Modelling and characterizing axon growth from in vivo data

Agustina Razetti

UNIVERSITÉ CÔTE D'AZUR

ÉCOLE DOCTORALE
SCIENCES ET TECHNOLOGIES DE L'INFORMATION ET DE LA COMMUNICATION

THÈSE DE DOCTORAT

Modélisation et caractérisation de la croissance des axones à partir de données *in vivo*

Agustina RAZETTI

Équipe MORPHEME – INRIA Sophia Antipolis/I3S/iBV

Présentée en vue de l'obtention
du grade de docteur en Automatique, traitement du signal et des images
d'Université Côte d'Azur
Dirigée par : Xavier Descombes
Soutenue le : 13 Avril 2018

Devant le jury, composé de :
Alin Achim, PhD, Bristol University
Florence Besse, PhD, Institute Biology Valrose
Kristian Franze, PhD, Cambridge Univeristy
Caroline Medioni, PhD, Institute Biology Valrose
Jonas Ranft, PhD, École Normale Supérieur
Anuj Srivastava, PhD, Florida State University
Michèle Studer, PhD, Institute Biology Valrose

UNIVERSITÉ CÔTE D'AZUR

SIGNALIFE

UNIVERSITY CÔTE D'AZUR

DOCTORAL SCHOOL STIC

SCIENCES ET TECHNOLOGIES DE L'INFORMATION ET DE LA COMMUNICATION

PHD THESIS

to obtain the title of

PhD of Science

of the University Côte d'Azur

Specialty : CONTROL SYSTEMS, SIGNAL AND IMAGE PROCESSING

Defended by:

Agustina RAZETTI

Modelling and characterizing axon growth from *in vivo* data

Thesis advisor: Xavier DESCOMBES

Prepared at MORPHEME Team, INRIA Sophia Antipolis/I3S/iBV

Defended on April 13, 2018

Jury :

Reporters :	Alin ACHIM	—	Bristol University
	Kristian FRANZE	—	Cambridge University
	Anuj SRIVASTAVA	—	Florida State University
Director :	Xavier DESCOMBES	—	Inria Sophia Antipolis Méditerranée
Examinators :	Florence BESSE	—	Institute Biologie Valrose
	Caroline MEDIONI	—	Institute Biologie Valrose
	Jonas RANFT	—	École Normale Supérieure
	Michèle STUDER	—	Institute Biologie Valrose

UNIVERSITÉ CÔTE D'AZUR

ÉCOLE DOCTORALE STIC

SCIENCES ET TECHNOLOGIES DE L'INFORMATION ET DE LA COMMUNICATION

T H È S E

pour l'obtention du grade de

Docteur en Sciences

de la Université Côte d'Azur

Mention : AUTOMATIQUE, TRAITEMENT DU SIGNAL ET DES IMAGES

présentée et soutenue par:

Agustina RAZETTI

Modélisation et caractérisation de la croissance des axones à partir de données *in vivo*

Thèse dirigée par: Xavier DESCOMBES

préparée dans l'Équipe MORPHEME, INRIA Sophia Antipolis/I3S/iBV

Prévu pour la défense le 13 Avril, 2018

Jury :

Rapporteurs :	Alin ACHIM	—	Bristol University
	Kristian FRANZE	—	Cambridge University
	Anuj SRIVASTAVA	—	Florida State University
Directeur :	Xavier DESCOMBES	—	Inria Sophia Antipolis Méditerranée
Examinateurs :	Florence BESSE	—	Institute Biologie Valrose
	Caroline MEDIONI	—	Institute Biologie Valrose
	Jonas RANFT	—	École Normale Supérieure
	Michèle STUDER	—	Institute Biologie Valrose

Abstract

How the brain wires up during development remains an open question in the scientific community across disciplines. Fruitful efforts have been made to elucidate the mechanisms of axonal growth, such as pathfinding and guiding molecules. However, recent evidence suggests other actors to be involved in neuron growth *in vivo*. Notably, axons develop in populations and embedded in mechanically constrained environments. Thus, to fully understand this dynamic process, one must take into account collective mechanisms and mechanical interactions within the axonal populations. However, techniques to directly measure this from living brains are today lacking or heavy to implement.

This thesis emerges from a multidisciplinary collaboration, to shed light on axonal development *in vivo* and how adult complex axonal morphologies are attained. Our work is inspired and validated from images of single wild type and mutated *Drosophila* γ axons, which we have segmented and normalized.

We first proposed a mathematical framework for the morphological study and classification of axonal groups. From this analysis we hypothesized that axon growth derives from a stochastic process, and that the variability and complexity of axonal trees result from its intrinsic nature, as well as from elongation strategies developed to overcome the mechanical constraints of the developing brain. We designed a mathematical model of single axon growth based on Gaussian Markov Chains with two parameters, accounting for axon rigidity and attraction to the target field. We estimated the model parameters from data, and simulated the growing axons embedded in spatially constraint populations to test our hypothesis.

We dealt with themes from applied mathematics as well as from biology, and unveiled unexplored effects of collective growth on axonal development *in vivo*.

Key words: axon growth, morphogenesis, stochastic modelling, Gaussian Markov chains, mechanical interactions, axon branching, elongation strategies

Résumé

La construction du cerveau et de ses connexions pendant le développement reste une question ouverte dans la communauté scientifique. Des efforts fructueux ont été faits pour élucider les mécanismes de la croissance axonale, tels que la guidance axonale et les molécules de guidage. Cependant, des preuves récentes suggèrent que d'autres acteurs seraient impliqués dans la croissance des neurones *in vivo*. Notamment, les axones se développent dans des environnements mécaniquement contraints. Ainsi, pour bien comprendre ce processus dynamique, il faut prendre en compte les mécanismes collectifs et les interactions mécaniques au sein des populations axonales. Néanmoins, les techniques pour mesurer directement cela à partir de cerveaux vivants sont aujourd'hui insuffisantes ou lourdes à mettre en œuvre.

Cette thèse résulte d'une collaboration multidisciplinaire, pour faire la lumière sur le développement axonal *in vivo* et les morphologies complexes des axones adultes. Notre travail a été inspiré et validé à partir d'images d'axones γ individuels chez la drosophile, de type sauvage et modifiés génétiquement, que nous avons segmentés et normalisés.

Nous avons d'abord proposé un cadre mathématique pour l'étude morphologique et la classification des groupes axonaux. A partir de cette analyse, nous avons émis l'hypothèse que la croissance axonale dérive d'un processus stochastique et que la variabilité et la complexité des arbres axonaux résultent de sa nature intrinsèque, ainsi que des stratégies d'élongation développées pour surmonter les contraintes mécaniques du cerveau en développement. Nous avons conçu un modèle mathématique de la croissance d'un axone isolé fondé sur des chaînes de Markov gaussiennes avec deux paramètres, représentant la rigidité axonale et l'attraction du champ cible. Nous avons estimé les paramètres de ce modèle à partir de données réelles et simulé la croissance des axones à l'échelle de populations et avec des contraintes spatiales pour tester notre hypothèse.

Nous avons abordé des thèmes de mathématiques appliquées ainsi que de la biologie, et dévoilé des effets inexplorés de la croissance collective sur le développement axonal *in vivo*.

Mots clés: croissance axonale, morphogenèse, modélisation stochastique, chaînes de Markov gaussiennes, interactions mécaniques, ramification axonale, stratégies d'élongation

To my favourite women; my mother and her mother

What you are doing does not matter so much as what you are learning from doing it. It is better not to know and to know that one does not know, than presumptuously to attribute some random meaning to symbols.

Temple of Amen of Ipet-isut (Karnak), 1391-1351 BC.

Acknowledgements

I would first like to thank my director Xavier Descombes, for his wise advise and support during this thesis. You have always been available for discussion when needed, but also let me the opportunity of making my own decisions and freedom to chose the direction of this work, which I very much appreciated. I am also enormously grateful to Florence Besse and Caroline Medioni for their guidance and constant commitment with this thesis, as well as for their personal counselling.

Very importantly, I want to specially thank the members of the jury for having accepted to evaluate this thesis: Alin Achim, Kristian Franze, Anuj Srivastava, Michele Studer and Jonas Ranft. Thank you a lot for your time when reading it, your valuable comments and corrections and interesting questions.

My everyday life during this thesis would have not been the same without the company of the amazing people that constitute the Morpheme team. Thank you for all the shared lunchtimes, coffee breaks, jokes and good moments, but also for your support and help when I needed it. A special mention for my fellow PhD students during these years, Alejandro Mottini, Alexis Zubiolo, Emmanuel Soubies, Gael Michelin, Lola Bautista, Arnauld Bletterer and Anca Grapa. Also, for the permanent members of Morpheme, thank you for helping to create and transmit this very special friendly environment in the team.

Last -but not least- I want to thank the people who stood besides me through all this process, since I decided to do my PhD in France to the very last day of my defence, in particular my mother, father and brother, as well as my closest friends and Flo, for the everyday encouragement.

Table of contents

List of figures

List of tables

Chapter 1

Introduction

1.1 Motivation and goals

In the nervous system, neurons transmit information to others by projecting long and branched extensions called axons. This thesis explores the vast and complex subject of neuron development and axonal morphogenesis *in vivo*, using diverse domains such as applied mathematics, probability theory, modelling and informatics. Intrinsically interdisciplinary, this work proposes new hypotheses and answers on this fundamental subject, and intends to close the existent gap between plausible experimentation and the complex reality of biological mechanisms.

During the last decades, important advances have been achieved in recognizing the main factors implicated in neuron development. External guiding cues involved in axonal pathfinding have been largely described, as well as the internal molecular machinery that allows axonal elongation, including many of the biochemical pathways regulating these processes. Interestingly, recent work shows that axons are not only capable of sensing and responding to external chemical signals, but are also guided by local changes in the mechanical properties of their surroundings. While most of these studies consider as mechanical environment mainly the extracellular matrix and the general mechanical properties of the tissue, little is known about intra-population axon-axon mechanical interactions and space confinement. In consequence, the effect of these physical phenomena on axon development and how they influence their final morphologies are still unknown. Furthermore, to measure mechanical interactions between individual growing neurons in the developing brain is experimentally not trivial, as tools to reproducibly visualize and manipulate axon-axon interactions *in vivo* in populations of growing neurons are either lacking or heavy to implement.

In this work we aim to better understand the origin of the complexity and variability of adult axonal morphologies as a consequence of their developmental process *in vivo*. Thus,

we consider and integrate the mentioned intra-population mechanical interactions in the otherwise better known individual axonal growth process. To do so, we propose *in silico* approaches validated on available *in vivo* data to expand the information given by the latter. As a result, we show the importance of considering axonal development as a dynamic and complex process of collective morphogenesis, rather than the sum of the growth of isolated individual axons.

Specifically, we first proposed an original classification framework to better characterize axonal morphologies and understand the similarities and differences between genotypically different axon groups. The framework consists of the stochastic modelling of relevant morphological features, the estimation of the model parameters from data and the classification of axons using likelihood analysis, enriched with statistics for a global description of axonal groups. Next, we designed a mechanistic individual axon elongation model based on Gaussian Markov chains with biologically related parameters, capable of generating realistic axonal morphologies. The estimation of the model parameters from real data happened to be not trivial, so we explored different approaches to solve this question. Using this individual axon growth model, we simulated the development of entire axonal populations considering external constraints such as an attractive guiding source, realistic geometrical limits and volume exclusion between the growing axons. These simulations allowed us to analyse the emergent effects of considering a population of axons represented as interacting self propelled elements. The validation of the simulation results with real data showed that our model is able to reproduce realistic morphologies. Interestingly, our simulations reproduced morphological characteristics of the axonal population that had not been included in the model, but rather emerged from its rules. Because our framework is based on the real biological process, it allowed us to better understand some aspects of axonal development *in vivo*. Finally, we could reproduce different mutation phenotypes by changing related aspects in our model. In consequence, our approach nicely assembles the three desirable aspects of a model: to be generative, explicative and predictive.

This thesis is the result of a solid collaboration with biologists from the Besse team at the Institute de Biologie de Valrose, Nice, France. They have provided us with rich image datasets from which this work was built, as well as biological expertise and advice. The stress of this thesis has been set on the inter collaboration and continuous feedback between mathematical modelling techniques and biological pertinence. Our dataset, composed of 3D confocal images of fixed *in vivo* single *Drosophila* γ axons at adult stage, as well as *live* 4D sequences of single or a group of *Drosophila* γ axons growing during metamorphosis, not only inspired our model, but was also used to estimate its parameters, as well as to validate

its results and predictions. Wild-type axons as well as different mutant genotypes where considered to enrich the study.

1.2 Main objectives

In summary, our two main goals are:

- To develop a methodological framework for the analysis of axonal shapes based on morphometric features and validate it with maximum likelihood classification of real axons.
- To propose a model for axonal development that considers the stochastic path generation and branching mechanisms with parameters that can be estimated from data; and that allows a straightforward implementation of population growth considering external constraints such as an attractive/repellent source, geometrical limits and volume exclusion with other axons. Ideally, this model should be generative, explicative and predictive of different real axon genotypes.

1.3 Main Contributions

The major contributions of this work are:

In modelling

- Tree-like structures classification method based on the stochastic modelling of morphological features.
- Two-parameter stochastic model to represent trajectories in space that allows renormalization (equivalence between different spatial scales) and parameter estimation from data.
- General framework with two opposite approaches to detect jump points and estimate the parameters from piecewise homogeneous Markov chains.

In biology

- Deep morphological analysis of wild-type and mutant γ axons which resulted in new observations.

- Showcase the importance of mechanical interactions between growing axons and confinement on *in vivo* neuronal development.
- Proposal of a novel elongation mechanism where branching is intrinsic to the axonal growth process.
- Original interpretation of the *imp* mutant phenotype.

1.4 Manuscript Organization

This manuscript is organised in three sequential blocks: i-general axon growth biological and modelling background, ii- stochastic framework for the study of axonal morphology and iii- the development of the individual axon growth model and the simulation of developing populations.

Chapter 2 presents the background of this work (i). We first summarize the state of the art on axonal growth as well as on axonal growth mathematical and computational models, highlighting the aspects relevant for this thesis. We then describe the used biological model: the *Drosophila* mushroom body γ neurons (wild-type and mutant) and the image database used throughout this work.

We start our contributions in Chapter 3, where we mainly aim at a deep understanding of the axon morphologies in our database (ii). We first describe the data treatment scheme that allowed us to obtain 3D axonal reconstructions from the images, and present a general stochastic framework for neuronal morphological comparison. In particular, it consists in applying stochastic models for each axon morphological feature, estimate their parameters from data, to finally contrast and compare the axons using likelihood analysis and statistics (Razetti et al., 2016, 2017).

Chapter 4 represents the main part of this thesis. It consists of the development of a mechanistic and dynamic mathematical model of individual axon growth, and its implementation in a realistic environment considering mechanical interactions between the growing axons (iii). The chapter starts by describing the data treatment that allows considering isolated axons as a population. Then we describe the proposed mathematical model, and present different approaches to estimate its parameters from data. Further, we perform simulations and analyse the results, comparing and contrasting them with data. We propose different branching occurrence hypothesis, and unveil new plausible elongation mechanisms. Finally, we further study the significance of the parameter values obtained from data and their implications on the collective growth of space-constraint tubular structures.

In the end, we present a general discussion of the contributions of this thesis and propose perspectives on different areas (Chapter 5).

1.5 List of Publications

Conference

- Razetti, A., Descombes, X., Medioni, C., & Besse, F. (2016, February). Statistical Characterization, Modelling and Classification of Morphological Changes in imp Mutant Drosophila Gamma Neurons. In BIOSTEC 2016-The 9h International Joint Conference on Biomedical Engineering Systems and Technologies.

Book chapter

- Razetti, A., Descombes, X., Medioni, C., & Besse, F. (2017). A Stochastic Framework for Neuronal Morphological Comparison: Application to the Study of imp Knockdown Effects in Drosophila Gamma Neurons. In: Fred A., Gamboa H. (eds) Biomedical Engineering Systems and Technologies. BIOSTEC 2016. Communications in Computer and Information Science, vol 690. Springer, Cham

Conference with abstract

- Razetti, A., Medioni, C., Malandain, G., Besse, F., & Descombes, X. (2017, May). Modelling collective axon growth from in vivo data reveals the importance of physical axon-axon interactions. In Cell biology of the neuron: Polarity, plasticity and regeneration.

- Razetti, A., Medioni, C., Malandain, G., Besse, F., & Descombes, X. (2018). Understanding in vivo axonal development as an inherently asymmetric process. First European Asymmetry Symposium.

Chapter 2

Background

If the human brain were so simple that we could understand it, we would be so simple that we couldn't.

Emerson M. Pugh

2.1 Axon growth: state of the art

The first historic known mention of the human brain, «skull-offal», was found in a series of Egyptian documents on medicine and surgery called the Edwin Smith Papyrus, written in 1600 BC (but probably copied from earlier texts, from 3000 BC). Even though bounds in the head were at that time actually associated with different serious pathologies, like speech or walk loss, ancient Egyptians believed that intelligence and sensation were driven by the heart, and not brain, and thus despised the importance of this organ. Almost 5000 years later, Ramon y Cajal (1852-1934) first described the neuron as the functional unit of the brain. Since then, the complexity of this marvellous organ has not stopped intriguing and challenging scientists of different disciplines across the world. In this thesis, we explore one of the most captivating open questions in neurobiology: «How» is the brain created. In particular, we focus on neuron growth and how its complex final morphology is attained. Specifically, we address the question about how the physical presence of other neurons as well as spatial confinement are involved in the growth process.

During development, neurons in our nervous system extend their axonal trees in order to reach target neurons or structures, and establish complex networks that allow the correct functioning of the body and the mind (see Fig. 2.1 for a general scheme of the neuron). To do so, they need to *recognize* the spatial position of their final target as well as the path they should follow until they reach it; and complete the growth process in a limited time.

Furthermore, axons can also create branches that emerge from their shafts, enhancing the number of possible connections. However, when and where these branches are formed, as well as how long each of them should grow is still poorly understood.

When neurons somehow fail to decode or correctly follow these instructions, they develop corrupted adult morphologies that give rise to a large spectrum of diseases. Engle (2010) reviews human syndromes known to be caused by aberrant axon connectivity. Corpus Callosum Dysgenesis, L1 Syndrome and Joubert Syndrome and Related Disorders arise when axons fail to cross the brain mid-line. They can cause a variety of symptoms, such as severe mental retardation. Other serious syndromes are caused by guidance defects of cranial nerves. For example, the Kallmann Syndrome is believed to be caused by errors in growth and guidance of olfactory axons. Another example is Albinism, in which visual deficiencies arise from altered pattern of axonal projections in the visual system, due to axonal misguidance.

Thus, both to understand how the brain achieves its complex wiring to function in normal conditions, as well as to unveil the causes of severe syndromes, the study of axonal growth during development is highly relevant. During this section, we introduce the basic mechanisms of axonal development that motivated this work. We first explain how axons find their paths guided by chemical gradients to achieve elongation. Then, we explore different new findings on the role of the mechanical environment of the developing brain and how it also constitutes an important cue on axonal growth, as well as refines axonal final morphologies. Finally, we stress the importance of considering axonal elongation as a collective process, where axons do not grow isolated, but within populations, and explain main aspects of axonal branching.

Axon elongation is possible thanks to the structure present in their tip during development named the growth cone (GC). GCs present highly dynamic behaviours, and are subdivided into three main regions: i) a proximal core region enriched in microtubules ii) a flat area called lamellipodia made of inter-crossed actin fibres, and iii) finger-like protrusions named filopodia, which are held by elongated actin fibres (Fig. 2.2). The growth cone senses its chemical environment and decides the direction of the neurite growth. Many signalling molecules have been described to interact with the GC during development. They can be found on other cells surfaces, attached to the extra cellular matrix (short-rate cues) or diffusing in the extracellular domain (long-rate cues). Directional guiding cues set gradients that indicate the GC preferred paths or directionality of growth, while positional cues indicate when to stop. These molecules determine the outgrowth as well as the guidance of the axon shaft, and can be attractive or repellent, providing the essential information to guide the axon towards their specific target, even during long trajectories (Bixby and Harris, 1991; Guan and

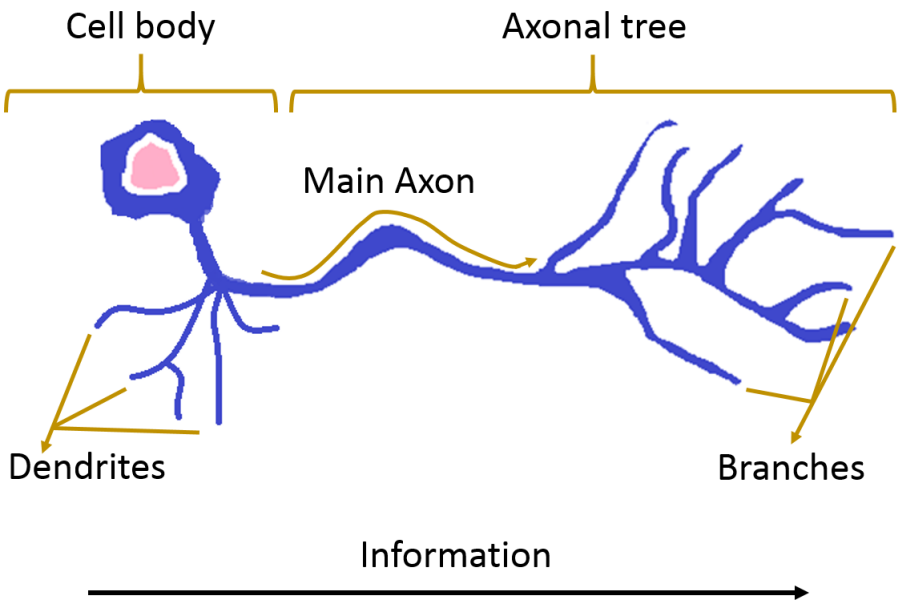

Fig. 2.1 Scheme of a unipolar neuron. Input from one or more neurons is received by the dendrites, and the information travels along the axon shaft towards its branches. These terminals form synapses with other neurons to which the information is transmitted. The morphology of the axonal tree determines the neurons connectivity patterns.

Rao, 2003; Kolodkin and Tessier-Lavigne, 2011; Plachez and Richards, 2005; Song and Poo, 2001; Tessier-Lavigne et al., 1996). Interestingly, guiding cues are multifunctional, meaning that the same molecule can be attractive or repellent, as well as act in a short or long rate manner. In addition, GCs respond not only to extrinsic but also to intrinsic factors, and the combination of these mechanisms may be at the origin of the diversity and complexity of GC behaviour that wires up the brain (Dickson, 2002).

Growth cone chemotaxis (i.e. movement in response to the influence of chemical stimulation) occurs in two basic actions: «directional sensing» and «motility», which together occur in the context of «polarization». «Directional sensing» means that the growth cone is capable of sensing and amplifying noisy gradients of molecules to decode the spatial information of growth directionality (Fig. 2.2A). «Motility» is achieved thanks to the coordinated dynamics of the cytoskeleton components (actin and tubulin) in the growth cone, and it means that the GC can reshape itself to be oriented towards the maximum attraction gradient. This remodelling combines two actions: i) the extension of protrusions towards the gradient directionality and ii) the retraction of the rest of the structure (Fig. 2.2B). These actions lead to the GC polarization, meaning that it maintains a *leading edge*, or front, and a *tail*. Interestingly, GCs can present different degrees of polarization. For example, "paused" GCs were observed to present varied and complex shapes, to maximize their sensitivity, while CGs growing along a well defined gradient adopt simpler bullet-shapes, related with less sensitivity and faster movement (Dent et al., 2011; Dickson, 2002; Mortimer et al., 2008; Plachez and Richards, 2005). Once the direction of the movement has been sensed, the shaft has to elongate (Fig. 2.2C). The most accepted theory of how this elongation is achieved is called "the clutch" hypothesis (Mitchison and Kirschner, 1988). It suggests that the adhesion of the growth cone to adhesive substrates in other cells surfaces or in the extracellular matrix creates a "clutch", or anchor, that prevents retrograde flow and thus promotes actin-based outgrowth of the growth cone (Lowery and Van Vactor, 2009).

Interestingly, even though bibliography on axon outgrowth is vast and chemotaxis has been largely described, recent work shows that axons are not only capable of sensing and responding to external chemical signals, but are also guided by local changes in the mechanical properties of their surrounding (Francisco et al., 2007; Franze, 2013; Kerstein et al., 2015; Koser et al., 2016; Sagasti et al., 2005). Francisco et al. (2007) study the growth of embryonic chicken primary sensory neurons both *in vitro* and *in vivo*. Their results highlight the importance of the 3D mechanical environment in axonal elongation, defined by the extracellular matrix and spatial confinement. They show *in vivo* that growth cones have mechanisms to sense mechanical constraints and respond by the redirection of the growth, and thus propose that mechanical constraints may act complementary to chemical repellent

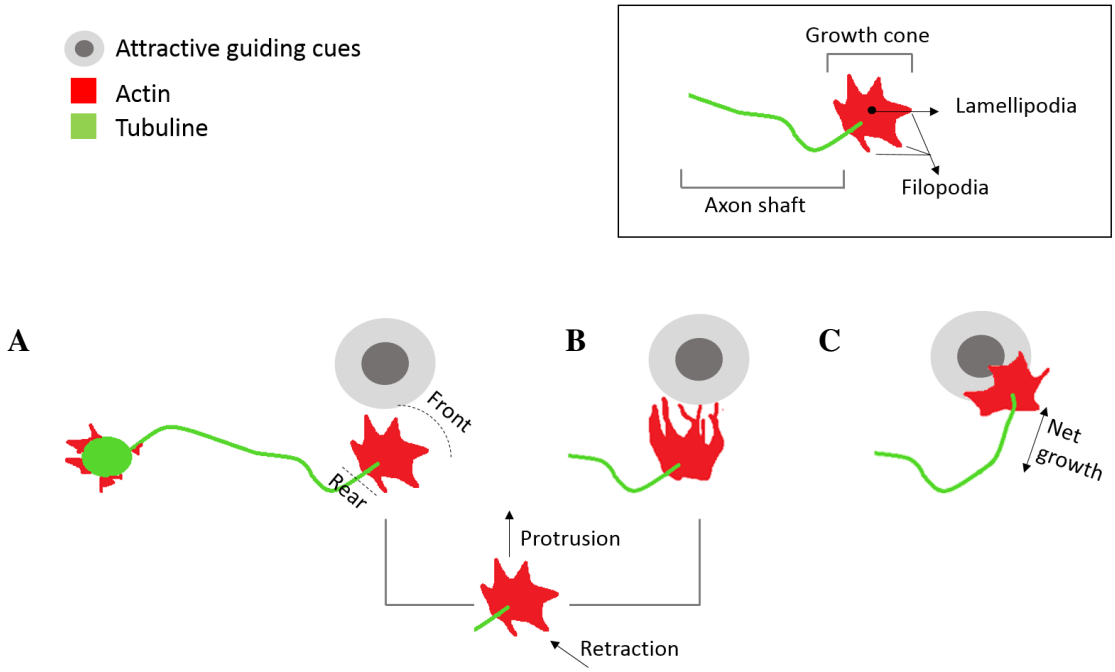

Fig. 2.2 Scheme of growth cone chemotaxis. (A) The growth cone senses a gradient set by external guiding cues. (B) It modifies its shape in order to direct its front toward the direction of the attraction field. (C) The axon shaft grows towards the established direction.

cues. Moreover, *in vitro*, they observe that the number of neurons establishing an axon was affected by the constrained space, proposing a direct link between available space and net axon growth. From dynamic observations, they report that axons encountering mechanical barriers spend time interacting with them, and propose this is the reason they stay shorter compared to axons without or encountering less mechanical constraints during their growth. This last conclusion gives rise to an interesting proposal, specifically when considering the developing brain, where the growth and developmental programs are time-limited. Kerstein et al. (2015) review the last findings on how the mechanical micro-environment of the GCs control their trajectory. They highlight the role of mechanical signals sent by immobilized molecules attached to the 3D extracellular matrix and of forces generated between the growth cone and its environment. They also propose that mechanical barriers may influence axonal elongation, as it is observed for the loss of regeneration capacity of peripheral axon with age, which is mainly caused by the mechanical barriers imposed by glial and nerve debris. Finally, they highlight the importance of durotaxis (i.e. movement in response to the influence of mechanical gradients), also largely studied by Koser et al. (2016). Through *in vitro* and *in vivo* experiments, this work shows the importance of local tissue stiffness as well as the role of mechanosensitive ion channels during neuron growth in the *Xenopus* retinal ganglion cells. They find out that axons exhibit a faster and directionally persistent net growth in stiffer tissues, while in soft tissues the growth cone moves locally faster exploring its surrounding but the net axonal elongation is slower. Thus, in general, axons growing in softer tissues grow less coherently and cross each other more frequently. Also, they have observed axons extending from stiffer to softer regions which were guided mainly by this gradient of density. As well as Francisco et al. (2007), Koser et al. (2016) mention the importance of space availability on the mechanical control of axon growth. All this works underline the importance of including physical interactions to our understanding of neuron growth and conclude that a major challenge and open question is to understand how the mechanical environment of the developing neurons affects their growth and final morphology *in vivo*. While most of them consider as mechanical environment mainly the extracellular matrix, spatial confinement and other tissue components, Sagasti et al. (2005) interestingly highlights the importance of inter-axon contact interactions for the morphogenesis of axonal trees. Through *live* imaging of *Zebrafish* peripheral sensory axons, they conclude that direct repulsion caused upon axon-axon contact is the predominant force sculpting the final trees. As mentioned in the Introduction, inter-axon mechanical interactions, as well as space confinement, are main topics of study of this thesis.

Taking into account inter-axon interactions during development carries wider considerations. In particular, it means that we do not further consider neurons growing and relating

to their environment *individually*, as in classical studies, but as part of a population in a *collective* way. During the last few years, the study of tissue development and cell migration in general has started to consider group effects and collective models to explain cell behaviour (Haeger et al., 2015; Schumacher et al., 2016; Vedula et al., 2013; Weijer, 2009; Woods et al., 2014). Not only cell-environment but also cell-cell interactions were shown to lead from local chaos to a general gobal order as a function of cell density (Mehes and Vicsek, 2014; Vedula et al., 2013); phenomenon known as collective motion. Even though classically cell collective motion is defined as a group of moving cells that are connected via cell-cell junctions (Haeger et al., 2015), recent works proposed to extend the definition scope to other cell-cell interactions. Chick cranial neural crest cells, for example, control their displacements autonomously, but still rely on cell-cell contacts for overall group navigation. This behaviour is called *multicellular streaming* or *loose* collective cell migration (Friedl et al., 2012; Rørth, 2012; Schumacher et al., 2016), and it is similar to the mechanism described by Sagasti et al. (2005) to shape the final axonal trees in *Zebrafish* peripheral sensory axons. Here, neurons respond to internal thus individual developmental programs, but rely on the contact with others to adopt their adult morphology. Furthermore, other important forces than attractive and repulsive local interactions are described to be needed to order and give directionality to the collective system; such as space confinement and external chemoattractants (Haeger et al., 2015; Theveneau and Mayor, 2012; Vedula et al., 2013).

Different strategies to optimize growth in the context of increasing numbers of neurons competing for space have been developed across evolution. These include organization of neurons into populations that coordinate to grow and innervate target territories. Knowing the relative position of neighbour growing axons enormously helps populations of neurons to find they targets. Similarly, axon-axon interactions can be helpful to refine target innervation previously indicated by chemical cues. For example, in the fly, repulsive axon-axon interactions between the olfactory receptor neurons assure them to innervate the correct area of the antennal lobe (Hong and Luo, 2014; Luo and Flanagan, 2007). Similar observations were done also in vertebrates (Petrovic and Schmucker, 2015). This kind of target independent interactions play a fundamental role in axonal development and organization. On its side, also branching can be considered as an axonal strategy to increase the number of targets and thus the circuit complexity (Schmidt and Rathjen, 2010). However, even though *in vitro* studies have already identified branch-inducing factors, the mechanisms governing axonal branching *in vivo* remain an open question (Petrovic and Schmucker, 2015).

On their side, branches are an essential aspect of axonal morphology and functionality. During development, they can emerge by two distinct mechanisms: splitting of the terminal GC, or extend interstitially. In the latter case, branches develop from growth cone remnants

that accumulate on the axon shaft in regions where GC paused (Kalil and Dent, 2014; Schmidt and Rathjen, 2010), and was reported to be the most frequent branch occurrence mechanism *in vivo* (Schmidt and Rathjen, 2010). Different signalling pathways exist that regulate the different branching triggering mechanisms, even in the same type of axon, and it is the role of the GC to interpret and integrate simultaneously several extracellular cues that activate different internal cascades and may indicate to start a branching event. In addition, neuronal activity can also trigger or regulate branch occurrence, by generating transient fluctuations on intracellular calcium, which is an essential second messenger involved in the process (Gibson and Ma, 2011; Kalil and Dent, 2014). The question of axon branching *in vivo* is a relevant subject of this thesis, and is further studied on Chapter 4.

In summary, recent studies have highlighted the importance of the mechanical environment of the developing brain, proposing that physical cues and axon-axon contact need to be equally considered as chemical guidance in the neuron growth process. However, up to now, most attention has been paid to the mechanical interaction with the extracellular matrix and other neighbouring tissues, rather than direct contact between growing neurons, which has already been described for other cell types as *loose* collective phenomena (Friedl et al., 2012; Rørth, 2012; Schumacher et al., 2016). Although studies have revealed that inter-neuron coordination and interactions are particularly important in the context of a developing population (Demyanenko and Maness, 2003; Goodhill and Richards, 1999; Hua et al., 2005; Luo and Flanagan, 2007; Sagasti et al., 2005; Wang and Marquardt, 2013), axon growth has so far been mainly studied *in vitro* on isolated neurons, or *in vivo* on entire neuronal populations. However, to understand population growth as a whole, one needs to understand the behaviour of single constituent neurons, and how they interact and influence themselves to produce global growth. In this thesis we propose to consider axons growing as a population, where elongation and tree-shaping emerge from internal growth rules as well as from external interactions. In particular, we study the mechanical interactions between developing axons within a dense growing population, and explore their influence on adult morphology.

2.2 Axon growth models: state of the art

The bibliography regarding mathematical modelling of individual axon growth is vast. It is beyond the scope of this introduction to review in detail on the subject, but rather to mention its main areas and advances during the last years. Simpson et al. (2009) discuss how the field of axon growth models is fragmented -*i.e.* lacks of a unified framework- and propose a classification scheme into phenomenological, mechanistic and abstract. Phenomenological

models propose mathematical formulations to mimic behaviours observed experimentally, but without relying directly on biological mechanisms. Mechanistic models, on the other hand, rely closely on biological or biophysical processes to control the growth cone pathfinding and/or axonal elongation. Finally, abstract models describe the effects or results of these mechanisms into neuron growth in a more general way, and without necessarily modelling the biological mechanisms. This classification is quite general and can be applied to every biologically-inspired mathematical model, and results useful as it sets the limits of a model to be explicative and predictive of the biological mechanisms under study. However, one must remember that each model results useful to better understand different aspects of axonal growth, and a model including every known mechanism involved in neuron development would not only be heavy to implement, but also difficult to analyse and tune its parameters. Here, we present and briefly discuss a different way of classifying axonal growth models, better suited to this work. We consider four main types of models: i) *in vitro*-inspired, ii) 3D realistic, iii) inter-neuron interaction/competition and iv) 3D space-embedded models. We describe and cite relevant bibliography on groups i) to iii), and then go into more detail on group iv, which results the most pertinent for this work. Finally, we discuss models that take into account branch generation and their particular approaches and limits.

i) In their origin, axonal growth models have been mainly concerned by the study of GC guidance and neurite (axon/dendrite) elongation, and were mostly based on *in vitro* experiments, proposing isolated neurons growing in 2D environments. They include models describing overall morphology generation based on the diffusion of chemical substances in the environment and axon/branch shafts (Graham and Van Ooyen, 2001), on stochastic rules regarding branch occurrence probabilities for different neuron types (van Pelt and Schierwagen, 2004), or based on the specific role of molecules regulating internal calcium concentration (Hely et al., 2001). Others focused on microtubule or actin dynamics (De Rooij et al., 2017; McLean et al., 2004), as well as growth cone guidance, path-finding and elongation (Aeschlimann and Tettoni, 2001; Kobayashi et al., 2010; Mortimer et al., 2009; Nguyen et al., 2016; Segev and Ben-Jacob, 2000). For example, Nguyen et al. (2016) describes a stochatic model of axonal guidance based on persistence, bias and noise to better understand how GCs react by turning to different molecular gradients, *in vitro* and *in vivo*. These models are helpful to understand axonal growth and path-finding mostly at the molecular level, to investigate the role of the specific mechanism under study.

ii) Another important kind of models are those that aim at generating 3D realistic neurons (Ascoli and Krichmar, 2000; Eberhard et al., 2006; Koene et al., 2009). Nevertheless, their approach is mainly generative and thus generally based on descriptive and phenomenological rules from deep data analysis and their goal is to generate morphologically realistic neurons

and neuronal networks and do not intend to explain or better understand the generative process itself.

iii) Models taking into account inter-neuron interactions as well as different aspects of axon competition have been also largely described in the literature (Deppmann et al., 2008; Goodhill, 2007; Triplett et al., 2011; Van Ooyen, 2011; van Ooyen and Ribchester, 2003). Goodhill (2007) and Van Ooyen (2011) review models of competition for post target sites and neurotrophins (a guidance molecule) during neural map formation, largely called synaptic competition. van Ooyen and Ribchester (2003), for example, focus on how synaptic competition lead to the reduction of synaptic connections of target neurons. We should notice that these types of competitions mainly consider chemical resources, or space limitation but linked directly with innervation of targets and not during the whole process of axonal elongation and morphogenesis.

iv) Big advances have been done regarding the mathematical and computational modelling of neuron development, from the first mostly simple models, in 2D and focused to explore a single or a few mechanisms (i), to more complex and general ones, reproducing entire realistic neuronal 3D networks (ii) or taking into account some aspects of neuron-neuron competition (iii). However, up to date, 3D models considering spatial constraints and mechanical neuron-neuron interactions -3D physically space-embedded- are so far rare (Luczak, 2006; Torben-Nielsen and De Schutter, 2014; Vanherpe et al., 2016; Zubler and Douglas, 2009; Zubler et al., 2013). Luczak (2006) proposes a generative approach based on diffusion-aggregation process for dendritric tree generation. He early mentions the idea and tries to quantify the consequences of resource and space competition between developing neurons in confinement. Even though this work presents novel and interesting concepts, the mathematical approach is quite abstract, thus difficult to relate with the biological process beyond. Some years later, Zubler and Douglas (2009) proposed a model of neuron growth based on physical forces between objects and diffusion of substances through the extracellular domain. Their approach is very complex, as they model in detail the physical forces acting on the axons. They also take the neurons soma into account, modelled as spheric structures, which occupy most of the free space and thus represent the main mechanical obstacle for the growing tips. In addition, this work does not consider an internal intrinsic growth program, thus elongation depends exclusively on the exerted forces, without considering any random component in the mathematical formulation. Zubler et al. (2013) use this platform and presents a *complete* model of cortical neuron generation. They highlight the importance of considering the physical environment in morphogenesis, for example to correctly reproduce branch tortuosity. They also test and comment different plausible branch generation mechanisms. However, their work is exclusively cortex-oriented, thus their results are difficult to translate to other

neuron types. On its side, Torben-Nielsen and De Schutter (2014) developed a framework for context-aware neuron development considering realistic growth rules, and highlights the difference between neurons growing in isolation and within a more realistic mechanical context. They represent growing tips as cellular automata that, at each time point, can grow, stop or branch depending on their surroundings. The directionality of growth is stochastic, and depends on attraction and retraction from nearby structures. However, their approach is mainly phenomenological and they actually mention the need to implement further branching generation and stopping rules, maybe based on more mechanistic aspects. In particular, they propose that the simple branching generation and termination rules in their work are not sufficient, and that the branching probability should depend on the integration of certain intrinsic as well as extrinsic rules. Finally, Vanherpe et al. (2016) features the problem of axon-like structures growing in limited spaces considering volume exclusion, and develops an interesting framework for the development of non-intersecting tubular-like structures in confined spaces. Their neurons are modelled by a spherical soma and branches that consist on sequences of tubes, and consider structures of different complexity: from single random walks to trees. Regarding volume exclusion, they study soft and hard boundaries, as well as different levels of intersection avoidance. The mathematical formulation of the walks takes into account the axon rigidity, as well as random noise and attraction-repulsion rules. Their work highlights the dependence of axon elongation and final morphology on the particular spatial boundaries (soft/hard), as well as on neurite density and the exact mathematical growth model, which depends on the particular parameter tuning (and is relative to the walk noise). This last work proposes relevant measurements and physic results, relating elongation with volume exclusion. However, their scope and experiments are still far from the real biological process, as they do not compare their results with, or estimate their parameters from, real data.

Even if recent bibliography has already started to address the problem of more realistic 3D space-embedded neuron development models, most of them still under-consider the question of branch formation mechanisms. In these models, branch occurrence is considered abstractly (Luczak, 2006) or, as in most cases, in a phenomenological way (Ganguly et al., 2016; Koene et al., 2009; Torben-Nielsen and De Schutter, 2014). Vanherpe et al. (2016) proposes branch formation to be space-related, but in a cortex-related scenario, where the (otherwise phenomenological) generation rules change layer by layer. Zubler et al. (2013) and Zubler and Douglas (2009) are also focused on cortical neurons, and take into account external signals as well, which are chemical gradients. Interestingly, Torben-Nielsen and De Schutter (2014) proposed simple branching origin and ending mechanisms, and highlighted that they are not sufficient to explain axonal development and morphological variability, and

suggest that branch probability should depend on extrinsic as well as intrinsic rules. On the other hand, Suleymanov et al. (2013) presented a model oriented to simulate branching point creation along the shaft of a developing axon, where the guiding cues that promote branching come from neighbour axons. However, the approach was basically deterministic, the population numbers are quite small (≤ 20) and they did not consider mechanical but only chemical interactions. Cuntz et al. (2010) presented a model that generates dendritic arbours based on Ramón y Cajal′s laws of conservation of cytoplasm and conduction time in neural circuitry. Their model allows to generate realistic morphologies of different tested neuron types, as well as neuron networks. Remarkably, they highlight the importance of sharp physical boundaries of the tissue, as well as of the competition for inputs between neighbouring neurons. In addition, they showed that a direct consequence of growing under competitive conditions was spatial tiling, which might determine the variability in neuronal branching observed in data. Similar comments on morphological observed variability were done in Torben-Nielsen and De Schutter (2014). Hannezo et al. (2017) study branching morphogenesis in different organs (mouse mammary gland, kidney, and human prostate), and propose that branching morphogenesis emerges from the spatial competition between growing tips, which themselves grow and expand following stochastic laws. This work underlines that, if the same process occurred without competition for space, branched organs would be characterized by stereotypical rounds of purely symmetric branching, with the number of branches increasing exponentially with the branch level. Thus, that complex branched structures develop as a self-organized process, based on simple rules, without the need of a deterministic system of genetically programmed events.

In conclusion, models have evolved in complexity and in realism, reaching the category we called 3D embedded models, which takes into account not only the path-finding but also the environment of the growing axons. These works have all stressed the importance of considering space-embedded processes and interactions with the cellular environment when studying neuronal morphologies. However, they have not, or could not, use parameters estimated from real data. Large number of parameters make them difficult to estimate, so the models somehow loose contact with data. Another observation is that they rarely showcase explicative or predictive aspects of their approaches. Regarding branching, the occurrence mechanisms remain quite simple and mostly phenomenological rather than mechanistic. In this work (Chapter 4), we develop a model that is simple enough to allow biologically meaningful parameters that can be estimated from data, but that is still mechanistic, taking into account the main aspects of axonal development *in vivo* at the cellular scale. Our model can be also easily embedded in realistic environments, considering other axons, external physical limits and chemical gradients. Also, we test branching mechanisms that are not

pre-set or phenomenological, but that depend on the evolution of the whole system. With this, we propose to give a step forward, and obtain not only morphology reproduction, but also new hypothesis regarding relevant biological mechanisms.

