-8043-5. doi:10.1007/978-3-0348-8043-5_15.

[9] Minki Hwang, Marc Garbey, Scott A. Berceli and Roger Tran-Son-Tay. "Rule-Based Simulation of Multi-Cellular Biological Systems – A Review of Modeling Techniques". *Cellular and Molecular Bioengineering*, 2(3):285–294, 2009. doi:10.1007/s12195-009-0078-2.

[10] Dirk Drasdo and Stefan Höhme. "A single-cell-based model of tumor growth in vitro: monolayers and spheroids". *Physical Biology*, 2(3):133–147, Jul 2005. doi:10.1088/1478-3975/2/3/001.

[11] David A. Head, Alex J. Levine and Frederick C. MacKintosh. "Distinct regimes of elastic response and deformation modes of cross-linked cytoskeletal and semiflexible polymer networks". *Physical Review E*, 68(6), 2003. doi:10.1103/PhysRevE.68.061907.

[12] Jan Wilhelm and Erwin Frey. "Elasticity of Stiff Polymer Networks". *Physical Review Letters*, 91(10):108 103, Sep 2003. doi:10.1103/PhysRevLett.91.108103.

[13] Jan A. Åström, P. B. Sunil Kumar, Ilpo Vattulainen and Mikko Karttunen. "Strain hardening, avalanches, and strain softening in dense cross-linked actin networks". *Physical Review E*, 77(5):051 913, May 2008. doi:10.1103/PhysRevE.77.051913.

[14] Mo Bai, Andrew R. Missel, Alex J. Levine and William S. Klug. "On the role of the filament length distribution in the mechanics of semiflexible networks". *Acta Biomaterialia*, 7(5):2109–2118, May 2011. doi:10.1016/j.actbio.2010.12.025.

[15] Ali Shahsavari and R. Catalin Picu. "Model selection for athermal cross-linked fiber networks". *Physical Review E*, 86(1):011 923, Jul 2012. doi:10.1103/PhysRevE.86.011923.

[16] Ricardo Alonso, Jennifer Young and Yingda Cheng. "A Particle Interaction Model for the Simulation of Biological, Cross-Linked Fiber Networks Inspired From flocking Theory". *Cellular and Molecular Bioengineering*, 7:58–72, 2014. doi:10.1007/s12195-013-0308-5.

[17] Diane Peurichard, Fanny Delebecque, Anne Lorsignol, Corinne Barreau, Jacques Rouquette, Xavier Descombes, Louis Casteilla and Pierre Degond. "Simple mechanical cues could explain adipose tissue morphology". *Journal of Theoretical Biology*, 429:61–81, Jun 2017. doi:10.1016/j.jtbi.2017.06.030.

[18] Stefano Guido and Robert T. Tranquillo. "A methodology for the systematic and quantitative study of cell contact guidance in oriented collagen gels. Correlation of fibroblast orientation and gel birefringence". *Journal of Cell Science*, 105(2):317–331, Jun 1993. doi:10.1242/jcs.105.2.317.

[19] Dewi Harjanto and Muhammad H. Zaman. "Modeling Extracellular Matrix Reorganization in 3D Environments". *PLOS ONE*, 8(1):e52 509, Jan 2013. doi:10.1371/journal.pone.0052509.

[20] Michael Mak. "Impact of crosslink heterogeneity on extracellular matrix mechanics and remodeling". *Computational and Structural Biotechnology Journal*, 18:3969–3976, 2020. doi:10.1016/j.csbj.2020.11.038.

[21] James D. Murray. *Mathematical Biology*. Springer, 3rd edition, Jun 2013. ISBN 978-1-4757-7709-3. doi:10.1007/b98868.

[22] Nicola Bellomo. *Modeling Complex Living Systems: A Kinetic Theory and Stochastic Game Approach*. Birkhäuser, Basel, 1st edition, 2007. ISBN 0817645101.

[23] Larry A. Taber, Yunfei Shi, Le Yang and Philip V. Bayly. "A poroelastic model for cell crawling including mechanical coupling between cytoskeletal contraction and actin polymerization". *Journal of Mechanics of Materials and Structures*, 6:569–589, Jun 2011. doi:10.2140/jomms.2011.6.569.

[24] Jean-François Joanny, Frank Jülicher, Karsten Kruse and Jacques Prost. "Hydrodynamic theory for multi-component active polar gels". *New Journal of Physics*, 9(11):422–422, Nov 2007. doi:10.1088/1367-2630/9/11/422.

[25] Mohd Suhail Rizvi, Prasoon Kumar, Dhirendra S. Katti and Anupam Pal. "Mathematical model of mechanical behavior of micro/nanofibrous materials designed for extracellular matrix substitutes". *Acta Biomaterialia*, 8(11):4111–4122, Nov 2012. doi:10.1016/j.actbio.2012.07.025.

[26] John C. Dallon, Jonathan A. Sherratt and Philip K. Maini. "Mathematical Modelling of Extracellular Matrix Dynamics using Discrete Cells: Fiber Orientation and Tissue Regeneration". *Journal of Theoretical Biology*, 199(4):449–471, Aug 1999. doi:10.1006/jtbi.1999.0971.

[27] Brian Anthony DiDonna and Alex J. Levine. "Filamin Cross-Linked Semiflexible Networks: Fragility under Strain". *Physical Review Letters*, 97(6):068 104, Sep 2006. doi:10.1103/PhysRevLett.97.068104.

[28] Claus Heussinger and Erwin Frey. "Floppy Modes and Nonaffine Deformations in Random Fiber Networks". *Physical Review Letters*, 97(10):105 501, Sep 2006. doi:10.1103/PhysRevLett.97.105501.

[29] Eric Bertin, Hugues Chaté, Francesco Ginelli, Shradha Mishra, Anton Peshkov and Sriram Ramaswamy. "Mesoscopic theory for fluctuating active nematics". *New Journal of Physics*, 15(8):085 032, Aug 2013. doi:10.1088/1367-2630/15/8/085032.

[30] Emmanuel Boissard, Pierre Degond and Sebastien Motsch. "Trail formation based on directed pheromone deposition". *Journal of Mathematical Biology*, 66:1267–1301, 2013. doi:10.1007/s00285-012-0529-6.

[31] Michael J. Unterberger and Gerhard A. Holzapfel. "Advances in the mechanical modeling of filamentous actin and its cross-linked networks on multiple scales". *Biomechanics and Modeling in Mechanobiology*, 13(6):1155–1174, Nov 2014. doi:10.1007/s10237-014-0578-4.

[32] Elisabeth M. Huisman, Teun van Dillen, Patrick R. Onck and Erik Van der Giessen. "Three-Dimensional Cross-Linked F-Actin Networks: Relation between Network Architecture and Mechanical Behavior". *Physical Review Letters*, 99(20):208 103, Nov 2007. doi:10.1103/PhysRevLett.99.208103.

[33] Chase P. Broedersz, Michael Sheinman and Frederick C. MacKintosh. "Filament-Length-Controlled Elasticity in 3D Fiber Networks". *Physical Review Letters*, 108(7):078 102, Feb 2012. doi:10.1103/PhysRevLett.108.078102.