2.3 Biological model: Drosophila γ neurons

To better understand the mechanisms of individual axon growth embedded in neuronal populations that give birth to adult neural morphologies, we chose to work with the fruit fly *Drosophila melanogaster*. An important advantage of this model organism is that imaging of fluorescently-labelled individual axons in fixed tissue or in real time can easily be combined with genetic manipulations, to obtain images of single (or a few) axons that can be wild-type or present mutations.

Specifically, we study the γ neurons, a population forming with companion populations - the $\alpha\beta$ and $\alpha'\beta'$ neurons - bilateral structures located in the central *Drosophila* brain, called the Mushroom Bodies (MBs) (Fig. 2.3A and B). These structures, each composed of roughly 2,000 cells (Aso et al., 2009), are high-order integration centres involved in olfactory learning and memory functions (Heisenberg, 2003). The population of MB γ neurons consists of about 650 cells, and constitutes a model particularly suitable to study axon growth in a population context, as it is the only one in the Mushroom body that goes through the process of remodelling (pruning followed by regrowth). This means that, after an initial growth at larval stages, distal γ axons degenerate during metamorphosis and regrow synchronously, in a short time frame, to establish adult axonal projection patterns (Rabinovich et al., 2016; Watts et al., 2003; Yaniv and Schuldiner, 2016) (Fig. 2.4 C). At this stage, axons regrow in a complex and dense environment constituted of surrounding neuronal populations and glial cells (Aso et al., 2014; Hakim et al., 2014). Furthermore, axon growth is associated with branching, such that adult γ axons typically exhibit a varied number of side branches (Lee et al., 1999). Axon arborization patterns are however not completely stereotypic, as an interesting but unexplained wide range of morphologies is observed within populations of γ neurons (see Fig. 2.5).

As many invertebrate neurons, γ neurons are unipolar (Fig. 2.1). Their cell bodies are clustered at the dorsal posterior surface of the brain, and their dendrites project just beneath, in a structure termed calyx (Fig. 2.3B). Proximally, MB axons fasciculate to form a dense fiber projecting ventrally: the peduncle (Fig. 2.3B). More distally, axons de-fasciculate to innervate the so-called medial lobe (ML), following either straight paths or more tortuous trajectories (Fig. 2.3C). In the medial lobe, they establish a dense network of branches (Figs. 2.3B and C and 2.5).

As described, γ axons grow through two very different physical environments. Firstly in the peduncle, where the axon shafts elongate rigidly in a tight and ordered parallel bundle. Secondly, they enter the ML. In contrast, the ML is a large structure, where the neurites are physically less constrained and develop long terminal branches, giving birth to a big variety of morphologies. This particular conditioning of the neuronal morphology by the physical environment through different stages of neural growth is explored in Koser et al. (2016) for diverse neuron types.

In conclusion, γ neurons represent an accurate biological model to study mechanical constraints and inter-axon interactions during axon development, as they grow as a dense population, forming unstereotypical morphologies in a limited brain area.

2.3.1 Imp as a regulator of axon regrowth and branching

Actin is a fibrous elongated protein composed by smaller subunits (G-actin) that have the ability to assemble and disassemble dynamically. As mentioned in Section 2.1, it is found on the axonal extremities during development and its role is to guide the axon tip through the correct path following internal and external cues, as well as to provide with the force to make the axon extend. There are many actin regulators which, depending on the particular context, can favour or prevent actin fibers assembly or dismantling.

In this study we consider, in addition to wild-type individuals, axons that are mutant for the conserved mRNA transport protein Imp. Molecular and genetic analysis performed in the Besse laboratory have shown that Profilin mRNA, which encodes an actin cytoskeleton regulator (Luo, 2002; Schlüter et al., 1997; Verheyen and Cooley, 1994), is a direct and functional target of Imp and both are key regulators of the Drosophila γ neuron axonal remodelling process, acting on the same molecular pathway. Imp carries $profilin$ mRNA along the axon shaft towards the extremities where the regrowth occurs, enabling Profilin proteins to be synthetized *in situ*. Fig. 2.4 A and B present a scheme of Imp and Profilin implication during metamorphosis. Medioni et al. (2014) have shown that Imp is essential to allow axonal regrowth of γ neurons during metamorphosis. Our database (Section 2.4) includes *imp* mutant γ axons, which we used to validate the morphological analysis and different growth hypothesis in this thesis. Furthermore, in Section 4.6 we propose a novel interpretation of the role of Imp in the remodelling process, and its impact on adult γ axon morphology.

Regarding the *imp* mutant phenotype of a single mutated axon in an otherwise wild-type environment, Medioni et al. (2014) reported that, in adults, about 50% of *imp* mutants displayed shorter axons than wild-types and failed to reach their target. *imp* mutants also exhibit an overall loss of branch number and complexity. Interestingly, they observed that

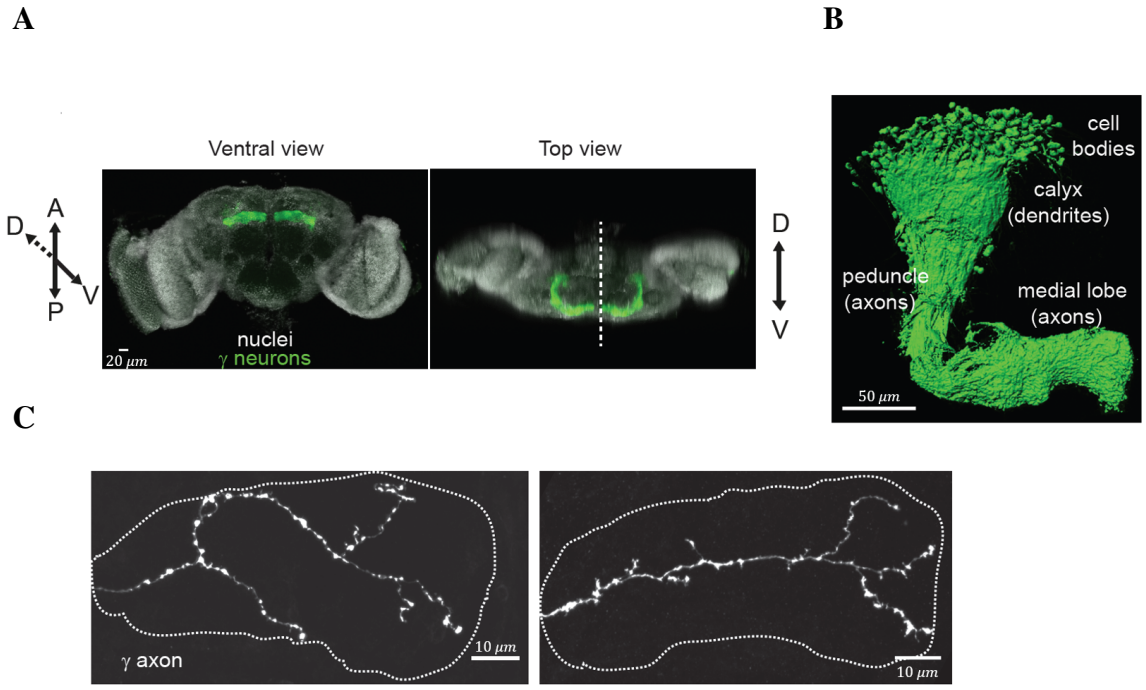

Fig. 2.3 Characteristics of adult Mushroom Body γ neurons. (A) Wild-type adult *Drosophila* brain expressing the membrane-tagged CD8-GFP construct in γ neurons, under the control of the MB009B-Gal4. Nuclei are labelled in white with DAPI. The dotted line on the top view image corresponds to the midline. This images were obtained using the 10x 0.6 NA air objective of a LSM780 NLO Zeiss microscope. Z sections were taken every 0.5 μm, with a xy pixel size of 0.57 μm. 3D projections have been created using FiJi (Schindelin et al., 2012).(B) 3D reconstruction of one Mushroom Body where γ neurons only were labelled. Genotype: MB009B-Gal4/UAS CD8-GFP. This image was obtained using the 40X 1.2 NA water objective and the pulsed laser of a LSM780 NLO Zeiss microscope. Z sections were taken every 0.4 μm, with a xy pixel size of 0.11 μm. The 3D reconstruction was generated using the surface mode of Imaris 8.1 Bitplane software. (C) Axonal arborizations of two individual adult γ neurons labelled by GFP using the MARCM technique. Confocal images taken along the z axis were projected (Maximum Intensity Projection). Dotted lines delimit the shape of the medial lobe. Scale bars: 20 μm in A, 50 μm in B and 10 μm in C.

the over-expression of Profilin in *imp* mutant backgrounds partially restored the main axon length, but not the branch complexity. *imp* mutant axons rescued by Profilin over-expression are also included in our database (Section 2.4). These results suggest that Imp controls axonal extension during remodelling at least partly by regulating *profilin* mRNA expression. However they also suggest that the branching process may be dependent on the regulation of other Imp mRNA targets, yet to be identified. In Section 4.6 we further investigate the role of Imp in γ neuron remodelling. We propose a novel hypothesis on the origin of the short phenotype, and explore the mechanistic reasons for this partial rescue observed after Profilin over-expression.

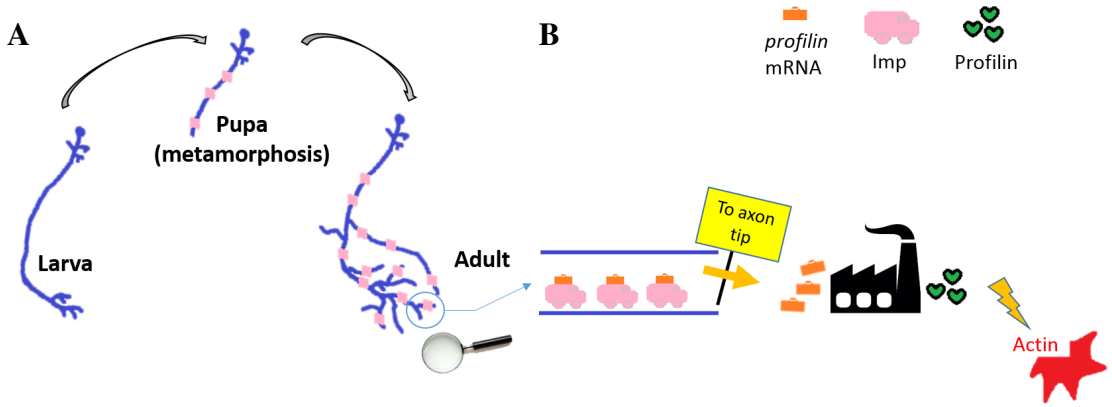

Fig. 2.4 The role of Imp and Profilin during remodelling. (A) Scheme of a γ neuron remodelling. Imp (pink) is transported to axons during metamorphosis. (B) Schematic representation of the role of Imp (pink) as a *profilin* mRNA (orange) transport protein. Imp carries *profilin* mRNA from the neuron body towards its extremities. This allows the production of Profilin proteins *in situ*, for a rapid control of actin polymerization in the growing tip.

2.4 Database Description

The development and validation of models proposed in this thesis rely on two databases, provided by Florence Besse and Caroline Medioni in the Besse Team at IBV (France). The first one consists of 3D confocal images of individually-labelled γ neurons from adult *Drosophila* brains, and the second one of 3D+t image sequences of growing γ axons in the *Drosophila* developing brain.

2.4.1 Adult single cell confocal *in vivo* images

This database consists of 3D confocal images, each comprising two channels. The first one displays a single labelled adult γ neuron, generated with the MARCM technique (Wu and Luo, 2006). This technique allows to image a single axon that has developed *in vivo* within the brain in the γ population. The second channel displays the $\alpha\beta$ entire population shaping the medial and dorsal lobes of the MB. The γ axons grow in the ML, which is highlighted with a white dotted line in Fig. 2.5. These images were used to understand the γ neuron morphology (Chapter 3), to propose the main growth model and branch generation mechanism, as well as to directly estimate the model parameters and to validate the results (Chapter 4).

The database consists of: 43 wild-type (WT), 45 *imp* mutants (Imp), 15 *imp* mutants rescued with Profilin over-expression (Prof Rescue), 15 *profilin* mutants (Prof), 42 *imp* mutants rescued by Imp over-expression (Imp Rescue) and 27 unf mutants rescued by TOR (Unf). Both *imp* and $profilin$ mutations used here are protein-null mutations. Wild-type and *imp* databases are used through all this work, while the others served for particular studies. For all the studied mutants, the labelled axon(s) in the images is (are) the only one(s) of the γ population that present the mutation. Thus, we observe mutant axons that have developed within an otherwise wild-type γ neuron environment.

MARCM clones were generated as described in Wu and Luo (2006), using the following fly stocks: hsp-flp, tub-Gal80, FRT19A; 201YGal4,UAScGFP; FRT19A + and FRT19A imp^7. Brains were dissected at the adult stage, and stained with anti-GFP (molecular probes life technology; ref A11122) and anti-FasciclinII (1D4, DSHB) primary antibodies, revealed by respectively anti-rabbit Alexa 546 and anti-mouse Cy5 secondary antibodies (see Medioni et al. (2014) for a detailed procedure). Brains were mounted in propyl-galate mounting medium, and imaged with an inverted Zeiss LSM 710 confocal microscope equipped with a 40X/1.1 NA water objective. Z sections were taken every 0.6 to 0.9 μm, with a xy pixel size of 0.09 μm. Fig. 2.5 shows the maximum intensity Z projection of both channels of different images in this database.

2.4.2 4D image sequences of the developing brain

This database consists of movies of the living developing brain, where a single γ axon or a group of axons are visualized via targeted expression of the GFP marker during metamorphosis (examples in Fig. 2.6). These data have been used for the branch classification regarding formation and functionality that we present on Chapter 4, as well as to analyse the dynamic aspects of axonal growth as a population.

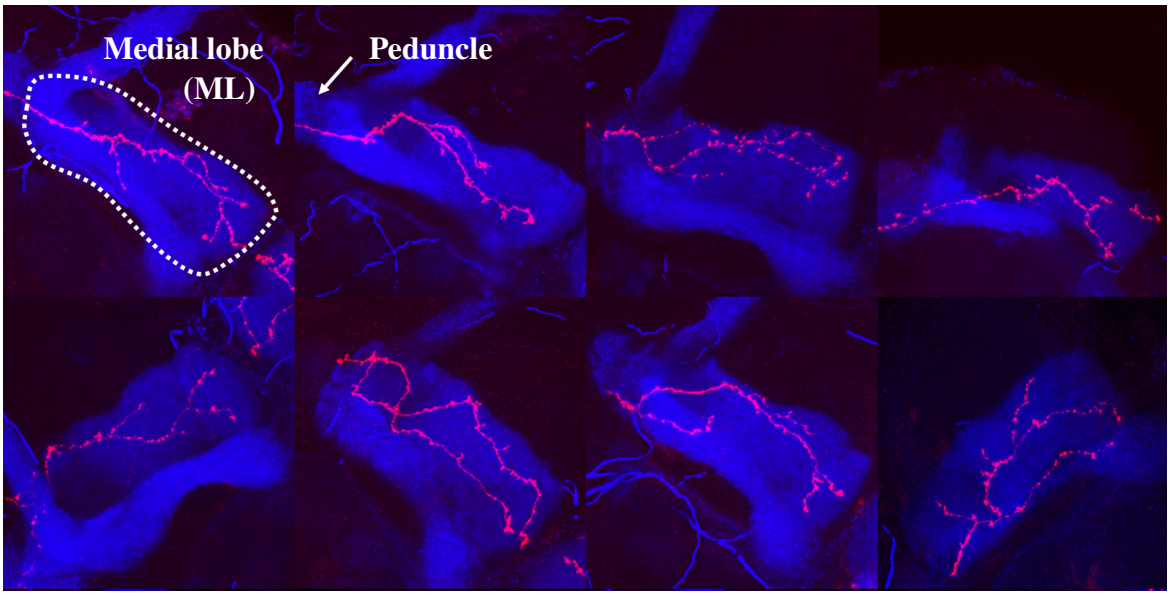

Fig. 2.5 Adult single cell confocal *in vivo* images. Examples of our database, exhibiting a single adult wild-type γ axon (red) and the dorsal and medial lobes of the MB (blue). The maximal intensity projection in Z is shown for simplicity. The white dotted line in the first image indicates the medial lobe.

To create these movies, brains were dissected out of pupae 24-30 APF (After Pupa Formation) and mounted in a Labtek II chambered coverglass (#155378, Fisher Scientific) in culture medium (Schneider medium, 10% FCS, 1% Antibiotic Antimycotic Solution (Sigma), $200\mu g/ml$ insulin (Sigma), $1\mu g/ml$ ecdysone (20HE; Sigma) (see Medioni et al. (2015b) for a detailed protocol). Brains were imaged using a Zeiss LSM780 NLO inverted two-photon microscope and a 40X/1.2 NA water objective. Medial lobe regions were imaged. Z sections were taken every $0.8\mu m$ and covered the entire medial lobe volume. Z stacks were acquired every 5 minutes over up to 15 hours. Pixel size is $0.13\mu m$.

Drift correction was carried out by Gregoire Malandain (see Medioni et al. (2015a)). For the single axon cases, a two-fold drift compensation was performed. First, maximum intensity (MIP) projections of the acquired images were computed, couple of successive MIPs (that are 2D images) were co-registered with 2D affine transformations, and these transformations were compounded to express all the transformations with respect to a reference image (say the first one). This first step allowed an in-plane drift compensation, but a motion along the Z direction may still exist. A residual translation along the Z direction was then estimated using the multiple transformation strategy described in Medioni et al. (2015a). The compounding of the 2D affine transformation and the Z translation yielded the final 3D drift compensation.

The segmentation of the single axon movies used to quantify the length distribution of dynamic branches (Fig. 4.35 in Section 4.5.1) was also carried out by Gregoire Malandain,

and proceeded in two steps. First, main axon and type I (see Chapter 4) branches were defined as segments that remain stable throughout the whole series and were extracted. Their centrelines were computed on the first time points, *via* classical image processing tools (thresholding, skeletonization, and manual selection of the skeleton parts corresponding to the main axon and type I branches). Then, they were propagated throughout the whole 3D+t series. The use of homotopy-preserving operations guarantees that the main axon and type I branches form a tree (in a topological sense). Type II branches (see Chapter 4) were then segmented by the means of an homotopic dilatation from main axon and type I branches. Centrelines were extracted and tracked throughout the series, and their maximal length was measured.

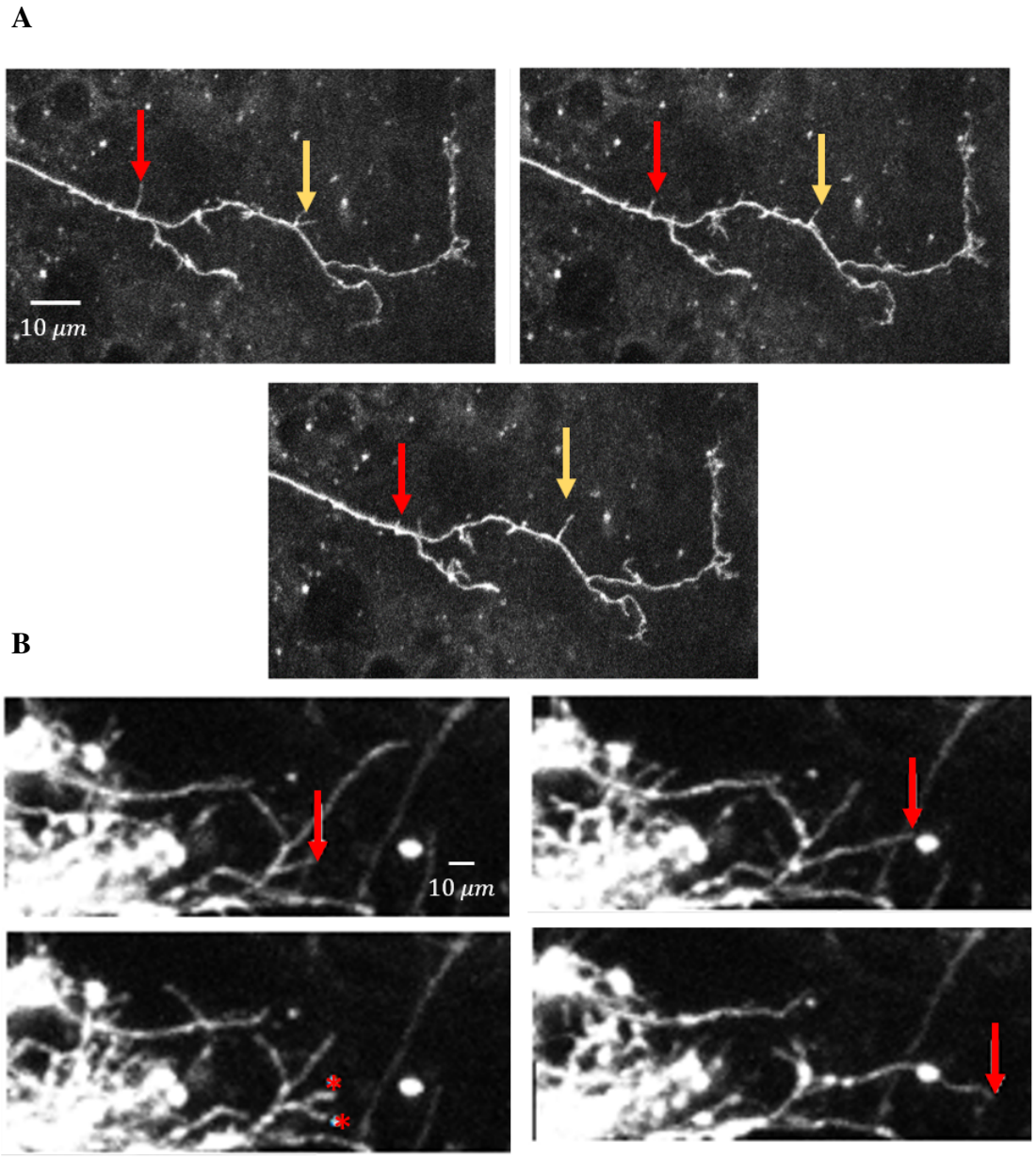

Fig. 2.6 Video of the developing *Drosophila* brain. (A) Image sequence extracted from one of the movies in our database, where a single γ neuron from a wild-type brain undergoing metamorphosis is followed, during the regrowth phase. Red arrows point at a branch that shortens with time, while yellow arrows point at a branch that is growing. (B) Detail from another movie. The first two frames show an elongation step, the third one retraction and branching event, and the last one growth again. Arrows shows axonal tips and asterisks, the formation of branches. Scale Bars: $10\mu m$.

Chapter 3

Morphological Study

Not everything can be explained, not everything has an answer, not everything makes sense, not everything is fair, not everything is logic. Learn to live with that.

Anonymous

In this chapter we propose a deep axonal morphological analysis, and apply it to the study of adult γ axons to understand the behaviour and relevance of main shape features; as well as to further understand the role of Imp and the importance of *profilin* mRNA expression regulation during remodelling.

To do so, we developed a stochastic framework to exhaustively compare the adult γ axon morphologies between the WT, Imp and Prof Rescue groups (Fig. 3.1). Our framework can be summarized in three steps: i) selection of relevant morphological features describing the axons, ii) stochastically model the behaviour followed by each of the chosen morphological features and estimate the parameters associated to each model from the data (for WT and mutant groups) and iii) classify each individual axon to a group (WT, Imp or Prof Rescue), considering the features separately and altogether using the maximum likelihood of the applied models. The classification results give a quantitative measure of the global similarity or difference between groups (through its confusion matrix). We also apply statistical tests under null hypothesis between the neuron groups for each morphological feature to enrich the analysis. This approach provides both a biological interpretation and a quantification of the resemblance between biological samples, detecting differences as well as similarities between the groups. This framework is general and can be applied to model and characterize different axon types (Fig. 3.2).

Neuron morphological automatic classification has already been addressed in the bibliography. Kong et al. (2005) proposed an unsupervised clustering of ganglion cells in the mouse retina by the k-means algorithm in order to define cell types. They initially disposed of 26

morphological parameters and found out that clustering with only three of them was the most effective way. Guerra et al. (2011) established the advantage of applying supervised classification methods regarding morphological feature based classification to distinguish between interneurons and pyramidal cells. They also conclude that reducing the number of features to an optimal number outperforms the classical approach of using all the available information. López-Cruz et al. (2014) built a consensus Bayesian multinet representing the opinions of a set of experts regarding the classification of a pool of neurons. The morphological parameters chosen by each expert to make their decisions are not considered. A different approach was proposed earlier in our team (Mottini et al., 2015), which consists in classifying different neuron types by reducing them to trees and calculating a distance, combining geometrical and topological information.

The different published approaches intend to accurately discriminate between different types of neurons, considering misclassification as a methodological error and consequently developing techniques to avoid these cases. However, similarities between populations are not necessarily to be excluded as they may reflect the properties of biological samples and provide useful information for their characterization. Furthermore, these methods do not intend to understand which morphological characteristic is discriminant between different species and at which level. A deeper multi-criteria statistical analysis is thus required. Our approach allows to assess the similarities and differences between the populations for each chosen morphological feature separately as well as considering them all together for a global analysis. Neurons are treated individually through the likelihood analysis as well as globally within the studied groups through the statistical comparisons, leading to an exhaustive analysis.

Thanks to this framework we show that *imp* mutant neurons can be divided in two phenotypical groups with a different aberrancy degree, and that Profilin over-expression partially rescues the main axon as well as branch development (new observation respect Medioni et al. (2014)), and thereby reduces the proportion of neurons with the strongest mutant phenotype.

3.1 Digital reconstruction of the axonal trees

In this section we describe the segmentation and data treatment process (see Fig. 3.2) using the adult single axon images described in Section 2.4. To avoid artificial jumps along the Z axis due to image anisotropy (*i.e.* voxel Z dimension is several times higher than that in XY), we applied the bicubic interpolation algorithm included in FIJI (the open source image analysis software developed by NIH, Maryland, USA) (Schindelin et al., 2012).

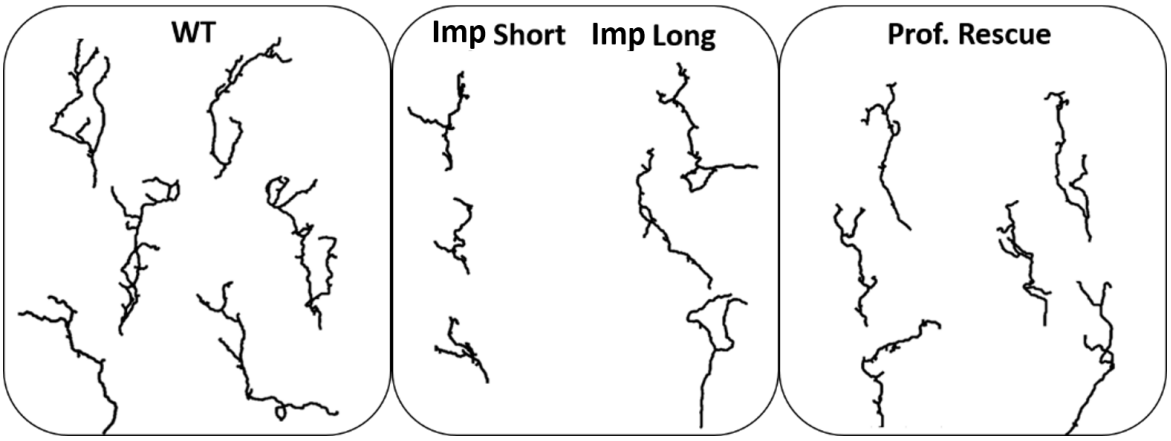

Fig. 3.1 Examples of adult γ axons of genetically distinct populations. From left to right: Wild-type (WT), *imp* mutant (Imp) and *imp* mutant rescued by Profilin (Prof. Rescue). *imp* mutants are divided into short and long species (named Imp Sh and Imp L respectively) as both phenotypes are observed.

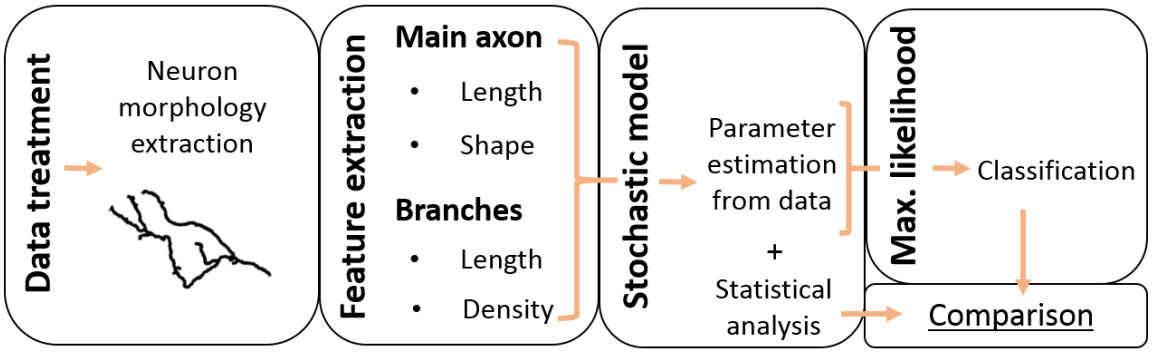

Fig. 3.2 Scheme of the proposed stochastic framework for the comparison of axonal morphologies between groups. It starts by data treatment and axon morphological extraction. Then the selection of the main axonal morphological features and their stochastic modelling and parameter estimation from data. Finally, a maximum likelihood analysis is developed for each individual neuron to quantify the similarity or difference between groups (see Section 3.3).

3.1.1 Segmentation

An automatic segmentation of these images is still not available regarding our needs, due to a noisy background and poorly defined and non-continuous neuron trace. Thus, we tested and compared two well-known cell segmentation software: Imaris ® and Neuromantic (Myatt et al., 2012). The first one is a data visualization and management software developed and commercialized by Oxford Instruments. Even though it counts with an interactive user interface and an automatic 3D segmentation tool, its performance is low when dealing with our images. It has difficulties to handle their size (≈200 MB) and the automatic mode fails to correctly segment the entire axon. The use the manual mode becomes mandatory, which is time demanding, not intuitive and inaccurate along the Z axis. On the other hand, Neuromantic is an open software (Myatt et al., 2012), specially developed to segment 2 or 3D neurons manually or semi-automatically. The performance of this last mode, even though time consuming as well, is more precise and easier to handle than Imaris®, as the user goes through the axon and branches paths while the software adapts the skeleton to the nearest trace in the image. In addition, posterior changes to the skeletons are easily handed and it has no problem with our images size, making it the most suitable option for our needs. The software output is a file (.swc extension) containing the nodes and edges of the resulting skeleton graph and the neuron radius in each point. A script to read .swc files into Matlab ® had already been developed by the team. As output, we obtained a set of points along the main axon and branches that we connected using a Bresenham-inspired 6-connectivity algorithm. We chose this connectivity to keep further measurements simple. After this process we obtained a tree-like set of numeric 3D curves describing the morphology of each axon (Fig. 3.3).

To ensure all the neurons to be similarly oriented we manually rotated the images to align the medial lobe with the horizontal axis (Fig. 3.4). We only considered a rigid transformation to avoid axon deformation. Conserved morphology was preferred rather than more accurate spatial location.

3.1.2 Tree hierarchy

When studying their morphology, it is necessary to understand how axonal trees are structured *i.e.* the segment hierarchy: main axon and first, second, third (etc.) order branches. For this purpose, we developed an automatic recursive algorithm capable of processing trees of any order. It starts by taking the whole tree and selects the main axon, according to specific criteria. After this process, a number of independent sub-trees of different orders is obtained. During the following steps each of the sub-trees is analysed. Their main path is assigned,

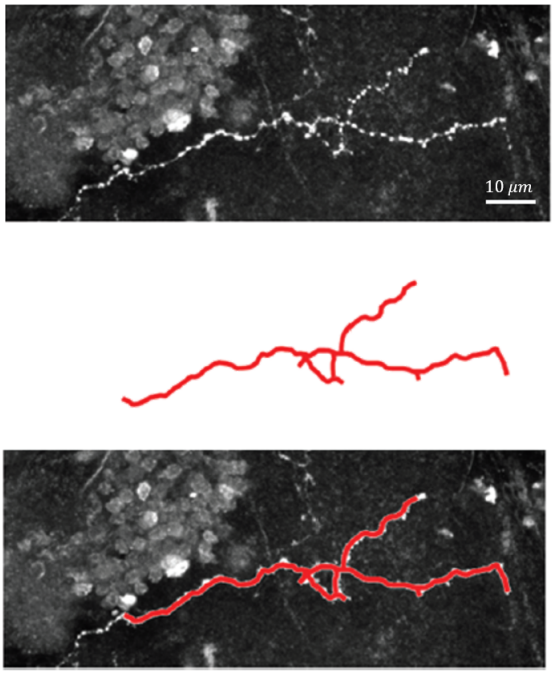

Fig. 3.3 Extraction of neuron morphology from confocal 3D images. Maximum intensity projection of a confocal image depicting a single γ neuron stained with GFP (top), its skeleton in red after segmentation (middle) and the overlay between the original axon and its reconstructed skeleton (bottom). The segmentation starts at the point where the axon enters the ML (present in the second channel of the image, see Section 2.4). Scale bar: 10 μm.

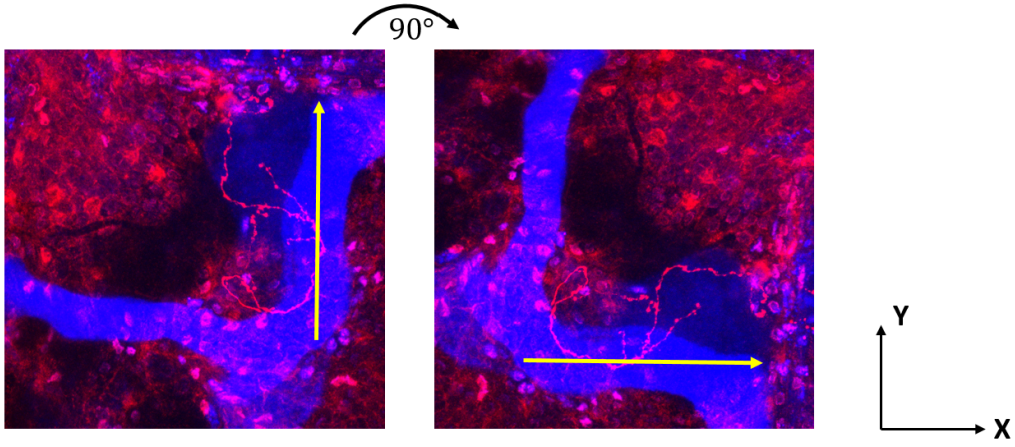

Fig. 3.4 Alignment of the medial lobe to the horizontal axis. The yellow arrow indicates the medial lobe main axis. In this example, a rotation of 90° was applied to align the medial lobe to the horizontal axis (X).

leaving again a set of independent sub-trees, and so on. The algorithm continues until no more untagged segments are left (Fig. 3.5).

In each step, the main path of the sub-tree is assigned following the same criteria used by experts when done visually: total length, directionality and sense coherence. To achieve this, we considered the points for each path between the root and the extremities of the tree (or sub-tree), and calculate a linear regression obtaining a straight guideline (see Fig. 3.5), used to determine directionality and sense coherence. For each path of the analysed tree/sub-tree, a cost function f_{cost}^{p} (Eq. 3.1) is computed that depends on the distance between each point in the path and the guideline (directionality), the parallelism between them (accounting for the sense coherence of the path) and the path total length. Finally the path that minimizes this cost function is selected as main axon in the case of the whole tree (first step), or main branch in the case of the different sub-trees (Fig. 3.5). The cost function f_{cost}^{p} for each path is defined by

$$f_{cost}^{p} = \frac{distance}{length^2 angle^2} \tag{3.1}$$

where $distance$ is the sum of the Euclidean distances from each point of the path p to the guideline normalized by the number of points (N_p), $length$ refers to the path length and $angle$ is defined as,

$$angle = \frac{N_1}{N_1 + N_2} \sum_{i=1}^{N_1} \cos \phi_i \tag{3.2}$$

where ϕ_i are the N_1 angles between each vector defined by two continuous points of the path p and the guideline whose cosine is positive (parallel direction). N_2 is the number of anti-parallel points i.e.: when the vector defined by a point and the previous one in the path forms an angle with the guideline whose cosine is negative ($N_1 + N_2 = N_p$).

Fig. 3.6A shows an example of the application of the hierarchy algorithm to one axon in our database. Fig. 3.6B highlights the advantage of this algorithm respect to the one previously used by our team. In the latter, the main axon label was assigned to the longest path from the starting point to any of the end points of the tree. The branch level was then assigned following the order of appearance in the .swc file (and not a morphological criteria), and was limited to 3^{rd} order branches. Even though this algorithm was effective in most cases, it occasionally committed errors regarding the main axon determination, as shown in Fig. 3.6B. In this example, the longest path from the root to the end, thus the one assigned as main axon by the previously used algorithm, turns completely backwards (light blue line Fig. 3.6B); making it unlikely to be the main axon. Thus, the criteria of sense coherence included

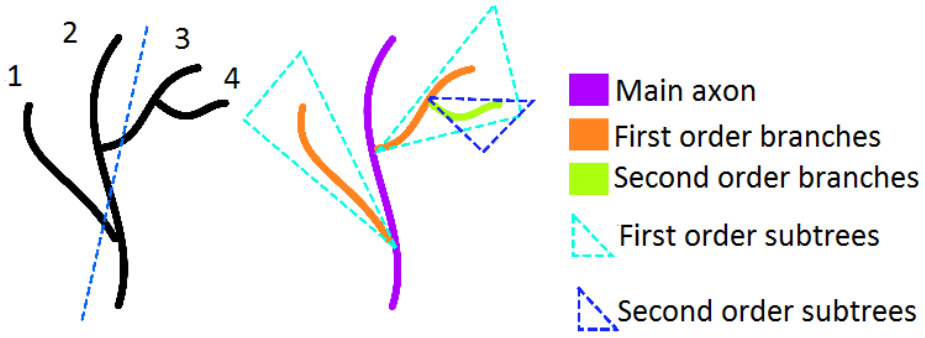

Fig. 3.5 Scheme of the three-hierarchy algorithm. For a given tree, a guideline is first calculated as the linear regression of all the points in the tree (dotted blue line, left tree). Secondly, the cost function for each possible path (1-4) is obtained. The one minimizing the cost function is labelled as "main axon" (here path 2). The algorithm is applied recursively to each sub-tree, resulting in the hierarchy of the entire tree.

in our new algorithm becomes necessary to accurately label the main axon (pink line Fig. 3.6B).

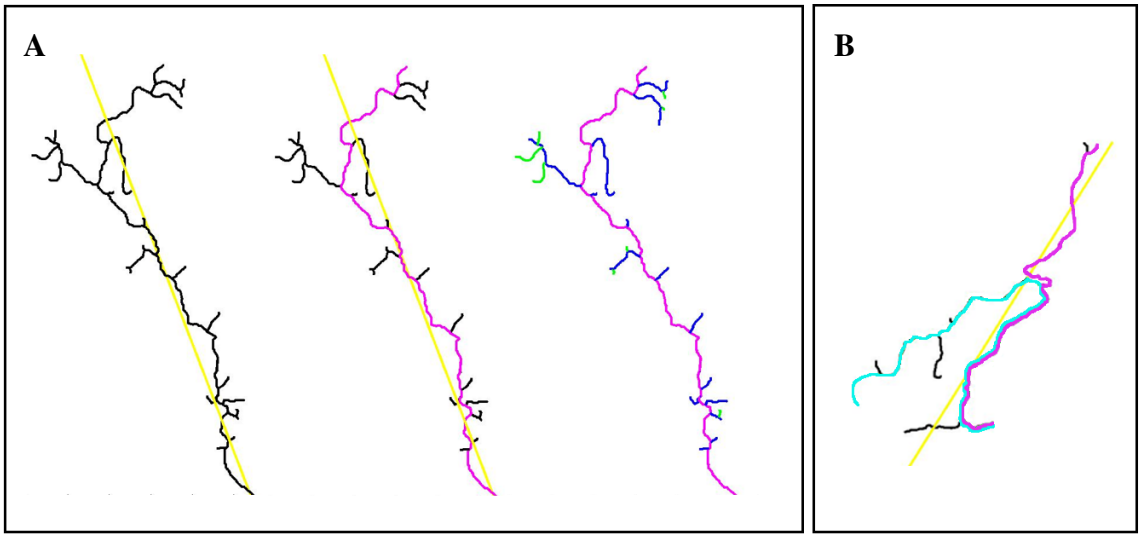

Fig. 3.6 Results of the hierarchy algorithm. (A) Basic steps followed by the automatic tree hierarchy algorithm. The guideline (yellow) is calculated for all the tree (left) and the main axon (pink) is assigned following the chosen criteria (middle). The same two steps are applied recursively to all the neuronal subtrees (right), to label first (blue) and second (green) order branches. (B) Previous approach, where the main axon is the longest path (main axon in light blue) vs. new approach, where the main axon is assigned following intuitive criteria (main axon in pink).

3.2 Stochastic modelling of the main morphological features

After segmentation, our data is composed of 3D tree-like skeletons of axons, where each 6-connected path (main axon or branch) is described by its round coordinates or pixels (Fig. 3.7). Taking this simple geometrical description into account, we define the main morphological features that best describe the individuals and present biological interest: i) main axon length and ii) sinuosity, as well as iii) branch density and iv) length distribution. Further, we develop the probabilistic models chosen to describe each feature, estimate their parameters and compute associated statistical tests under null hypothesis between the different studied groups (Feature extraction and Stochastic modelling in Fig. 3.2).

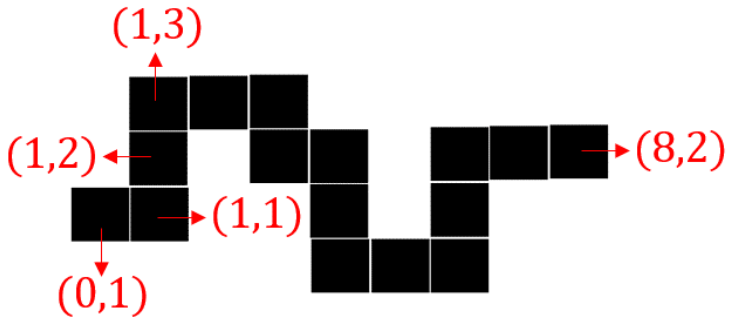

Fig. 3.7 Axon path scheme. Each neurite path (main axon or branch) is composed by 6-connected pixels, described by they coordinates in space (labelled in red as an example). For simplicity, a 4-connected skeleton in 2D is shown instead.

3.2.1 Main axon length

The main axon length was measured taking the total amount of pixels in the corresponding path and multiplying it by the pixel size (in μm). The length distribution was modelled as Gaussian, and the mean and standard deviation for each group X ($\mu^X_{m.a.}$, $\sigma^X_{m.a.}$) were calculated from data (see Table 3.1). We observed the bimodal behaviour in the Imp group reported by Medioni et al. (2014) (Fig. 3.8). Therefore, in order to make a more accurate modelling of this parameter, we separated *imp* mutant axons into two groups -with long (Imp L) and short axons (Imp Sh)- using the k-means algorithm. 54% of the neurons were assigned to Imp Sh and 46% to Imp L, which is consistent with the percentage reported by Medioni et al. (2014). Fig. 3.8 shows the main axon length histograms for each group, Imp divided into Imp L and Imp Sh. We decided to keep this division of the Imp group through all the

Table 3.1 Mean and standard deviation corresponding to the main axon length distribution per group (in μm).

	$\mu^{\mathbf{x}}_{\mathbf{m.a.}}$	$\sigma^{\mathbf{x}}_{\mathbf{m.a.}}$
WT	147	28
Imp L	130	24
Imp Sh	65	16
Prof Rescue	115	20

analysis in this Chapter, in the attempt to detect other morphological differences between the two genetically identical subgroups. To assess which groups present significant differences regarding the main axon length, the non-parametric Kruskal Wallis test was carried out between all the possible pairs of groups. We chose this test for the sake of consistency, as it can be applied to analyse all the features independently of each model. The only pair not presenting a significant difference (taken as $p > 0.05$) is WT vs. Imp L. It is interesting to highlight also that Prof Rescue distribution lies in between the distributions for Imp L and Sh and is more similar to Imp L, as it does not present extremely short axons.

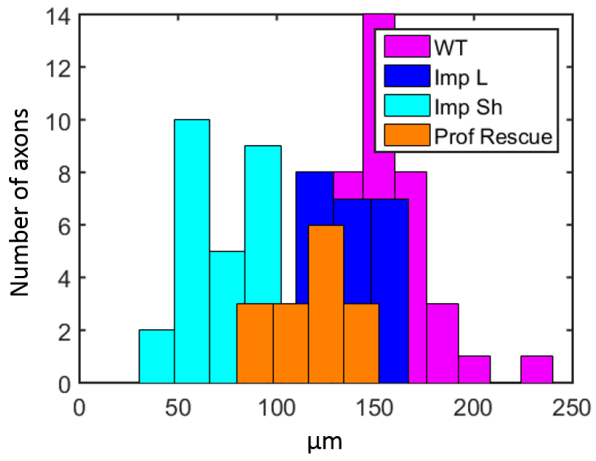

Fig. 3.8 Main axon length distributions for each biological sample.

3.2.2 Main axon morphology

To define the shape model, we considered as random variable the unit vector $\overrightarrow{x}_t$ that accounts for the shift of the axon tip between $t-1$ and t. Because we used the 6-connectivity and backwards moves are not allowed, each $\overrightarrow{x}_{t+1}$ (shift of the axon tip between t and $t+1$)

can take five different values, as shown in Fig. 3.9. Assuming the main axon development follows a second order Markov property, we have

$$P(\overrightarrow{x}_{t+1}|\overrightarrow{x}_j, j \leq t) = P(\overrightarrow{x}_{t+1}|\overrightarrow{x}_t, \overrightarrow{x}_{t-1}) \tag{3.3}$$

First approach: Markov chain model with 9 parameters

We first explored the possibility of assigning to each conditional probability P_i (Fig. 3.9 and eq. 3.3) a physical meaning, related to both the axon rigidity and the presence of an attractive field $\overrightarrow{F}$ generated by external guiding cues. Fig. 3.9 shows two basic configurations of a pair of (past) unit vectors $[\overrightarrow{x}_t, \overrightarrow{x}_{t-1}]$ (blue arrows in Fig. 3.9) and their corresponding five possible (future) steps $\overrightarrow{x}_{t+1}$ (red arrows in Fig. 3.9). The first one is one of the six possible cases where the vectors $\overrightarrow{x}_t$ and $\overrightarrow{x}_{t-1}$ are in line (in this case in the +z direction). In these cases, the elasticity parameters are two: one (α_1) for a totally rigid path (i.e. P_5 in Fig. 3.9) and a second parameter (α_2) for a more flexible one (i.e. P_{1-4} in Fig. 3.9). We assume isotropy for the elasticity leaving the same parameter accounting for a turn in any direction. The second configuration in Fig. 3.9 shows one of the twenty-four cases where the vectors $\overrightarrow{x}_t$ and $\overrightarrow{x}_{t-1}$ are not in line. In these cases, similar parameters regarding elasticity were assigned. One (α_3) accounts for the most rigid option (i.e. P_{i+4} in Fig. 3.9), two different ones ($\alpha_{4,5}$) to represent the cases where the future vector turns with respect to $\overrightarrow{x}_t$ and stays in the plane formed by $\overrightarrow{x}_t$ and $\overrightarrow{x}_{t-1}$, in the same or opposite direction as $\overrightarrow{x}_{t-1}$ (i.e. P_{i+3} and P_{i+1} in Fig. 3.9, respectively). Finally, a sixth parameter (α_6) accounts for the axon elasticity when changing plane with respect to $\overrightarrow{x}_t$ and $\overrightarrow{x}_{t-1}$ and is the same for any of both possible directions respecting the axon elasticity isotropy (P_i and P_{i+2} in Fig. 3.9).

Considering these parameters accounting for the axon rigidity [($\alpha_{1:6}$)] plus three parameters $[F_x, F_y, F_z]$ representing an external attracting field, we can define the value of each possible probability $P(\overrightarrow{x}_{t+1}|\overrightarrow{x}_t, \overrightarrow{x}_{t-1})$. For instance, the five conditioned probabilities represented by the first configuration in Fig. 3.9 are defined as

$$\begin{cases} P_1 = \frac{\alpha_2}{D} \\ P_2 = \frac{\alpha_2}{D} \\ P_3 = \frac{\alpha_2 + F_y}{D} \\ P_4 = \frac{\alpha_2 + F_x}{D} \\ P_5 = \frac{\alpha_1 + F_z}{D} \\ D = 4\alpha_2 + \alpha_1 + F_x + F_y + F_z \end{cases} \tag{3.4}$$

From data, we can naturally estimate the values of the 150 conditional probabilities P_i (30 configurations of two past vectors $[\overrightarrow{x}_t, \overrightarrow{x}_{t-1}]$ with 5 possible future vectors each) as the number of occurrences of the three unit vectors $[\overrightarrow{x}_{t+1}, \overrightarrow{x}_t, \overrightarrow{x}_{t-1}]_i$ divided by the number of occurrences of the two past vectors $[\overrightarrow{x}_t, \overrightarrow{x}_{t-1}]_j$.

We thus obtain an overdetermined system of 150 equations with nine unknowns ($[(\alpha_{1:6}, F_x, F_y, F_z)]$). However, because the average length of the axons from data is not enough to ensure good statistics for each one of the 150 equations (which is the same as saying that the frequency of some three unit vector configurations present high variation), the overdetermined system becomes inconsistent, thus impossible to solve. Because the parameters of the model cannot be estimated, we propose another approach, described in the next section.

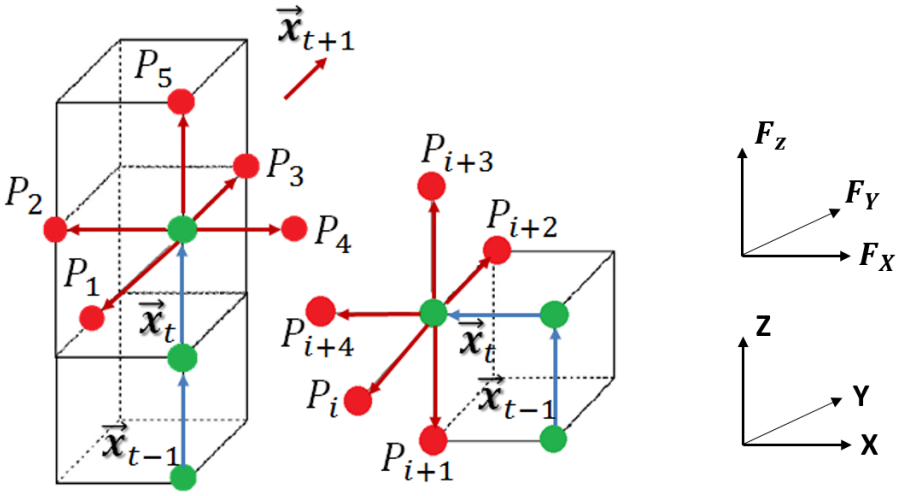

Fig. 3.9 Second order Markov model for axon morphology. Two examples of three vector configurations considering a 3D 6-connected path. Each future direction has a probability of occurrence conditioned by the present and past directions, and is numbered from 1 to 150 (30 possible configurations for past plus present and 5 possible future configurations for each of them). (left) One of the 6 possible configurations where the two past vectors are in line. (middle) One of the 24 configurations where the two past vectors are orthogonal. (right) Considered coordinate axis and external field components.

Second approach: non-parametric Markov chain model

Because of the problems raised by the nine parameters approach described in the previous section, we considered the alternative non-parametric model. Thus, trajectories are modelled by the set of conditional probabilities $P_i, i \in [0, 150]$, which describe the intrinsic geometry of the main axons in each group under study. Naturally, the explicit bond with the biological

Table 3.2 Number of parameters with $p < 0.05$ for the non-parametric Kruskal Wallis test.

	Imp L	**Imp Sh**	**Prof Rescue**
WT	12	22	28
Imp L		16	19
Imp Sh			24

nature of the problem is lost, thus the model becomes phenomenological. However, this approach presents no need of solving systems of equations, as each parameter is directly defined from data, avoiding the statistical problems raised by the nine parameters approach.

We performed the Kruskal Wallis non-parametric test between populations for each conditional probability P_i, $1 \leq i \leq 150$ (calculated per axon). Table 3.2 shows the number of parameters P_i that present $p < 0.05$ between each pair of populations. Note that the values in Table 3.2 are small considering that from the existent 150 parameters, the maximum amount of them showing a significant difference between groups is 28. In addition, regarding the probability of each combination of two past unit vectors $[\overrightarrow{x}_t, \overrightarrow{x}_{t-1}]_i$, the results of the estimation show that all the groups share the six most frequent configurations, representing together between 65 and 76% of the total. Moreover, no relevant differences where observed regarding these six frequent discrete distributions between the groups (Hellinger distances between them presented near zero values in every case).

3.2.3 Branching point occurrence

We propose a model that describes the branching point distribution independently of the axon length and that is inspired from the biological process of interstitial branch formation during development. We consider interstitial branch formation and not growth cone bifurcations, for two main reasons. The first one is that growth cone bifurcations usually occur to stablish well-known Y or T shapes in neurons, and are necessary for the nervous system organization (Davis et al., 1989; Gallo, 2011a; Schmidt et al., 2007). This description does not seem to fit our particular case where neurons present quite different morphologies between them (and not a particular structural organization) and there is no evident need to reach different anatomical regions with different neuron tips. The second one is that branches that are born from bifurcation processes grow exactly as their parallel neurite (Gallo, 2011a), and our database includes examples of phenotypes where branches are affected but not the main axon.

This process can be described in three simple steps (Fig. 3.10): i) the main axon grows following particular external and internal guiding cues. ii) When the growth cone senses external guiding cues indicating the formation of an interstitial branch, the main axon

decreases its growing speed until it totally stops while it accumulates molecular material in its tip (*i.e*. filopodial and lamellipodial protrusions which serve as precursors to branches (Kalil and Dent, 2014)). iii) After some time the main axon continues growing following its particular cues, leaving the accumulated material in a specific zone of its shaft. The left material is then organized into an independent growing tip and starts elongating an interstitial branch towards its particular target (adapted from (Szebenyi et al., 1998)).

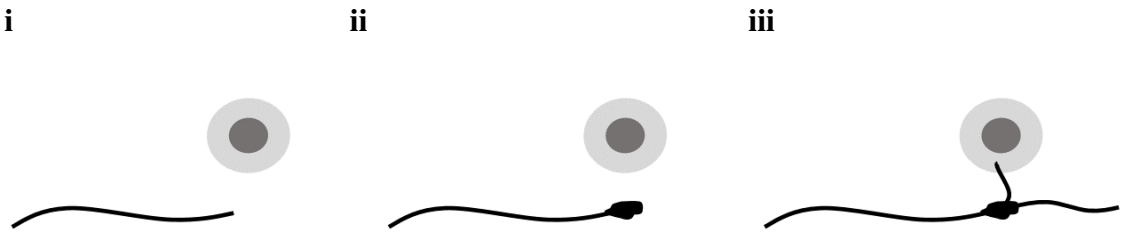

Fig. 3.10 Interstitial branch formation during axonal development described schematically in three main steps, adapted from Szebenyi et al. (1998). i) The main axon grows. ii) When the growth cone senses external guiding cues indicating to branch, the growth speed is decreased until it stops. iii) After some time the main axon continues growing, leaving accumulated material in a specific zone of its shaft. The left material is then organized into an independent growing tip and starts elongating an interstitial branch towards its particular target.

Modelling this process was initially complicated as none of two main actors (growing rate, guiding cues presence) can be measured from the adult stage confocal images in our data-base. Regarding these limitations, we focused our study on the behaviour of the axon growing rate v. We assumed that the axon starts growing with a certain initial speed v_0 and evolving until $v = 0$, when a new branch point appears.

We measured the number k of pixels between every two successive branching points along each axon. Then, we supposed that each of these pixels represented a differential progress in the axon growth, where the axon had a certain growing rate v. Our model assumes random decreases in speed which we call $\triangle v$, with a probability of occurrence p. When a certain number of decreases $\triangle v$ occurs, the speed v equals zero, thus the growing tip stops, allowing the material needed to form a branch to accumulate. After some time the process starts again, with initial speed v_0.

Because at each k pixel a decrease in v may or not happen independently, we described the problem using a Bernoulli distribution, where each success means the occurrence of a decrease in speed ($\triangle v$). We considered that the growing rate goes to zero after $A+1$ steps of speed decrease. The probability to reach $v = 0$ after k steps is given by an accumulation process written as follows

$$P(k) = \binom{k-1}{A} p^{A+1}(1-p)^{k-A-1}. \tag{3.5}$$

Equation 3.5 gives the probability of having A successes in $k-1$ trials and a success in the k^{th} trial. This means that the axon tip decreases its speed A times before stopping completely (which happens in $A+1$), or equivalently that the length between two branching points is k (Fig. 3.11). Thus, our accumulation Bernoulli-based, time-mimicking branching point distribution model has two parameters, A and p, to be estimated from data. Knowing all the distances k_i between successive branching points for every axon in each group, we can calculate their mean and variance μ_k and σ_k^2. From

$$\mu_k = \sum_i k_i \binom{k_i-1}{A} p^{A+1}(1-p)^{k_i-A-1} \tag{3.6}$$

and

$$\sigma_k^2 = \sum_i k_i^2 \binom{k_i-1}{A} p^{A+1}(1-p)^{k_i-A-1} - \mu_k^2. \tag{3.7}$$

It can be shown that $\mu_k(A,p)$ and $\sigma_k^2(A,p)$ have the simple following forms (proof described in A.1 in Appendix A):

$$\mu_k(A,p) = \frac{A}{p}; \tag{3.8}$$

$$\sigma_k^2(A,p) = \frac{(1-p)A}{p^2}. \tag{3.9}$$

Equations 3.8 and 3.9 allow to easily estimate A and p from data. Once A and p are estimated, A needs to be rounded as it has to be an integer. Then, p can be recalculated knowing the value of A as

$$p = \frac{\sqrt{A(\mu_k + \sigma_k^2)}}{\mu_k + \sigma_k^2}. \tag{3.10}$$

With this model, the number $A+1$ of needed $\triangle v$ in order to form a branch, and p their probability to happen define each axonal group regarding their branch density.

Tables 3.3 and 3.4 present the resulting values of the parameters for each group and the p values from the Kruskal Wallis test comparing all the distances between two consecutive branches k_i among groups, respectively. Interestingly, the only significant difference ($p < 0.05$) appears between Imp Sh and every other group (i.e. Imp Sh is more densely branched). While each group has the same value of A, Imp Sh presents the highest value of p. This

Table 3.3 p values from the non-parametric Kruskal Wallis test comparing the distances in pixels between consecutive branches between the studied groups. Significant differences appear in blue.

	Imp L	**Imp Sh**	**Prof Rescue**
WT	0.9	$4.2e-3$	0.6
Imp L		$2.2e-2$	0.7
Imp Sh			$1.3e-2$

Table 3.4 Value of the parameters A and p describing the branching point distribution.

	A	**p**	**p for A = 1**
WT	1.2	0.0087	0.0078
Imp L	1.0	0.0068	0.0067
Imp Sh	0.9	0.0080	0.0084
Prof Rescue	1.2	0.0074	0.0068

means that $\triangle v$ occurrence is more probable, thus it takes less time to reach $v = 0$, and consequently it is the most branched group.

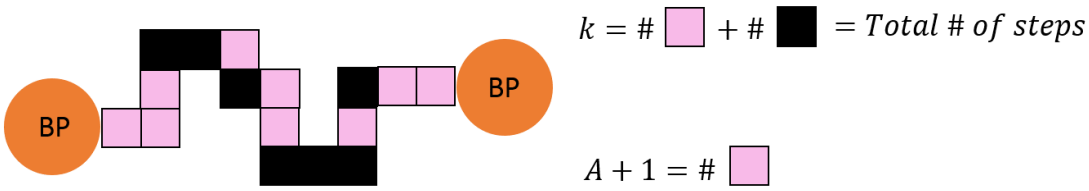

Fig. 3.11 Bernoulli model for branch occurrence. 2D 4-connected path showing an axonal trajectory between two branching points (3D not shown for simplicity). Light coloured pixels depict $\triangle v$ and occur with a probability p, decreasing the growing rate. When the number of light coloured pixels equals $A+1$, $v = 0$ and a new branching point appears.

3.2.4 Branch length

To study the branch length distribution within the groups, we established four length categories (μm); $L_1 : (0,1]$, $L_2 : (1,5]$, $L_3 : (5,10]$ and $L_4 : (10,\text{inf})$ following Tessier and Broadie (2008). In this work, the authors study the role of the $fmr1$ gene which, like imp, codes a mRNA transport protein whose function has been associated with pruning of γ axon branches during late development. They consider large branches (>10μm) and short ones (<5μm). One of their findings is that small branches go through a process of pruning at eclosion in

Table 3.5 Branch length distribution by length category per group (%).

	L_1	L_2	L_3	L_4
WT	10.6	49.2	11.7	28.5
Imp L	8.4	66.5	10.1	15
Imp Sh	19.8	48.2	14.5	17.5
Prof Rescue	19.5	48.3	10.2	22

Table 3.6 p values from the non-parametric Kruskal Wallis test comparing the branch length distribution in L_2 between the studied groups. Significant differences appear in blue.

	Imp L	**Imp Sh**	**Prof Rescue**
WT	$8.9e-5$	0.9	0.8
Imp L		$9.0e-4$	$1.0e-3$
Imp Sh			0.9

wild-type individuals, but not in $dfmr1$ null ones. The category L_1 was added to eliminate the influence of very small branches that may actually be artefacts. The length was measured in the same way as described for the main axon. The probability distribution modelling the relative amount of branches within these length categories and for each group was considered as Gaussian. For each group of axons X, we calculated the mean and standard deviation $(\mu_{bi}^X, \sigma_{bi}^X)$, $1 \leq i \leq 4$ of the relative number of first order branches per axon corresponding to each length category b_1-b_4 (Table 3.5). To know between which groups and for which length categories the differences are significant, we performed the Kruskal Wallis non-parametric test for the four length groups. Significant results ($p < 0.05$) are only present within L_2 and L_4 categories. The p values are shown in Table 3.6 and Table 3.7. Imp L presents significantly more branches in L_2 than any other group while WT has a bigger proportion of L_4 branches than Imp L and Sh, but not Prof Rescue. For further analysis, we took only these discriminant categories L_2 and L_4.

Table 3.7 p values from the non-parametric Kruskal Wallis test comparing the branch length distribution in L_4 between the studied groups. Significant differences appear in blue.

	Imp L	**Imp Sh**	**Prof Rescue**
WT	$3.5e-4$	$1.3e-4$	0.2
Imp L		0.7	0.1
Imp Sh			0.1

3.3 Likelihood Analysis

We first define the likelihood according to the model of each feature:

- Main axon length:

  The likelihood of a given axon n of length l_n to belong to a given group X is defined by the Normal probability density function

$$L_l(l_n|n \in X) = P(l_n|n \in X) = \frac{1}{\sigma^X_{m.a.}\sqrt{2\pi}} \exp \frac{-(l_n - \mu^X_{m.a.})^2}{2{\sigma^X_{m.a.}}^2} \tag{3.11}$$

- Main axon morphology:

  The computation of the Markov chain likelihood appears to lack of robustness to compare populations. This can be explained by the limited length of the axons in pixels ($\sim$1500) and the combinatorial of the problem (150 conditional probabilities). Indeed, some of the three vector configurations, even though with non-null probability, may not appear in the learning sample. When this is the case, if the axon to classify does present at least one time this configuration the likelihood becomes zero. This means that the likelihood is extremely sensible to fluctuations in the presence of low probable events, which is statistically inevitable with the size of our data. To overcome this inconvenience and add robustness to the likelihood analysis, an original approach was applied (Eq. 3.12). We assume the 30 discrete probability distributions $P(.|\overrightarrow{x}_t, \overrightarrow{x}_{t-1})$ as independent, and defined a multinomial Bernoulli distribution for the variable $\overrightarrow{x}_{t+1}$ for each given $[\overrightarrow{x}_t, \overrightarrow{x}_{t-1}]_j$. For each axon n, the likelihood of each group X according to the shape model of X, $P_{(1-150,X)}$, and the frequencies of appearance of three unit vectors corresponding to n, $\#_i$, is then defined as follows

$$\begin{aligned} L_{sh}(\#_{1:150}|n \in X) &= P(\#_1 - \#_{150}|n \in X) = \\ \prod_{s=1}^{30} \binom{N_s}{\#_{k+1}} P_{k+1}^{\#_{k+1}} \binom{N_s - \#_{k+1}}{\#_{k+2}} P_{k+2}^{\#_{k+2}} \quad &\cdots \\ \binom{N_s - \#_{k+1} - \#_{k+2} - \#_{k+3} - \#_{k+4}}{\#_{k+5}} P_{k+5}^{\#_{k+5}} & \\ k = 5(s-1); & \\ N_s = \sum_{k=1}^{5} \#_{s(s-1)+j} & \end{aligned} \tag{3.12}$$

- Branching point occurrence:

To calculate the likelihood of each neuron n to belong to the group X regarding this model, we use the Binomial probability density function considering the distances between each pair of consecutive branches $k_{n,m}$ independent between them, obtaining

$$L_{bp}(k_{n,m}|n \in X) = P(k_{n,m}|n \in X) = \prod_{m=1}^{M} P(k_{n,m}|n \in X) = \prod_{m=1}^{M} \binom{k_{n,m}-1}{A_X} p_X^{A_X+1} p^{k_{n,m}-1-A_X} \tag{3.13}$$

where M is the total number of pairs of branches.

- Branch length:

  To calculate the likelihood of each neuron n with each group X regarding the branch length distribution in L_2 and L_4 - $b_{n,2}$ and $b_{n,4}$ - we considered a bivariate Gaussian distribution with mean $\overrightarrow{\mu}_b^X = (\mu_{b2}^X, \mu_{b4}^X)$ and Σ^X the covariance matrix between μ_{b2}^X and μ_{b2}^X.

$$L_{bl}(\overrightarrow{b}_n|n \in X) = P(\overrightarrow{b}_n|n \in X) = \frac{1}{2\pi\sqrt{|\Sigma^X|}} \exp -\frac{1}{2}(\overrightarrow{b}_n - \overrightarrow{\mu}_b^X)^T \Sigma^{X^{-1}} (\overrightarrow{b}_n - \overrightarrow{\mu}_b^X), \tag{3.14}$$

  where $|\Sigma^X|$ is the determinant of the covariance matrix Σ^X.

Thus, for each axon n, we calculated its features and then computed the likelihood for each group X, ($X \in WT, Imp, ProfRescue$). The axon n was thus classified in the group that maximizes the global likelihood. All the classifications presented in this work were done using the "leave one out" technique, which consists in classifying an element of the sample that has been removed from the database to perform the learning step (i.e. the estimation of the model parameters). This maximum likelihood classification does not only provide an assessment about the discriminative properties of the proposed models, but is also useful to analyse the mixture of feature values between the populations.

Considering the four features independent, the global likelihood is given as follows

$$L(l_n, \#_{1:150,n}, k_n, \overrightarrow{b}_n|n \in X) = L_l(l_n|n \in X) L_{sh}(\#_{1:150,n}|n \in X) L_{bp}(k_n|n \in X) L_{bl}(\overrightarrow{b}_n|n \in X) \tag{3.15}$$

and the maximum likelihood estimation results

Table 3.8 Global likelihood analysis considering the four features.

		Predicted (%)		
		WT	Imp L	Imp Sh
Actual Class	WT	82.6	17.4	0
	Imp L	54.5	45.5	0
	Imp Sh	19.2	3.9	76.9
	Prof Rescue	40	26.7	33.3

$$n \in X_0 \longleftrightarrow X_0 = \underset{X}{\text{argmax}} L(l_n, \#_{1:150,n}, k_n, \overrightarrow{b}_n | n \in X), \tag{3.16}$$

$$X = WT, ImpL, ImpSh, ProfRescue. \tag{3.17}$$

Equation 3.16 allows to classify each axon by resemblance to each group considering the four morphological features (main axon length and shape, branch length distribution and branch point distribution) and their stochastic models. Table 3.8 presents the results of the global likelihood analysis.

These results suggest a relevant global difference between neurons belonging to Imp L and Imp Sh, as well as between WT and Imp Sh; while between WT and Imp L this difference is weaker. More than half of Imp L axons are likely to be WT while for Imp Sh this proportion is less than 20%. Some WT axons are classified as Imp L but none as Imp Sh. Interestingly, the percentage of Prof Rescue neurons likely to be WT lies in between the percentages for Imp Sh and Imp L. This result supports the idea of a partial rescue of the Imp neuron morphology.

To understand how each morphological feature contributes to the results in Table 3.8, we carried out the maximum likelihood analysis regarding each of them separately. For the main axon length, as expected from Fig. 3.8, WT axons are shared between WT and Imp L categories; and Imp L is mixed with WT. Imp Sh is completely separated from the rest of the groups (Table 3.9). Regarding Prof Rescue, our results agree with those in Medioni et al. (2014) about the main axon length being partially rescued by Profilin over-expression. Our study pointed out that 46% of *imp* mutant neurons present a conserved main axon length while 54% are significantly shorter than WT. Here we show that Prof Rescue axons are distributed by a 67% (Imp L + WT) vs. 33% (Imp Sh), moving the tendency towards a WT phenotype.

According to the main axon shape in Table 3.10, WT and Imp L look again similar and, interestingly, Imp Sh looks more similar to WT than to Imp L. Prof Rescue behaviour is opposite to that of Imp Sh.

Table 3.9 Likelihood analysis according to the **main axon length** feature.

		Predicted (%)		
		WT	Imp L	Imp Sh
Actual Class	WT	39.1	54.4	6.5
	Imp L	22.7	77.3	0
	Imp Sh	0	0	100
	Prof Rescue	6.7	60	33.3

Table 3.10 Likelihood analysis according to the **main axon shape** feature.

		Predicted (%)		
		WT	Imp L	Imp Sh
Actual Class	WT	54.3	43.5	2.2
	Imp L	50	50	0
	Imp Sh	61.5	38.5	0
	Prof Rescue	40	60	0

Table 3.11 presents the likelihood analysis results regarding the branch point density. It can be noticed that every group is mainly classified as Imp Sh, which our previous analysis revealed as the most branched group. The reason for this behaviour relies on the nature of the model. Even though the means of the distances between branches are different between the biological groups, axons frequently display one or more pairs of branches which are close. Because for close branches the likelihood is maximum for Imp Sh, with a significant difference from the other groups, the presence of near branches automatically classifies a neuron as Imp Sh. Nevertheless, the branch density coherence is respected for each group as the resemblance with Imp Sh is maximum for the most branched group (itself) and is followed in the correct order (regarding the statistical analysis): WT first, followed by Imp L and Prof Rescue.

Finally, according to the branch length distribution (Table 3.12) WT, Imp L and Imp Sh show a higher resemblance to their own groups, suggesting a significant difference between them regarding this feature. Prof Rescue has a slight preference for Imp Sh which

Table 3.11 Likelihood analysis according to the **branching point** feature.

		Predicted (%)		
		WT	Imp L	Imp Sh
Actual Class	WT	0	13	87
	Imp L	13.6	18.2	68.2
	Imp Sh	7.7	11.5	80.8
	Prof Rescue	6.7	26.7	66.7

Table 3.12 Likelihood analysis according to the **branch length distribution** feature.

		Predicted (%)		
		WT	Imp L	Imp Sh
Actual Class	WT	60.9	23.9	15.2
	Imp L	18.2	72.7	9.1
	Imp Sh	15.4	30.8	53.8
	Prof Rescue	33.3	20	46.7

Table 3.13 Global likelihood analysis considering the four features. Imp includes Imp Sh and Imp L.

		Predicted (%)	
		WT	Imp
Actual Class	WT	80.4	19.6
	Imp	37.5	62.5
	Prof Rescue	60	40

is understandable as both have the same proportion of branches in L_2 which is the most abundant group of branches. Nevertheless, its resemblance to WT regarding this feature is notoriously higher than those for Imp L and Imp Sh. This result reveals that Profilin over-expression partly rescues branch length distribution –i.e. it presents a bigger proportion of long branches- in addition to the main axon length.

In order to better understand the morphological changes induced by Profilin over-expression in *imp* mutant axons, we performed the global maximum likelihood analysis considering *imp* mutants altogether (i.e. Imp Sh + Imp L), and the possible classification groups either altogether (Table 3.13) or split between Imp L and Imp Sh (Table 3.14).

From the analysis in Table 3.13, we can highlight that while only 37.5% of *imp* mutant axons present a WT phenotype, 60% of Profilin rescue axons exhibit this behaviour. Moreover, it is interesting to analyse how Prof Rescue is classified regarding Imp L and Imp Sh (Table 3.14). The percentage of neurons classified as Imp Sh decreases compared to *imp* mutant from 42 to 33% while the tendency for Imp L and WT is increased in Prof Rescue.

Table 3.14 Global likelihood analysis considering the four features. Imp is split between L and Sh for possible classification groups.

		Predicted (%)		
		WT	Imp L	Imp Sh
Actual Class	WT	82.6	17.4	0
	Imp	35.5	23	41.5
	Prof Rescue	40	26.7	33.3

Finally we performed a brief comparison regarding our classification results with those in Mottini et al. (2013), where wild-type as well as *imp* mutant γ axons were analysed as well. The authors reported 85 and 91.7% of accurate classifications for WT and Imp respectively with the ESA curve distance method (based on the Elastic Shape Analysis Framework Mottini et al. (2015)), and 52.6 and 79.2% with RTED (an efficient solution for the tree edit distance problem (Pawlik and Augsten, 2011)). It is relevant to highlight that the goal in their work was to discriminate between populations, thus they considered exclusively highly discriminative parameters. On the contrary, our results -80.4 and 62.5% for WT and Imp respectively- show and value not only the differences but also the existing similarities between phenotypes, considering all the relevant morphological features (including those that may be known as not discriminative). Our work also allows to correlate the conclusions with biological parameters. In addition, our sample size doubles the one used in the previous work.

3.4 Simulation

The models behind the main axon length, morphology and branching point distribution are not only descriptive but also generative, meaning that new axons can be simulated using the models and their parameters, whose values were estimated from data.

The Markov chain model for the main axon path establishes the probability of the next position knowing the current position and the two before. This is why the first three points of the simulated paths are arbitrarily pre-established as (0,0,0), (0,1,0) and (0,2,0). The length of the main axon (in number of steps) is given by a random number from a Normal distribution with ($\mu^X_{m.a.}$, $\sigma^X_{m.a.}$) estimated from a particular group of neurons. Finally, each pixel of the resulting path is associated with a success or a failure following the Bernoulli p parameter for branching point distribution. When $A+1$ successes are reached, a branch point is placed and the success counter is reinitialized as zero.

Fig. 3.12 contrasts real and simulated main axons, one per group. The axons relative lengths as well as their branching point densities are conserved. However, simulated axons are shorter than real ones in a μm scale. The reason is that, unlike real axons, their trajectories present some looping and high frequency turns. These pixel-scale phenomena are artefacts of our model, and suggest that the considered scale may not be appropriated for a generative model. In particular, the results show that the pixel scale leads to data over-sampling, thus the estimated parameters capture not only the main axon shape but also the noise present at the pixel scale, due to the discretization and segmentation. To avoid this problem, one possible solution could be to consider Markov chains of higher order. However, this adds

complexity (*i.e.* number of parameters, or conditional probabilities P_i) and worsens the problem mentioned about the lack of representation of some conditional probabilities given the axon typical lengths in our data. Other solutions are to consider as variable other quantity than position, and to change the model scale to a larger and more adequate one. Taking these aspects into account, we developed a more accurate generative model, presented in Chapter 4.

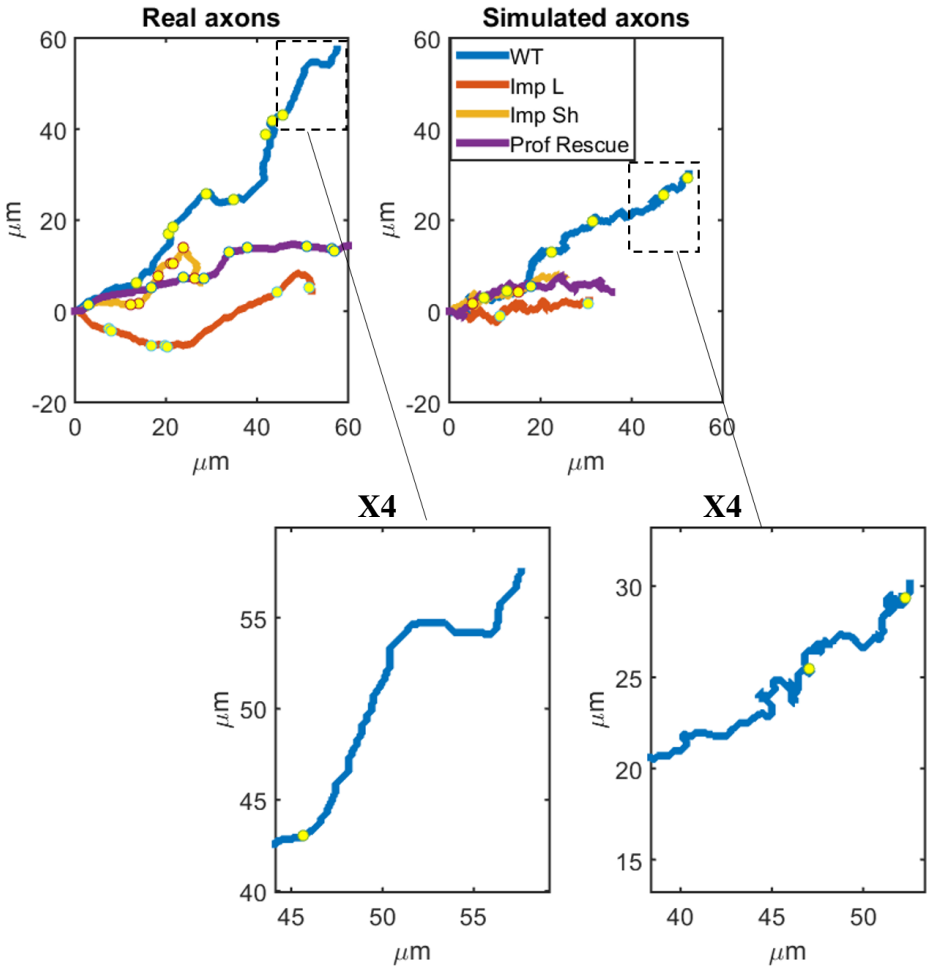

Fig. 3.12 Real vs. Simulated axons from the studied groups. Yellow spots represent branch points. Zoomed areas are also shown in both cases.

Fig. 3.13 displays 50 simulated axons for each neuron group in order to show their mass behaviour. WT presents more diagonal directions than the rest. Imp L, Imp Sh and Prof Rescue display similar growth directions, while Prof Rescue shows less variation between its individuals in this point than Imp L and Imp Sh.

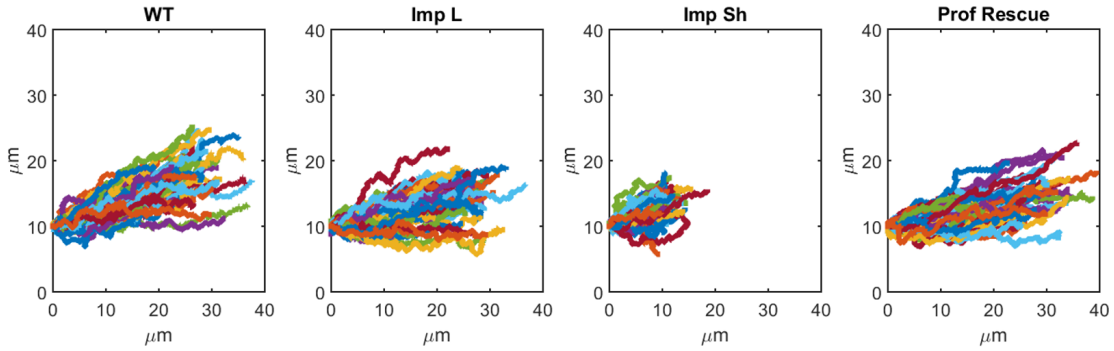

Fig. 3.13 Presentation of 50 simulated main axons for each studied population.

3.5 Discussion and Contributions

3.5.1 Axon growing rate and branch formation

The value of $A = 1$ indicates that, according to our model, the axon tip diminishes its growing speed only two times before stopping to create a branch, instead of doing it gradually. A biological interpretation could be that the first time may be related to the moment when the axon senses the external guiding cues. Then it continues growing more slowly, which may facilitate other cues detection, until it finally stops, consequence of the second and last speed lost. When this happens, branching material could be accumulated and after some time an interstitial branch is created. An increased value of p may indicate a higher sensibility to external cues and/or an increased concentration of internal cues triggering branching. Another interpretation can be that axons with a defective growing rate (i.e. slower speed, or high p) are more susceptible to stop independently from external cues, and therefore to branch more. All the groups presented the same value of A, indicating that this two-step behaviour may be conserved and therefore independent from Imp. Regarding p, Imp Sh is significantly more branched than the rest of the groups, including Imp L, even though they have the same genotype. This is coherent with the short phenotype, as in our model high p means overall growing speed decrease and thus less net growth during the same developmental time. More interestingly, Prof Rescue axons present the same value of p than Imp L. This suggests that the phenotype presenting high branch density (Imp Sh) may be rescued by *profilin* over-expression. Because $profilin$ also rescues the main axon length, one possibility is that the actual rescued feature is the main axon growing speed.

3.5.2 Wild-type axons are mostly differentiated by their branch length distribution

The global maximum likelihood analysis results in more than 80% of WT axons to be correctly classified (Table 3.8, Table 3.13 and Table 3.14). Nevertheless, when looking at each particular feature it becomes clear that WT shares most of those with Imp L. Regarding the main axon length (Table 3.9), 54% of WT neurons are likely to be Imp L and 43% for the main axon shape (Table 3.10). The analysis following the branch point density results in 13% of WT neurons likely to be Imp L, while no WT neuron was correctly classified. This results are validated by the p values for main axon and branch length distribution that do not show significant differences between WT and Imp L. Regarding the shape model the situation is similar, as between Imp L and WT the number of significantly different parameters is the minimum of all the group pairs and it is only of 12 in 150.

The maximum likelihood analysis considering only the branch length distribution is the sole to correctly classify WT axons (Table 3.12). While WT and Imp L present both 80% of branches in L_2 and L_4 altogether (Table 3.5), the difference between them is that WT shows statistically more branches in L_4 while Imp L in L_2. We can relate our results to those of Tessier and Broadie (2008) and Medioni et al. (2014). The first publication reports that a loss of L_2 branches by a late pruning process occurs in wild-type neurons and not in $dFMRP$ mutants ($dFMRP$ is also a Profilin regulator and mutations in $fmr1$ gene coding FMRP is the main cause of the Fragile X Syndrome, a mental retardation disorder and the second one concludes a defective development of long branches (L_4) in *imp* mutants.

The maximal percentage of correct classification for WT considering the features separately is 60% for the branch length distribution (Table 3.12), followed by 54, 39 and even 0% corresponding to main axon shape, length and branching point distribution (Table 3.10, Table 3.9 and Table 3.11). Interestingly, the global classification mixing the four features improves these percentages to 80% (Table 3.8, Table 3.13 and Table 3.14). This suggests that WT neurons are well defined and different from Imp, but it is necessary to consider all the morphological features together for a correct classification. This also highlights the advantages of our method as it goes beyond a simple statistical analysis, allowing to mix different features as well as to consider each neuron independently.

3.5.3 *imp* mutants present two different phenotypes

Medioni et al. (2014) reported that *imp* mutants could either present a conserved main axon length or an aberrant one, with $\sim 50\%$ of occurrence each. Following these results we applied the k-means automatic algorithm to separate the Imp population in Imp L and Imp Sh, and

obtained a 46 vs. 54% of incidence correspondingly. This bimodal behaviour can also be seen in the length distribution (Fig. 3.8). Surprisingly, we have found other relevant morphological differences between this two groups that had not been yet reported in the bibliography. The main one is the branching point distribution, as Imp Sh is significantly more densely branched than Imp L (Table 3.4 and Table 3.3). Also, the percentage of branches ranging from 1 to 5 μm, while aberrant in Imp L, is conserved in Imp Sh (which shows no differences from WT (Table 3.6)).

Regarding the global likelihood analysis (Table 3.8), while less than 20% of Imp Sh neurons can be considered to have a WT phenotype, 55% of Imp L do, allowing to conclude that Imp L presents a generally more conserved phenotype. Globally, we conclude that the penetrance of the *imp* phenotype is 63%, following our global likelihood analysis (Table 3.13 and Table 3.13).

These results are consistent with an essential role of Imp in main axon elongation and branch formation and elongation during the remodelling of γ neurons. Nevertheless, the phenotypical variability within *imp* mutants (i.e. from globally aberrant to completely WT-like neurons) indicates the existence of other –maybe Imp-independent- important actors with the capability of controlling also the remodelling process; and/or that the inhibition of the gene in this condition is not 100% efficient.

3.5.4 Profilin overexpression partially rescues the main axon length as well as the branch length distribution

The global likelihood analysis (Table 3.13) considering Imp altogether shows that Profilin decreases the percentage of the *imp* mutant phenotype from 63 to 40%. Regarding the main axon length, while aberrant neurons represent 54% of the Imp population, they represent only 33% in Prof Rescue (in Prof Rescue 67% of neurons present a conserved length (WT + Imp L) and only 33% do not). Following the branch length distribution likelihood analysis, 33% of Prof Rescue axons are classified as WT and represent the second maximum percentage after WT itself (only 18 and 15% correspond to Imp L and Sh, respectively). Looking at the p values between branch length categories (Table 3.6 and Table 3.7), we suggest that Profilin rescues the late pruning of small branches (Tessier and Broadie, 2008), showing a conserved percentage of L_2 branches and, more interestingly, also allows to develop more branches in the long category. Even though the mean percentage of branches in L_4 is slightly smaller for Prof Rescue than WT (Table 3.5), this difference is not significant in the statistical tests, suggesting a conserved percentage of long branches in Prof Rescue which is not seen in Imp

Sh nor in Imp L, and thus that branching is rescued by Profilin. This results are interesting and novel regarding the previous observations done by Medioni et al. (2014).

Finally, regarding the global likelihood analysis considering Imp L and Imp Sh separately (Table 3.14), we conclude that Profilin rescue diminishes the general morphological aberration, as it moves the tendency towards WT and Imp L phenotypes and lowers the percentage of neurons with an Imp Sh phenotype.

This study suggests that Profilin is also involved in branch formation and elongation, in addition to main axon elongation during remodelling. Nevertheless, because the phenotypical rescue is not complete, we can conclude that the localisation of $profilin$ mRNA in axons by *imp* and thus the local production of Profilin is still essential for elongation, and/or that other Imp RNA targets are involved. Regarding branching, the over-expression of Profilin in the neuron may lead to longer branches, but maybe is not reflecting the wild-type situation neither.