[34] Byoungkoo Lee, Xin Zhou, Kristin Riching, Kevin W. Eliceiri, Patricia J. Keely, Scott A. Guelcher, Alissa M. Weaver and Yi Jiang. "A Three-Dimensional Computational Model of Collagen Network Mechanics". *PLOS ONE*, 9(11):1–12, Nov 2014. doi:10.1371/journal.pone.0111896.

[35] Shengmao Lin and Linxia Gu. "Influence of Crosslink Density and Stiffness on Mechanical Properties of Type I Collagen Gel". *Materials*, 8(2):551–560, 2015. doi:10.3390/ma8020551.

[36] R. Catalin Picu and Ahmed Sengab. "Structural evolution and stability of non-crosslinked fiber networks with inter-fiber adhesion". *Soft Matter*, 14(12):2254–2266, Mar 2018. doi:10.1039/C7SM02555F.

[37] Diane Peurichard, Marielle Ousset, Jenny Paupert, Benjamin Aymard, Anne Lorsignol, Louis Casteilla and Pierre Degond. "Extra-cellular matrix rigidity may dictate the fate of injury outcome". *Journal of Theoretical Biology*, 469:127–136, Feb 2019. doi:10.1016/j.jtbi.2019.02.017.

[38] Louis-Pierre Chaintron and Antoine Diez. "Propagation of chaos: A review of models, methods and applications. I. Models and methods". *Kinetic and Related Models*, 15(6):895–1015, Dec 2022. doi:10.3934/krm.2022017.

[39] Louis-Pierre Chaintron and Antoine Diez. "Propagation of chaos: A review of models, methods and applications. II. Applications". *Kinetic and Related Models*, 15(6):1017–1173, Dec 2022. doi:10.3934/krm.2022018.

[40] Chase P. Broedersz, Martin Depken, Norman Y. Yao, Martin R. Pollak, David A. Weitz and Frederick C. MacKintosh. "Cross-Link-Governed Dynamics of Biopolymer Networks". *Physical Review Letters*, 105(23):238 101, Dec 2010. doi:10.1103/PhysRevLett.105.238101.

[41] Julien Barré, Pierre Degond and Ewelina Zatorska. "Kinetic Theory of Particle Interactions Mediated by Dynamical Networks". *Multiscale Modeling & Simulation*, 15(3):1294–1323, 2017. doi:10.1137/16M1085310.

[42] Pierre Degond, Fanny Delebecque and Diane Peurichard. "Continuum model for linked fibers with alignment interactions". *Mathematical Models and Methods in Applied Sciences*, 26(2):269–318, 2016. doi:10.1142/S0218202516400030.

[43] Seán I. O'Donoghue, Anne-Claude Gavin, Nils Gehlenborg, David S. Goodsell, Jean-Karim Hériché, Cydney B. Nielsen, Chris North, Arthur J. Olson, James B. Procter, David W. Shattuck, Thomas Walter and Bang Wong. "Visualizing biological data – now and in the future". *Nature methods*, 7(3):S2–S4, Mar 2010. doi:10.1038/nmeth.f.301.

[44] Christa Kelleher and Thorsten Wagener. "Ten guidelines for effective data visualization in scientific publications". *Environmental Modelling & Software*, 26(6):822–827, Jun 2011. doi:10.1016/j.envsoft.2010.12.006.

[45] Stephen R. Midway. "Principles of Effective Data Visualization". *Patterns*, 1(9):100 141, Dec 2020. doi:10.1016/j.patter.2020.100141.

[46] Elif Hilal Korkut and Elif Surer. "Visualization in virtual reality: a systematic review". *Virtual Reality*, 27(2):1447–1480, Jun 2003. doi:10.1007/s10055-023-00753-8.

[47] Mohamed El Beheiry, Sébastien Doutreligne, Clément Caporal, Cécilia Ostertag, Maxime Dahan and Jean-Baptiste Masson. "Virtual Reality: Beyond Visualization". *Journal of Molecular Biology*, 431(7):1315–1321, 2019. doi:10.1016/j.jmb.2019.01.033.

[48] Kimberly K. Arcand, Sara R. Price and Megan Watzke. "Holding the Cosmos in Your Hand: Developing 3D Modeling and Printing Pipelines for Communications and Research". *Frontiers in Earth Science*, 8:590 295, Nov 2020. doi:10.3389/feart.2020.590295.

[49] Christoph Bader, Dominik Kolb, James C. Weaver, Sunanda Sharma, Ahmed Hosny, João Costa and Neri Oxman. "Making data matter: Voxel printing for the digital fabrication of data across scales and domains". *Science Advances*, 4(5):eaas8652, 2018. doi:10.1126/sciadv.aas8652.

[50] Arindam Banerjee and Hanhuai Shan. "Model-Based Clustering". In *Encyclopedia of Machine Learning*, edited by Claude Sammut and Geoffrey I. Webb, pp. 686–689. Springer US, Boston (MA), 2010. ISBN 978-0-387-30164-8. doi:10.1007/978-0-387-30164-8_554.

[51] Martin A. Fischler and Robert C. Bolles. "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography". *Communication of the ACM*, 24(6):381–395, Jun 1981. doi:10.1145/358669.358692.

[52] James MacQueen. "Some methods for classification and analysis of multivariate observations". In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, volume 1, edited by Lucien M. Le Cam and Jerzy Neyman, pp. 281–297. University of California Press, 1996.

[53] Stuart P. Lloyd. "Least squares quantization in PCM". *IEEE Transactions on Information Theory*, 28(2):129–137, Mar 1982. doi:10.1109/TIT.1982.1056489.

[54] Frank Nielsen. "Hierarchical Clustering". In *Introduction to HPC with MPI for Data Science*, pp. 195–211. Springer, Cham, Feb 2016. ISBN 978-3-319-21902-8. doi:10.1007/978-3-319-21903-5_8.

[55] Hans-Peter Kriegel, Peer Kröger, Jörg Sander and Arthur Zimek. "Density-based clustering". *WIREs Data Mining and Knowledge Discovery*, 1(3):231–240, Apr 2011. doi:10.1002/widm.30.

[56] Martin Ester, Hans-Peter Kriegel, Jörg Sander and Xiaowei Xu. "A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise". In *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, edited by Evangelos Simoudis, Jiawei Han and Usama M. Fayyad, pp. 226–231, 1996.

[57] Mariá C. V. Nascimento and André C. P. L. F. de Carvalho. "Spectral methods for graph clustering - A survey". *European Journal of Operational Research*, 211(2):221–231, 2011. doi:10.1016/j.ejor.2010.08.012.

[58] Yang Xu, Arun Srinivasan and Lingzhou Xue. "A Selective Overview of Recent Advances in Spectral Clustering and Their Applications". In *Modern Statistical Methods for Health Research*, edited by Yichuan Zhao and (Din) Ding-Geng Chen, Emerging Topics in Statistics and Biostatistics, pp. 247–277. Springer International Publishing, Cham, 2021. ISBN 978-3-030-72437-5. doi:10.1007/978-3-030-72437-5_12.

[59] Jos B. T. M. Roerdink and Arnold Meijster. "The Watershed Transform: Definitions, Algorithms and Parallelization Strategies". *Fundamenta Informaticae*, 41(1,2):187–228, Apr 2000. doi:10.3233/FI-2000-411207.