3.5.5 Conclusion

In this chapter we proposed a framework to compare neuron groups based on their morphology. Our procedure consists of applying probabilistic models to describe the behaviour of selected morphological features (*i.e.* main axon length and shape as well as branch length and density), associated parameters estimation from data, and a likelihood analysis combined with statistical tests. We applied this framework to understand the effects of *imp* knock-down -as well as its rescue by Profilin- in Drosophila γ adult axon morphology. The similarities and differences highlighted between wild-type and mutant neurons allowed to better understand the role of Imp and Profilin during axonal remodelling, particularly on axon elongation and branch formation. While the Imp Long phenotype seems to differ from WT only by the absence of long (>10μm) branches, the short phenotype presents, in addition to length shortage, an overall increase in branch density. Regarding the branch occurrence model, this can be interpreted as Imp Sh growing slower, and may suggest that Profilin over-expression rescues the growth speed. Interestingly, the rescue with Profilin results in an homogeneous phenotype, with main axon lengths near to those in WT, as well as an increased number of long ($> 10\mu m$) branches. Finally, our analysis shows that Profilin reduces the general *imp* phenotype penetrance from 63% to 40%, suggesting the importance of Profilin in elongation and branching, but also the need of targeted synthesis (allowed by Imp) to obtain fully grown WT axons. An interesting overall observation is that main axon length rescue correlates with the rescue of the branch phenotype, both in length and number. This correlation is broadly studied in the following chapter, as well as the currently unexplained origin of Imp double phenotype.

We propose that this method consisting in feature selection, model application and likelihood analysis could be applied to any case of study between species where similarities are as important as differences. We can also conclude that the study of individuals is relevant and more enriching than just population analysis driven by ordinary statistics. In addition, we highlight the importance of combining different features to achieve a global result.

A final consideration is that the proposed models -particularly the main axon morphology one- performed better at discriminative and descriptive than generative tasks. In the next chapter we will focus on the generation of realistic axonal morphologies by a mechanistic model, to better understand the developmental process of wild-type axons, as well as the origin of the considered phenotypes.

Chapter 4

Space-Embedded Axon Growth Model

Chaos is merely order waiting to be deciphered.

José Saramago

In the previous chapter we analysed in detail the morphology of adult γ axons and how it is affected by *imp* knock-down and its rescue with Profilin over-expression. In this chapter we go further and try to understand, for each case, *how* the particular axonal tree-structure is attained considering the dynamic process of remodelling of the ML (see Fig. 2.3). We propose a simple mathematical model that captures the essential aspects of axonal elongation and guidance (see Chapter 2), and estimate its parameters directly from data. We then simulate the regrowth of the 650 γ axons synchronously within the ML, considering the retraction of each growing tip after contact with other axons or the ML geometrical limits. These simulations allowed us to study the relevance of considering mechanical interactions between the developing axons to explain their final morphology, as well as to understand certain dynamical aspects of their growth as a population (and not as independent individuals).

Further, we investigate the branching pattern of γ axons and include branch formation in our model. A very classical but still up-to-date approach on interstitial branch formation was considered in Chapter 3, and is described in Szebenyi et al. (1998). Here, branches appeared at points delimited by the growth cone after sensing particular guiding cues that indicate branching and stopping the growth for a while. Branch formation was thus, classically, considered regarding a single neuron growing isolated and interacting only with external chemical guiding cues. Kalil et al. (2000) further discussed interstitial branch formation and continued with the idea proposed by Szebenyi et al. (1998) that branching points appear in regions along the axon shaft where the primary growth cone had stopped its trajectory for a while. They also enhanced the idea that it is the growth cone itself that distinguishes the places of the future interstitial branches by cue recognition, and proposed that signalling

as well as internal pathways seem to be shared between main axon and branches growing tips. A decade later, Gallo (2011b) reviewed the signalling and internal mechanisms of branch formation, and ends up with some unsolved questions: What additional localized signalling events occur in axons? How are these signalling events initiated? What aspects of the mechanisms of branch formation are shared by those of axon extension and guidance?

These open questions inspired and oriented our work towards the study of branching origin, specifically involving mechanical and not only chemical localized signals and possible shared mechanistic and structural aspects between main axons and branches. While the importance of branching has mostly been considered in the context of establishing connections with partners (Armijo-Weingart and Gallo, 2017; Tsigankov and Koulakov, 2009), our approach suggests, strikingly, that they might play another fundamental role than just synapse formation with target cells, which has not been particularly reported before. Finally, this work proposes that the origin of the intra-population phenotypical differences observed within the genetically identical Imp population (largely studied on the previous chapter) is intrinsic to the axonal collective evolution, and one needs to consider mechanical interactions between the developing axons to explain these observations.

4.1 Data treatment

Our database is composed of images of a single axon each. In the previous chapter we analysed them individually and we considered the pixel as the measure unit for all the measurements and model design.

However, for this new step, we need to consider neurons within a population to take into account group effects as well as their realistic environment, so some corrections need to be applied. In particular, we need to take into account the possible differences between images i.e. pixel size, Mushroom Body geometry, dimensions, rotation and relative position of the axon within the ML. We thus implemented a normalization procedure that allowed us to place all the single axons (collected from different MB) as if they were part of the same ML. To do so, we chose a Standard ML (SML) from one of the images and drew two main curves signalling the entrance and the end of the ML (Fig. 4.1). Then, to position each neuron in the SML, we defined the point on the entrance (red) curve that better represents the actual axon Entrance to the Lobe (EL), taking into account the anatomy of the MB in each particular case. To correct the difference in scale and in ML size, we took the distance (in μm and parallel to the ML main axis) from the entrance (EL) to the end of the ML in both, the image we want to normalize and the SML. We named the quotient between these two values the Scale Factor (SF) by which we then multiply the points in the neuron segmentation. Finally, we measured

the angle between the ML main direction and the image horizontal axis to have the Rotation Angle (RA). We considered that the RA in the SML is zero. All these magnitudes (EL, SF and RA) are illustrated in Fig. 4.1B and C. We chose this procedure instead of a classic image registration method because a rigid one would not have been accurate to account for differences in the morphology of the MLs and a non-rigid one, on the other hand, would have also deformed the neuron geometry. Moreover, the size of our data set makes it possible to apply a handmade procedure analysing each image separately. We also decided to consider a higher scale than the pixel size to describe the axon paths, *via* a simple polygonization of the trajectories. The procedure to normalize an already segmented axon is then to multiply the points of its path by the SF and then re-sample (polygonize the paths) considering the new scale (distance between points). Throughout this work we used as sampling scale the general axonal diameter observed in the images: 1 μm. Then we rotated every point by the angle RA and subtracted the initial point to all the points in the tree (take the axon to the origin). The last step is to add the EL point to place the neuron in the starting point.

Fig. 4.2 shows examples of single axons before (A) and after (B) the normalization, as well as the entire WT population placed in the SML (C). An interesting observation is that not every axon reaches the ML end (Fig. 4.1 A). We thus defined the "stopping region", as the region where wild-type axons stop growing in our data after registration. It goes from $X = 62\mu m$ to the end of the ML. In other words, the "travelled distance" of every wild-type axon is higher or equal than $62\mu m$. The schemes representing the stopping region and travelled distance are shown on Fig. 4.2 D. In the rest of this work, we consider that an axon is elongated if it stops within the defined stopping region (and non-elongated if it does not). This classification does not pretend that every γ axon has as unique goal to reach this region, but becomes useful to develop underlying hypothesis of axonal length shortage, generally considered aberrant in the bibliography (Medioni et al., 2014; Yaniv et al., 2012).

Fig. 4.3 shows all the axons in our dataset normalized to the SML following this procedure. As it can be seen, the wild-type γ axons spread over the entire medial lobe, suggesting that our database is representative of the entire population.

4.2 Mathematical model formulation

To model the growth of adult γ axons that occurs *in vivo* during metamorphosis, we consider the growth of individual axon tips as a succession of discrete steps, and used a Markov chain with an artificial constant time unit to model the angles of these incremental steps. At each artificial time point t_j, individual growing tips advance of a certain number n_{max} of discrete steps ρ_i with a constant size $\Delta\rho = 1\mu m$ fixed according to data sampling (see 4.1). We

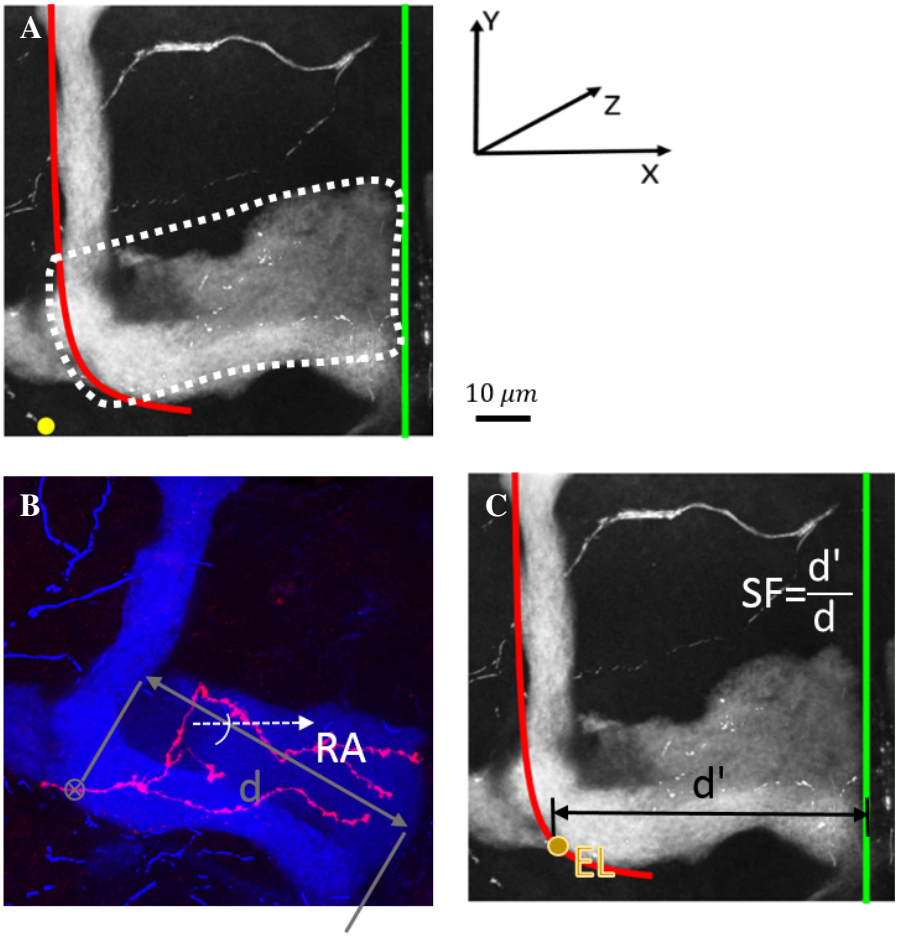

Fig. 4.1 γ axon normalization to the same standard medial lobe (SML). (A) shows the entrance of the neurons to the ML (red) as well as the end of the ML (green) and in yellow the image origin point (origin of coordinates). The dotted white line represents the ML shape. (B-C) shows how we defined RA, SF and EL for each one of our images. The cross in a circle (in B) shows the point where the neuron enters the dorsal lobe, which is represented as the EL point in the SML. The Rotation Angle RA is taken between the ML and the X axis (in the SML this angle is equal to zero) and the Scaling Factor (SF) is the quotient between the distances from EL to the end of the ML that is parallel to the lobe from the image and the SML. Scale bar: $10\mu m$.

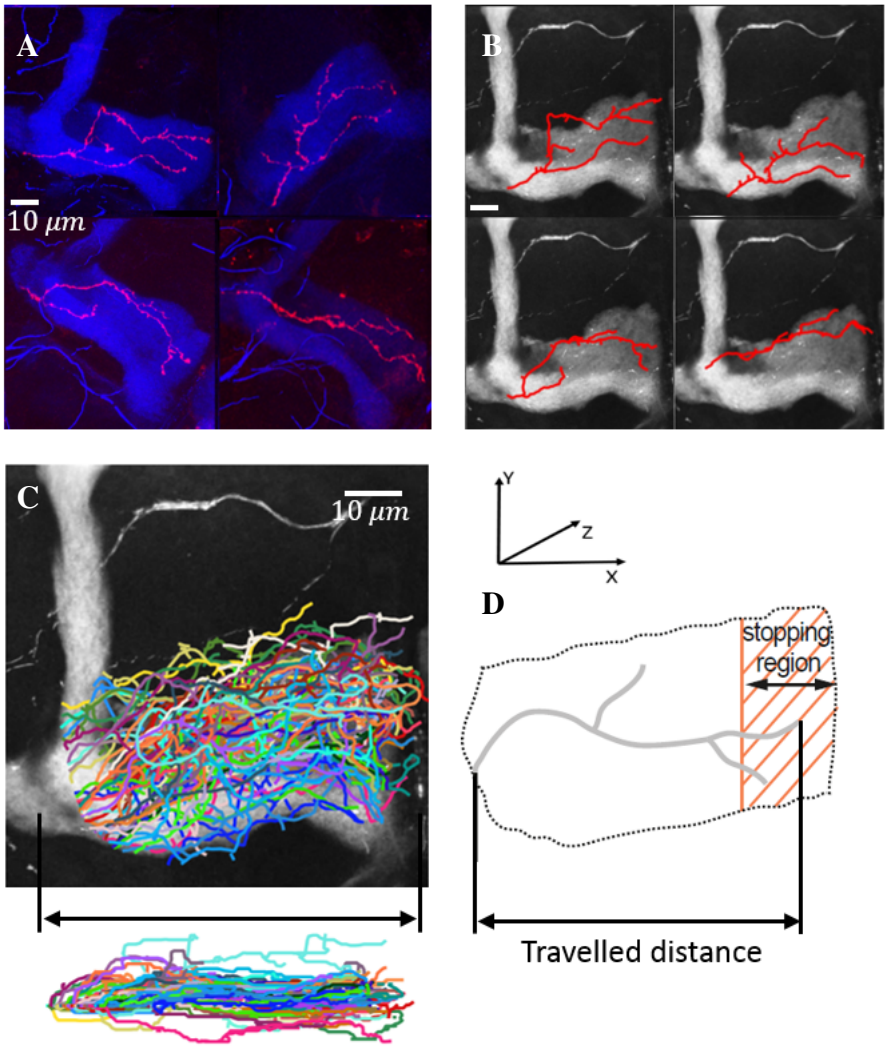

Fig. 4.2 γ axon normalization: results. (A) Maximum intensity projections of some wild-type single axon images in our dataset. (B) The same axons in (A), segmented and placed in the SML. (C) The entire wild-type γ population in our database, segmented and placed altogether in the SML. In Z, all the axons start from the origin ($Z = 0$). (D) ML scheme indicating the definitions of the stopping region (i.e. where wild-type axons stop) and travelled distance (i.e. projection on the X axis). Scale bars: $10\mu m$.

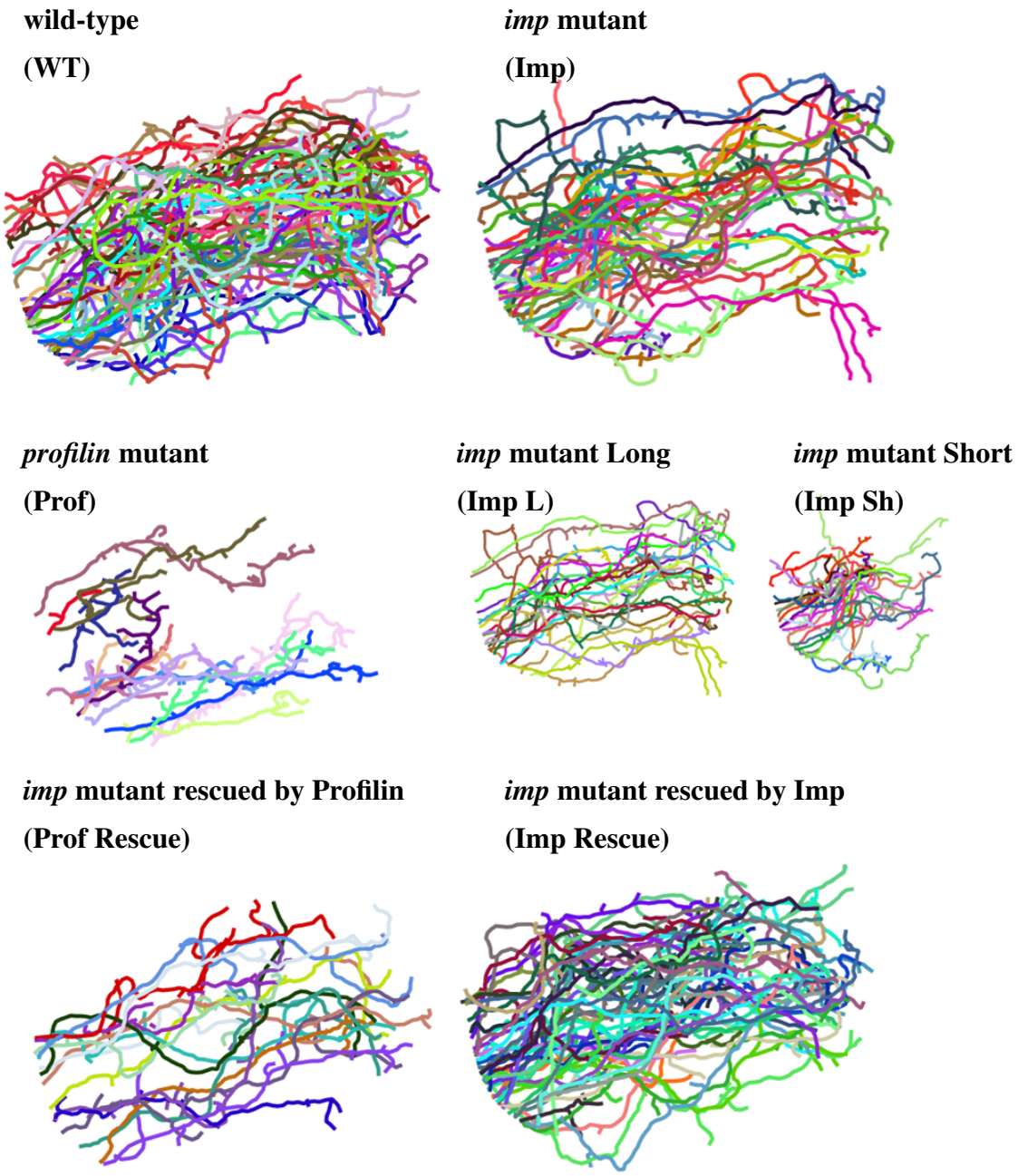

Fig. 4.3 Normalized γ axon populations. We present all the populations in our database, normalized to the SML. *imp* mutant axons are also split between the short and long phenotypes.

hypothesized that the direction of each step ρ_i is influenced both by an internal mechanical constraint (the axon shaft stiffness), and by an external attractive gradient field that sets the axon preferred growth direction (Fig. 4.4 A). These two features are represented in our model by the α and β parameters, respectively. We use spherical coordinates $(\Delta\rho, \phi_{x,y}, \phi_z)$ to describe the i^{th} step vector, where $\phi_{x,y}$ represents the angle in the xy plane and ϕ_z the elevation along the z axis. The step size $\Delta\rho$ is set constant, thus the model is reduced to two variables although we consider 3D trajectories. Moreover, we consider that both angles are independent, reducing the full model to two sub-models with only one variable each: ϕ. The i^{th} step vector is thus described by its spherical coordinates $(\Delta\rho, \phi_{x,y}, \phi_z)_i$ and placed at the end of the previous step $i-1$, at coordinates $(x,y,z)_{i-1}$.

The probability distribution of the i^{th} axonal tip Cartesian position $(x,y,z)_i$, knowing the previous one, $(x,y,z)_{i-1}$ is thus directly defined by the probability distribution of ϕ_i, which is conditioned by ϕ_{i-1} and defined by

$$P(\theta_i|\theta_{i-1}) \propto exp - [\alpha(\theta_i - \theta_{i-1})^2 + \beta(\theta_i)^2], \tag{4.1}$$

where

$$\theta_i = tan(\frac{\phi_i - \phi_{0,i}}{2}), \tag{4.2}$$

and $\phi_{0,i}$ is the direction of the external attractive field at the position $(x,y,z)_{i-1}$. The transformation in Eq. 4.2 establishes the domain of the model to $(-\infty, \infty)$, allowing a direct implementation of the Normal distribution and avoiding the Bessel functions for the normalization of the usual von Mises distribution for the angular domain (Fig. 4.4B). With this definition, the variable θ_i is locally defined in reference to the attractive field direction. Furthermore, α weights the difference between the future and current direction (*i.e.* stiffness), and β the difference between the i^{th} step vector direction and the direction of the attractive force (*i.e.* attractiveness to the external field). Thus, high values of α favor a straight axonal trajectory while $\alpha \to 0$ results in a very tortuous one (Fig. 4.4 C). Similarly, if the external attraction is very high ($\beta \to \infty$), axons tend to align to the external field gradient lines, while if this attraction is low ($\beta \to 0$) they do not follow any preferential direction (Fig. 4.4 C).

Notably, the model can be equivalently written as follows:

$$\theta_i = \frac{\alpha}{\alpha + \beta}\theta_{i-1} + \xi_i, \tag{4.3}$$

where

$$\xi_i \sim \mathcal{N}(0, \frac{1}{2(\alpha+\beta)}). \tag{4.4}$$

With this formulation, it becomes clear that at each step of the chain θ_i conditioned to the step before, θ_{i-1}, follows a Normal distribution with $\mu = \frac{\alpha}{\alpha+\beta}\theta_{i-1}$ and $\sigma^2 = \frac{1}{2(\alpha+\beta)}$ (Fig. 4.4 A).

To feed our model with biologically meaningful parameter values, we estimated the model parameters α and β from the reconstructed axon paths obtained from real data. The nature of the axon trajectories from our data add a particular difficulty in order to estimate the single pair of parameters that describes the whole population. A deep study on the subject and a review of estimation methods is presented in Section 4.3.

Finally, one interesting characteristic about this model is that it allows an instant transformation of the parameter values in case of information loss (e.g. to have only one over two data points)(Fig. 4.5 B). In other words, we can calculate $P(\theta_i|\theta_{i-2})$ knowing $P(\theta_i|\theta_{i-1})$ (*i.e.* renormalization by decimation) as

$$P(\theta_i|\theta_{i-2}) = \int_{-\infty}^{\infty} P(\theta_i|\theta_{i-1})P(\theta_{i-1}|\theta_{i-2})d_{\theta_{i-1}} \tag{4.5}$$

to obtain the expression

$$P(\theta_i|\theta_{i-2}) \propto exp - [\alpha'(\theta_i - \theta_{i-2})^2 + \beta'(\theta_i)^2], \tag{4.6}$$

with

$$\alpha' = f_1(\alpha,\beta) = \frac{(\alpha+\beta)\alpha^2}{(\alpha+\beta)^2+\alpha^2} \tag{4.7}$$

$$\beta' = f_2(\alpha,\beta) = \frac{(\alpha+\beta)[(\alpha+\beta)^2-\alpha^2]}{(\alpha+\beta)^2+\alpha^2}. \tag{4.8}$$

Thus, α' and β' are the parameters defining the probability of θ_i conditioned by θ_{i-2}[1]. Interestingly, these equations are invertible such as

$$\alpha = f_1^{-1}(\alpha',\beta') = \alpha'' = \sqrt{\frac{(2\alpha'+\beta')^3 - (2\alpha'+\beta')^2(\alpha'+\beta')}{\alpha'+\beta'}} \tag{4.9}$$

$$\beta = f_2^{-1}(\alpha',\beta') = \beta'' = 2\alpha'+\beta'-\alpha''. \tag{4.10}$$

[1] α' and β' can also be obtained by recursively applying Eq. 4.3

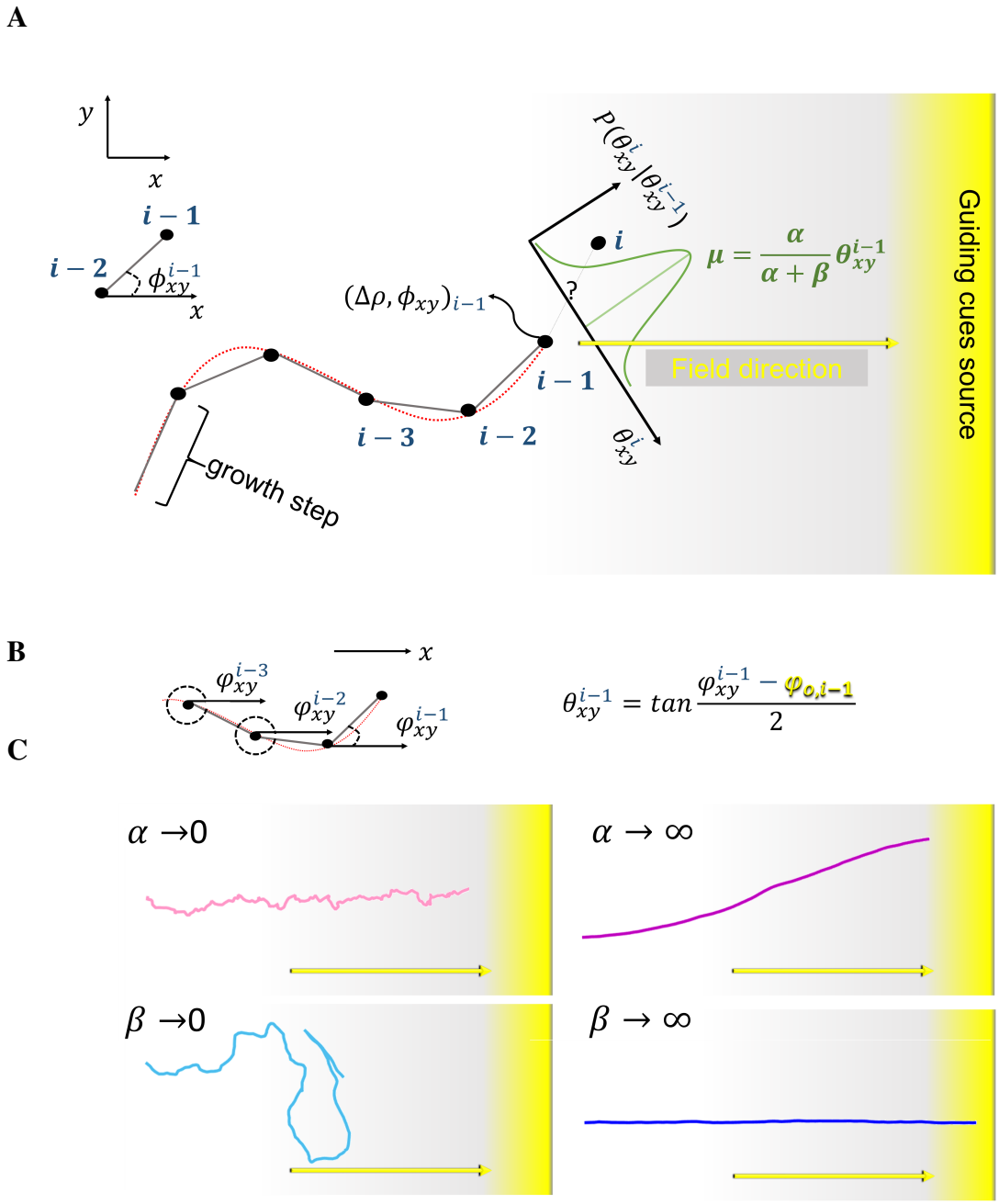

Fig. 4.4 3D Mathematical model of individual axon growth. (A) The axon elongates step by step (grey segments), each step being delimited by the current point in space i and the previous one, $i-1$. Each i^{th} position in space is described by its spherical coordinates $(\Delta\rho, \phi_{x,y}, \phi_z)_i$ (represented in 2D for simplicity). $\phi_{x,y}^{i}$ is thrown from a conditional Normal probability distribution (in green). The most probable value (mean of the distribution (μ)) considers the directions of the last step and of the attractive field, each contribution weighted by α and β respectively. For each time unit t_j, a maximum number of steps is allowed. (B) Detail of how angles $\phi_{x,y}^{i-1:3}$ are measured and the value of $\theta_{x,y}^{i-1}$. (C) Influence of α and β values on axon trajectories. α represents the axon stiffness; values near zero result in very tortuous trajectories, while higher values leads to the formation of straight axons. β represents the axon sensitivity to the external field (represented by an arrow). Values near zero indicate no perception of the field direction. Axons with high β grow straight to follow with high fidelity the field direction.

Thus, we can extrapolate the values of the parameters from the current scale to any other, including a smaller one (e.g. to have two or more data points instead of one, Fig. 4.5 C), to compute the equivalent models to a certain scale to any other.

This means that we can estimate the parameters from our data at the sampling scale of 1 μm (Fig. 4.5 A), and then simulate continuous neurons (*i.e.* scale as small as desired, e.g. 0.1 μm) without losing formalism nor data fidelity (Fig. 4.5 C). This is particularly interesting because if we wanted to estimate the parameters from our data directly from a smaller scale sampling, the results would loose fidelity. In particular, they would be biased by oversampling and data noise at the small scale. Thus, simulations with parameters estimated this way would not reproduce realistic, but much more noisy trajectories. This is the same situation as observed and analysed for the Markov chain model at the pixel level discussed in Section 3.4. Fig. 4.6 A shows the main axons in our wild-type database sampled with a step size of $1.44\mu m$. Fig. 4.6 B shows the same number of axons simulated using the model described above in this section, $\Delta\rho = 1.44\mu m$ and the parameters estimated from the trajectories in Fig. 4.6 A. As it can be seen, the shape of the trajectories is accurately reproduced by the model with this step size. However, with a x16 zoom the discrete steps are observable. On the other hand, Fig. 4.7 A is identical as Fig. 4.6 A, but with the trajectories sampled with $\Delta\rho = 0.09\mu m$, which corresponds to the pixel size of the images (Section 2.4). If we estimate the parameters from these trajectories and simulate using the same step size, the obtained axons are noisy and thus shorter in the μm scale (Fig. 4.7 B, top). This is the effect of oversampling. However, if we renormalise the parameters estimated from the trajectories in Fig. 4.6 A using the inverse functions of 4.7 and 4.8 four times ($\log_2 \frac{1.44}{0.09} \simeq 4$) and then simulate axons using these new values and a step size $\Delta\rho = 0.09\mu m$, the shape of the trajectories is accurately reproduced and the noise is avoided, generating smoother trajectories. In these 2D simulations, the model parameters α and β were estimated using the first approach, in Subsection 4.3.1, the number of steps that define the axon length are thrown for each axon from a Gaussian distribution with parameters estimated from data and the external guiding cues direction is constant and equal to the angle of the linear regression of all the points in data. Another advantage of the normalization is that it is possible to estimate parameters from images with different quality (i.e. different sampling scale) and still transform the parameters to compare them in an equivalent scale.

4.3 Parameter estimation

First, in Subsection 4.3.1 we present and compare two approaches to estimate the parameters α and β of Markov chains following Eq. 4.1 .

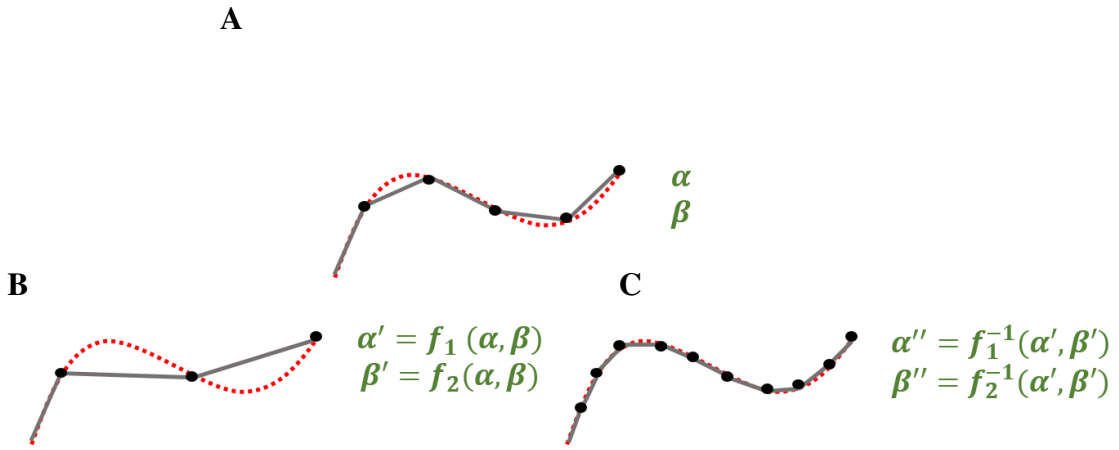

Fig. 4.5 Concept of model renormalization. (A) Extract of an axonal trajectory with a certain sampling and the correspondent values of α and β. (B) The same extract of axonal trajectory as in (A), but with one over two points missing. The parameters describing the model of these sequence of steps, α' and β' are a function of the parameters in (A). (C) The same extract of axonal trajectory as in (B), but with double the points. The parameters describing the model of these sequence of steps, α'' and β'' are also a function of the parameters in (A).

Next, in Section 4.3.2, we focus on the problem of estimating the parameters from a population of real axons, and review the different approaches we used throughout this work to estimate the parameters from our database. The task is not trivial due to the typical lengths of our axons, the intra-population variability and the branching process (as branches may be born with angles that do not follow the Markov model and then be included as part of the main axon, as shown in Fig. 4.40 C).

4.3.1 Parameter estimation of Markov chains

First approach

We first approach the problem using the definition of the variance σ_0^2 of the conditional probability of the model $P(\theta_i|\theta_{i-1})$ (with $\frac{\alpha}{\alpha+\beta}\theta_{i-1}$ being the mean of the conditional law). From Eq. 4.4 we know that $\sigma_0^2 = \frac{1}{2(\alpha+\beta)}$. Thus, for a chain of length M composed by the sequence $\theta_{i=1:M}$ we have the expression for the empirical variance estimation

$$\hat{\sigma^2} = \frac{1}{M}\sum_{i=1}^{M}(\theta_i - \frac{\alpha}{A}\theta_{i-1})^2; \tag{4.11}$$

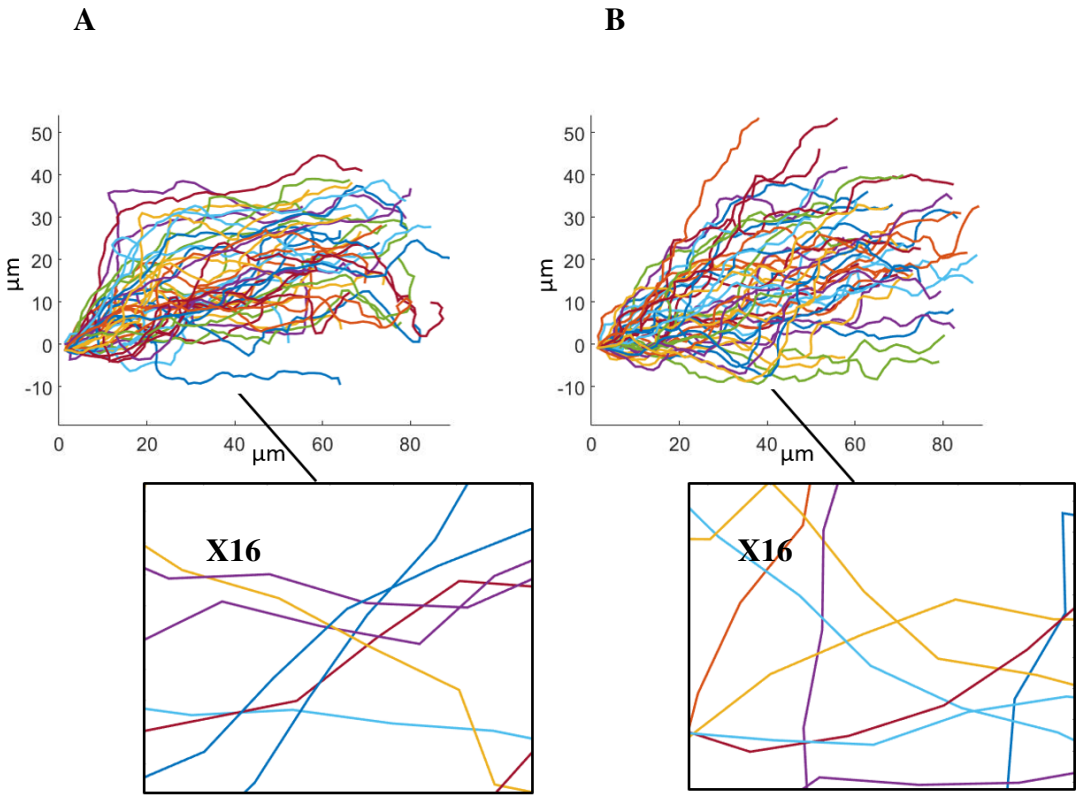

Fig. 4.6 Data reconstruction and simulation with $\Delta\rho = 1.44\mu m$. (A) Main axon trajectories from wild-type data, with a step size of $1.44\mu m$. The black-bordered square shows the image 16 times bigger. (A) Simulated axon trajectories with parameters estimated from data in (B) and with a step size of $1.44\mu m$. The black-bordered square shows the image 16 times bigger.

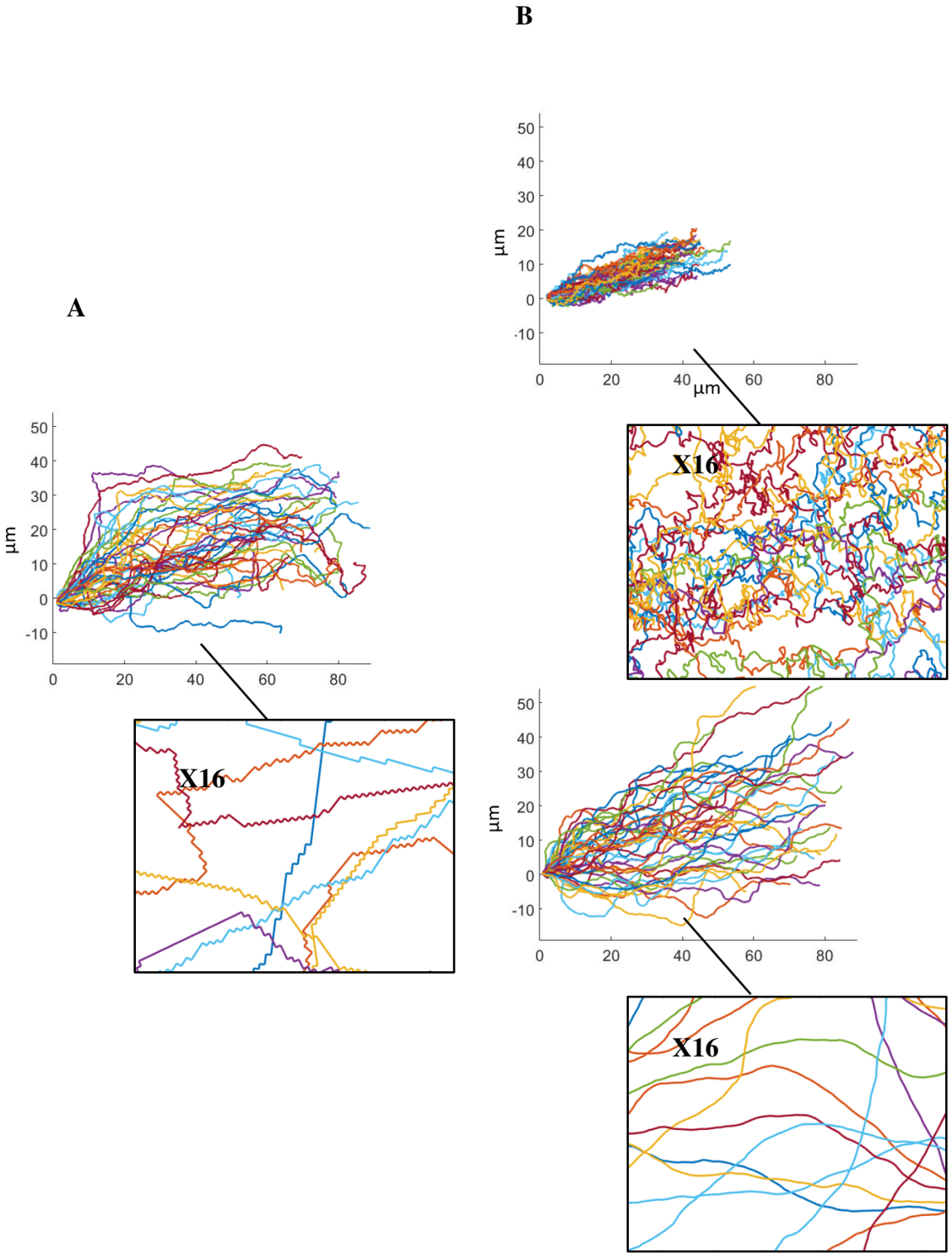

Fig. 4.7 Data reconstruction and simulation with $\Delta\rho = 0.09\mu m$ (pixel size). (A) Main axon trajectories from wild-type data, with a step size of $0.09\mu m$. The black-bordered square shows the image 16 times bigger. (B) (Top) Simulated axon trajectories with parameters estimated from data in (A) and with a step size of $0.09\mu m$. The black-bordered square shows the image 16 times bigger. (Bottom) Simulated axon trajectories with parameters estimated from data in Fig. 4.6 A renormalized 4 times, and with a step size of $0.09\mu m$. The black-bordered square shows the image 16 times bigger.

with

$$A = \alpha + \beta. \tag{4.12}$$

When $M \to \infty$, Eq.4.11 converges to

$$\hat{\sigma^2} = \frac{1}{M}\sum_{i=1}^{M}(\theta_i - \frac{\alpha}{A}\theta_{i-1})^2 \to \frac{1}{2A}. \tag{4.13}$$

Working out 4.11 and 4.13 we obtain

$$A = 2A^2 < \theta_i^2 > +2\alpha^2 < \theta_{i-1}^2 > -4\alpha A < \theta_i\theta_{i-1} >, \tag{4.14}$$

thus a quadratic equation $aA^2 + bA + C = 0$ with unknown A and coefficients

$$a = 2 < \theta_i^2 >; \tag{4.15}$$

$$b = -4\alpha < \theta_i\theta_{i-1} > -1; \tag{4.16}$$

$$c = 2\alpha^2 < \theta_{i-1}^2 > . \tag{4.17}$$

In the case that Eq. 4.14 has two real and different roots, only the solution adding the discriminant is plausible. Indeed, for simplicity, lets consider

$$< \theta_{i-1}^2 > \simeq < \theta_i^2 > = J \quad \text{and} \quad < \theta_i\theta_{i-1} > = L. \tag{4.18}$$

Then Eq. 4.14 has the solutions

$$A_{+,-} = \frac{4\alpha L + 1 \pm \sqrt{(4\alpha L+1)^2 - 16\alpha^2 J^2}}{4J}. \tag{4.19}$$

Both solutions are real and different if

$$\sqrt{(4\alpha L+1)^2 - 16\alpha^2 J^2} > 0 \tag{4.20}$$

from where

$$L = J - \frac{1}{4\alpha} + M \quad \text{with} \quad M \in \mathbb{R}^+. \tag{4.21}$$

In addition, we consider an attractive model ($\beta > 0$), thus from Eq. 4.12 the solutions must follow

$$A_{+,-} > \alpha \tag{4.22}$$

Thus, replacing L with Eq. 4.21, the solution A_- is only admissible if

$$4\alpha M > \sqrt{16\alpha^2 M^2 + 32\alpha^2 JM}. \tag{4.23}$$

That equals

$$32\alpha^2 JM < 0, \tag{4.24}$$

which is an absurd as both J and M are positive by definition, proving that only A_+ is an admissible solution.

Equation 4.14 allows to estimate A from a chain. However, we have only one equation and two unknowns. Thus, we add the equation equivalent to 4.14 (when $M \to \infty$), but θ_i conditioned by θ_{i-2} (see Eqs. 4.7 and 4.8)

$$\frac{1}{A'}\sum_{i=1}^{M}(\theta_i - \frac{\alpha'}{A'}\theta_{i-2})^2 \to \frac{1}{2A'} \tag{4.25}$$

with

$$A' = \alpha' + \beta'. \tag{4.26}$$

Assuming M big enough, Eq. 4.25 is equivalent to

$$A' = 2A'^2 < \theta_i^2 > +2\alpha'^2 < \theta_{i-1}^2 > -4\alpha' A' < \theta_i \theta_{i-1} > . \tag{4.27}$$

From Equations 4.7, 4.8 and 4.26 we have

$$A' = \frac{A^3}{A^2 + \alpha^2}. \tag{4.28}$$

The final system from 4.14, 4.27, 4.28, 4.7 and 4.8 results in

$$\begin{cases} A = 2A^2 < \theta_i^2 > +2\alpha^2 < \theta_{i-1}^2 > -4\alpha A < \theta_i \theta_{i-1} > \\ \frac{A^3}{A^2+\alpha^2} = 2(\frac{A^6}{(A^2+\alpha^2)^2}) < \theta_i^2 > +2(\frac{A^2\alpha^4}{(A^2+\alpha^2)^2}) < \theta_{i-1}^2 > -4(\frac{A^4\alpha^2}{(A^2+\alpha^2)^2}) < \theta_i \theta_{i-1} > . \end{cases} \tag{4.29}$$

This system is highly non-linear and does not have an analytic solution. Thus, we proposed an iterative approach described in Algorithm 1.

input : Arbitrary initial value α_1
output : Estimated values of $(\alpha, A = \alpha + \beta)$
while $|\alpha_{it} - \alpha_{it-1}| < \varepsilon \wedge |A_{it} - A_{it-1}| < \varepsilon$ **do**
    it=it+1;

    'STEP 1: '

    solve A_{it} using $(< \theta_i^2 >, < \theta_{i-1}^2 >, < \theta_i \theta_{i-1} >, \alpha_{it-1})$ from Eq. 4.14 ;
    solve α'_{it} using (A_{it}, α_{it-1}) from Eq. 4.7;

    'STEP 2: '

    solve A'_{it} using $(< \theta_i^2 >, < \theta_{i-2}^2 >, < \theta_i \theta_{i-2} >, \alpha'_{it})$ from Eq. 4.27 ;
    solve α_{it} using (A_{it}, A'_{it}) from Eq. 4.28;
end

Algorithm 1: Iterative method algorithm.

Second approach

While the previous approach used the definition of the variance of the conditional probability of the model P($\theta_i|\theta_{i-1}$), this approach uses the variance of θ_i. It can be shown (see A.2in Appendix A) that the variance of $\theta_{i=1:N}$ for a chain of length N is

$$\sigma_{\theta_i}^2 = \sigma_0^2 \sum_{i=1}^{N-1} \gamma^{2i} = \sigma_0^2 \frac{1-\gamma^{2N}}{1-\gamma^2} \tag{4.30}$$

where

$$\gamma = \frac{\alpha}{\alpha+\beta}; \sigma_0^2 = \frac{1}{2(\alpha+\beta)}. \tag{4.31}$$

When $i \to \infty$ we obtain the expression

$$\sigma_{\theta_\infty}^2 = \frac{\sigma_0^2}{1-\gamma^2}. \tag{4.32}$$

It can also be shown (see A.3 in Appendix A) that the variance of the difference $\theta_i - \theta_{i-1}$ (for $i \to \infty$) is

$$\sigma_{\Delta\theta_\infty^n}^2 = \frac{2\sigma_0^2(1-\gamma)}{1-\gamma^2}. \tag{4.33}$$

To obtain the estimations $\hat{\alpha}$ and $\hat{\beta}$, we assume that the chain is long enough, and calculate the empirical estimates $\hat{\sigma^2_{\theta_\infty^k}}$ and $\hat{\sigma^2_{\Delta\theta_\infty^k}}$, from where we get $\hat{\gamma}$ and $\hat{\sigma_0^2}$ to finally apply (4.34)

$$\hat{\alpha} = \frac{\hat{\gamma}}{2\hat{\sigma_0^2}}, \quad \hat{\beta} = \frac{1}{2\hat{\sigma_0^2}} - \hat{\alpha}. \tag{4.34}$$

and obtain the estimates of the model parameters.

Comparison

To analyse and compare the performance of the proposed approaches, we applied both methods to estimate the parameters of the same simulated chains, with varying lengths and parameter values.

Figs. 4.8 and 4.9 show the boxplots of the results of the estimations with varying α and β respectively, for 50 chains of lengths 10, 100 and 1000 and a heating period of 100 steps. As it can be observed, the second approach performs better, the main difference being the variance and the number of outliers, particularly for short chains.

Regarding the first approach, Algorithm 1 results robust respect to the initialization and converges within a large interval of α_0 values (see Fig. B.1 in Appendix B) in a short number of iterations. To illustrate its convergence, we estimated the parameters of 1000 simulated chains of length 100, with $\varepsilon = 0.00001$, $\alpha = \beta = 10$ and maximum number of iterations of 100. We observed that 81% of the trials terminated before the 12 iterations, 17.8% between 12 and 20 and 0.8% between 20 and 30. Only a 0.4% was stopped at 100 iterations due to lack of convergence (Fig. B.2 in Appendix B).

4.3.2 Parameter estimation from a population of real axons

We applied both methods to estimate the parameters of the chains that describe the axons from our wild-type database. Under the hypothesis that the axons of a same group behave the same, thus have the same parameter values, we first considered the M axons altogether (a single sequence $\theta_{i=1:N_n}^{n=1:M}$). However, we obtained parameter values of $\alpha = 0.023$ and $\beta = 0.021$ using the first method and $\alpha = 0.002$ and $\beta = 0.052$ using the second one. We know that these values do not represent the axon morphology. Fig. 4.10 shows 20 simulations of 100 steps of size $1\mu m$ using each pair of thee estimated values, where the trajectories are clearly noisier than axons in data.

To better understand these results, we estimated the parameters from each axon individually. Fig. 4.11 presents the obtained distributions. As it can be seen, the estimated values present large variations for both parameters, including values near zero. This could explain the low values obtained considering the axons altogether. In particular, while simulated chains with the same length distribution as real axons and parameters equal to the mean of

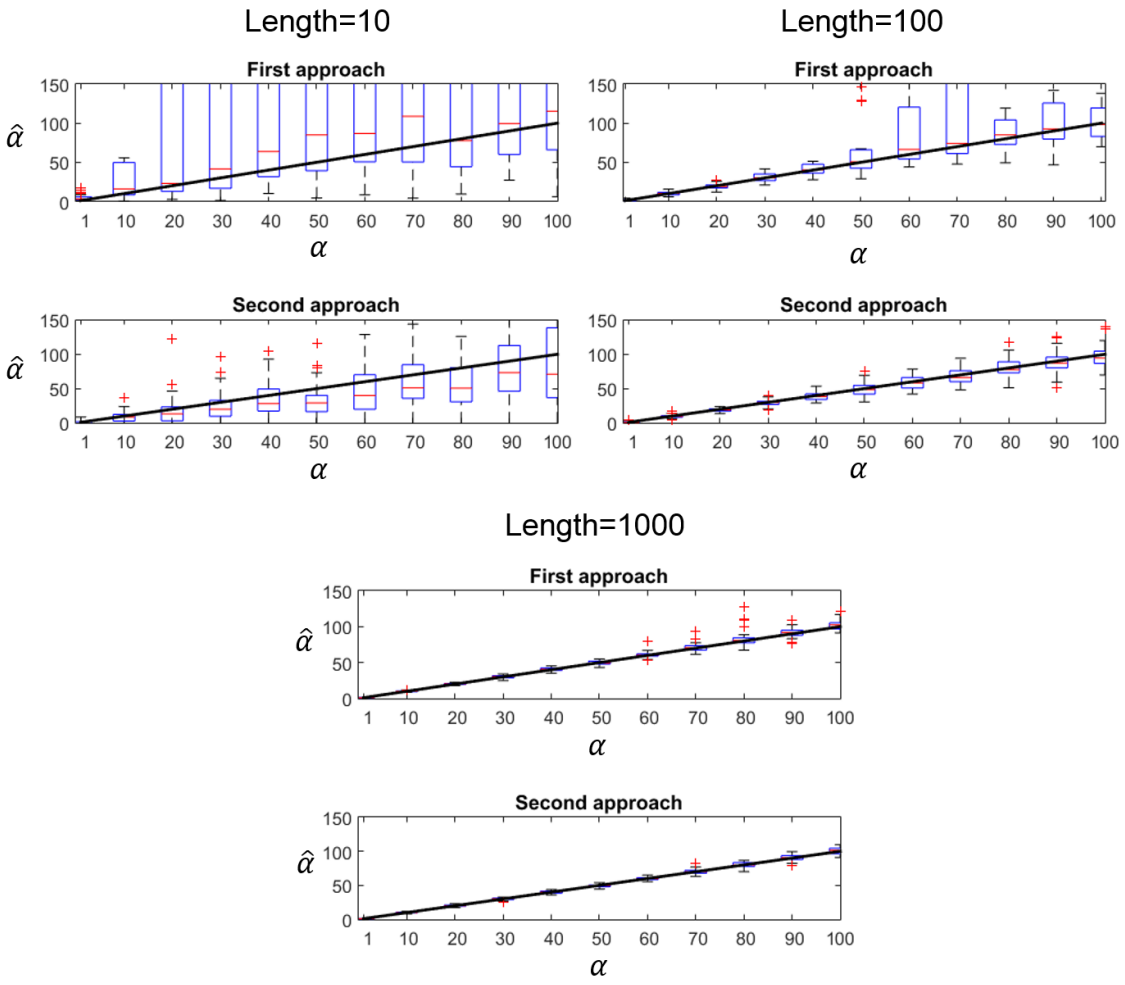

Fig. 4.8 Parameter estimation method performance (α). Distributions of the estimates of the parameter α ($\hat{\alpha}$), for 50 chains of lengths 10, 100 and 1000 and varying values of α. The central mark indicates the median, and the bottom and top edges of the box indicate the 25^{th} and 75^{th} percentiles, respectively. The whiskers extend to the most extreme data points not considered as outliers, and the outliers are plotted individually using the $+$ symbol. Black lines represent the line $\hat{\alpha} = \alpha$. For all cases, $\beta = 10$. For the first approach, $\varepsilon = 0.00001$, $\alpha_0 = 100$ and the maximum number of iterations was 100.

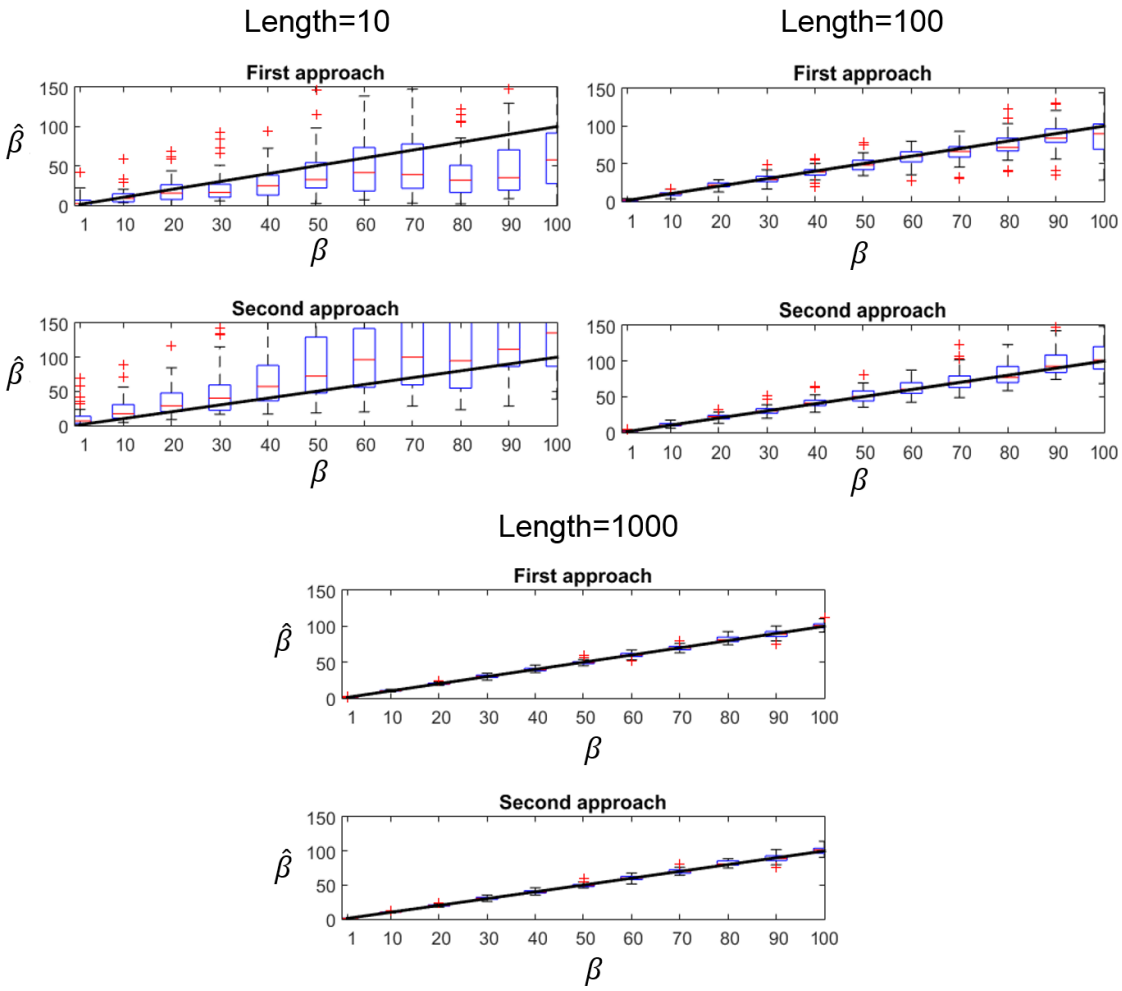

Fig. 4.9 Parameter estimation method performance (β). Distributions of the estimates of the parameter β ($\hat{\beta}$), for 50 chains of lengths 10, 100 and 1000 and varying values of β. The central mark indicates the median, and the bottom and top edges of the box indicate the 25^{th} and 75^{th} percentiles, respectively. The whiskers extend to the most extreme data points not considered as outliers, and the outliers are plotted individually using the $+$ symbol. Black lines represent the line $\hat{\beta} = \beta$. For all cases, $\alpha = 10$. For the first approach, $\varepsilon = 0.00001$, $\alpha_0 = 100$ and the maximum number of iterations was 100.

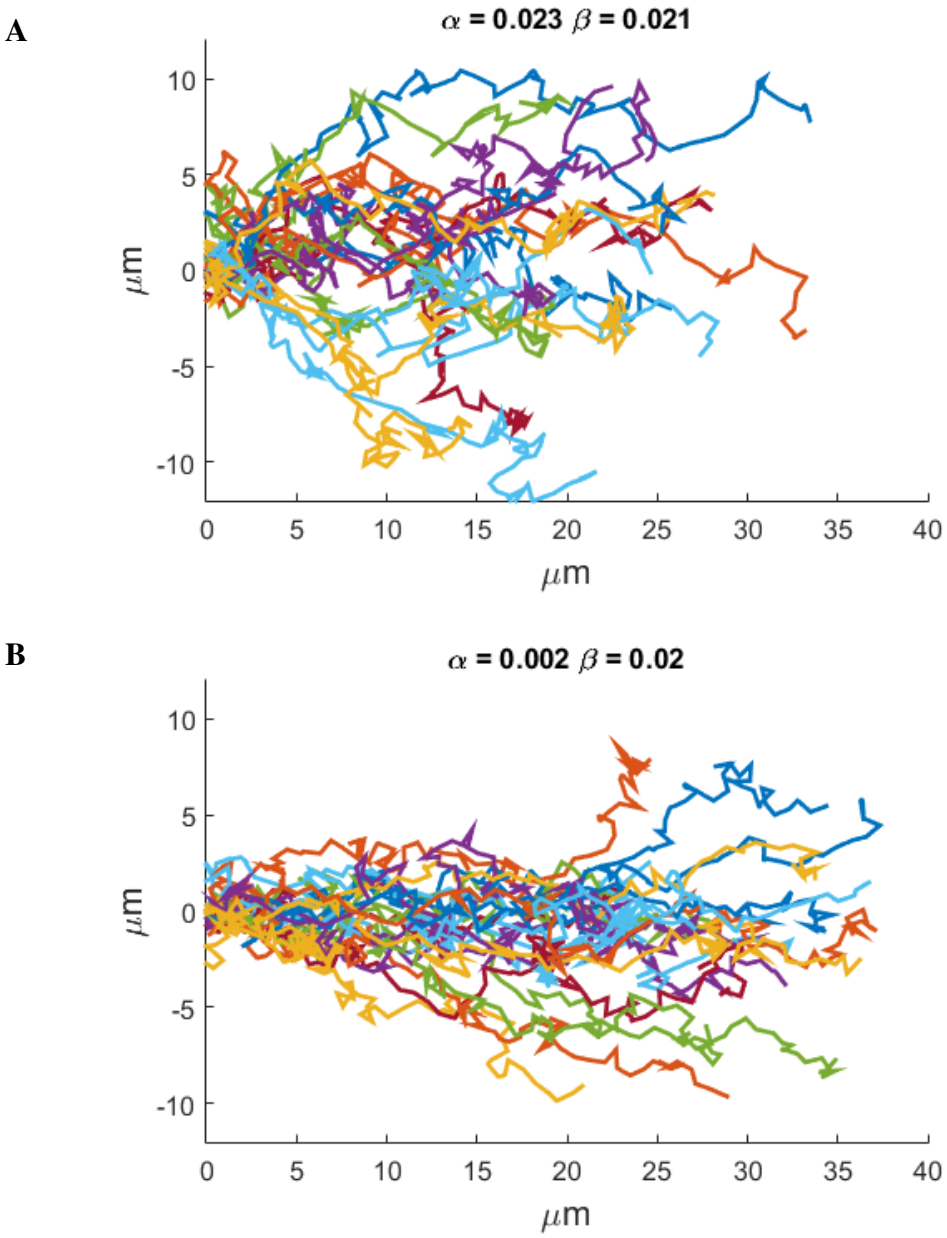

Fig. 4.10 Simulations using parameters estimated from data considering the axons altogether. 20 simulations of length 100. (A) α and β estimated using the first method. (B) α and β estimated using the second method.

the histograms in Fig. 4.11 presented standard deviations within the estimations of α and β equal to $(\sigma_\alpha = 1.0, \sigma_\beta = 0.7)$, the standard deviations of the estimates from data are much higher $(\sigma_\alpha = 3.0, \sigma_\beta = 1.9)$.

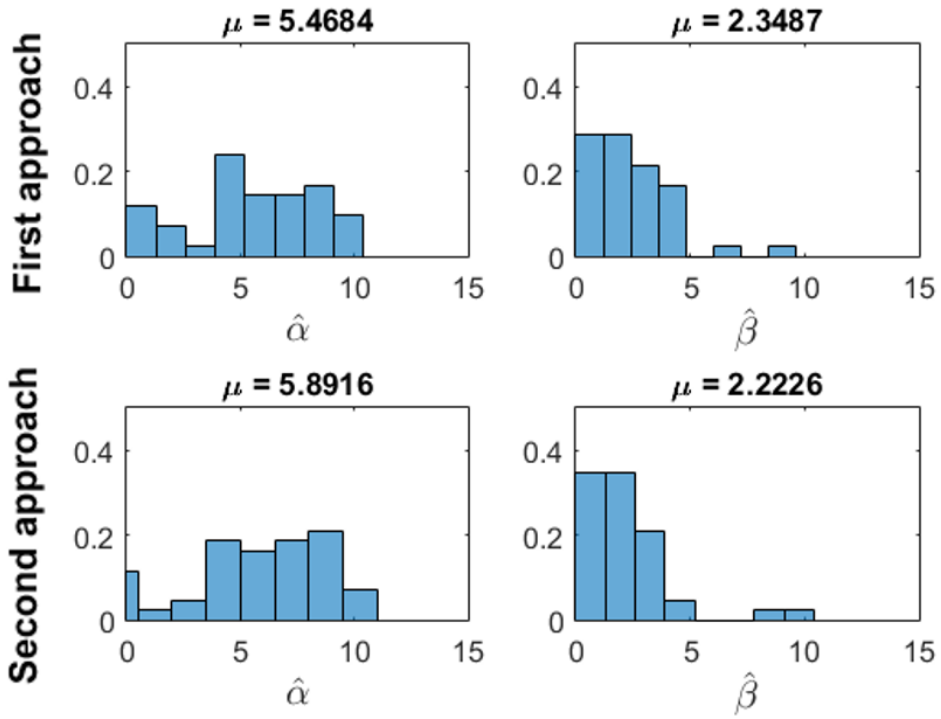

Fig. 4.11 Distributions of the values of α and β estimated axon per axon from real data. (top) Using the first method and (bottom) using the second one.

At a first glance, the shape of the obtained histograms could be explained in different ways. For example, it may mean that actually each axon has different parameter values. However, the typical axonal length is not enough to obtain a precise estimation individually. In addition, values of α and β near zero generate trajectories very different to those observed in reality (see Fig. 4.10). A second possibility could be that our estimation methods are not robust enough, which seems in contradiction with the previous study on their performance and the fact that both present very similar results when applied to data. A third possibility could be, unfortunately, that the proposed model is not suitable to describe real axons.

During the rest of this section, we study this problem and finally propose that the variability in the estimates actually emerges from the nature of the problem itself, which is the modelling of axons growing *in vivo*. Real axons do not grow isolated, but in constant interaction with their environments. Thus, they encounter obstacles during their development,

which may force them to change their trajectories, creating broken points in their otherwise Markovian paths. Next, we display the evidence that lead us to these ideas, and propose two different approaches to estimate the hidden parameters that govern the entire population of axons.

Regarding the two presented estimation methods, even though they both perform very similarly regarding the estimation of real axons, we have already shown that the second one presents less variance and outliers on the estimations from chains with lengths similar to those of our axons (~ 100). In addition, it is faster to apply (no iterations). Thus, it is the one chosen to use for the rest of the study.

Histogram study

Fig. 4.12 (top) shows the distributions of the estimated values of α and β axon per axon for the wild-type population. A first estimation of the global value of the parameters could be given by the mean values of the distributions: $\overline{\alpha} = 5.9$ and $\overline{\beta} = 2.2$. However, because these histograms do not seem to follow any known distribution, it is not trivial to decide which single value of α and β is the best one to describe the entire population. To find out the origin of this behaviour, we simulated $N = 45$ chains using our model (Eq. 4.1), with lengths thrown from a Gaussian distribution with parameters estimated from data. We then estimated the parameters (α, β) per simulated chain, to compare the obtained distributions with those from real data. As it can be seen from Fig. 4.12 (top + middle), the obtained parameter distributions do not particularly resemble those for real axons, as they are more Gaussian-like. However, it can be noticed that their variance is still relevant. This behaviour is inherent to the model and directly related to the axonal typical length (~ 100 steps). For longer simulated axons, this variance is dramatically reduced (see Fig. B.3 in Appendix B for comparison). The question is, therefore, which other aspect of our axons, apart from their length, contributes to the observed histograms (Fig. 4.12 (top))? One can notice that a relevant difference between the distributions for real and simulated axons is the presence of particularly low parameter values in the first case. This may mean that for some axons, the paths do not seem to accurately follow the Markov model. We hypothesize that this comes from the fact that we are working with real axons that grew up embedded in mechanically-constrained environments. We propose that during their growth they follow the Markov model with (α, β), unless some mechanical event occurs that forces them to behave otherwise. To test this idea, we simulated paths as before but including random jumps in the Markov chains such that at each step the model may be broken -i.e. next step is thrown from the distribution $\mathcal{N}(0, \frac{\pi}{2})$ - with probability $P = \frac{1}{N_n}$. This is consistent with the birth of a branch (that appears with a random angle, as we further study in Section 4.5). The distributions of the estimated

parameter values of the latter case are shown in 4.12 (bottom). These histograms do resemble those from real axons, presenting high variances and a flattened shape. This supports the hypothesis of the occurrence of random jumps in the chains may be true, and accurately describe our data. Figs. 4.13 and 4.14 show the same analysis, but using space-embedded simulated axons instead of real data. This means that axons grow altogether in a constraint environment and mechanically interact between them (this will be developed in detailed in Section 4.4) instead of real data, to further validate this hypothesis. Fig. 4.13 depicts space-embedded main axons vs. simulated classic chains, and the histograms result similar (but not similar to those from real axons). However, Fig. 4.14 shows the histograms for space-embedded simulated axons that form branches vs. classic chains and classic chains with breakdowns. These embedded simulated axons present histograms that look similar to those from data, as well as to those for simulated axons with breakdown. This suggests that the histogram shape from real axons may be due, mainly, to the existence of branches.

First approach: Histogram comparison

From now on, we thus consider our real axons as Gaussian Markov chains with parameters (α, β) plus random jumps. So, one way to estimate the values of α and β is to simulate populations with the same size and length distribution considering different parameter values and adding jump points. Re-estimate the parameters of these simulated chains, and observe the resulting histograms. Thus, the values of α and β that most likely represent our population are those for which their histograms resemble the most those from real data. We defined the estimation scheme as follows:

1) Estimate the parameters for each individual axon in the dataset, save the distributions of estimated parameters and consider an interval around the mean value of each parameter.

2) Simulate Markov Chains with the same sample size and length distribution as real data, and with parameters within the range from **1)**, and include random broken points (from the distribution $\mathcal{N}(0, \frac{\pi}{2})$, with probability $P = \frac{1}{N_n}$).

3) Estimate the parameters from the resulting simulations and save the distributions for each tested pair of parameters;

4) Finally, select the pair of parameters whose distribution is most similar to that obtained from data using the Kruskal Wallis test for the distribution comparison.

In order to accomplish step **4)**, we added the p values of each test for α and β, and obtained the 3D function $f(\alpha, \beta)$ shown in Fig. 4.15 A. We then approximated this function with a polynomial surface and chose the pair of parameters for which this surface reaches its maximum value (Fig. 4.15 B). The estimated parameters values for different populations in our data are presented in Table 4.1.

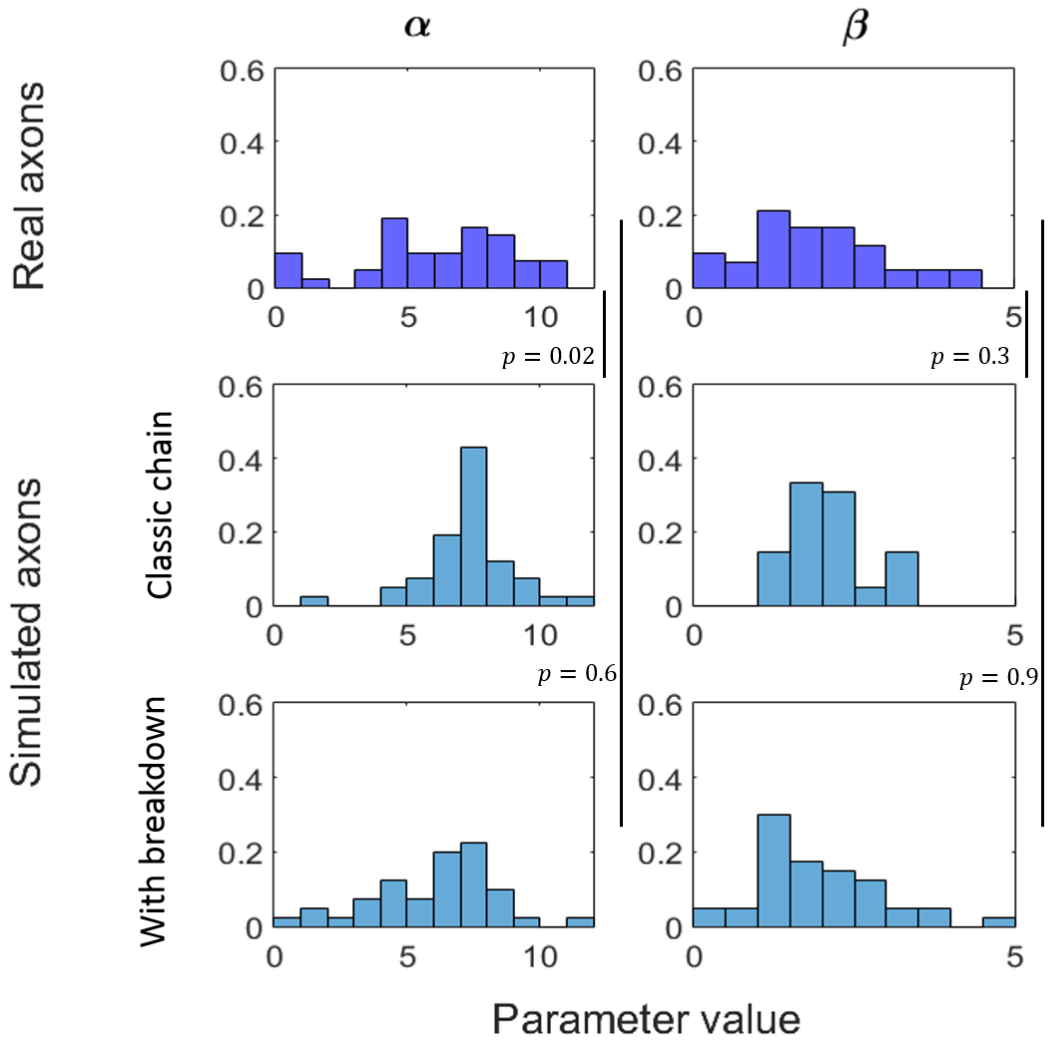

Fig. 4.12 Axon by axon parameter estimation - Real wild-type and simulated axons. (top) Histograms of the parameter values estimated axon per axon from the wild-type population in our database. (middle) Histograms of the parameter values estimated chain per chain from $N = 45$ simulated paths using $\alpha = 7.5$, $\beta = 1.7$, and lengths thrown from a Gaussian distribution with parameters estimated from data. (bottom) Histograms of the parameter values estimated chain per chain from $N = 45$ simulated paths using $\alpha = 7.5$, $\beta = 1.7$, and lengths thrown from a Gaussian distribution with parameters estimated from data. At each step the model may be broken -i.e. next step comes from a uniform the distribution $\mathcal{N}(0, \frac{\pi}{2})$ - with probability $P = \frac{1}{M_n}$. All the indicated p values come from the Kruskal Wallis test.

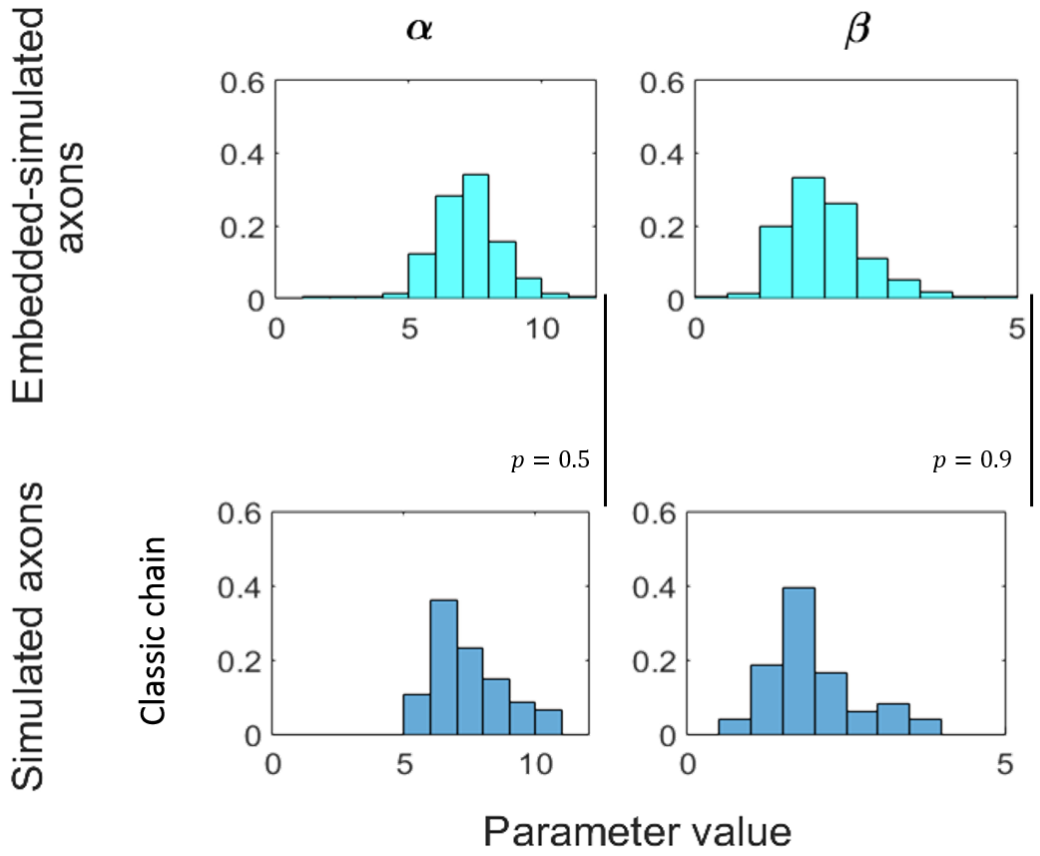

Fig. 4.13 Axon by axon parameter estimation - Embedded-simulated and simulated axons, no branching. (top) Histograms of the parameter values estimated axon per axon from N=45 axons from space-embedded simulations as described in Section 4.4. (bottom) Histograms of the parameter values estimated axon per axon from N=45 simulated chains using $\alpha = 7.5$, $\beta = 1.7$, and lengths thrown from a Gaussian distribution with parameters estimated from the space-embedded simulated axons. All the indicated p values come from the Kruskal Wallis test.

Table 4.1 Values of the parameters estimated for different groups in our database using the histogram comparison method.

	α	β
WT	7.5	1.7
Imp L	7.1	1.7
Imp Sh	6.0	1.6
Imp	6.7	1.6
Imp Rescue	7.4	1.6
Prof	5.1	3.7

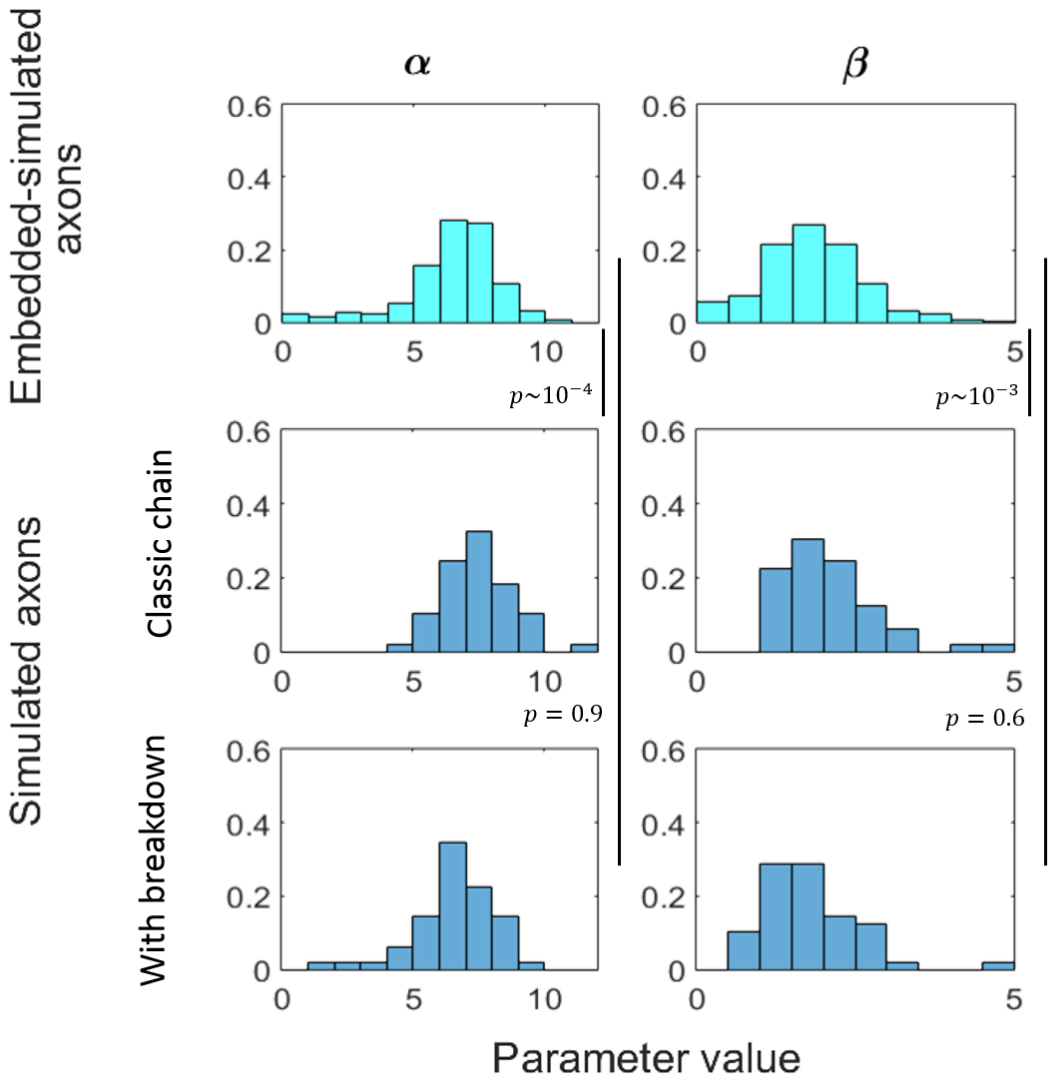

Fig. 4.14 Axon by axon parameter estimation - Embedded-simulated and simulated axons, with branches. (top) Histograms of the parameter values estimated axon per axon from N=45 axons from space-embedded simulations with mechanical-induced branches (further described in in Section 4.5.2). (middle) Histograms of the parameter values estimated axon per axon from N=45 simulated chains using $\alpha = 7.5$, $\beta = 1.7$, and lengths thrown from a Gaussian distribution with parameters estimated from the space-embedded simulated axons. (bottom) Histograms of the parameter values estimated axon per axon from N=45 simulated chains using $\alpha = 7.5$, $\beta = 1.7$, and lengths thrown from a Gaussian distribution with parameters estimated from the space-embedded simulated axons. At each step the model may be broken -i.e. next step comes from a uniform distribution- with probability $P = \frac{1}{M_n}$. All the indicated p values come from the Kruskal Wallis test.

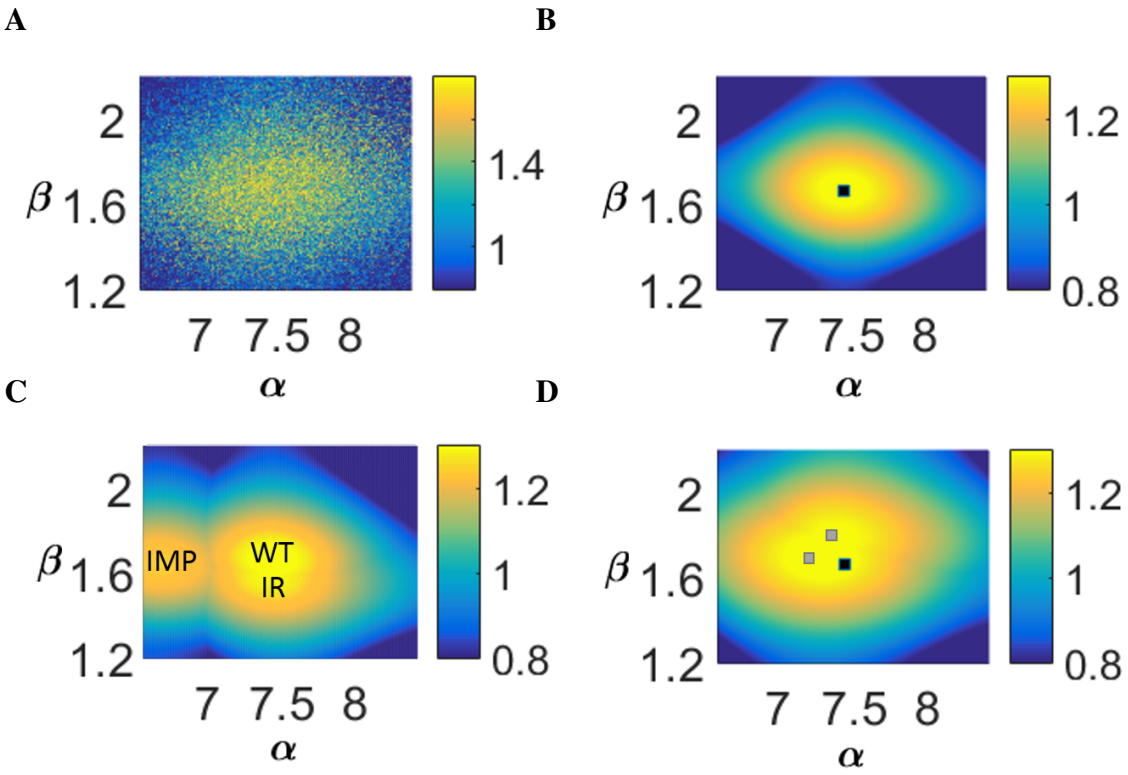

Fig. 4.15 Parameter estimation results using the histogram comparison method. (A) 3D function ($f(\alpha,\beta)$) representing the added p values comparing the distributions of α and β estimated from real data *vs*. estimated from simulations using different values of parameters. (B) Polynomial surface that approximates the function in (A). The black square indicates the maximum value. (C) Parameter estimation surfaces for different populations in data: WT, Imp and Imp Rescue. (D) Parameter estimation surfaces for two random halves of the real wild-type axons. The black square indicates the pair of parameters that maximizes the surface for the entire population (the same as in (B)), and the smaller grey ones the values that maximize the surfaces corresponding to each random half of the population.

Fig. 4.15 C depicts the superposition of the approximated polynomial surfaces for the wild-type (WT), *imp* knock-down (Imp) and *imp* knock-down rescued by Imp (Imp Rescue) populations in our data. An important advantage of this method is that it provides a way of quantitatively comparing and contrasting the parameter values for different populations. In the depicted case, we can easily appreciate that WT and Imp Rescue seem to have similar parameter values, while the Imp surface is clearly apart. This result is interesting as it suggests an effective rescue of the *imp* knock-down phenotype towards WT.

To assess the robustness of the method, we separated the WT population into two random halves and estimated the parameters separately. Fig. 4.15 D shows the superposition of the surfaces for both cases. We can appreciate that the estimated values are conserved, showing that the method is robust, as well as the morphological coherence within the WT population.

The results obtained with this method are the ones used for simulations through the rest of this work.

Second approach: Breakdown take-out

After the presented analysis, we found out that axons from data accurately follow the proposed Markov model with constant parameters, with the exception of some random positions in the chain where the model is broken and another law is followed instead (jump points). Here, we go further with this logic, and propose two approaches for the joint estimation of the Markov chain parameters and the jump point location. We first consider a Bayesian framework, embedding a Bernoulli prior on the jump points distribution. In the second approach we derive a statistical test on the likelihood ratio between the Markov chain model and an alternative one. In both cases, the aim is to detect the jump points and take them out, in order to estimate the parameters using Eq. 4.34 directly from the entire population, instead of axon per axon. As a reminder, when doing this directly we obtained very low parameter values (which do not represent the axonal path morphology, Fig. 4.10).