[60] J. Gordon Betts, Kelly A. Young, James A. Wise, Eddie Johnson, Brandon Poe, Dean H. Kruse, Oksana Korol, Jody E. Johnson, Mark Womble and Peter DeSaix. "The Tissue Level of Organization". In *Anatomy and Physiology*. OpenStax Publishing, Houston (Texas), 2nd edition, Apr 2022.

[61] Trevor A. Nezwek and Matthew Varacallo. *Physiology, Connective Tissue*. StatPearls Publishing, Treasure Island (FL), updated 19 September 2022.

[62] Payvand Kamrani, Geoffrey Marston, Tafline C. Arbor and Arif Jan. *Anatomy, Connective Tissue*. StatPearls Publishing, Treasure Island (FL), updated 5 March 2023.

[63] Achilleas D. Theocharis, Spyros S. Skandalis, Chrysostomi Gialeli and Nikos K. Karamanos. "Extracellular matrix structure". *Advanced Drug Delivery Reviews*, 97:4–27, 2016. doi:10.1016/j.addr.2015.11.001.

[64] Christian Frantz, Kathleen M. Stewart and Valerie M. Weaver. "The extracellular matrix at a glance". *Journal of Cell Science*, 123(24):4195–4200, Dec 2010. doi:10.1242/jcs.023820.

[65] Beatrice Yue. "Biology of the Extracellular Matrix: An Overview". *Journal of Glaucoma*, 23:S20–S23, Oct 2014. doi:10.1097/IJG.0000000000000108.

[66] Sylvie Ricard-Blum and Florence Ruggiero. "The collagen superfamily : from the extracellular matrix to the cell membrane". *Pathologie Biologie*, 53(7):430–442, 2005. doi:10.1016/j.patbio.2004.12.024.

[67] Peter Fratzl. "Collagen: Structure and Mechanics, an Introduction". In *Collagen: Structure and Mechanics*, edited by Peter Fratzl, pp. 1–13. Springer US, Boston (MA), 2008. ISBN 978-0-387-73906-9. doi:10.1007/978-0-387-73906-9_1.

[68] Wenyu Kong, Cheng Lyu, Hongen Liao and Yanan Du. "Collagen crosslinking: Effect on structure, mechanics and fibrosis progression". *Biomedical Materials*, 16, Nov 2021. doi:10.1088/1748-605X/ac2b79.

[69] Peter K. Mays, Robin J. McAnulty, Juan S. Campa and Geoffrey J. Laurent. "Age-related changes in collagen synthesis and degradation in rat tissues. Importance of degradation of newly synthesized collagen in regulating collagen production". *Biochemical Journal*, 276(2):307–313, Jun 1991. doi:10.1042/bj2760307.

[70] Andrea Malandrino, Xavier Trepat, Roger D. Kamm and Michael Mak. "Dynamic filopodial forces induce accumulation, damage, and plastic remodeling of 3D extracellular matrices". *PLOS Computational Biology*, 15(4):1–26, Apr 2019. doi:10.1371/journal.pcbi.1006684.

[71] Kandice R. Levental, Hongmei Yu, Laura Kass, Johnathon N. Lakins, Mikala Egeblad, Janine T. Erler, Sheri F. Fong, Katalin Csiszar, Amato Giaccia, Wolfgang Weninger, Mitsuo Yamauchi, David L. Gasser and Valerie M. Weaver. "Matrix crosslinking forces tumor progression by enhancing integrin signaling". *Cell*, 139(5):891–906, 2009. doi:10.1016/j.cell.2009.10.027.

[72] Jeremy Herrera, Craig A. Henke and Peter B. Bitterman. "Extracellular matrix as a driver of progressive fibrosis". *The Journal of Clinical Investigation*, 128(1):45–53, Jan 2018. doi:10.1172/JCI93557.

[73] Kevin Dzobo and Collet Dandara. "The Extracellular Matrix: Its Composition, Function, Remodeling, and Role in Tumorigenesis". *Biomimetics*, 8(2):146, Apr 2023. doi:10.3390/biomimetics8020146.

[74] Nikos K. Karamanos, Achilleas D. Theocharis, Zoi Piperigkou, Dimitra Manou, Alberto Passi, Spyros S. Skandalis, Demitrios H. Vynios, Véronique Orian-Rousseau, Sylvie Ricard-Blum, Christian E. H. Schmelzer, Laurent Duca, Madeleine Durbeej, Nikolaos A. Afratis, Linda Troeberg, Marco Franchi, Valentina Masola and Maurizio Onisto. "A guide to the composition and functions of the extracellular matrix". *The FEBS Journal*, 288(24):6850–6912, 2021. doi:10.1111/febs.15776.

[75] Robert B. Diller and Aaron J. Tabor. "The Role of the Extracellular Matrix (ECM) in Wound Healing: A Review". *Biomimetics*, 7(3):87, 2022. doi:10.3390/biomimetics7030087.

[76] Samantha L. Wilson, Marie Guilbert, Josep Sulé-Suso, Jim Torbet, Pierre Jeannesson, Ganesh D. Sockalingum and Ying Yang. "A microscopic and macroscopic study of aging collagen on its molecular structure, mechanical properties, and cellular response". *The FASEB Journal*, 28(1):14–25, 2014. doi:10.1096/fj.13-227579.

[77] Allen J. Bailey, Robert Gordon Paul and Lynda Knott. "Mechanisms of maturation and ageing of collagen". *Mechanisms of Ageing and Development*, 106(1–2):1–56, 1998. doi:10.1016/s0047-6374(98)00119-5.

[78] Katherine Cianflone, Magdalena Maslowska and Allan D. Sniderman. "Acylation stimulating protein (ASP), an adipocyte autocrine: new directions". *Seminars in Cell & Developmental Biology*, 10:31–41, Feb 1999. doi:10.1006/scdb.1998.0272.

[79] J. Saleh, N. Al-Wardy, H. Farhan, M. Al-Khanbashi and K. Cianflone. "Acylation stimulating protein: a female lipogenic factor?" *Obesity Reviews*, 12:440–448, Jun 2011. doi:10.1111/j.1467-789X.2010.00832.x.

[80] M. D. Klok, S. Jakobsdottir and M. L. Drent. "The role of leptin and ghrelin in the regulation of food intake and body weight in humans: a review". *Obesity Reviews*, 8(1):21–34, Jan 2007. doi:10.1111/j.1467-789X.2006.00270.x.

[81] Antonio La Cava and Giuseppe Matarese. "The weight of leptin in immunity". *Nature Reviews Immunology*, 4:371–379, May 2004. doi:10.1038/nri1350.

[82] M. Maumus, J. A. Peyrafitte, R. D'Angelo, C. Fournier-Wirth, A. Bouloumié, Louis Casteilla, Coralie Sengenès and Philippe Bourin. "Native human adipose stromal cells: localization, morphology and phenotype". *International journal of obesity*, 35:1141–1153, Sep 2011. doi:10.1038/ijo.2010.269.

[83] Stephane Gesta and C. Ronald Kahn. "White Adipose Tissue". In *Adipose Tissue Biology*, edited by Michael E. Symonds, pp. 149–199. Springer International Publishing, Cham, 2017. ISBN 978-3-319-52031-5. doi:10.1007/978-3-319-52031-5_5.

[84] Paul Trayhurn and John H. Beattie. "Physiological role of adipose tissue: white adipose tissue as an endocrine and secretory organ". *Proceedings of the Nutrition Society*, 60(3):329–339, 2001. doi:10.1079/PNS200194.