To detect the points that do not follow the Markov model described by Eq. 4.1 an easy way is to go through the axon and calculate, for each point i, its conditional likelihood regarding the model (L_m^i)

$$L_m^i = \sqrt{\frac{\alpha_m + \beta_m}{\pi}} e^{-(\alpha_m+\beta_m)[\theta_i - \frac{\alpha_m}{\alpha_m+\beta_m}\theta_{i-1}]^2} \tag{4.35}$$

and another null hypothesis (in case of jumps),

$$L_0^i = \sqrt{\frac{\alpha_0 + \beta_0}{\pi}} e^{-(\alpha_0+\beta_0)[\theta_i - \frac{\alpha_0}{\alpha_0+\beta_0}\theta_{i-1}]^2}; \tag{4.36}$$

with parameters (a_0, b_0). The parameters governing the jump hypothesis are assumed to be known, or directly obtained from the signal parameters: $a_0 = f(a_m, b_m)$, $b_0 = g(a_m, b_m)$. Here we will consider in particular two cases: $(a_0 = 0, b_0 = b_m)$ (*i.e.* loss of the Markov property) and $(a_0 = a_m, b_0 = 0)$ (loss of the external attraction).

Then, the i^{th} point is broken if $L_0^i > L_s^i$. However, this method is too simple and may over estimate the real number of broken points. A second obstacle is that one needs to know the values of (a_m, b_m) (in addition to the null hypothesis) to calculate the likelihoods, which are mostly generally unknown. Applying this analysis to the WT population with the null hypothesis given by $\alpha_0 = 0$ and $\beta_0 = \beta_m$ ($H_{0:\alpha_0=0,\beta_0=\beta_m}$), and $(a_m = 7.5, b_m = 1.7)$, we obtained that, in average, 12% of the points of each axon are broken. Moreover, if we do the same analysis using simulated chains instead of real axons -*i.e.* theoretically unbroken- the percentage is the same. This means that this method is too simple and effectively overestimates the number of broken points. Similar results are obtained if we choose the null hypothesis $H_{0:\alpha_0=\alpha_m,\beta_0=0}$ instead.

Bayesian Framework

We first propose a Bayesian approach, as follows

$$P(\theta_i, y_i|\theta_{i-1}) = P(\theta_i|\theta_{i-1}, y_i)P(y_i, \theta_{i-1}) = P(\theta_i|\theta_{i-1})P(y_i). \tag{4.37}$$

Where $y_i = 1$ when the chain is broken and 0 otherwise. We consider a Bernoulli prior such that $P(y_i = 1) = \lambda$ and $P(y_i = 0) = 1 - \lambda$. Therefore,

$$L_m^i = \sqrt{\frac{\alpha_m + \beta_m}{\pi}} e^{-(\alpha_m+\beta_m)[\theta_i - \frac{\alpha_m}{\alpha_m+\beta_m}\theta_{i-1}]^2}(1-\lambda) \tag{4.38}$$

and

$$L_0^i = \sqrt{\frac{\alpha_0 + \beta_0}{\pi}} e^{-(\alpha_0+\beta_0)[\theta_i - \frac{\alpha_0}{\alpha_0+\beta_0}\theta_{i-1}]^2}\lambda. \tag{4.39}$$

To estimate the values of $(\alpha_m, \beta_m, \lambda)$ from a data set, we propose Algorithm 2. It starts with arbitrary values of the parameters and a proposed set (α_0, β_0). It detects the broken points using the likelihood ratio, takes them out and re-estimates (α_m, β_m) until convergence, λ being estimated by the detected jump point frequency.

input : Arbitrary initial values $(\alpha_m, \beta_m, \lambda)$ and a jump model likelihood L_0, with parameters (α_0, β_0)
output : Estimated values of $(\alpha_m, \beta_m, \lambda)$
while $|a_m^{it} - a_m^{it-1}| < \varepsilon_1 \wedge |b_m^{it} - b_m^{it-1}| < \varepsilon_2$ **do**
  it=it+1;
  for $i = 2 : M$ **do**
  end
  Calculate L_0^i and L_m^i from eqs. 4.39 and 4.38;
  if $\frac{L_0^i}{L_m^i} > 1$ **then**
    $y_i = 1$;
  end
  if $\frac{L_0^i}{L_m^i} \leq 1$ **then**
    $y_i = 0$;
  end
end
Calculate $(\alpha_m^{it}, \beta_m^{it})$ using $\theta_i / y_i = 0$ from eq. 4.34. $\lambda^{it} = \sum \frac{y_i}{M}$;
Compute $\alpha_0^{it} = f(\alpha_m^{it}, \beta_m^{it})$ and $\beta_0^{it} = g(\alpha_m^{it}, \beta_m^{it})$;

Algorithm 2: Bayesian-based parameter estimation algorithm for a signal of length M.

Statistical Framework

In the second approach we design a statistical test of hypothesis based on the likelihood ratio. Given two Gaussian distributions $\mathcal{N}(\mu_m, \sigma_m^2)$ and $\mathcal{N}(\mu_0, \sigma_0^2)$ with variable θ_i, the likelihood ratio is given by

$$R(\theta_i) = \frac{L_0}{L_m} = \frac{\sigma_m}{\sigma_0} e^{-\frac{(\theta_i - \mu_0)^2}{2\sigma_0^2} + \frac{(\theta_i - \mu_m)^2}{2\sigma_m^2}}. \tag{4.40}$$

For each θ_i, the test considers that θ_i was preferably thrown from the alternative distribution $(H_0) : P(\theta_i|\theta_{i-1}) = P_0(\theta_i) = \mathcal{N}(\mu_0 = \frac{\alpha_0}{\alpha_0+\beta_0}\theta_{i-1}, \sigma_0 = \frac{1}{2(\alpha_0+\beta_0)})$ rather than from the model $(H_1) : P(\theta_i|\theta_{i-1}) = P_1(\theta_i) = \mathcal{N}(\mu_1 = \frac{\alpha_m}{\alpha_m+\beta_m}\theta_{i-1}, \sigma_1 = \frac{1}{2(\alpha_m+\beta_m)})$ if $R(\theta_i) > T, T \in \mathbb{R}_+$. The so-called p value of the test is thus given by the probability p that this occurs randomly under hypothesis P_1: $p = P(R(\theta_i) > T)$.

To calculate this probability, the inequality

$$R(\theta_i) > T \tag{4.41}$$

can be rewritten as a quadratic equation in θ_i:

$$[-\frac{(\theta_i-\mu_0)^2}{2\sigma_0^2}+\frac{(\theta_i-\mu_m)^2}{2\sigma_s^2}] > log(T\frac{\sigma_0}{\sigma_m}) \tag{4.42}$$

If (m,M) are the real roots of equation 4.42, its solution is $\theta_i \in (-\infty,m]\cup[M,\infty)$ or $\theta_i \in [m,M]$ depending of the sign of $-\frac{1}{2\sigma_0^2}+\frac{1}{2\sigma_m^2}$. Thus,

$$p=\int_{-\infty}^{m} P_1(\theta_i)d_{\theta_i}+\int_{M}^{\infty} P_1(\theta_i)d_{\theta_i} \text{ or } p=\int_{m}^{M} P_1(\theta_i)d_{\theta_i} \tag{4.43}$$

depending on the case. For a given set $(\alpha_m,\beta_m,\alpha_0,\beta_0,\theta_{i-1})$, T can thus be tabulated in function of the chosen value of p. Fig. 4.16 shows as an example the surface relating p with T and θ_{i-1} for the null hypothesis $H_{0:\alpha_0=0,\beta_0=\beta_m}$ and $H_{0:\alpha_0=\alpha_m,\beta_0=0}$ with $(\alpha_m=7.5,\beta_m=1.7)$.

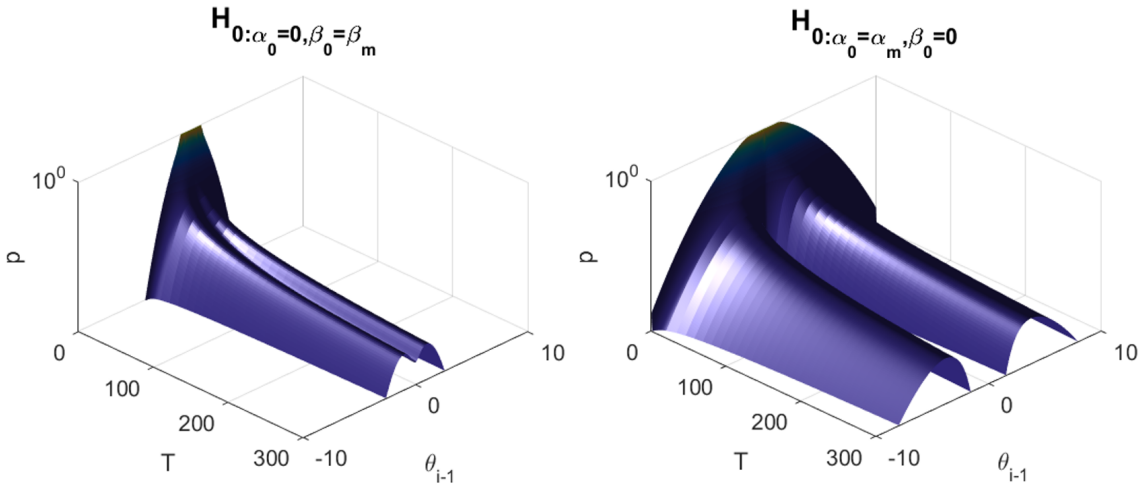

Fig. 4.16 Design of a statistical test of hypothesis. The surfaces represent the probability $p=P(R(\theta)>T)$ of acceptance of the null hypothesis H_0 under hypothesis H_1 in function of T and θ_{i-1} for two different null hypothesis: $H_{0:\alpha_0=0,\beta_m=\beta_m}$ and $H_{0:\alpha_0=\alpha_m,\beta_0=0}$ with $(\alpha_m=7.5,\beta_m=1.7)$.

To estimate the values of (α_m,β_m) from a data set, we consider Algorithm 3. It starts with arbitrary values of the parameters (α_m,β_m) and a chosen p value. It detects the broken points using Eq. 4.41, takes them out and re-estimates (α_m,β_m) until convergence.

input : Arbitrary initial values (α_m, β_m), an alternative jump model likelihood H_0 with parameters (α_0, β_0) and p
output : Estimated values of (α_m, β_m)
while $|\alpha_m^{it} - \alpha_m^{it-1}| < \varepsilon_1 \wedge |\beta_m^{it} - \beta_m^{it-1}| < \varepsilon_2$ **do**
    it=it+1;
    for $i = 1 : M^k$ **do**
        Calculate $R(x_i)$ using eq. 4.40;
        Obtain T in function of p solving eq. 4.43;
        if $R(x_i) > T$ **then**
            $y_i = 1$;
        end
        if $R(x_i) \leq T$ **then**
            $y_i = 0$;
        end
    end
    Calculate $(\alpha_m^{it}, \beta_m^{it})$ using $\theta_i / y_i = 0$ from eq. (4.34).;
    Compute $\alpha_0^{it} = f(\alpha_m^{it}, \beta_m^{it})$ and $\beta_0^{it} = g(\alpha_m^{it}, \beta_m^{it})$;
end

Algorithm 3: Statistic test of hypothesis-based parameter estimation algorithm a signal of length M.

Validation

We tested Algorithms 2 and 3 with chains of length 1000 and generated with parameters $(\alpha_m = 8, \beta_m = 2)$. The algorithms were set with initial parameter values of $(\alpha_m^1 = 20, \beta_m^1 = 20$ and $\lambda^1 = 0.1)$, and different jump point probabilities P_j (from $P_j = 0.005$ to 0.11). For every case, we used $\varepsilon_1 = \varepsilon_2 = 0.0001$.

We first analyze the jump point detection. Fig.4.17 A shows a single chain with $P_j = 0.02$ and jump points thrown from the conditional law (H_0) with parameters $(\alpha_0 = 0, \beta_0 = 0.125)$. The bayesian method detected 89% of the jump points with no false positives. The statistical method (with $p = 0.001$) detected 90% of the jumps from which 15% were false positives.

To test the improvement of the parameter estimates using both methods and compared with the direct estimation (without taking out the jump points, thus applying eq. 4.34 directly) we applied the algorithms to 100 chains, with jump points thrown from the conditional law (H_0) with parameters $(\alpha_0 = 0, \beta_0 = \beta_m)$ and Algorithm 3 with $p = 0.001$. Fig. 4.17 B shows the mean errors in the parameters estimates for the algorithms and with the direct estimation.

Both methods reduce the error in the estimates. Notably, the statistical method performs slightly better than the bayesian one.

Fig. 4.17 C shows as an example one of the previously simulated chains (from Fig. 4.17 B, with $P_j = 0.05$) highlighting the good detections, false positives and undetected jumps. As it can be observed, in this case the jump points are harder to detect than the case presented in Fig.4.17 A, and the methods detect only about 20% of the jump points. However, this amount results sufficient to reduce the error in the parameter estimates (Fig. 4.17 B).

Both algorithms are robust regarding their number of iterations needed for convergence and initial values. Fig. 4.18 shows the errors of the estimates (as in Fig. 4.17 B) considering 100 chains of length 1000, parameters $(\alpha_m = 8, \beta_m = 2)$, jumps following the conditional law (H_0) with parameters $(\alpha_0 = 0, b_0 = b_m)$ and with probability $P_j = 0.05$; in function of varying initial values α_m^1, β_m^1 and λ_m^1. As it can be observed, the method results are invariant to the large variation of initial values. In every case, the algorithms stopped within iterations number 3 and 14 despite the small value of $\varepsilon_1 = \varepsilon_2 = 0.0001$. Regarding computing time, the statistical method results longer as it calculates T for every i^{th} point which involves the numeric solution of the integral in eq. 4.43.

Application to data analysis

Further, we applied both presented approaches to estimate the parameters from different populations in our database. Tables 4.2 and 4.6 show the results from the bayesian method considering $H_{0:\alpha_0=0,\beta_0=\beta_m}$ and $H_{0:\alpha_0=\alpha_m,\beta_0=0}$ respectively (initial values $(\alpha_m = \beta_m = 20, \lambda = 0.1)$). Tables 4.3 and 4.7 show the results from the statistical method considering $H_{0:\alpha_0=0,\beta_0=\beta_m}$ and $H_{0:\alpha_0=\alpha_m,\beta_0=0}$ respectively and $p = 0.001$ (initial values $(\alpha_m = \beta_m = 20)$). Tables 4.4 and 4.8 are equivalent but considering $p = 0.0001$. Finally, Tables 4.5 and 4.9 are the same, with $p = 0.01$. For the null hypothesis $H_{0:\alpha_0=0,\beta_0=\beta_m}$, the bayesian method and the statistical one with $p = 0.001$ and $p = 0.0001$ present similar results, both in terms of parameter estimates and P_j, with around 2% of detected broken points. Naturally, the statistical method with $p = 0.01$ detects more broken points ($\sim 5\%$), thus the parameters estimates are higher. Interestingly, with the null hypothesis $H_{0:\alpha_0=\alpha_m,\beta_0=0}$, the bayesian approach (as well as the statistical one with $p = 0.0001$) detects less jump points ($\sim 0.5\%$) than the statistical one with $p = 0.001$ and $p = 0.01$, and obtains estimates very close to those in Table 4.1 with the histogram comparison method. This makes sense as for that method, we assumed that each point of an axon length N_n had $P_j = \frac{1}{N_n}$ of being broken, which for the WT axons means $P_j = 0.01$, and is close to that obtained in Table 4.6 for the bayesian

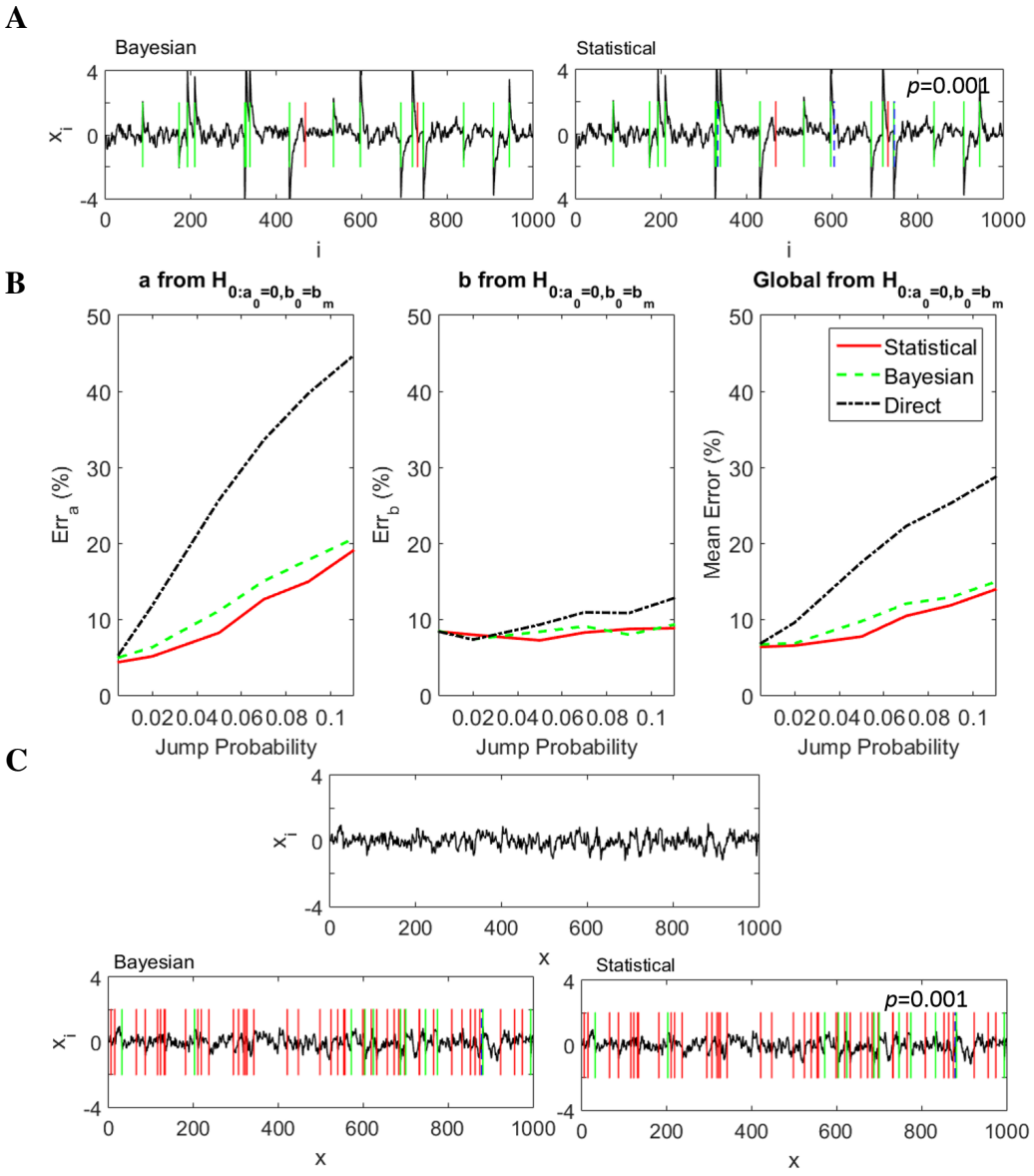

Fig. 4.17 Validation of the bayesian and statistical approaches from chains of length 1000 generated with parameters $(\alpha_m = 8, \beta_m = 2)$. (A) Jump point detection of a chain where the jump points thrown from the conditional law (H_0) with parameters $(\alpha_0 = 0, \beta_0 = 0.125)$. Non-detected jumps (red), detected jumps (green) and false positives (blue) are shown for the bayesian and statistical methods. (B) Mean error in function of the jump probability P_j in the parameter α (left), β (centre) and their average (right) regarding the two proposed algorithms and the direct method (directly applying eq. 4.34). 100 chains are considered and the jumps follow (H_0) with parameters $(\alpha_0 = 0, \beta_0 = \beta_m)$. (C) Jump point detection of a chain generated as in (B) with $P_j = 0.05$. The colour code for the jump detection is the same as in (A).

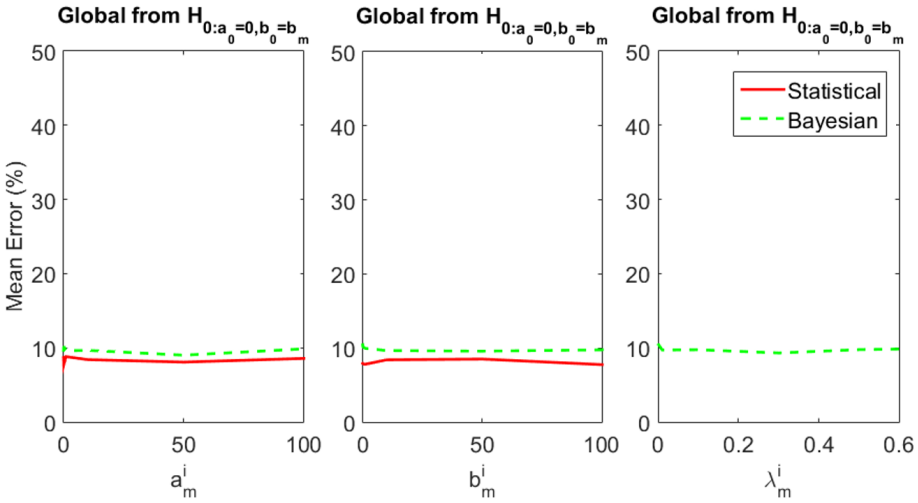

Fig. 4.18 Independence on initial values. Mean error in function of varying initial values α_m^1 (left), β_m^1 (centre) and λ_m^1 (right) for both algorithms. In each case, 100 chains of length 1000, parameters $(\alpha_m = 8, \beta_m = 2)$, jumps following (H_0) with parameters $(\alpha_0 = 0, \beta_0 = \beta_m)$ and probability $P_j = 0.05$ were considered. When not varied, $\alpha_m^1 = 20$, $\beta_m^1 = 20$ and $\lambda_m^1 = 0.5$.

Table 4.2 Values of the parameters and jump probability P_j estimated for different groups in our database using the breakdown take-out method, with $H_{0:\alpha_0=0,\beta_0=\beta_m}$ and initial values $(\alpha_m = \beta_m = 20, \lambda = 0.1)$. Bayesian approach.

	α	β	$\mathbf{P_j}$
WT	9.30	1.92	0.023
Imp	9.33	1.84	0.029
Imp Rescue	8.44	1.75	0.013
Prof	4.40	2.30	0.004

method. Another observation is that, in general, for the jump hypothesis $H_{0:\alpha_0=\alpha_m,\beta_0=0}$ less jump points are detected than with $H_{0:\alpha_0=0,\beta_0=\beta_m}$.

To better understand the exact effects that the jump take-out algorithms have on the axonal population, Fig. 4.19 shows the distributions of the parameters from each of 50 individual simulated paths of length 100 and $(\alpha_m = 8, \beta_m = 2)$ for two different cases: unbroken and broken with $P_j = 0.05$ and $H_{0:\alpha_0=0,\beta_0=\beta_m}$; plus the WT axons from data. We interpose the histograms after having applied Algorithm 3 with $p = 0.001$. As it can be seen, in both simulated and real data the algorithms *move* the histograms towards higher values, consequence of taking out the jump points that naturally lower the parameter estimates. The resulting histograms resemble more those for unbroken chains. This analysis finally unveils

Table 4.3 Values of the parameters and jump probability P_j estimated for different groups in our database using the breakdown take-out method, with $H_{0:\alpha_0=0,\beta_0=\beta_m}$ and initial values $(\alpha_m = \beta_m = 20)$. Statistical approach - $p = 0.001$.

	α	β	$\mathbf{P_j}$
WT	9.24	1.90	0.022
Imp	9.05	1.82	0.026
Imp Rescue	8.95	1.77	0.018
Prof	5.39	2.42	0.010

Table 4.4 Values of the parameters and jump probability P_j estimated for different groups in our database using the breakdown take-out method, with $H_{0:\alpha_0=0,\beta_0=\beta_m}$ and initial values $(\alpha_m = \beta_m = 20)$. Statistical approach - $p = 0.0001$.

	α	β	$\mathbf{P_j}$
WT	9.0	1.90	0.020
Imp	8.70	1.80	0.022
Imp Rescue	8.44	1.75	0.013
Prof	4.59	2.34	0.005

Table 4.5 Values of the parameters and jump probability P_j estimated for different groups in our database using the breakdown take-out method, with $H_{0:\alpha_0=0,\beta_0=\beta_m}$ and initial values $(\alpha_m = \beta_m = 20)$. Statistical approach - $p = 0.01$.

	α	β	$\mathbf{P_j}$
WT	12.53	2.05	0.058
Imp	12.88	1.97	0.068
Imp Rescue	12.46	1.91	0.058
Prof	8.02	3.08	0.052

Table 4.6 Values of the parameters and jump probability P_j estimated for different groups in our database using the breakdown take-out method, with $H_{0:\alpha_0=\alpha_m,\beta_0=0}$ and initial values $(\alpha_m = \beta_m = 20, \lambda = 0.1)$. Bayesian approach

	α	β	$\mathbf{P_j}$
WT	7.20	1.79	0.006
Imp	6.57	1.67	0.005
Imp Rescue	7.51	1.68	0.004
Prof	4.26	2.28	0.003

Table 4.7 Values of the parameters and jump probability P_j estimated for different groups in our database using the breakdown take-out method, with $H_{0:\alpha_0=\alpha_m,\beta_0=0}$ and initial values ($\alpha_m = \beta_m = 20$). Statistical approach - $p = 0.001$.

	α	β	$\mathbf{P_j}$
WT	8.42	1.98	0.014
Imp	8.50	2.0	0.020
Imp Rescue	8.50	1.89	0.013
Prof	5.51	2.84	0.020

Table 4.8 Values of the parameters and jump probability P_j estimated for different groups in our database using the breakdown take-out method, with $H_{0:\alpha_0=\alpha_m,\beta_0=0}$ and initial values ($\alpha_m = \beta_m = 20$). Statistical approach - $p = 0.0001$.

	α	β	$\mathbf{P_j}$
WT	7.95	1.87	0.010
Imp	7.77	1.85	0.013
Imp Rescue	8.04	1.77	0.008
Prof	4.47	2.42	0.006

Table 4.9 Values of the parameters and jump probability P_j estimated for different groups in our database using the breakdown take-out method, with $H_{0:\alpha_0=\alpha_m,\beta_0=0}$ and initial values ($\alpha_m = \beta_m = 20$). Statistical approach - $p = 0.01$.

	α	β	$\mathbf{P_j}$
WT	9.55	2.32	0.030
Imp	9.76	2.43	0.040
Imp Rescue	9.85	2.23	0.035
Prof	7.01	4.10	0.053

the origin of the dispersion of the estimates from real data, both in a mathematical and in a mechanistic point of view.

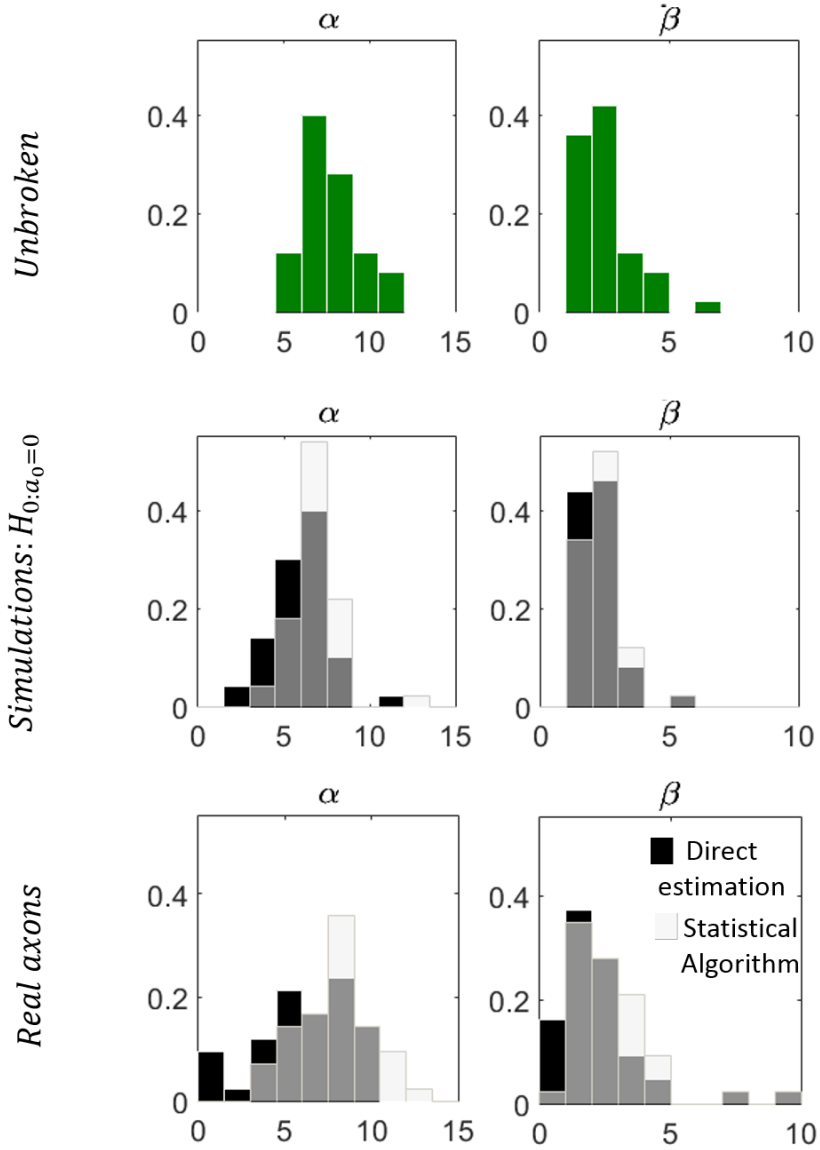

Fig. 4.19 Effect of the breakdown take-out algorithm on the parameter histograms. Distributions of the parameters from each of 50 individual simulated paths of length 100 and $(\alpha_m = 8, \beta_m = 2)$ for two different cases: unbroken (top) and broken with $P_j = 0.05$ and $H_{0:\alpha_0=0,\beta_0=\beta_m}$ (middle); plus the WT axons from data (bottom). For the broken simulations and real data, the histograms obtained directly from Eq. 4.34 are in black, and after having applied the statistical algorithm with $p = 0.001$ are in light grey. Dark grey areas correspond to intersections.

This approach has the advantage of being versatile, as the user can experiment with different null hypothesis representing different physical phenomena, and with both frameworks and different p values. One must notice that in general when more jump points are detected, the estimates are higher. This is not problematic nor a nonsense, but just represent different ways of reading data and allows different interpretations. It is important that, when simulating, the same hypothesis is followed (*i.e.* include jump points with the same law and probability of occurrence than when estimated) to respect the overall energy (variance) of the system.

Finally, this framework allows to detect the exact places where the axons present jumps, which can be interpreted axon per axon (Fig. 4.20), as well as globally in the case of a registered population as we have (see Fig. 4.68A in Section 4.8). Fig. 4.20 presents some axons from our database and their jump points detected with the statistic method with $p = 0.001$ and null hypothesis $H_{0:\alpha_0=0,\beta_0=\beta_m}$. As it can be seen, they are generally located on branching points ($\frac{1}{2}$ axons present jump points on or very close -one step away- to branching points), or on regions where the axons turn sharply or go backwards. The statistical approach, as well as the bayesian with a Bernoulli prior, do not take into account the density of broken points (or the distance between two continuous jump points), and may assign several neighbour points as jump points, as it can be seen in some examples of Fig. 4.20. This is of interest to detect every point that does not follow the Markov model, which indicates in our application if the axon encountered a single constraint or several consecutive ones (for example indicating a very dense region in space). However, one may consider a prior following a Poisson distribution or an accumulation process to impose constraints in the distance between jump points.

4.4 Space-embedded simulations

We have developed a mathematical model to generate individual axon paths and estimated its intrinsic parameters from real data. In this section we model a complete γ neuron population growing inside the ML during remodelling. But before, we need to define how we simulate the dynamic aspects, as well as the medial lobe shape, external guiding cue field and volume exclusion.

4.4.1 Dynamic simulation framework

In vivo, adult γ axons grow as a group, in a confined environment defined by surrounding neuronal and glial cells (Fig. 4.22 and Aso et al. (2014)). Thus, to take into account the

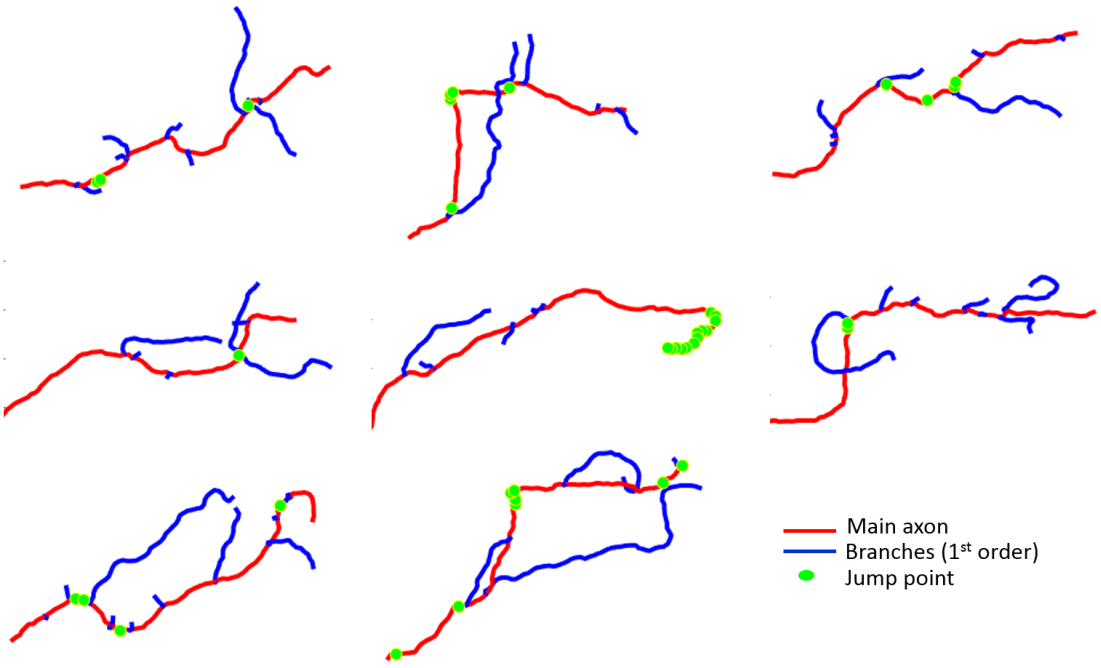

Fig. 4.20 Jump point detection on real axons. Examples of real axons (red) and first order branches (blue), where the jump points (green) are highlighted. Jump points were detected with the statistic method with $p = 0.001$ and null hypothesis $H_{0:\alpha_0=0,\beta_0=\beta_m}$.

mechanical constraints that underlie axon growth in the crowded environment of a maturing brain, we implemented constraints imposed by neighbouring cells in our Markov model. At each time-point t_j, every neurite is given the possibility of growing the same number of steps, n_{max}, following the Markov model for their spatial orientation. If no mechanical constraint is encountered, axon tips grow with a maximal speed of $v_{max} = \frac{n_{max}\Delta\rho}{\Delta t}$ during t_j ($\Delta\rho$ being the step size). If, on the contrary, axon tips encounter another γ neuron from the same population, or the medial lobe geometrical limits before accomplishing n_{max} steps (Fig. 4.21 II), they retract the last few steps (n_R: retraction rate) realized during t_j (Fig. 4.21 III), similarly to other growth and repulsion models (Simpson et al., 2009; Torben-Nielsen and De Schutter, 2014; Vanherpe et al., 2016). Then, they try to regrow in another direction until n_{max} steps are accomplished (Fig. 4.21, case A). However, if a second obstacle is encountered within the same time frame t_j (Fig. 4.21, case B), the axon tip stops its growth after retraction, and tries again only during the next time point (t_{j+1}). For the sake of computational time, as well as to mimic real timing constraints imposed by the developmental program, each axon tip has a limited global number of trials (after mechanical interactions) before it stops growing (defined by the *counter*). Axons reaching the extremity of the medial lobe before this limit will stop growing. The *counter* is augmented by 2 at each time the tip fails to elongate n_{max} steps (representing the encounter of two consecutive mechanical obstacles, CASE B in Fig. 4.21). The proposed growth and repulsion mechanism follows observations done in the consulted bibliography. Francisco et al. (2007) observed that neurons growing with obstacles sample repeatedly the encountered constraint, until they found a free way. Similar observations were done by Sagasti et al. (2005), who also mention contact with other neurites during development and the so-called growth and repulsion mechanism. Regarding the growth stop conditions, how to exactly stop the growth of a neurite is still under discussion in recent modelling bibliography. For example, Zubler et al. (2013)and Zubler and Douglas (2009) reduce the diameter of the new branches until a critical value is trespassed, which indicates the neurite to stop. However, *Drosophila* axon diameter has not been described to vary from the soma towards the axonal terminations. More recently, Vanherpe et al. (2016) also associates a *counter* to each neurite in a similar way as we do.

Axons start developing at the same time, and continue growing if i) they have not reached the end of the medial lobe (Fig. 4.1), ii) their *counter* is smaller or equal to a fixed maximum value, and, iii) no other branch from the same neuron has reached the extremity of the medial lobe (type I, when branches are considered, see Section 4.5.2). The value of the *counter* is incremented by two (accounting for the two encountered mechanical obstacles, steps II and IV B in Fig. 4.21) at each time point t_j the axon fails to elongate n_{max} steps (case B in Fig. 4.21). Axons finding too many mechanical obstacles along their way will thus reach the

maximum *counter* value before reaching the lobe extremity. The simulation is completed when there is no growing axon left. At each time-point t_j, axons grow sequentially to avoid overlapping (which is computationally difficult to solve if all next positions are decided simultaneously).

Unless the contrary is specified, for every simulation presented in this work $n_{max} = 6$, and the *counter* equals 140. Furthermore, we defined n_R as the rate of retraction after encountering an obstacle. Upon contact, axons not only loose this particular step, but also the n_R steps before, provided the neurite has already elongated more than n_R steps during t_j. n_R equals 2 in our experiments, unless the contrary is specified. Simulations showed that our results are not significantly sensible to variations in n_{max} nor n_R (see Fig. B.4 in Appendix B).

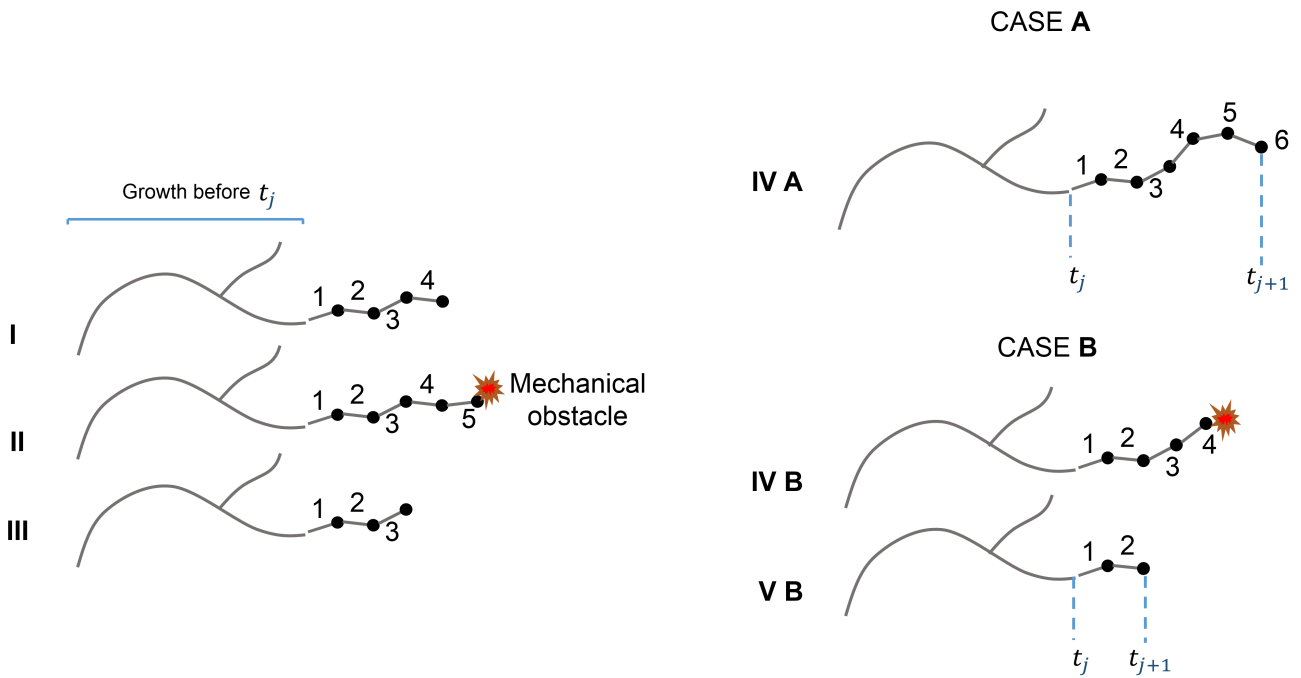

Fig. 4.21 Simulation of γ axon growth in the context of a population of physically interacting axons. Step by step axon growth scheme during t_j considering physical interactions. The axon tries to elongate n_{max} steps (I), until it encounters a mechanical constraint (II) and retracts (III). During the same time point (t_j), it will re-try to reach n_{max} steps. If no other mechanical constraint is encountered (CASE A), the axon will advance to reach n_{max} steps. If, on the contrary, another mechanical constraint is encountered (CASE B), the axon will retract again and stop its growth until the next time point (tj+1). The total growth during t_j is then smaller than n_{max} steps.

4.4.2 Medial lobe morphology reconstruction

One of the main constraints under study in this thesis is spatial limitation. The axons cannot grow *wherever* the model indicates, but must stay within the medial lobe. We thus designed an artificial lobe with hard limits that the axons cannot trespass. To do so, we used the lobes in the second channels of the confocal images in our data (Section 2.4), as well as the MB description in Aso et al. (2014). Fig. 4.22A depicts the limits of the ML considered as mechanical constraints in the simulations (blue lines). Here, a simplified geometry of the ML is represented. The purple line is a transversal cut of the peduncle placed just before the entrance to the ML, from where the simulated axons emerge. As mentioned in Section 2.3, the γ axon shafts elongate rigidly in a tight and ordered parallel bundle along the peduncle before entering the ML, where they spread and develop (see Section 2.3 and Fig. 2.3). To reproduce this behaviour, the simulated axons are initially ordered (Fig. 4.22B), placed in an upright square lattice on the plane transversal to the peduncle.

4.4.3 External Field determination

Another important contributor in our model is the external attractive field that sets the directionality of axon growth. If axons were simply put altogether with their initial point set to the origin of coordinates (*i.e.* not registered to the ML), a straightforward way to set the attractive direction would be to use the directionality of the linear regression of all the points, and consider it constant in space (as done in the simulations in Figs. 4.6 and 4.7). In this context, a constant field indicating the preferred direction is sufficient to create realistic guided trajectories. However, when the axons are registered to the ML (Section 4.1), more information is required as the spatial configuration is complex and the direction of the guiding cues attractiveness may not be constant in space. Thus, the challenge is to find the vector field $\mathbb{R}^2 \to \mathbb{R}^2$ to describe the field direction at each point in space. For simplicity, we considered it constant in Z. To give direct biological meaning to the attractive field, we describe it as the gradient field of a function that is maximal in the regions that correspond to the axon targets, and thus the highest concentration of attractive cues and thus axon targets (red segments Fig. 4.24). In this model, we use only the directionality of the cue attractiveness and not its magnitude.

However, neither the identity nor the source of the cue(s) guiding the *in vivo* growth of γ axons are currently known. Thus, we calibrated the external attractive field based on the observed directionality of real axons, and some assumptions. From data we obtained the mean directionality distribution of the WT axons (Fig. 4.23). This information can be considered as a sample of the attractive field plus some noise. As it can be observed, γ main

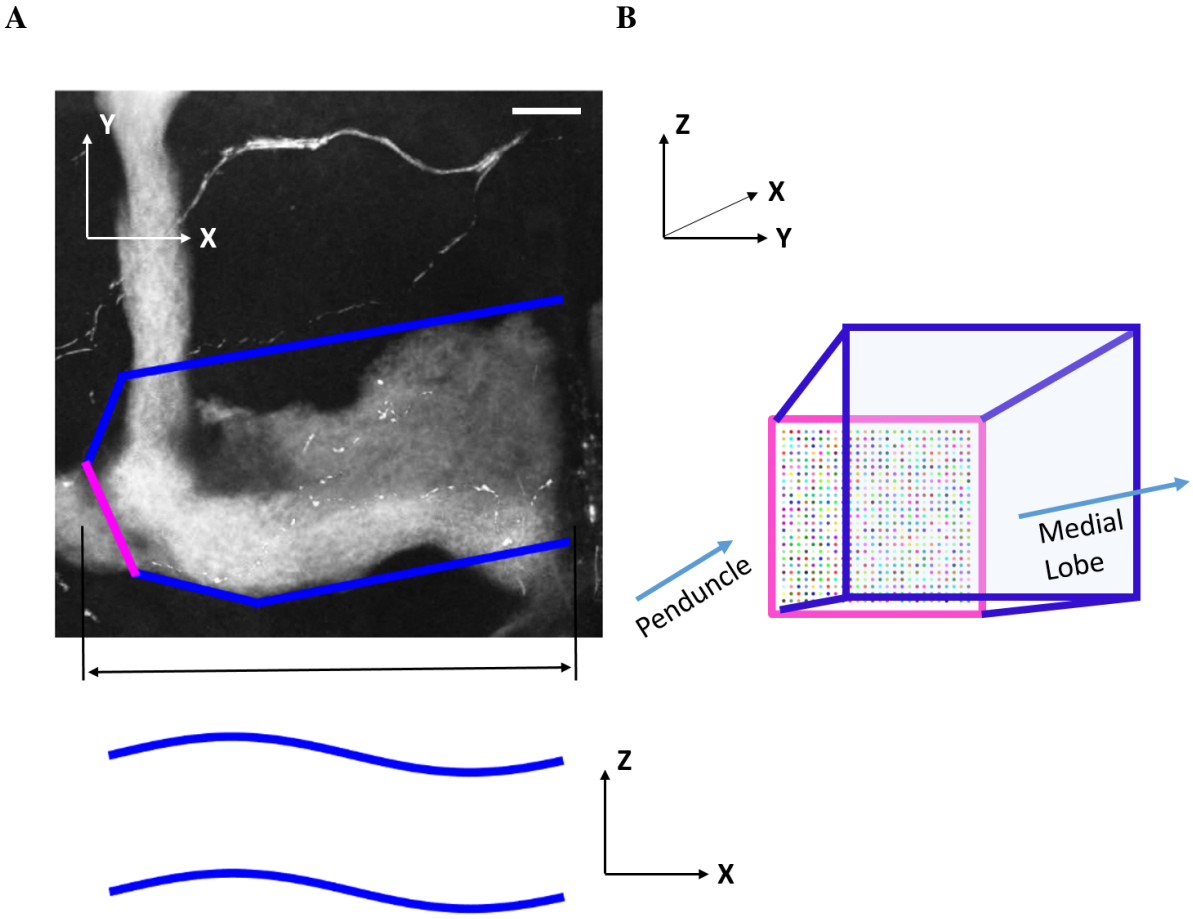

Fig. 4.22 Medial lobe geometry reconstruction. (A) 3D geometry of the medial lobe used in the simulations. Axons start their growth perpendicular to the purple plane, representing the peduncle transversal section, and their growth is constrained to the volume delimited by surfaces represented by the blue lines. (B) Detail of the initial position of axons at the purple plane. They start placed in an upright square lattice orthogonal to the peduncle.

axons are strongly oriented toward the mid-line (dotted line, Fig. 2.3 A), and exhibit a small bias toward the dorsal distal and ventral distal part of the medial lobe (Fig. 4.24, Real Axons). In addition, we assumed that the guiding cues are concentrated at the end of the ML (see Fig. 4.1 A), which makes sense regarding the MB connectivity and the axonal directionality (Fig. 4.23). However, one must take into account that this is a modelling assumption and an abstraction, and does not claim to represent an exact description of the biological model.

We tested different vector fields, result of simple guiding cue concentration configurations (Fig. 4.24), and compared the obtained maps with the directionality of real data. While a punctual source of attractive cue generates exclusively convergent attraction forces, a single segment source generates forces that are convergent at the edges of the segment, and parallel in its middle. Fig. 4.24 shows the distributions in the ML of the directionality of real axons, as well as for the different proposed field configurations. It can be appreciated how real axons enter the lobe with positive angles, and then become regressively orthogonal to the mid-line. Even though this behaviour is more or less captured by all the proposed configurations, the main difference between them relies on the region near the end of the ML (distal part). Real axon trajectories are biased positively in the dorsal region, and negatively in the ventral one. While the Punctual and Single segment configurations behave in the exact the opposite way, the Inverted C captures this behaviour.

To tune the shape of the Inverted C that sets the most suitable gradient field, we evaluated each candidate with a similarity index. We defined it as the average of the cosine of the difference between the angle of each segment composing the axons with the corresponding field directionality. This index thus goes from -1 (opposite direction) to 1 (perfect match), and we naturally chose the function with the maximum score. To systematically test a range of possible configurations, we considered two variables (D,V) accounting for the dorsal and ventral extension of the attractive source (Fig. 4.25A). The initial point $(D,V) = (0,0)$ represents a punctual source placed at the centre of the lobe extremity. We then varied (D,V) from 0 to 60 μm with intervals of 5 μm and calculated all corresponding similarity indexes (Fig. 4.25B). Three field configurations where found to maximize the similarity index: $(D,V) = (45,30)$ (Fig. 4.25C), $(D,V) = (45,25)$ and $(D,V) = (40,25)$ (which corresponds to an inverted L shape, *i.e.* the ventral segment orthogonal to the mid-line is absent). From these three equally fitting options, we chose the one depicting an inverted C ($(D,V) = (45,30)$) as it is the only one that simulates the ventral distal bias observed in data (Fig. 4.23).

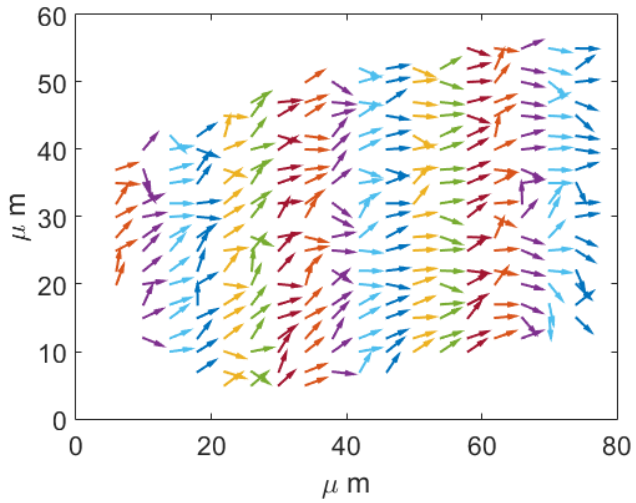

Fig. 4.23 γ main axon directionality. Local mean directionality of the WT reconstructed and registered γ axons along the medial lobe. The lobe was divided into rectangles of 2x1 μm in $XY, \forall Z$ and the mean directionality in each region calculated. Then the map was interpolated with the Matlab function *griddata*, and the arrows placed in a 2x2 μm grid.

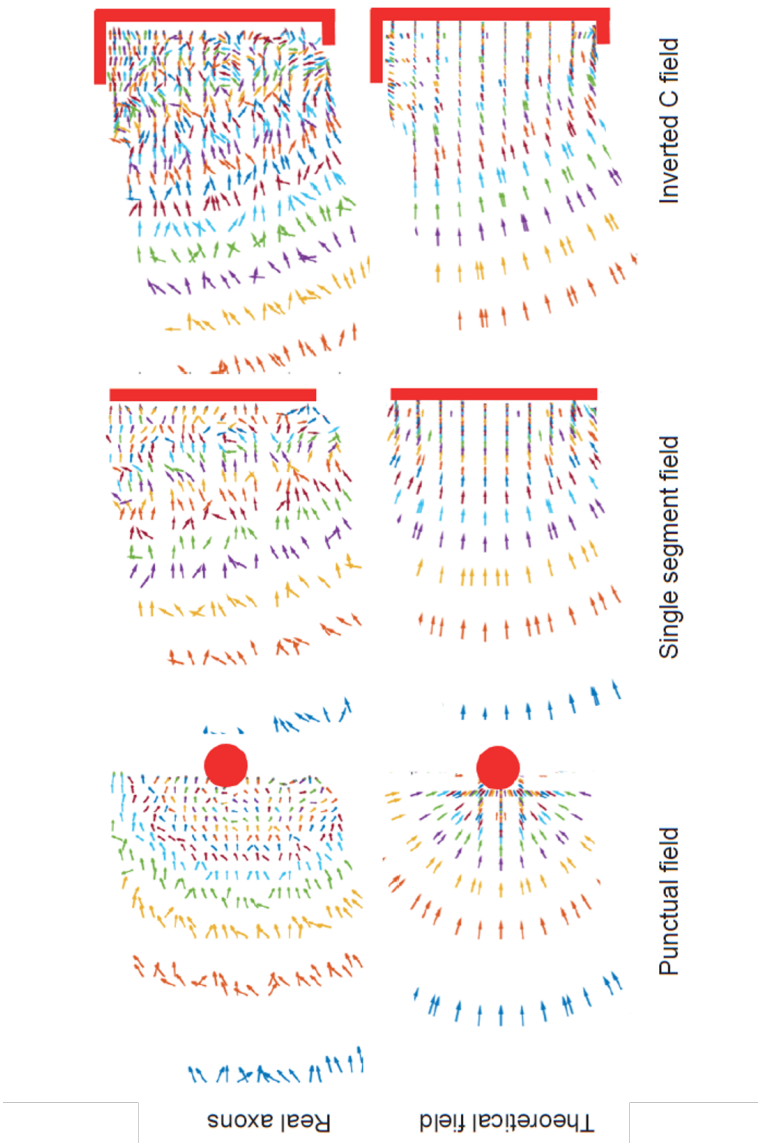

Fig. 4.24 Alternative attractive field configurations and chosen field. Upper panels: axonal directionality vectors obtained from real data, placed on the level lines of each considered field configuration. Lower panels: force directionality vectors of each theoretical attractive field, placed on the corresponding level lines.

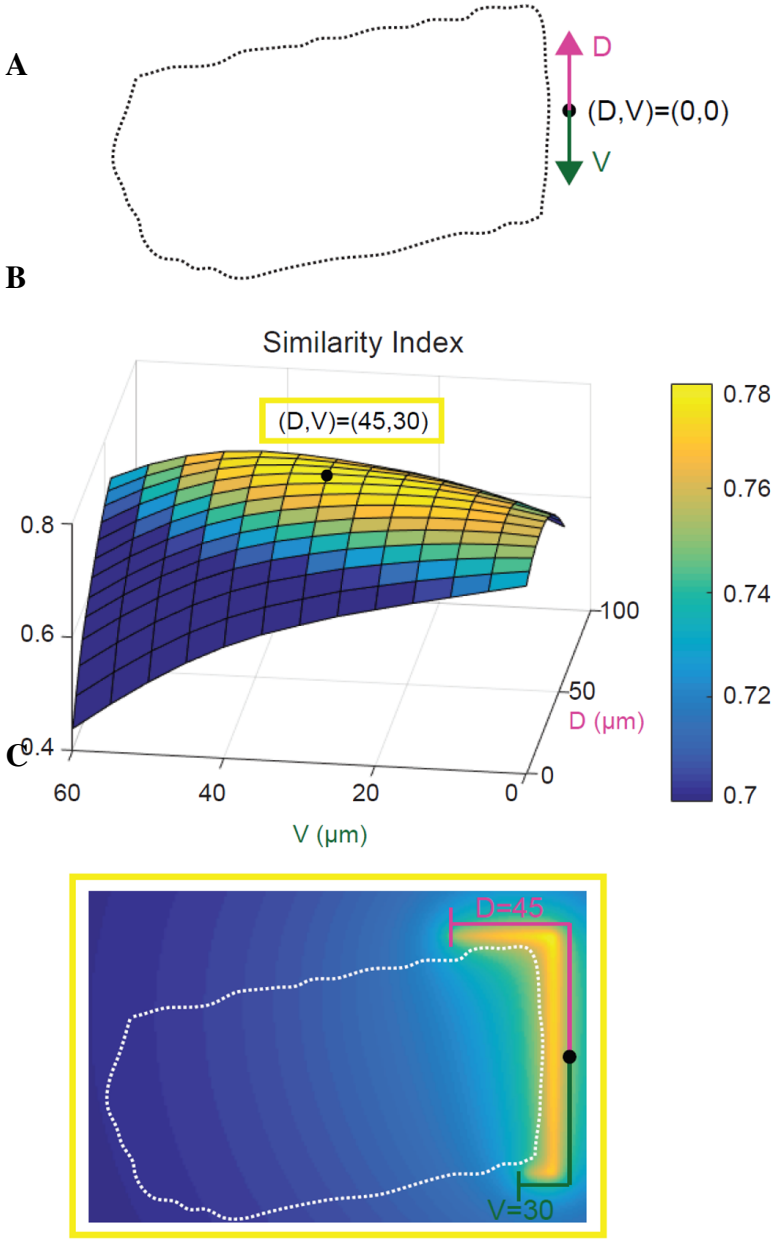

Fig. 4.25 Attractive field calibration. (A) Schematic representation of the D and V distances used to generate different attractive field configurations. The attractive source is positioned at the end of the medial lobe, and depends on two variables (D,V). D describes the extent of the source along the dorsal part of the lobe, and V its extent along the ventral part. The origin point $(D,V) = (0,0)$ is also shown. (C) Similarity index in function of D and V. The attractive field configuration used in this study ($(D,V) = (45,30)$) is represented by a black dot. Its similarity value is 0.779. (C) Attractive field configuration maximizing the similarity index and used for the simulations.

With this field configuration, the trajectories in the lobe following the Markov model, and with different parameter values are shown in Fig. 4.26.

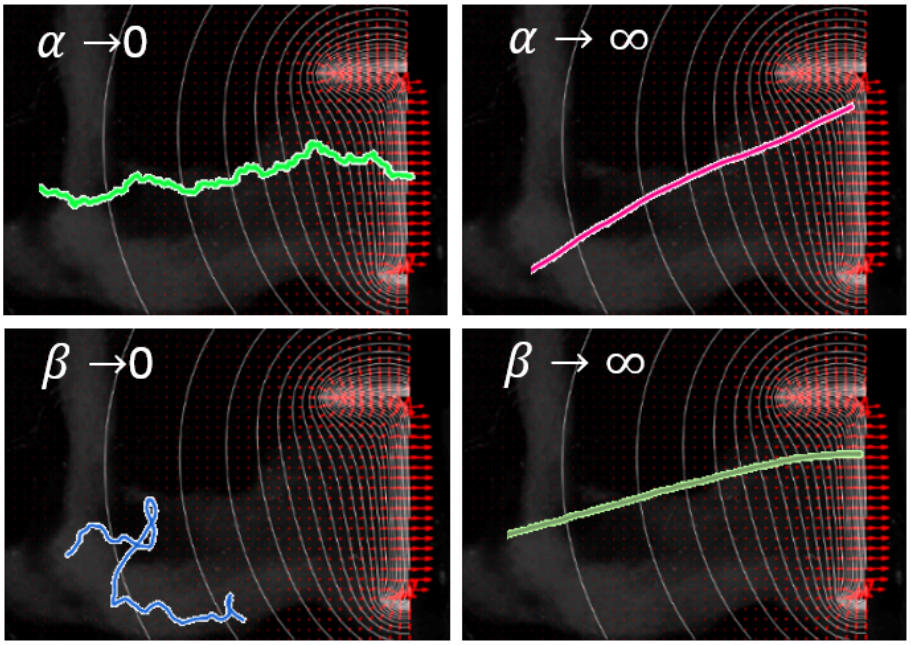

Fig. 4.26 Influence of the chosen field configuration on axon trajectories in function of α and β values. The field is represented in the SML (see Section 4.1) by its directionality (red arrows) and level lines (white curves). α represents the axon stiffness; values near zero result in very tortuous trajectories that are only guided by the external field, while higher values lead to straight axons that mostly ignore the external field and tends to follow its initial direction. β represents the axon sensitivity to the external field. Values near zero indicate no perception of the field direction, leading to undirected axons. Axons with high β grow straight to follow the field direction (*i.e.* its trajectory stays orthogonal to the level lines of the field).

4.4.4 Volume exclusion

Inter-axon interactions are taken into account *via* volume exclusion, considering the same axon diameter d for all the neurites (main axon and branches).

Volume exclusion in the simulations is performed considering that the *distance* between the growing tip and any other existing neurite in the lobe must be equal or higher than d (Fig. 4.27 A). To calculate this *distance*, we divide each step ρ_i ($\Delta\rho = 1\mu m$)of each existing neurite in k points such as $k = int(\frac{\Delta\rho}{d})$, and calculate the Euclidean distance between the growing tip end and each of these points (Fig. 4.27 B). Geometrically, this means that we approximate the neurite volumetric shape as a sequence of spheres (Fig. 4.27 C). This approximation

allows a simple algorithmic implementation, and takes into account observations showing that the neurites diameter may not be perfectly constant (images present irregular traces (Figs. 2.3 C and 2.5).

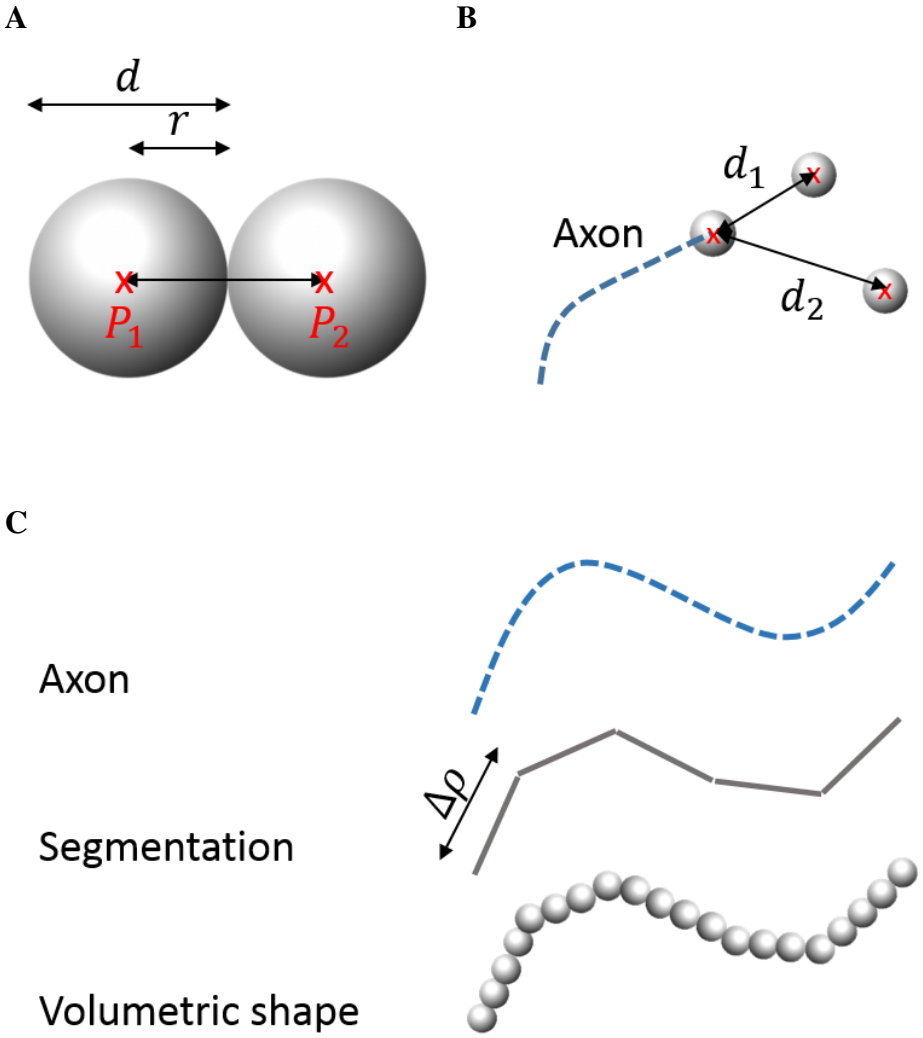

Fig. 4.27 Neurite 3D geometry and volume exclusion. (A) Minimum Euclidean distance between two spheres considering hard core volume exclusion. (B) The dotted line represents the growing axon. The volume exclusion condition imposes that all the distances $(d_1, d_2, ..)$ between its tip and neighbour neurites are higher or equal to d. (C) The dotted curve represents the unidimensional trajectory in space of the axon, the grey lines show the segmentation in the segments of length $\Delta\rho$ and the spheres its 3D geometry modelling.

Unfortunately, d cannot be measured directly from the confocal images in our data, as the axonal trace diameter varies enormously (they can go up to $d \sim 1\mu m$, as well as sometimes disappear completely).

With this considerations, the value of d was estimated from the electron microscopy images in Watts et al. (2004) (depicting diameters of the order of $10^{-1}\mu m$), and then sharpened by simulations. We simulated the γ population considering different possible diameter values and chose the one that minimizes the ratio between the percentage of non-elongated axons (*i.e.* axons that do not reach the stopping region, see Fig. 4.2 D) and the ML neurite density (function of the diameter). In other words, we selected the maximum diameter with the minimum elongation error as possible. The obtained diameter value results $d = 0.23\ \mu m$, and the obtained curves are shown in Fig. 4.28.

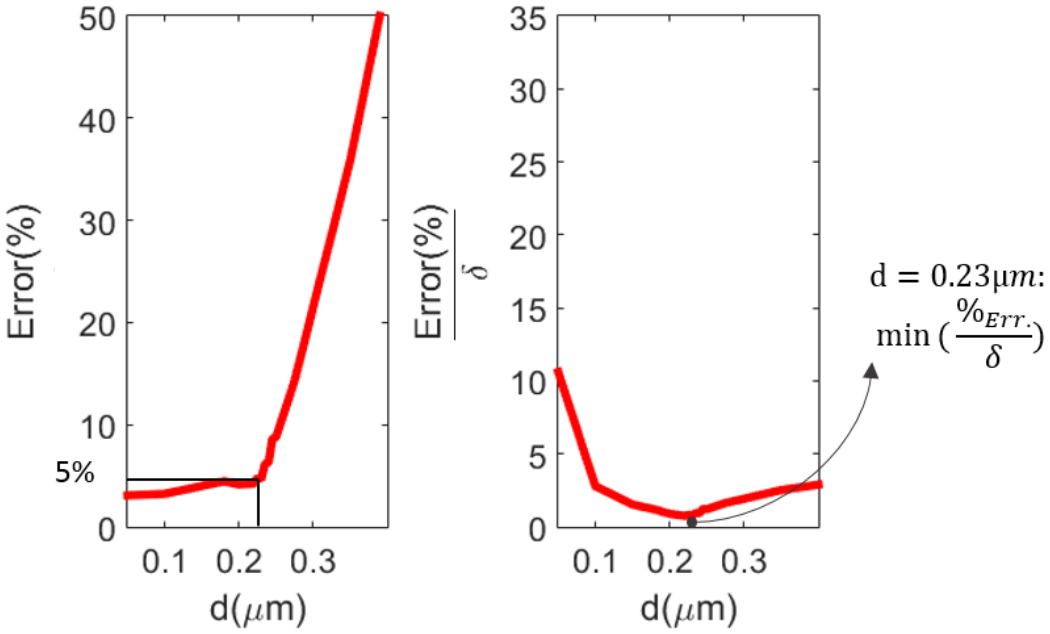

Fig. 4.28 Diameter determination. (left) Percentage of non-elongated axons (see Fig. 4.2 D for definition) from simulations with different values of neurite diameter. The value just before the elbow, $d = 0.23\mu m$, represents an error of 5%. The simulations were performed considering branching upon contact (see Section 4.5.2). (right) Percentage of non-elongated axons divided by the planar density δ from simulations with different values of neurite diameter ($\delta = \frac{650\pi(d/2)^2}{A_{ML}}$, where A_{ML} corresponds to the largest area of the ML cut by a plane $X = cte$, δ represents the occupied fraction of the ML). This function is minimum for $d = 0.23\mu m$. This value optimizes the ratio between the error and the maximum diameter.

As it can be seen, these curves seem to follow a logistic law such that the percentage of non-elongated axons grows very slowly for small values of d, and then grows very fast from $d = 0.23\mu m$. Fig. 4.29 presents the empirical curves of non-elongated axon percentage in function of varying diameter values (d) and for different ML depths. For each curve, an

approximation by a theoretical sigmoid function is also shown. As it can be observed, the approximations result accurate for every case. Nevertheless, the exact dependence of the sigmoid function parameter value with the physical properties of the problem is unknown. Thus, we can only obtain these curves by simulation.

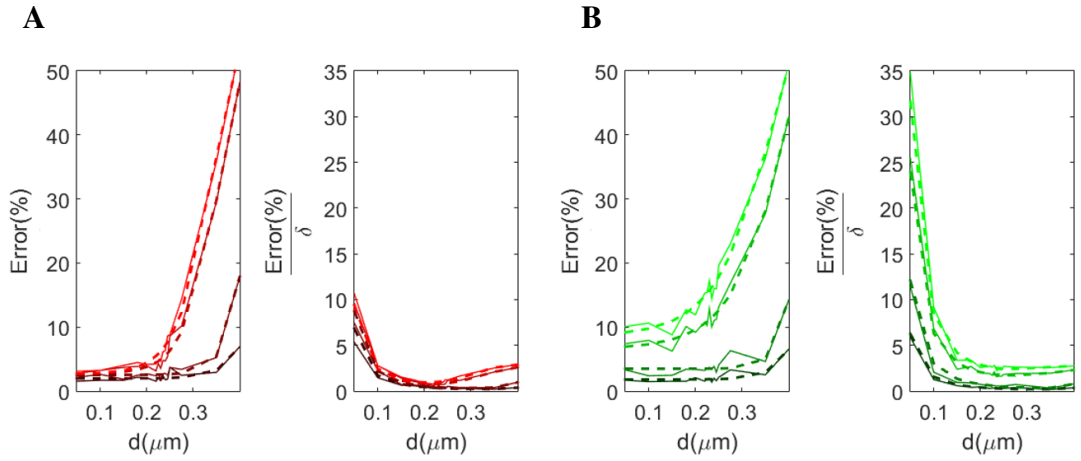

Fig. 4.29 Error (%) (percentage of non-elongated axons) in function of the neurite diameter: Study. (A) The filled curves (left and right) are the same as in Fig. 4.28, but for the percentage of axons stopping at different depths (X) of the ML: $X = 5, X = 15, X = 50$ and $X = 62\mu m$ from darker to lighter. Dotted lines correspond to approximated sigmoid functions (left) and the same approximated sigmoid functions divided by δ (right). This theoretical functions nicely follow the empirical ones, suggesting a logistic behaviour of the latter. (B) Same curves as in (A), but for simulations without considering branching (only main axons). Interestingly, the curves $\frac{Error(\%)}{\delta}$ do not present a minimum value, but decrease $\forall d$ in contrast to those considering branches, for both empirical and theoretical curves.

One last comment is that axon elongation depends on the total occupied volume (consequence of axon-axon interaction), which is determined by a certain function of the diameter d (and other variables that can be considered invariant for this problem). Thus, when choosing a particular way to model the 3D shape of the axons, one sets this function. It is not the ambition of this work to give a precise estimation of the diameter of γ neurons, or their volumetric shape, but rather design a simple and coherent method to develop the space-embedded simulations. Other 3D shapes would have changed this function, but the procedure to set d would have stayed the same. If, for example, the neurites were modelled as cylindrical, then probably the estimated value of d would have been lower (as with the same maximum diameter, a cylinder has a higher volume than a sequence of spheres).

4.4.5 Resulting trajectories

We now present the results of simulating 650 interacting main γ axons in the ML, considering the external field and volume exclusion. To assess the accuracy of our simulations, we first observed the morphology of the resulting axons. As shown in Fig. 4.30, simulated trajectories strongly resemble real γ main axons. This similarity can be observed in the high frequency noise of the paths (h.f.n, *i.e.* how the trajectories behave along a small distance, see Fig.4.30), as well as their global morphology (g.m., see Fig.4.30). Furthermore, the population of simulated axons filled the entire medial lobe, the same way as populations of *in vivo* neurons do (Fig. 4.30).

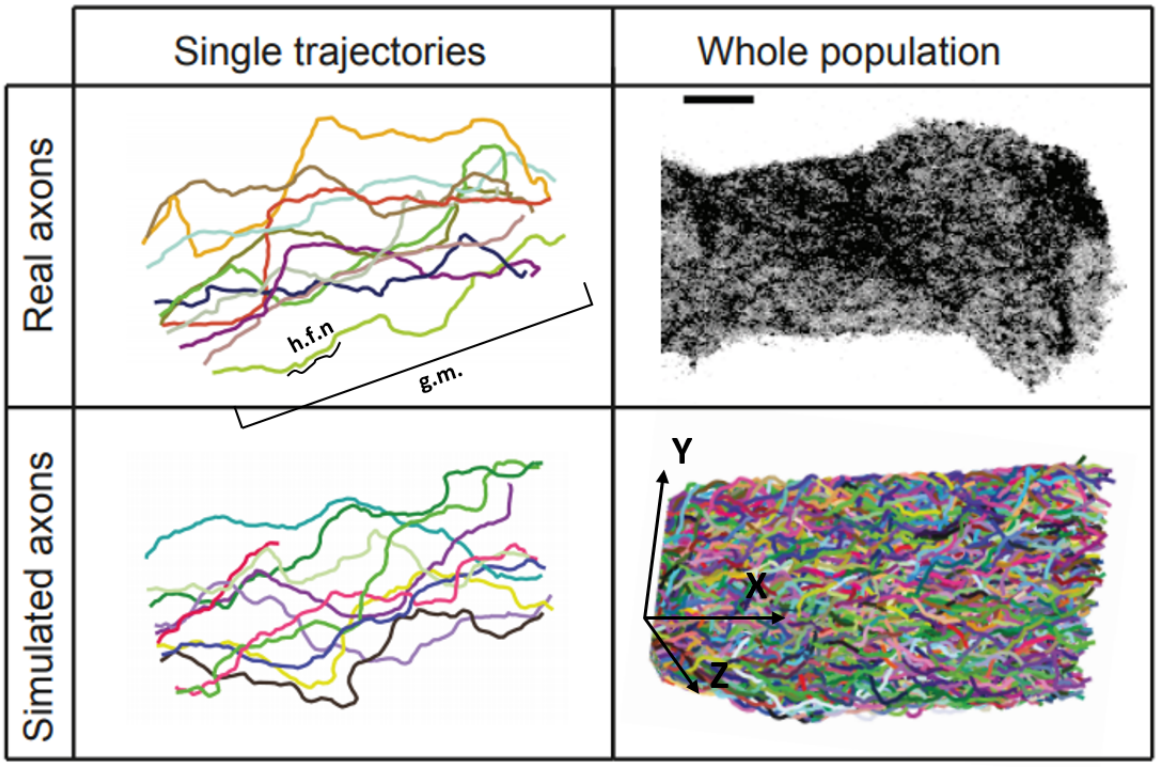

Fig. 4.30 Space-embedded simulated vs. real axon trajectories. (left) examples of reconstructed real wild-type γ axons (upper panel) and simulated γ axons (lower panel). The high frequency noise (h.f.n) and global morphology (g.m.) are indicated.. (right) populations of *in vivo* γ axons visualized by mCD8 fused to GFP expressed specifically in γ neurons (upper panel; scale bar corresponds to $10\mu m$), and simulated γ axons (lower panel).

To more precisely assess the similarity between real and simulated axons, we also compared the distributions of main axon lengths and travelled distances (distance travelled in the ML parallel to its main axis, see Fig. 4.2 D in Section 4.1). The main axon length distribution profiles of real and simulated axons match very well, particularly regarding

the data dispersion (see Fig. 4.31, left panels). This indicates that the sinuosity as well as path variability is effectively well respected by the model. One must remember that the final length of the simulated axons is not previously set, but is an emergent feature of the model. However, the simulated axons are generally longer than real ones ($\mu_{real} = 93\mu m$ and $\mu_{simu} = 118\mu m$). This difference can be explained by the fact that the axons in the confocal images are flattened (in the Z axis), which is a known consequence of the tissue fixation technique. In particular, from the channel two of the images in our data (see Section 2.4), the ML has an average width (in Z) of $10\mu m$. However, Aso et al. (2014) performs an accurate 3D reconstruction of this structure, where the ML reported width is $\sim 20\mu m$ (or double our measurement). Thus, the 3D reconstructions of the real axons segmented from the confocal images may present a bias in the increments parallel to the Z axis, while the simulated axons are allowed to grow freely in Z. For a fair comparison with real data, the reported values μ_{simu} and σ_{simu} were calculated without the tail in the distribution, considering only the axons that reach the stopping region (only lengths higher than $89\mu m$ were considered, see Fig. 4.31)).

One way of avoiding this bias and compare real *vs*. simulated axons is thus to measure the lengths of the axons projected to the XY plane. Fig. 4.32 shows the results of considering the axonal length in 2D and contrast with the 3D case. Nicely, we can observe that when considering projected lengths, the distribution of simulated and real axons become more similar, and the bias is avoided. As for the 3D case, the reported value μ_{simu} was calculated without the tail in the distribution, considering only the axons that reach the stopping region (lengths higher than $58\mu m$ were considered, see Fig. 4.31)).

Regarding the travelled distance (Fig. 4.31, right), 85% of the simulated axons are correctly elongated (definition in Section 4.1 and Fig. 4.2 D). This is the first clear impact of taking into account the mechanical environment of the growing neurons, as an important percentage of them fail to elongate. This showcases the importance of considering axon development in a realistic environment, and shows that mechanical competition may stop some axons from reaching the stopping region, thus from elongating accurately. The distribution of travelled distances for simulated axons also shows that non elongated axons stop all along the ML, and not in a specific region.

In addition, interesting dynamic aspects also emerge in our simulations. Even though simulated γ axons start growing simultaneously, they reach the stopping region asynchronously. Interestingly, this finding is very consistent with observations made on populations of γ axons growing in the context of real brains during metamorphosis. Fig. 4.33 shows a group of growing axons through three time frames from simulations as well as from a *live* video in our data (Section 2.4). In both cases, the first time point (t_1) depicts a group of axons where

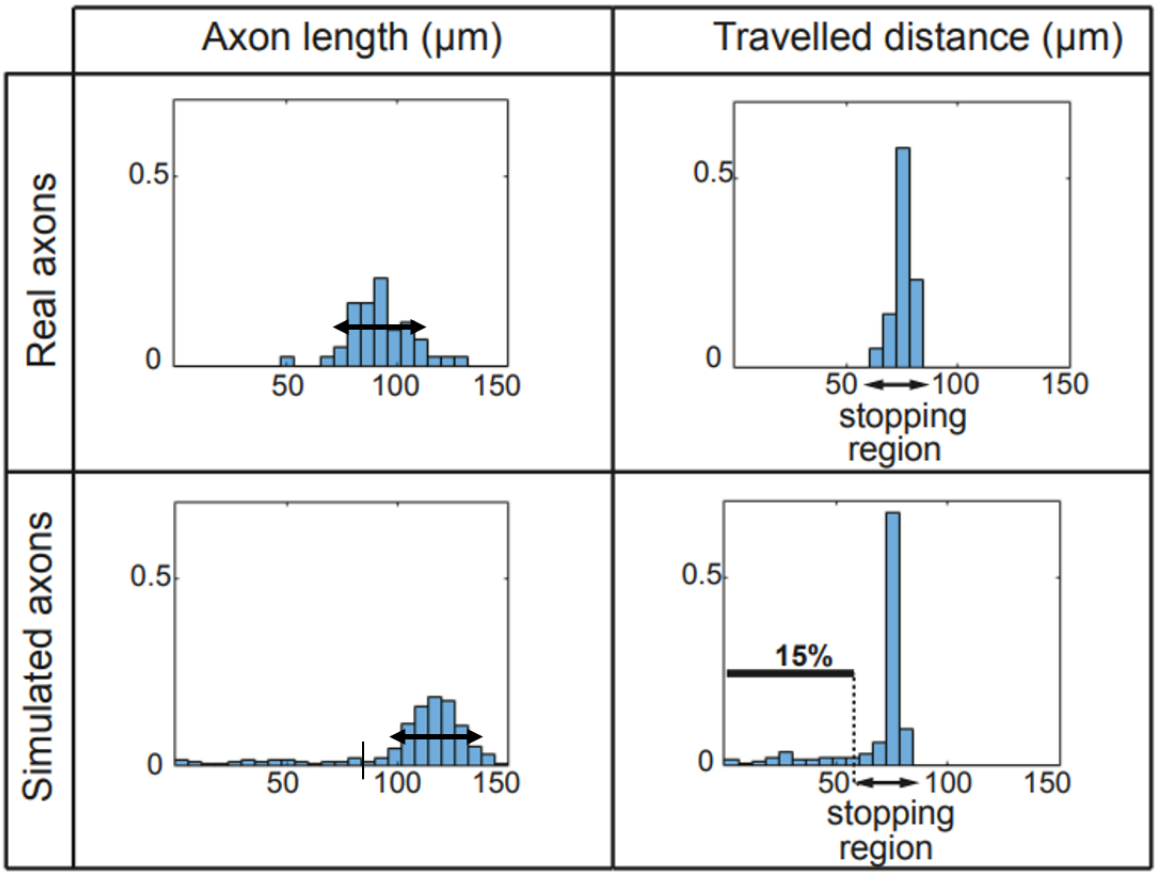

Fig. 4.31 Space-embedded simulated vs. real axon lengths and travelled distances. (left) Frequency distributions of main axon lengths. The same double arrow is placed on both distributions, showing the similarity in their typical widths. The black bar on the simulated histogram indicates from where the distribution was considered to calculate the values of mean (μ_{simu}) and standard deviation (σ_{simu}) reported. $\sigma_{real} = 14\mu m$ and $\sigma_{simu} = 11\mu m$. (right) Travelled distances within the medial lobe. Data from real wild-type γ axons is shown in the upper panels, and data from simulated ones in the lower panels.

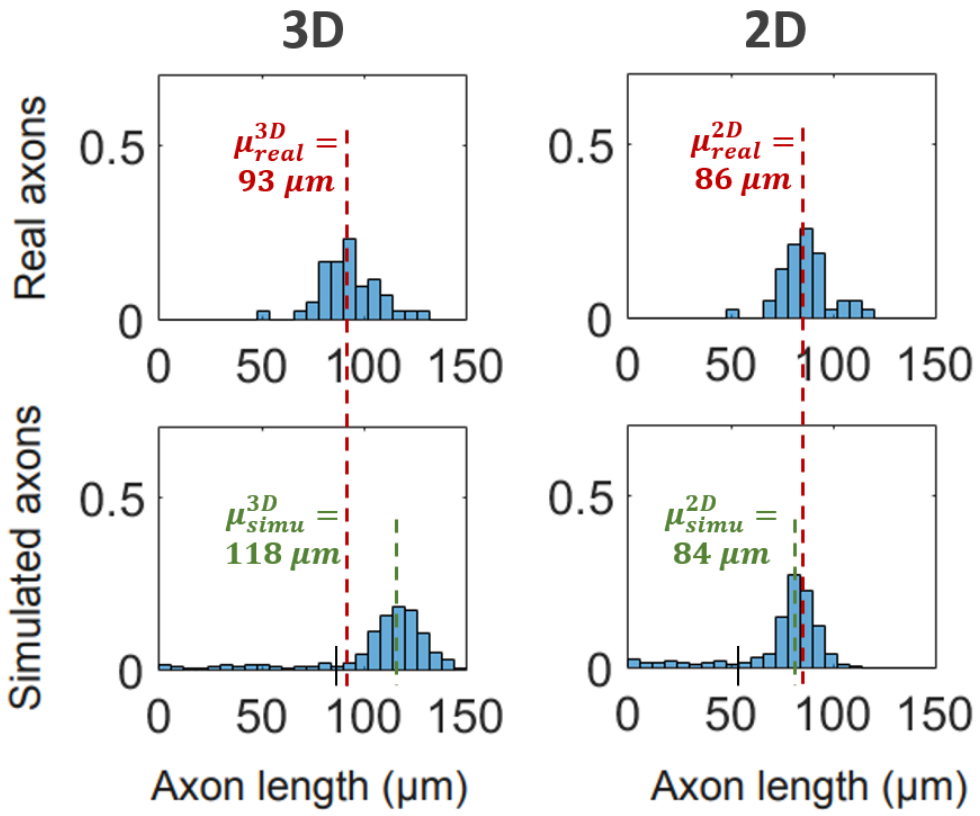

Fig. 4.32 Space-embedded simulated vs. real axon lengths and travelled distances: 3D *vs.* 2D projections. (left) Frequency distributions of main axon lengths measured in 3D. The dotted lines are placed at the means of the real (μ^{3D}_{real}, in red) and simulated (μ^{3D}_{simu}, in green). The black bar on the simulated histogram indicates from where the distribution was taken to calculate the value of the mean. (right) Frequency distributions of main axon lengths measured in 2D. The dotted lines are placed at the means of the real (μ^{3D}_{real}, in red) and simulated (μ^{3D}_{simu}, in green). The black bar on the simulated histogram indicates from where the distribution was taken to calculate the value of the mean.

none of them has fully elongated yet. In the next one (t_2), only one/two real/simulated axon/s reach the stopping region, while the rest of them continue growing behind. Such differences in the growth speed reflect not only the stochasticity of axon paths, but most importantly the competition within the population, which arises in our model as a consequence of spatial constraints. Thus, some axons may be *lucky* and encounter just a few obstacles, while others may encounter more, thus advance slower and loose time with respect to the *lucky* ones. Finally, in the last frame (t_3) and in both cases, a few more have also reached the stopping region while the others continue growing. A similar observation from another *live* video is shown in Figs. B.5 and B.6 in Appendix B.

4.5 Branching mechanisms

We have presented a mathematical model for axon elongation during development with biologically-interpretable parameters estimated directly from data. In addition, we proposed a computational framework that allows realistic dynamic space-embedded simulations of the growth of an entire γ population within the medial lobe of the *Drosophila* mushroom body. We showed that our model produces realistic main axon morphologies as well as reproduces and explains the asynchronicity between growing axons observed in *live* data.

However, γ axons are not composed by a sole main axon, but form trees displaying a wide variety of morphologies and complexities (Figs. 2.3 C, 2.5 and 3.1). The branches composing this non-stereotypical trees display different lengths and orientations, and their origin/end points are generally dispersed through the lobe. The developmental mechanism behind the generation of this complex network has not yet been fully described, and therefore deserves attention.

Moreover, given the number and size of γ axon branches, the combined lengths of the branches of a single axonal tree generally represent more than the length of the main axon. This means that the volume occupied by a complete axon (main axon plus its branches) may double (or even more) the one we considered until now (*i.e.* only the main axon). In our study this fact acquires a new central importance, as more occupied volume in the ML also means more mechanical constraints for all the axons during development, thus more difficulty to attain elongation at the population level. We have shown that with our framework, 15% of the γ axons fail to reach the stopping region due to mechanical obstacles (Fig. 4.31). Does this suggest that by adding the corresponding branches this percentage would increase? This would represent a paradox, as our model suggests the existence of a non-negligible percentage of non elongated axons while, actually, no *short* (*i.e.* non-elongated) wild-type axons were observed in the real data.