[85] Alison J. Richard, Ursula White, Carrie M. Elks and Jacqueline M. Stephens. *Adipose Tissue: Physiology to Metabolic Dysfunction*. MDText.com, Inc., South Dartmouth (MA), Apr 2020.

[86] Jan Nedergaard and Olov Lindberg. "The Brown Fat Cell". *International Review of Cytology*, 74:187–286, 1982. doi:10.1016/S0074-7696(08)61173-0.

[87] Daniel Ricquier. "Molecular biology of brown adipose tissue". *Proceedings of the Nutrition Society*, 48(2):183–187, 1989. doi:10.1079/PNS19890028.

[88] Paula Mota de Sá, Allison J. Richard, Hardy Hang and Jacqueline M. Stephens. "Transcriptional Regulation of Adipogenesis". *Comprehensive Physiology*, 7:635–674, Apr 2017. doi:10.1002/cphy.c160022.

[89] Elinor Arbuthnott. "Brown adipose tissue: structure and function". *Proceedings of the Nutrition Society*, 48:177–182, Jul 1989. doi:10.1079/PNS19890027.

[90] Dragutin Lončar. "Convertible adipose tissue in mice". *Cell and Tissue Research*, 266:149–161, Oct 1991. doi:10.1007/BF00678721.

[91] Béatrice Cousin, Saverio Cinti, Manrico Morroni, Serge Raimbault, Daniel Ricquier, Luc Pénicaud and Louis Casteilla. "Occurrence of brown adipocytes in rat white adipose tissue: molecular and morphological characterization". *Journal of Cell Science*, 103(4):931–942, Dec 1992. doi:10.1242/jcs.103.4.931.

[92] Tomas B. Waldén, Ida R. Hansen, James A. Timmons, Barbara Cannon and Jan Nedergaard. "Recruited vs. nonrecruited molecular signatures of brown, "brite", and white adipose tissues". *American Journal of Physiology - Endocrinology and Metabolism*, 302(1):E19–E31, Jan 2012. doi:10.1152/ajpendo.00249.2011.

[93] P. Young, J. R. S. Arch and Margaret Ashwell. "Brown adipose tissue in the parametrial fat pad of the mouse". *FEBS Letters*, 167(1):10–14, 1984. doi:10.1016/0014-5793(84)80822-4.

[94] Dragutin Lonçar, Björn A. Afzelius and Barbara Cannon. "Epididymal white adipose tissue after cold stress in rats I. Nonmitochondrial changes". *Journal of Ultrastructure and Molecular Structure Research*, 101(2):109–122, 1988. doi:10.1016/0889-1605(88)90001-8.

[95] Siyu Wang, Min-Hsiung Pan, Wei-Lun Hung, Yen-Chen Tung and Chi-Tang Ho. "From white to beige adipocytes: therapeutic potential of dietary molecules against obesity and their molecular mechanisms". *Food & Function*, 10:1263–1279, 2019. doi:10.1039/C8FO02154F.

[96] Adriana Mika, Filippo Macaluso, Rosario Barone, Valentina Di Felice and Tomasz Sledzinski. "Effect of Exercise on Fatty Acid Metabolism and Adipokine Secretion in Adipose Tissue". *Frontiers in Physiology*, 10, 2019. doi:10.3389/fphys.2019.00026.

[97] Jeff Ishibashi and Patrick Seale. "Beige Can Be Slimming". *Science*, 328(5982):1113–1114, 2010. doi:10.1126/science.1190816.

[98] Hyun Cheol Roh, Linus T. Y. Tsai, Mengle Shao, Danielle Tenen, Yachen Shen, Manju Kumari, Anna Lyubetskaya, Christopher Jacobs, Brian Dawes, Rana K. Gupta and Evan D. Rosen. "Warming induces significant reprogramming of beige, but not brown, adipocyte cellular identity". *Cell Metabolism*, 27(5):1121–1137, May 2018. doi:10.1016/j.cmet.2018.03.005.

[99] Pouneh K. Fazeli, Mark C. Horowitz, Ormond A. MacDougald, Erica L. Scheller, Matthew S. Rodeheffer, Clifford J. Rosen and Anne Klibanski. "Marrow Fat and Bone – New Perspectives". *The Journal of Clinical Endocrinology & Metabolism*, 98(3):935–945, Mar 2013. doi:10.1210/jc.2012-3634.

[100] Saverio Cinti. "Pink Adipocytes". *Trends in Endocrinology & Metabolism*, 29:651–666, Sep 2018. doi:10.1016/j.tem.2018.05.007.

[101] Kirsi A. Virtanen, Martin E. Lidell, Janne Orava, Mikael Heglind, Rickard Westergren, Tarja Niemi, Markku Taittonen, Jukka Laine, Nina-Johanna Savisto, Sven Enerbäck and Pirjo Nuutila. "Functional Brown Adipose Tissue in Healthy Adults". *New England Journal of Medicine*, 360(15):1518–1525, Apr 2009. doi:10.1056/NEJMoa0808949.

[102] Brooks P. Leitner, Shan Huang, Robert J. Brychta, Courtney J. Duckworth, Alison S. Baskin, Suzanne McGehee, Ilan Tal, William Dieckmann, Garima Gupta, Gerald M. Kolodny, Karel Pacak, Peter Herscovitch, Aaron M. Cypess and Kong Y. Chen. "Mapping of human brown adipose tissue in lean and obese young men". *Proceedings of the National Academy of Sciences*, 114(32):8649–8654, 2017. doi:10.1073/pnas.1705287114.

[103] Saverio Cinti. "The adipose organ at a glance". *Disease Models & Mechanisms*, 5:588–594, Sep 2012. doi:10.1242/dmm.009662.

[104] Dorothy B. Hausman, M. DiGirolamo, T. J. Bartness, G. J. Hausman and R. J. Martin. "The biology of white adipocyte proliferation". *Obesity Reviews*, 2(4):239–254, 2001. doi:10.1046/j.1467-789X.2001.00042.x.

[105] Saverio Cinti. "The adipose organ: morphological perspectives of adipose tissues". *Proceedings of the Nutrition Society*, 60:319–328, Aug 2001. doi:10.1079/PNS200192.

[106] Saverio Cinti. "The adipose organ". In *Adipose Tissue and Adipokines in Health and Disease*, edited by G. Fantuzzi and T. Mazzone, pp. 3–19. Humana Press, Totowa (NJ), 2007. ISBN 978-1-59745-370-7. doi:10.1007/978-1-59745-370-7_1.

[107] F. Wassermann. "The development of adipose tissue". *Comprehensive Physiology*, pp. 87–100, 2011. doi:10.1002/cphy.cp050110.

[108] Corinne Barreau, Elodie Labit, Christophe Guissard, Jacques Rouquette, Marie-Laure Boizeau, Souleymane Gani Koumassi, Audrey Carrière, Yannick Jeanson, Sandra Berger-Müller, Cécile. Dromard, Franck Plouraboué, Louis Casteilla and Anne Lorsignol. "Regionalization of browning revealed by whole subcutaneous adipose tissue imaging". *Obesity (Silver Spring)*, 24(5):1081–1089, 2016. doi:10.1002/oby.21455.

[109] Jules Dichamp, Corinne Barreau, Christophe Guissard, Audrey Carrière, Yves Martinez, Xavier Descombes, Luc Pénicaud, Jacques Rouquette, Louis Casteilla, Franck Plouraboué and Anne Lorsignol. "3D analysis of the whole subcutaneous adipose tissue reveals a complex spatial network of interconnected lobules with heterogeneous browning ability". *Scientific Reports*, 9(1), Apr 2019. doi:10.1038/s41598-019-43130-9.

[110] Qiong A. Wang, Caroline Tao, Rana K. Gupta and Philipp E. Scherer. "Tracking adipogenesis during white adipose tissue development, expansion and regeneration". *Nature Medicine*, 19:1338–1344, Oct 2013. doi:10.1038/nm.3324.

[111] Yun-Hee Lee, Emilio P. Mottillo and James G. Granneman. "Adipose tissue plasticity from WAT to BAT and in between". *Biochimica et Biophysica Acta*, 1842:358–369, Mar 2014. doi:10.1016/j.bbadis.2013.05.011.

[112] NCD Risk Factor Collaboration (NCD-RisC). "Worldwide trends in body-mass index, underweight, overweight, and obesity from 1975 to 2016: a pooled analysis of 2416 population-based measurement studies in 128.9 million children, adolescents, and adults". *The Lancet*, 390:2627–2642, Dec 2017. doi:10.1016/S0140-6736(17)32129-3.

[113] World Health Organization. "Fact sheet – Obesity and overweight", 2018. Updated 9 June 2021.

[114] Meriem Abdennour, Sophie Reggio, Gilles Le Naour, Yuejun Liu, Christine Poitou, Judith Aron-Wisnewsky, Frederic Charlotte, Jean-Luc Bouillot, Adriana Torcivia, Magali Sasso, Veronique Miette, Jean-Daniel Zucker, Pierre Bedossa, Joan Tordjman and Karine Clement. "Association of Adipose Tissue and Liver Fibrosis With Tissue Stiffness in Morbid Obesity: Links With Diabetes and BMI Loss After Gastric Bypass". *The Journal of Clinical Endocrinology & Metabolism*, 99(3):898–907, Mar 2014. doi:10.1210/jc.2013-3253.

[115] Rocío Guzmán-Ruiz, Carmen Tercero-Alcázar, Yoana Rabanal-Ruiz, Alberto Díaz-Ruiz, Rajaa El Bekay, Oriol A. Rangel-Zuñiga, M. Carmen Navarro-Ruiz, Laura Molero, Antonio Membrives, Juan F. Ruiz-Rabelo, Abhay Pandit, José López-Miranda, Francisco J. Tinahones and María M. Malagón. "Adipose tissue depot-specific intracellular and extracellular cues contributing to insulin resistance in obese individuals". *The FASEB Journal*, 34(6):7520–7539, Jun 2020. doi:10.1096/fj.201902703R.

[116] Geneviève Marcelin, Emmanuel L. Gautier and Karine Clément. "Adipose Tissue Fibrosis in Obesity: Etiology and Challenges". *Annual Review of Physiology*, 84(1):135–155, 2022. doi:10.1146/annurev-physiol-060721-092930.

[117] Tayeba Khan, Eric S. Muise, Puneeth Iyengar, Zhao V. Wang, Manisha Chandalia, Nicola Abate, Bei B. Zhang, Paolo Bonaldo, Streamson Chua and Philipp E. Scherer. "Metabolic Dysregulation and Adipose Tissue Fibrosis: Role of Collagen VI". *Molecular and Cellular Biology*, 29(6):1575–1591, 2009. doi:10.1128/MCB.01300-08.

[118] An Huang, Yi-Shiuan Lin, Ling-Zhen Kao, Yu-Wei Chiou, Gang-Hui Lee, Hsi-Hui Lin, Chih-Hsing Wu, Chin-Sung Chang, Kuo-Ting Lee, Yuan-Yu Hsueh, Pei-Jane Tsai, Ming-Jer Tang and Yau-Sheng Tsai. "Inflammation-induced macrophage lysyl oxidase in adipose stiffening and dysfunction in obesity". *Clinical and Translational Medicine*, 11(9):e543, 2021. doi:https://doi.org/10.1002/ctm2.543.

[119] Simon Lecoutre, Mélanie Lambert, Krzysztof Drygalski, Isabelle Dugail, Salwan Maqdasy, Mathieu Hautefeuille and Karine Clément. "Importance of the Microenvironment and Mechanosensing in Adipose Tissue Biology". *Cells*, 11(15), 2022. doi:10.3390/cells11152310.

[120] Kenneth J. Arrow, Leonid Hurwicz and Hirofumi Uzawa. *Studies in linear and non-linear programming*. Stanford University Press, 1958.

[121] Xiaoyu Liao, Xin Li and Rui Liu. "Extracellular-matrix mechanics regulate cellular metabolism: A ninja warrior behind mechano-chemo signaling crosstalk". *Reviews in Endocrine and Metabolic Disorders*, 24(2):207–220, Apr 2023. doi:10.1007/s11154-022-09768-z.

[122] Oliver Lieleg, Mireille M. A. E. Claessens and Andreas R. Bausch. "Structure and dynamics of cross-linked actin networks". *Soft Matter*, 6(2):218–225, Jan 2010. doi:10.1039/B912163N.

[123] Claus Heussinger and Erwin Frey. "Role of architecture in the elastic response of semiflexible polymer and fiber networks". *Physical Review E*, 75(1):011 917, Jan 2007. doi:10.1103/PhysRevE.75.011917.

[124] A. S. Abhilash, Prashant K. Purohit and Shailendra P. Joshi. "Stochastic rate-dependent elasticity and failure of soft fibrous networks". *Soft Matter*, 8(26):7004–7016, 2012. doi:10.1039/c2sm25450f.

[125] P. Ciarletta and M. Ben Amar. "A finite dissipative theory of temporary interfibrillar bridges in the extracellular matrix of ligaments and tendons". *Journal of the Royal Society Interface*, 6:909–924, Oct 2009. doi:10.1098/rsif.2008.0487.

[126] Heinrich Hertz. "Ueber die Berührung fester elastischer Körper". *Journal für die reine und angewandte Mathematik (Crelles Journal)*, 92:156–171, 1882.

[127] Heinrich Hertz. "On the contact of elastic solids". In *Miscellaneous Papers*, edited by Philipp Lenard, chapter V, pp. 146–162. Macmillan, London, 1896.

[128] Olinde Rodrigues. "Des lois géométriques qui régissent les déplacements d'un système solide dans l'espace, et de la variation des coordonnées provenant de ces déplacements considérés indépendants des causes qui peuvent les produire". *Journal de Mathématiques Pures et Appliquées*, 5:380–440, 1840.

[129] Naama Shoham, Pinhas Girshovitz, Rona Katzengold, Natan T. Shaked, Dafna Benayahu and Amit Gefen. "Adipocyte stiffness increases with accumulation of lipid droplets". *Biophysical Journal*, 106:1421–1431, Mar 2014. doi:10.1016/j.bpj.2014.01.045.

[130] Peter E. Kloeden and Eckhard Platen. *Numerical Solution of Stochastic Differential Equations*. Springer Berlin, Heidelberg, Aug 1992. ISBN 978-3-540-54062-5. doi:10.1007/978-3-662-12616-5.

[131] Ruihuan Cai, Lei Xu, Jinyang Zheng and Yongzhi Zhao. "Modified cell-linked list method using dynamic mesh for discrete element method". *Powder Technology*, 340:321–330, 2018. doi:10.1016/j.powtec.2018.09.034. URL https://www.sciencedirect.com/science/article/pii/S0032591018307629.

[132] Shaoyun Wang and Chaohui Tong. "Cell Lists Method Based on Doubly Linked Lists for Monte Carlo Simulation", 2020. doi:10.48550/arXiv.2003.05581.

[133] Zhen-Xi Zhao, Hua Liu and Zhao-Xin Gong. "A high-efficiency smoothed particle hydrodynamics model with multi-cell linked list and adaptive particle refinement for two-phase flows". *Physics of Fluids*, 33(6), Jun 2021. doi:10.1063/5.0052030.

[134] Esteban Vélez Ramírez and Christer Elvingson. "An efficient linked list for molecular simulations on a spherical surface". *Journal of Physics A: Mathematical and Theoretical*, 55(38), Aug 2022. doi:10.1088/1751-8121/ac852f.

[135] Sathish Sanjeevi. "spherocylinder (https://www.mathworks.com/matlabcentral/fileexchange/66904-spherocylinder)", 2020. Retrieved 2020-04-22.

[136] Fernand Meyer. "Topographic distance and watershed lines". *Signal Processing*, 38:113–125, Jul 1994. doi:10.1016/0165-1684(94)90060-4.

[137] Johannes Schindelin, Ignacio Arganda-Carreras, Erwin Frise, Verena Kaynig, Mark Longair, Tobias Pietzsch, Stephan Preibisch, Curtis Rueden, Stephan Saalfeld, Benjamin Schmid, Jean-Yves Tinevez, Daniel James White, Volker Hartenstein, Kevin Eliceiri, Pavel Tomancak and Albert Cardona. "Fiji: an open-source platform for biological-image analysis". *Nature Methods*, 9(7):676–682, Jul 2012. doi:10.1038/nmeth.2019.

[138] Carsen Stringer, Tim Wang, Michalis Michaelos and Marius Pachitariu. "Cellpose: a generalist algorithm for cellular segmentation". *Nature Methods*, 18(1):100–106, 2021. doi:10.1038/s41592-020-01018-x.

[139] Marius Pachitariu and Carsen Stringer. "Cellpose 2.0: how to train your own model". *bioRxiv*, 19(12):1634–1641, Dec 2022. doi:10.1038/s41592-022-01663-4.

[140] Stephan Preibisch, Stephan Saalfeld and Pavel Tomancak. "Globally optimal stitching of tiled 3D microscopic image acquisitions". *Bioinformatics*, 25(11):1463–1465, Apr 2009. doi:10.1093/bioinformatics/btp184.

[141] Yuriy Mishchenko. "3D Euclidean Distance Transform for Variable Data Aspect Ratio)", 2022. Retrieved 2022-06-28.

[142] Evren Özarslan, Baba C. Vemuri and Thomas H. Mareci. "Generalized scalar measures for diffusion MRI using trace, variance, and entropy". *Magnetic Resonance in Medicine*, 53(4):866–876, Apr 2005. doi:10.1002/mrm.20411.

[143] Pauline Chassonnery, Jenny Paupert, Anne Lorsignol, Childérick Sévérac, Marielle Ousset, Pierre Degond, Louis Casteilla and Diane Peurichard. "Fiber crosslinking drives the emergence of order in a 3D dynamical network model". *arXiv preprint*, 2023. doi:10.48550/arXiv.2307.10859.

[144] Alexandra S. Piotrowski-Daspit, Bryan A. Nerger, Abraham E. Wolf, Sankaran Sundaresan and Celeste M. Nelson. "Dynamics of Tissue-Induced Alignment of Fibrous Extracellular Matrix". *Biophysical Journal*, 113(3):702–713, 2017. doi:10.1016/j.bpj.2017.06.046.

[145] Ross P. Wohlgemuth, Sarah E. Brashear and Lucas R. Smith. "Alignment, cross-linking, and beyond: A collagen architect's guide to the skeletal muscle extracellular matrix". *American Journal of Physiology – Cell Physiology*, 325(4):C1017–C1030, Sep 2023. doi:10.1152/ajpcell.00287.2023.

[146] Stephen Williams and Albert Philipse. "Random Packings of Spheres and Spherocylinders Simulated by Mechanical Contraction". *Physical Review E : Statistical, nonlinear, and soft matter physics*, 67:051 301, Jun 2003. doi:10.1103/PhysRevE.67.051301.

[147] Caroline Bonnans, Jonathan Chou and Zena Werb. "Remodelling the extracellular matrix in development and disease". *Nature Reviews Molecular Cell Biology*, 15(12):786–801, Dec 2014. doi:10.1038/nrm3904.

[148] Christopher J. Philp, Ivonne Siebeke, Debbie Clements, Suzanne Miller, Anthony Habgood, Alison E. John, Vidya Navaratnam, Richard B. Hubbard, Gisli Jenkins and Simon R. Johnson. "Extracellular Matrix Cross-Linking Enhances Fibroblast Growth and Protects against Matrix Proteolysis in Lung Fibrosis". *American Journal of Respiratory Cell and Molecular Biology*, 58(5):594–603, 2018. doi:10.1165/rcmb.2016-0379OC.

[149] Carly M. Garrison and Jean E. Schwarzbauer. "Fibronectin fibril alignment is established upon initiation of extracellular matrix assembly". *Molecular Biology of the Cell*, 32(8):739–752, 2021. doi:10.1091/mbc.E20-08-0533.

[150] Yu-Han Jiang, Ying-Yue Lou, Teng-Hai Li, Bing-Zhang Liu, Kang Chen, Duo Zhang and Tian Li. "Cross-linking methods of type I collagen-based scaffolds for cartilage tissue engineering". *American journal of translational research*, 14(2):1146–1159, 2022.

[151] Xu Xue, Yan Hu, Sicheng Wang, Xiao Chen, Yingying Jiang and Jiacan Su. "Fabrication of physical and chemical crosslinked hydrogels for bone tissue engineering". *Bioactive Materials*, 12:327–339, 2022. doi:10.1016/j.bioactmat.2021.10.029.

[152] Kirsty L. Spalding, Erik Arner, Pål O. Westermark, Samuel Bernard, Bruce A. Buchholz, Olaf Bergmann, Lennart Blomqvist, Johan Hoffstedt, Erik Näslund, Tom Britton, Hernan Concha, Moustapha Hassan, Mikael Rydén, Jonas Frisén and Peter Arner. "Dynamics of fat cell turnover in humans". *Nature*, 453:783–787, May 2008. doi:10.1038/nature06902.

[153] Erik Arner, Pål O. Westermark, Kristy L. Spalding, Tom Britton, Mikael Rydén, Jonas Frisén, Samuel Bernard and Peter Arner. "Adipocyte Turnover: Relevance to Human Adipose Tissue Morphology". *Diabetes*, 59(1):105–109, 2010. doi:10.2337/db09-0942.

[154] Stephen C. Cowin. "Tissue Growth and Remodeling". *Annual Review of Biomedical Engineering*, 6(1):77–107, 2004. doi:10.1146/annurev.bioeng.6.040803.140250.

[155] Tatsuo Ushiki. "Collagen Fibers, Reticular Fibers and Elastic Fibers. A Comprehensive Understanding from a Morphological Viewpoint". *Archives of Histology and Cytology*, 65(2):109–126, 2002. doi:10.1679/aohc.65.109.

[156] Anastasia Pacary, Diane Peurichard, Laurence Vaysse, Paul Monsarrat, Clémence Bolut, Christophe Guissard, Anne Lorsignol, Valérie Planat-Benard, Jenny Paupert, Marielle Ousset and Louis Casteilla. "A digital tissue repair model identifies an early transient decrease in fibre cross-linking that unlocks regeneration in adult mammals". *submitted*, 2023.

[157] Giovanna Mazzoleni, Diego Di Lorenzo and Nathalie Steimberg. "Modelling tissues in 3D: the next future of pharmaco-toxicology and food research ?" *Genes & nutrition*, 4(1):13–22, Mar 2009. doi:10.1007/s12263-008-0107-0.

[158] Guillaume Jacquemet, Hellyeh Hamidi and Johanna Ivaska. "Filopodia in cell adhesion, 3D migration and cancer cell invasion". *Current Opinion in Cell Biology*, 36:23–31, 2015. doi:10.1016/j.ceb.2015.06.007.

[159] Jong Bin Kim. "Three-dimensional tissue culture models in cancer biology". *Seminars in Cancer Biology*, 15(5):365–377, 2005. doi:10.1016/j.semcancer.2005.05.002.

[160] Francesco Pampaloni, Emmanuel G. Reynaud and Ernst H. K. Stelzer. "The third dimension bridges the gap between cell culture and live tissue". *Nature Reviews Molecular Cell Biology*, 8(10):839–845, Oct 2007. doi:10.1038/nrm2236.

[161] Peter Bolhuis and Daan Frenkel. "Numerical study of the phase diagram of a mixture of spherical and rodlike colloids". *The Journal of Chemical Physics*, 101(11):9869–9875, Dec 1994. doi:10.1063/1.467953.

[162] Marie Adams and Seth Fraden. "Phase Behavior of Mixtures of Rods (Tobacco Mosaic Virus) and Spheres (Polyethylene Oxide, Bovine Serum Albumin)". *Biophysical Journal*, 74(1):669–677, 1998. doi:10.1016/S0006-3495(98)77826-9.

[163] Swetlana Jungblut, Kurt Binder and Tanja Schilling. "Suspensions of rod-like colloids and a depleting agent under confinement". *Journal of Physics: Condensed Matter*, 20(40):404 223, Sep 2008. doi:10.1088/0953-8984/20/40/404223.

[164] Tomonori Koda, Manabu Numajiri and Susumu Ikeda. "Smectic-A Phase of a Bidisperse System of Parallel Hard Rods and Hard Spheres". *Journal of the Physical Society of Japan*, 65(11):3551–3556, 1996. doi:10.1143/JPSJ.65.3551.

[165] Santiago Lago, Alejandro Cuetos, Bruno Martínez-Haya and Luis F. Rull. "Crowding effects in binary mixtures of rod-like and spherical particles". *Journal of Molecular Recognition*, 17(5):417–425, 2004. doi:10.1002/jmr.704.

[166] James J. Tomasek, Giulio Gabbiani, Boris Hinz, Christine Chaponnier and Robert A. Brown. "Myofibroblasts and mechano-regulation of connective tissue remodelling". *Nature Reviews Molecular Cell Biology*, 3(5):349–363, May 2002. doi:10.1038/nrm809.

[167] Paul Van Liedekerke, Johannes Neitsch, Tim Johann, Enrico Warmt, Steffen Grosser, Ismael González Valverde, Josef Käs, Stefan Hoehme and Dirk Drasdo. "Quantifying the mechanics and growth of cells and tissues in 3D using high resolution computational models". *bioRxiv*, Nov 2019. doi:10.1101/470559.

[168] S. Majid Hosseini and James J. Feng. "A particle-based model for the transport of erythrocytes in capillaries". *Chemical Engineering Science*, 64(22):4488–4497, 2009. doi:10.1016/j.ces.2008.11.028.

[169] Paul Van Liedekerke, Pieter Ghysels, Engelbert Tijskens, Giovanni Samaey, Dirk Roose and Herman Ramon. "Mechanisms of soft cellular tissue bruising. A particle based simulation approach". *Soft Matter*, 7:3580–3591, Feb 2011. doi:10.1039/C0SM01261K.

[170] Angelika Manhart, Dietmar Oelz, Christian Schmeiser and Nikolaos Sfakianakis. "An extended Filament Based Lamellipodium Model produces various moving cell shapes in the presence of chemotactic signals". *Journal of Theoretical Biology*, 382:244–258, 2015. doi:10.1016/j.jtbi.2015.06.044.

[171] Sara Christensen and Peter P. Purslow. "The role of matrix metalloproteinases in muscle and adipose tissue development and meat quality: A review". *Meat Science*, 119:138–146, Sep 2016. doi:10.1016/j.meatsci.2016.04.025.

[172] Moritz Mercker, Dirk Hartmann and Anna Marciniak-Czochra. "A Mechanochemical Model for Embryonic Pattern Formation: Coupling Tissue Mechanics and Morphogen Expression". *PLOS ONE*, 8(12):1–6, Dec 2013. doi:10.1371/journal.pone.0082617.

[173] Hedi A. Soula, Hanna Julienne, Christophe O. Soulage and Alain Géloën. "Modelling adipocytes size distribution". *Journal of Theoretical Biology*, 332:89–95, Sep 2013. doi:10.1016/j.jtbi.2013.04.025.

[174] Satoru Okuda, Yasuhiro Inoue, Tadashi Watanabe and Taiji Adachi. "Coupling intercellular molecular signalling with multicellular deformation for simulating three-dimensional tissue morphogenesis". *Interface focus*, 5(2):20140 095, Apr 2015. doi:10.1098/rsfs.2014.0095.

[175] Diane Peurichard. "Macroscopic Model for Cross-Linked Fibers with Alignment Interactions: Existence Theory and Numerical Simulations". *Multiscale Modeling & Simulation*, 14(4):1175–1210, 2016. doi:10.1137/15M1026729.

[176] Chad M. Topaz, Andrea L. Bertozzi and Mark A. Lewis. "A Nonlocal Continuum Model for Biological Aggregation". *Bulletin of Mathematical Biology*, 68(7):1601–1623, Oct 2016. doi:10.1007/s11538-006-9088-6.

[177] Pierre-Emmanuel Jabin and Zhenfu Wang. "Mean Field Limit for Stochastic Particle Systems". In *Active Particles, Volume 1 : Advances in Theory, Models, and Applications*, edited by Nicola Bellomo, Pierre Degond and Eitan Tadmor, pp. 379–402. Springer International Publishing, Cham, 2017. ISBN 978-3-319-49996-3. doi:10.1007/978-3-319-49996-3_10.

[178] Maria Bruna, Stephen J. Chapman and Martin Robinson. "Diffusion of Particles with Short-Range Interactions". *SIAM Journal on Applied Mathematics*, 77(6):2294–2316, Feb 2017. doi:10.1137/17M1118543.

[179] Maria Bruna, Stephen J. Chapman and Markus Schmidtchen. "Derivation of a macroscopic model for Brownian hard needles". *Proceedings of the Royal Society A*, 479(2274), Jun 2023. doi:10.1098/rspa.2023.0076.

[180] Marie Doumic, Sophie Hecht, Benoit Perthame and Diane Peurichard. "Multispecies cross-diffusions: from a nonlocal mean-field to a porous medium system without self-diffusion", 2023. doi:10.48550/arXiv.2306.01777.

Mots clefs : Biologie mathématique ; Modèles individu-centrés ; Visualisation et segmentation en 3D ; Tissus conjonctifs ; Réseaux de fibres dynamiques en 3D ; Tissus adipeux.

Résumé : Cette thèse porte sur l'hypothèse que des interactions mécaniques locales simples entre un nombre limité de composants puissent régir l'émergence de l'architecture 3D des tissus biologiques. Pour explorer cette possibilité, nous développons deux modèles mathématiques. Le premier, ECMmorpho-3D, vise à reproduire un tissu conjonctif non-spécialisé réduit à la matrice extra-cellulaire, c'est-à-dire à un réseau 3D de fibres interconnectées dynamiquement. Le second, ATmorpho-3D, est obtenu par ajout de cellules sphériques apparaissant et croissant spontanément dans ce réseau de fibres afin de modéliser la morphogenèse du tissu adipeux, un tissu conjonctif spécialisé ayant une grande importance sur le plan biomédical. Pour analyser les données produites par ces deux modèles, nous construisons un outil générique permettant de visualiser en 3D des systèmes composés d'un mélange d'éléments sphériques (cellules) et de bâtonnets (fibres) et de détecter automatiquement dans de tels systèmes des amas d'objets sphériques séparés par des bâtonnets. Cet outil peut également être utilisé pour traiter des images biologiques issues de microscopie en 3D, permettant ainsi une comparaison directe entre les structures *in vivo* et *in silico*. L'étude des structures produites par le modèle ECMmorpho-3D via des simulations numériques montre que ce modèle peut générer spontanément différents types d'architectures, que nous identifions et caractérisons grâce à notre outil d'analyse. Une analyse paramétrique approfondie nous permet d'identifier une variable émergente, le nombre de liens par fibre, qui explique et, dans une certaine mesure, prédit le devenir du système modélisé. Une analyse temporelle révèle que l'échelle de temps caractéristique de ce processus d'auto-organisation est fonction de la vitesse de remodelage du réseau et que tous les systèmes suivent la même trajectoire évolutive. Enfin, nous utilisons le modèle ATmorpho-3D pour explorer l'influence de cellules sphériques sur l'organisation d'un réseau de fibres dynamique, en prenant comme référence le modèle ECMmorpho-3D. Nous montrons que le nombre de cellules influence l'alignement local des fibres mais pas l'organisation globale du réseau. Par ailleurs, les cellules s'organisent spontanément en amas entourés de feuillets de fibres, dont les caractéristiques morphologiques sont très proches de celles des structures cellulaires *in vivo*. De plus, la distribution des différentes morphologies d'amas cellulaires est similaire dans les systèmes *in silico* et *in vivo*. Ceci suggère que le modèle est capable de produire des morphologies réalistes non seulement à l'échelle d'un amas mais aussi à l'échelle du système entier, en reproduisant les variabilités structurelles observées dans les échantillons biologiques. Une analyse paramétrique révèle que la proportion de chaque morphologie dans un système *in silico* est gouvernée principalement par les capacités de remodelage du réseau de fibres, pointant le rôle essentiel des propriétés de la matrice extra-cellulaire dans l'architecture et le fonctionnement du tissu adipeux (ce qui concorde avec plusieurs constatations biologiques ainsi que des résultats antérieurs en 2D). Le fait que ces modèles mathématiques très simples puissent générer des structures réalistes corrobore notre hypothèse selon laquelle l'architecture des tissus biologiques pourrait émerger spontanément à partir d'interactions mécaniques locales entre les composants du tissu, indépendamment des phénomènes biologiques complexes se déroulant dans ce tissu. Ce travail ouvre de nombreuses perspectives quant à notre compréhension des principes fondamentaux gouvernant la manière dont l'architecture d'un tissu émerge durant l'organogenèse, est maintenue au cours de la vie et peut être affectée par diverses pathologies. Les applications potentielles vont de l'ingénierie tissulaire à la possibilité de promouvoir la régénération chez les mammifères adultes.

Keywords : Mathematical biology ; Individual-based models ; 3D visualisation and segmentation ; Connective tissues ; 3D dynamical fibre networks ; Adipose tissues.

Abstract : In this thesis, we investigate whether simple local mechanical interactions between a reduced set of components could govern the emergence of the 3D architecture of biological tissues. To explore this hypothesis, we develop two mathematical models. The first one, ECMmorpho-3D, aims at reproducing a non-specialised connective tissue and is reduced to the extra-cellular matrix component, that is a 3D dynamically connected fibre network. The second, ATmorpho-3D, is built by adding to this network spherical cells which spontaneously appear and grow in order to mimic the morphogenesis of adipose tissue, a specialised connective tissue with major biomedical importance. We then construct a unified analysis framework to visualise, segment and quantitatively characterise the fibrous and cellular structures produced by our two models. It constitutes a generic tool for the 3D visualisation of systems composed of a mixture of spherical (cells) and rod-like (fibres) elements and for the automatic detection of in such systems of clusters of spherical objects separated by rod-like elements. This tool is also applicable to biological 3D microscopy images, enabling a comparison between *in vivo* and *in silico* structures. We study the structures produced by the model ECMmorpho-3D by performing numerical simulations. We show that this model is able to spontaneously generate different types of architectures, which we identify and characterise using our analysis framework. An in-depth parametric analysis lead us to identify an intermediate emerging variable, the number of crosslinks per fibre, which explains and partly predicts the fate of the modelled system. A temporal analysis reveals that the characteristic time-scale of the organisation process is a function of the network remodelling speed, and that all systems follow the same, unique evolutionary pathway. Finally, we use the model ATmorpho-3D to explore the influence of round cells over the organisation of a fibre network, taking as reference the model ECMmorpho-3D. We show that the number of cells can influence the local alignment of the fibres but not the global organisation of the network. On the other hand, the cells inside the network spontaneously organise into clusters with realistic morphological features very close to those of *in vivo* structures, surrounded by sheet-like fibre bundles. Moreover, the distribution of the different morphological types of clusters is similar in *in silico* and *in vivo* systems, suggesting that the model is able to produce realistic morphologies not only on the scale of one cluster but also on the scale of the whole system, reproducing the structural variability observed in biological samples. A parametric analysis reveals that the proportion in which each morphology is present in an *in silico* system is governed mainly by the remodelling characteristic of the fibres, pointing to the essential role of the extra-cellular matrix properties in adipose tissue architecture and function (in agreement with several biological results and previous 2D findings). The fact that these very simple mathematical models can produce realistic structures supports our hypothesis that biological tissues architecture could emerge spontaneously from local mechanical interactions between the tissue components, independently of the complex biological phenomena taking place around them. This opens many perspectives regarding our understanding of the fundamental principles governing how biological tissue architecture emerges during organogenesis, is maintained throughout life and can be affected by various pathological conditions. Potential applications range from tissue engineering to therapeutic treatment inducing regeneration in adult mammals.