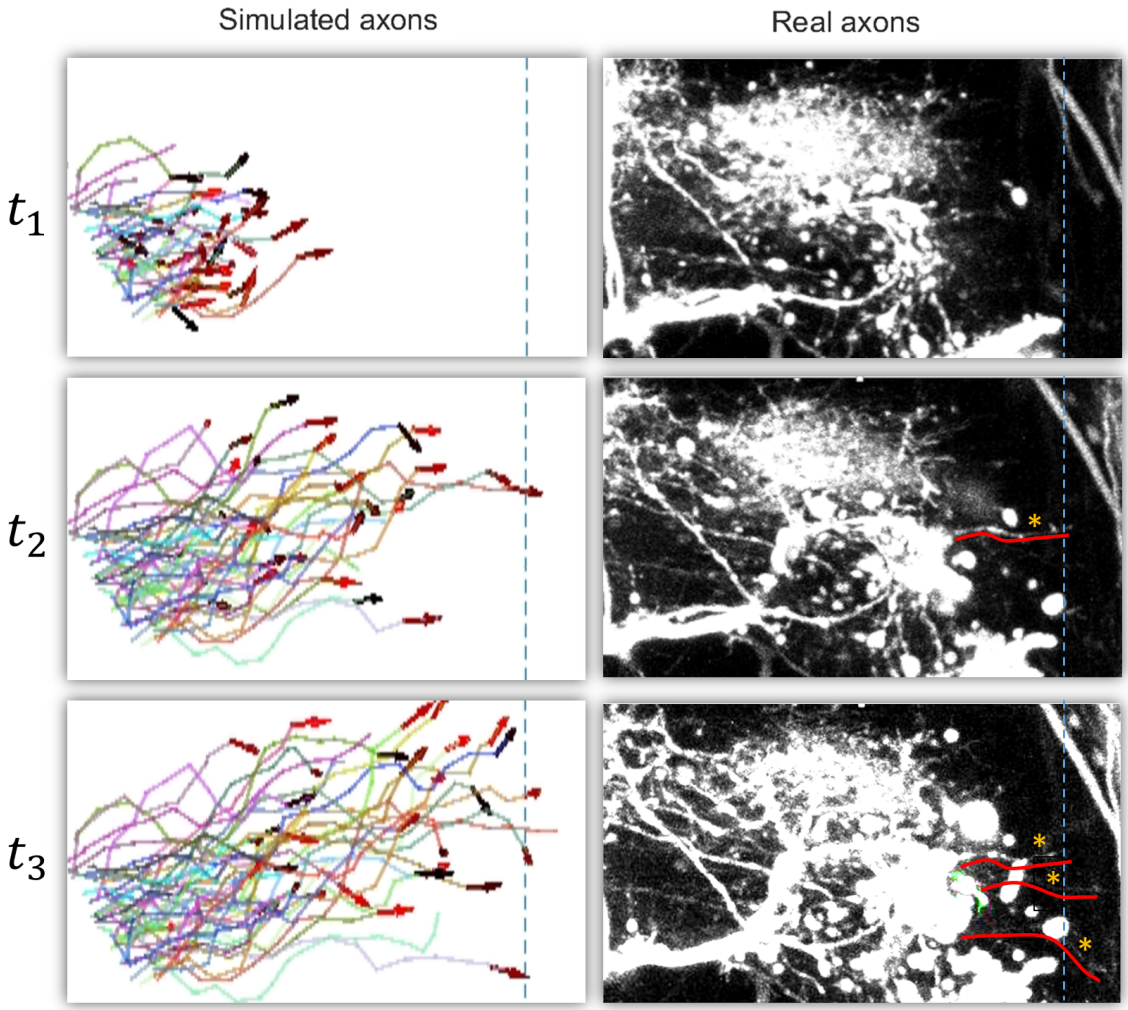

Fig. 4.33 Emergent asynchronicity within the γ population. Panels in the left show a subgroup of 50 simulated axons, and the red arrows represent the velocity of the axon tip growth at the specific time point. The absence of arrow means that the axon encountered mechanical obstacles and could not grow. Panels in the right correspond to a *live* film in our data ($t_1 = 0min$, $t_2 = 90min$ and $t_3 = 215min$). The dotted lines depict the beginning of the stopping region (Fig. 4.2 D). In both cases (real and simulated), the first screen-shot depicts a moment in the beginning of the growth when no axon reached the stopping region. The second one shows the first axon that arrives (yellow asterisk and red line for real video), while in the third one several axons do (yellow asterisks and red lines for real video).

In this section, we analyse in detail the morphological characteristics of γ axonal branches from data, and propose different branch formation mechanisms that we test through modelling. The analysis of the simulation results suggest interesting branching functionalities in axon development that had not been previously described, particularly related to the mechanical environment.

4.5.1 γ neurons present two types of branches

In order to include a developmental branching mechanism in our model, we first analysed from our real data the adult branch morphology. Within the medial lobe, γ axons extend side branches of different length, as we can observe from the images in our database. Fig. 4.34A shows an example, where one longer branch is differentiated from other smaller ones. Fig. 4.34B presents the lengths for first and second order branches (see Fig. 3.5 for a schema of branch hierarchy) measured from the WT axons in data. While there are 8.8 branches per axon ($\frac{b}{ax.}$) in the first group, only 3 are observed in the second one. In addition, these branches are quite small.

To understand how these branches are generated, we observed the *live* videos in our database (Section 2.4). As shown in Fig. 4.35, from the videos we could also distinguish between longer (purple arrows) and shorter (asterisks) branches. Interestingly, the difference between both was not only their length, but also their dynamic behaviour. While longer branches stayed quite fixed over time (and presented a dynamic activity restricted to their tips), the shorter ones presented very unstable behaviours, appearing and disappearing from one time frame to another. Measuring the lengths of the branches with high dynamic behaviour from two different movies, we observed that these branches measure between 2 and 10 μm long (Fig. 4.35B). To further analyse the dynamic behaviour of these branches, we superposed all the frames of the movie in Fig. 4.35A, as shown in Fig. 4.36A. Then, we extracted the stable tree (one main axon plus two long branches in this case) and quantified how many new branches appeared along its skeleton (Fig. 4.36B). We observed that these events happened all along the structure, and are more frequent towards the tips (Fig. 4.36C). Despite their unstable behaviour, it can be considered that at least some of these dynamic branches finally stabilise, to generate the small ($\leq 10\mu m$) branches observed in adult axons. This idea is supported by Leondaritis and Eickholt (2015), as they observed that stable branches in dendrites and axons can be initiated by the maturation of small and transient membrane protrusions which are highly dynamic actin-enriched structures. Given the morphological (length) and functional (dynamics) difference between γ axon branches, we classified them in two groups: type I ($> 10\mu m$, stable) and type II ($\leq 10\mu m$, dynamic).

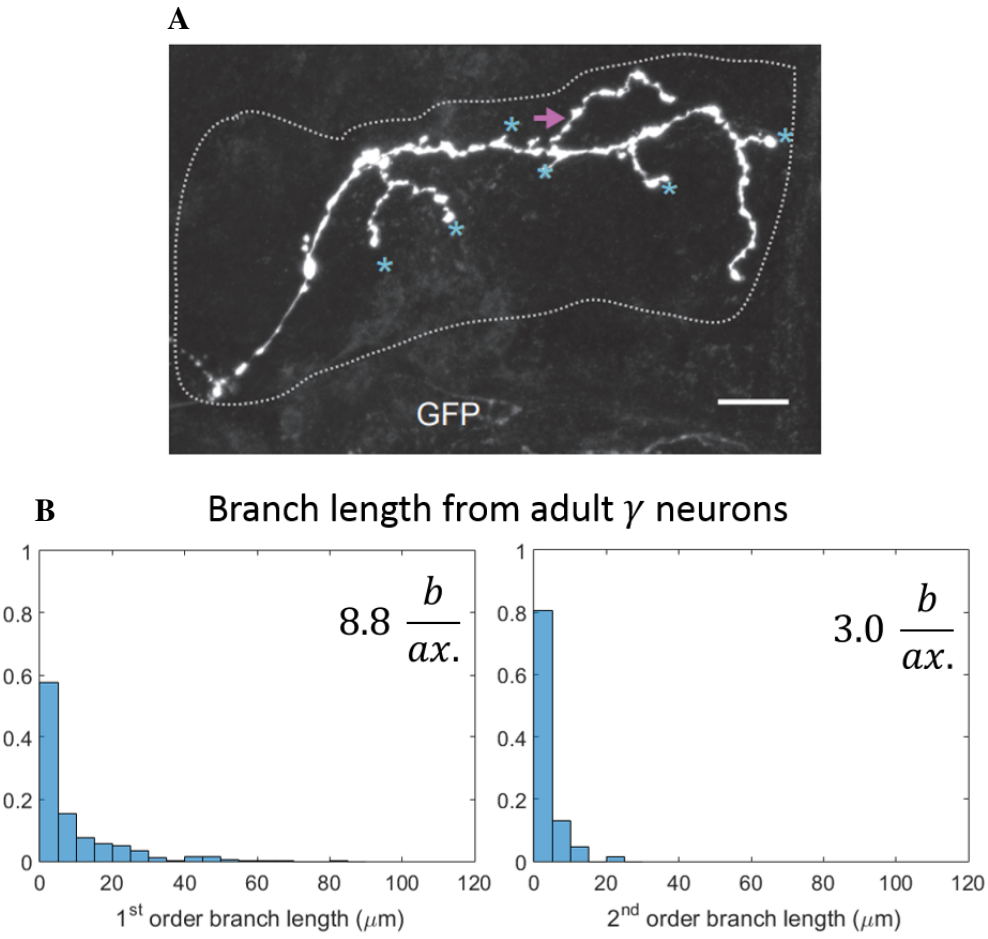

Fig. 4.34 Adult γ branch length distribution. (A) Confocal image depicting a single γ axon labelled by GFP. The purple arrow shows a long branch, and the asterisks smaller ones. Scale bar: $10\mu m$. (B) Length distributions for first and second order branches. In average, γ axons present 8.8 branches per axon of first order and 3 of second order.

Now that the branch classification is established, a new question arises: are there other functional differences than dynamism between type I and II branches? To answer to this question, we observed the orientation of branches in their final state (adult). Fig. 4.37A shows all the branches in our database (normalised to the SML, Section 4.1), classified by length. Interestingly, we observed that even though small ones are oriented quite uniformly, the longer they grow the more their orientation is biased towards the distal end of the medial lobe. Fig. 4.37B shows them altogether, to highlight this observation.

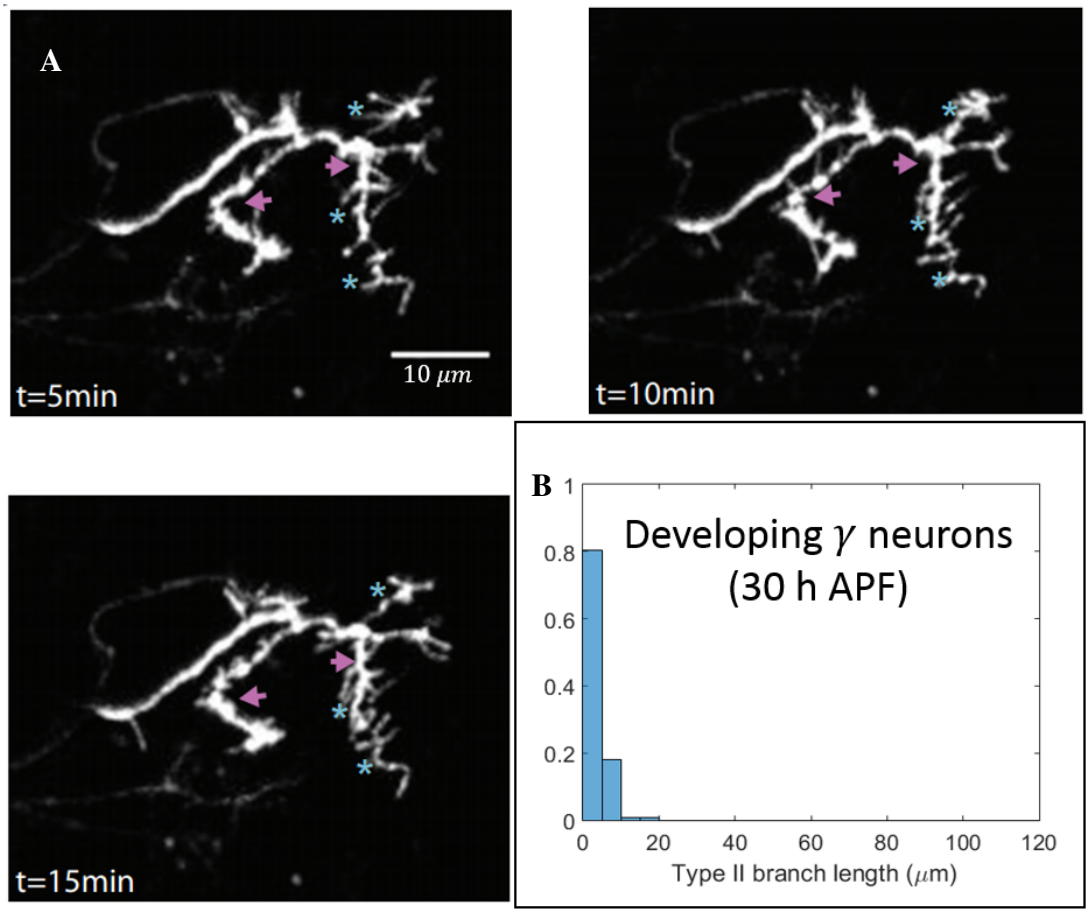

Fig. 4.35 γ axon branch dynamic behaviour during development. (A) Time frames from a video depicting a single γ axon during its regrowth phase. Purple arrows indicate long stable branches and asterisks short dynamic ones that appear, disappear or change their position over time. Scale bar: $5\mu m$. (B) Length distribution of the dynamic branches observed from two independent videos. The total number of branches analysed is 484.

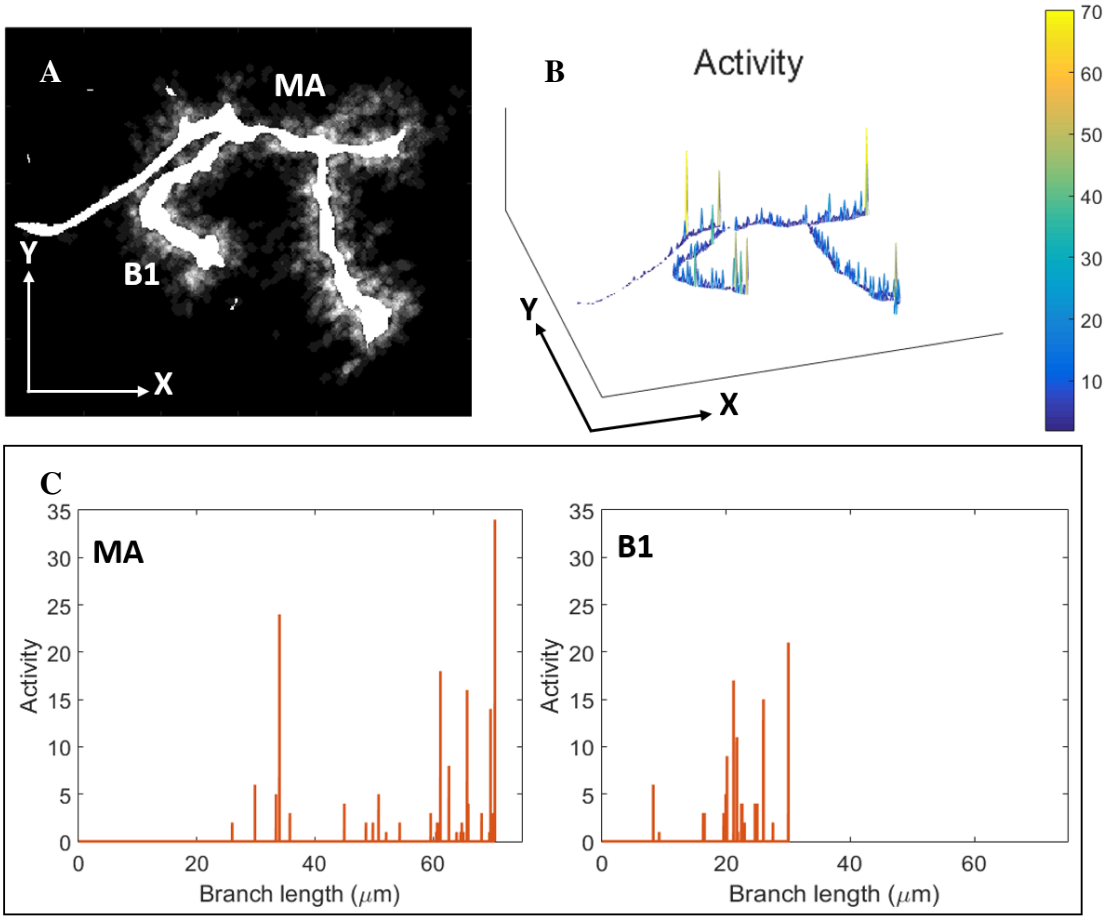

Fig. 4.36 γ axon branch dynamic behaviour during development -bis. (A) Superposition of all the time frames of the movie shown in Fig. 4.35 and normalized to the maximum value. The white central structure is present through all the movie, thus is stable. The grey area around indicates the unstable presence of dynamic branches. (B) Number of times a new branch appears along the skeleton of the stable structure. (C) Detail of the activity (only peaks of activity over 30 branch appearances are shown) for the main axon (MA) and a branch (B1) extracted from (B).

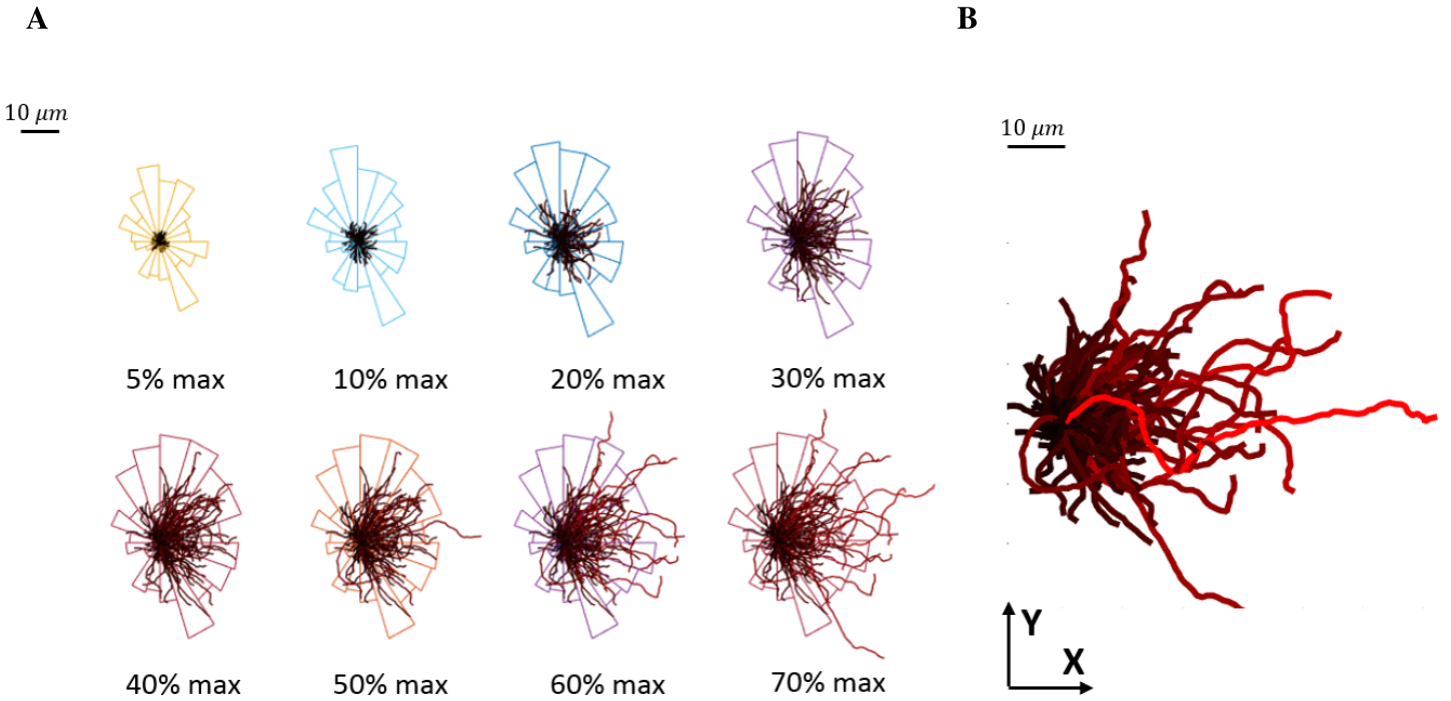

Fig. 4.37 γ axon branch orientation in the medial lobe. (A) Branches from WT axons in data normalised to the SML (Section 4.1) and put altogether, classified by length (in percentage of the maximum observed length: 83 μm). The histogram is superposed in each case. (B) All the branches from our database normalised to the SML and put together. Colour code: from black (shortest branch) to red (longest branch). Scale bars: 10μm.

In order to gain insight on what this difference in orientation means, and find out if there is a link with our branch classification (type I and II), we analysed and compared other relevant orientation distributions for type I and type II branches. Fig. 4.40A shows the angles between the whole branch and the corresponding main axon for both types. While type II branches are oriented in average orthogonally with respect to their main axons, type I branches tend to be more parallel to them.

A similar observation as in Fig. 4.40A can be done regarding the angle between the whole branch and the main axon only locally around the branching point (Fig. 4.38, middle), as well as for the initial segment of the branch and the main axon locally (Fig. 4.39 middle). In addition, not only entire branches are oriented towards the end of the medial lobe as they grow longer (Fig. 4.37), but also are their initial segments (as shown in Fig. 4.39, top). Finally, the difference in orientation between type I and II is not only evident regarding the lobe axis (Fig. 4.37) or relative to their main axons (Figs. 4.38 and 4.39), but also to the external field (described in Fig. 4.25C). Fig. 4.39 (bottom) shows the distributions of the angles between the initial segment of the branches and the local field directionality. While type II branches are oriented independently of the field orientation, an important bias towards the field direction is observed for type I branches. This is interesting, as the field was designed to match the main axon directionality and still makes sense regarding the branches initial angle. But even more importantly, it suggests information about the branching mechanism. While the stabilization of type II branches does not seem to depend on their initial orientation relatively to the field, type I ones are more successful when they are born aligned to it. This angle analysis allows to conclude that type I branches behave similarly to main axons, regarding orientation in the lobe and their behaviour respect the external field. On their side, type II branches seem to behave independently from main axons.

Another interesting observation is that axons reach the stopping region with between one and four tips, considering the main axon and type I branches (Fig. 4.40 B). The fact that in some cases only the main axon arrives, suggests that it is sufficient to create a functional axon. Thus, the variability on the number of arriving tips within the group looks like a redundant mechanism. Based on this observation and on the conclusions obtained from the branch orientation analysis, we proposed the hypothesis that type I branches grow exactly as main axons, following the external cue field and trying to reach the end of the ML. If this hypothesis is accurate, the definition of the main axon needs to be revised. While experts usually take into account aspects as length and orientation to define it (Section 3.1), the lack of more information (e.g. varying diameter) leads to possible confusions. In particular, we may label as main axon a segment that mixes the neurite that originally entered the lobe and a type I branch that reached the stopping region. Fig. 4.40C shows as an example two

possible segment classification for the same axonal tree. Even though this new hypothesis on branch formation questions the concept of main axon as it is classically conceived, we continue using this tree hierarchy (main axon and first, second order branches; as defined in Section 3.1.2) to describe real as well as simulated axons for simplicity and uniformity, and to allow a direct comparison with other results.

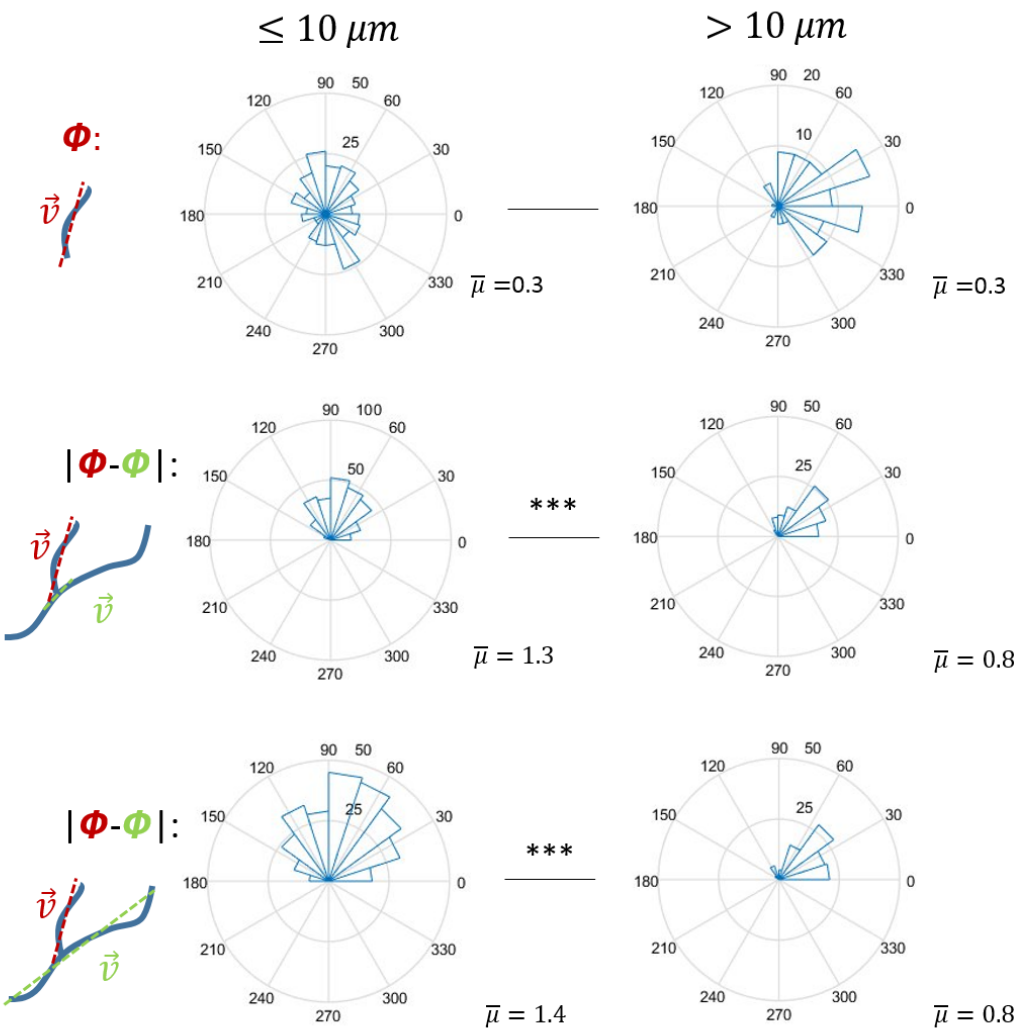

Fig. 4.38 Relevant branch angle distributions from data. Comparison between angles for branches $\leq 10\mu m$ and $> 10\mu m$. (top) Whole branch angle relative to the X axis in the SML (Fig. 4.1) (same information as in Fig. 4.37). (middle) Angle between the entire branch and the main axon locally at the branch point. (bottom) Angle between the whole branch and the entire main axon (equivalent to Fig. 4.40 A). $***$ accounts for $p < 0.001$ from the Kruskal Wallis test.

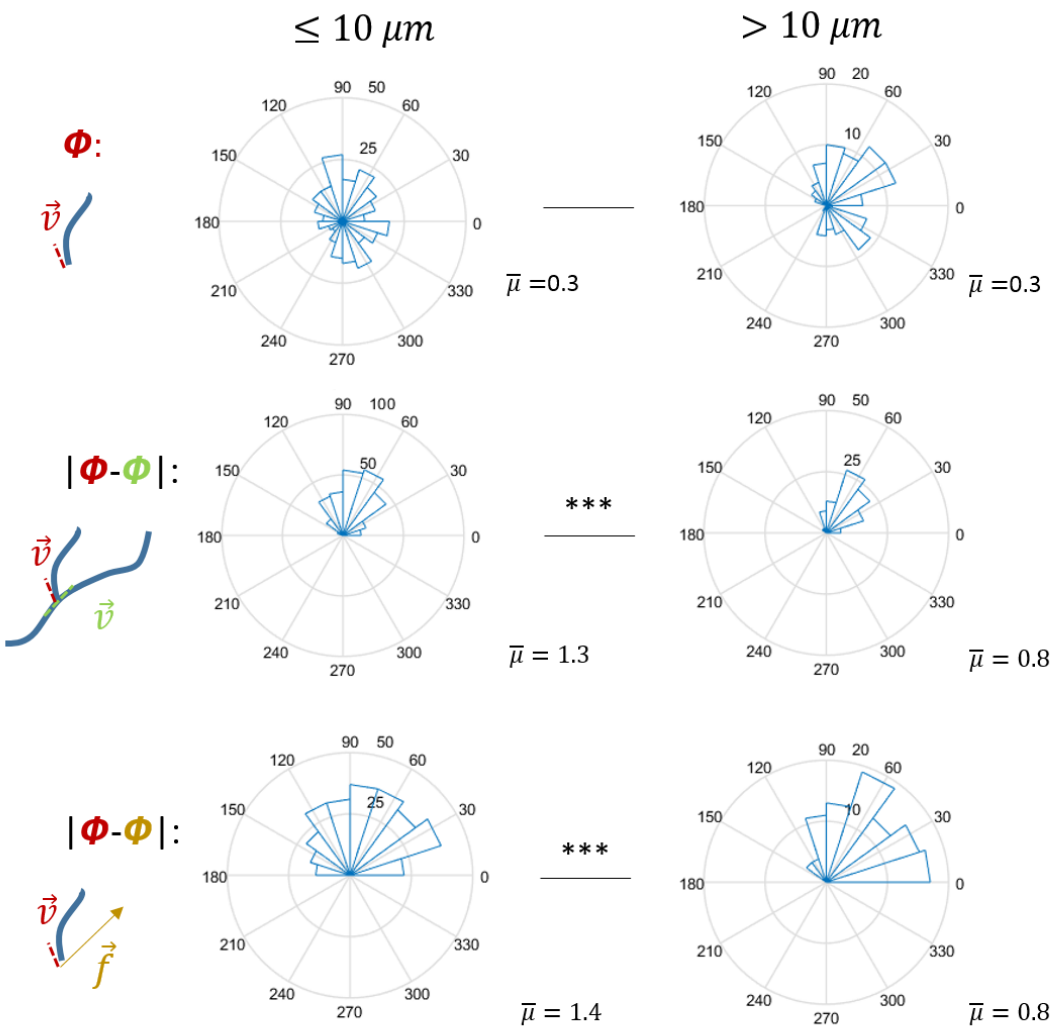

Fig. 4.39 Relevant branch angle distributions from data - bis. Comparison between angles for branches $\leq 10\mu m$ and $> 10\mu m$. (top) Initial branch angle relative to the X axis in the SML (Fig. 4.1). (middle) Angle between the initial branch segment and the main axon locally at the branch point. (bottom) Angle between the initial branch segment and the direction of the external field (Section 4.4). $***$ accounts for $p < 0.001$ from the Kruskal Wallis test.

Together, these observations suggest that type I and type II branches have different properties, and that their growth is controlled *via* distinct mechanisms. In particular, type I branches seem to behave as main axons regarding growth and connectivity, while type II branches may be useful to sense the nearest environment of the axons.

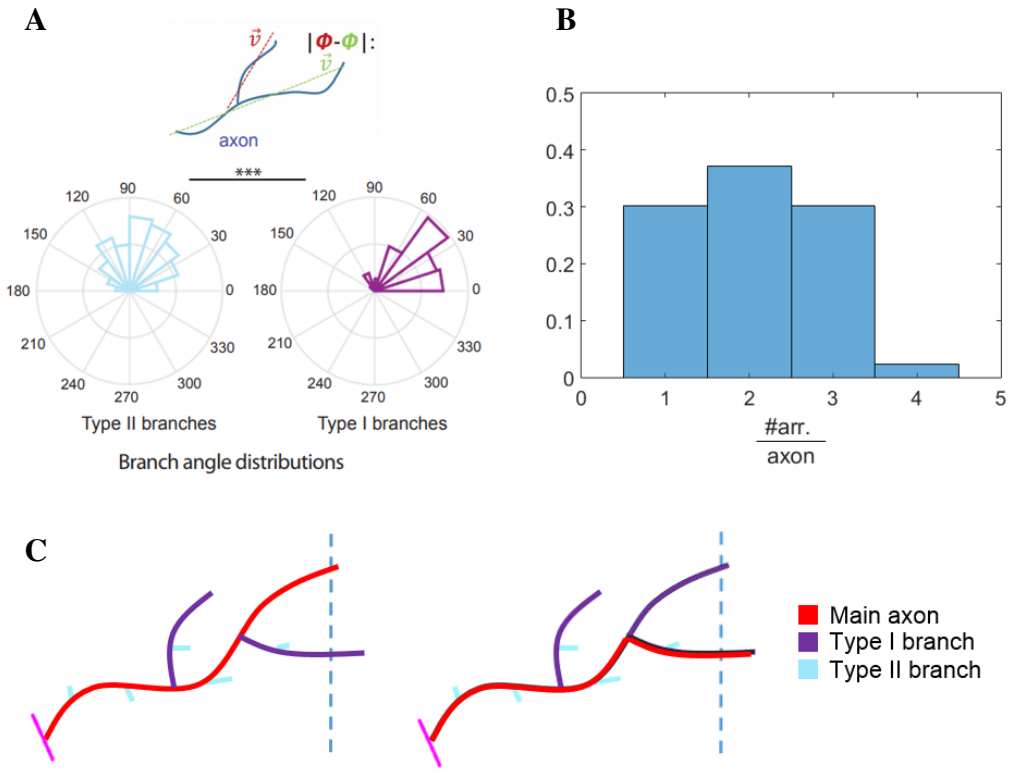

Fig. 4.40 Type I branches behave as main axons. (A) Distributions of the angles between the entire branch and its main axon for type I and II branches. (B) Distribution of the number of arriving tips per axon ($\frac{\#arr.}{axon}$). (C) Two possible segment hierarchy classifications of the same axonal tree. The pink segment indicates the transversal cut of the peduncle (see Fig. 4.22), from where the neurite that originally enters the lobe emerges.

4.5.2 Modelling the branch process

In order to implement the occurrence of type I and II branches in our model, we considered that branching may occur at the end of every time point t_j, after elongation. For any type of branch, we establish that the distance between two branches should respect a Poisson distribution with parameter λ estimated from data ($\lambda = \mathcal{N}(6.2, 0.1)$, Fig. B.7 in Appendix

B). This avoids two branches to be too near, and simulates the need to accumulate biological material before creating a new branch (Ganguly et al., 2016; Szebenyi et al., 1998).

Two scenarios can be considered at the end of each t_j. In the first one, the criteria for the formation of a type I branch is evaluated (*CONDITION* in Fig. 4.41). The distinct considered criteria used in this work are described in the following two sections. A type I branch effectively emerges if *CONDITION* is verified and a random number from zero to one ($rand_{0-1}$) is smaller or equal to the Poisson probability regarding the distance to the previous branch (*DBP* in Fig. 4.41). In the second one, if a type I branch is not placed after t_j, a type II branch tries to emerge. A random distance from the last branch point is thrown from the Poisson distribution ($rand_{Poisson}$) and, if it fits ($rand_{Poisson} \leq DBP$), the branch is placed at that position.

Without any known prior information on the branch emergence distribution, both types of branches initially emerge with a random uniform angle. Type I branches emerge as a segment of two steps and if they fit (*i.e.* no intersection with other neurite) they continue growing following the Markov model. Type II branches measure between 2 and 10 μm (with length distribution following Fig. 4.35B), and appear and disappear randomly during all the simulation until they contact another branch tip or branching point and get stabilized. If they do not stabilize by the end of the entire simulation they disappear. Type I branches grow following the same rules as main axons, and may also form type II branches (but not type I). The *counter* for new type I branches is set to zero (see Section 4.4). As type II branches are quite short and dynamic (and their exact dynamic behaviour is unknown), we did not consider their volume in our model, but only their formation and stabilization.

To analyse simulated and real axons exactly the same way, the hierarchy of the simulated trees was automatically determined after the simulations using the same algorithm as for our data (Section 3.1).

First approach: Random branching

We now need to fix a *CONDITION* (Fig. 4.41) for the creation of type I branches. Unfortunately, we do not have any extra information on the events that trigger branching in γ axons *in vivo*. Thus, we first chose a uniform random mechanism. Therefore, we simulated the birth of type I branches with a certain uniform branch probability, P_b. Thus, at the end of each time unit t_j, the axon will create a type I branch if a random number from zero to one is higher or equal to P_b, and another random number from zero to one is smaller or equal to the Poisson probability regarding the distance (DBP) from the previous branch to a randomly chosen position within those performed during t_j (in Fig. 4.42 the random chosen position is number two, and DBP is indicated).

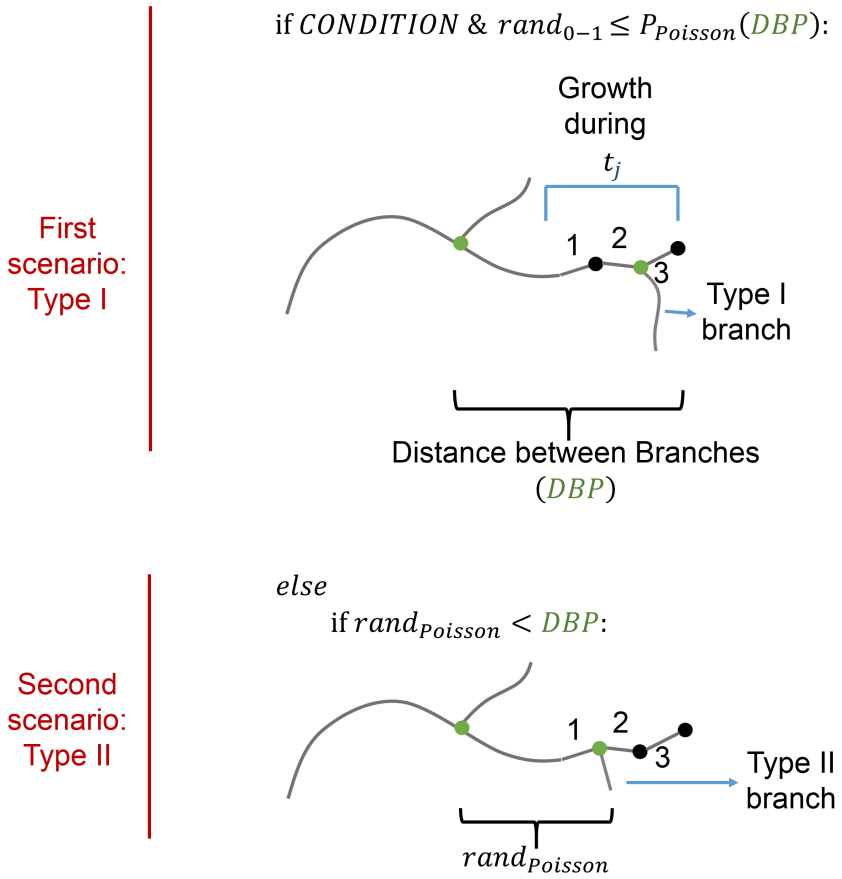

Fig. 4.41 Modelling of the branch process. Schematic representations illustrating branch occurrence. At the end of each time point and after elongation, a type I branch is created if a certain branching *CONDITION* is fulfilled, and a random number from zero to one ($rand_{0-1}$) is smaller or equal to the Poisson probability regarding the distance to the previous branch (DBP). If a type I branch is not created, then a type II may appear if a random number from the estimated Poisson distribuion ($rand_{Poisson}$) is smaller or equal than *DBP*. The new branch is thus placed at a distance $rand_{Poisson}$ from the previous branch.

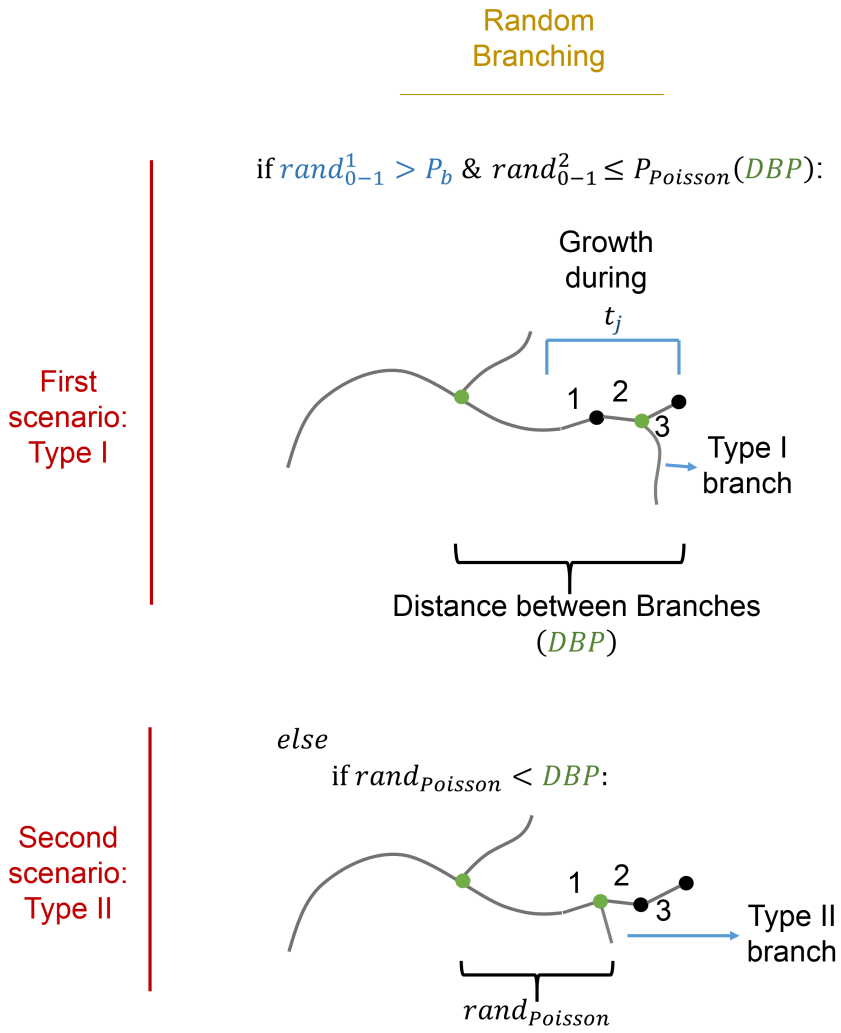

Fig. 4.42 Modelling of the branch process - Random branching. Schematic representation illustrating branch occurrence. At the end of each time point and after elongation, a type I branch is created if a random number from zero to one ($rand^{1}_{0-1}$) is smaller or equal to the branching probability (P_b), and another random number from zero to one ($rand^{2}_{0-1}$) is smaller or equal to the Poisson probability regarding the distance to the previous branch (DBP). The branching point is placed at a step performed during t_j and selected randomly. If a type I branch is not created, then a type II may appear if a random number from the estimated Poisson distribuion ($rand_{Poisson}$) is smaller or equal than DBP. The new branch is thus placed at a distance $rand_{Poisson}$ from the previous branch.

We simulated entire γ populations with different values of P_b. Naturally, varying the branching probability resulted in axonal trees with different number of type I branches. Fig. 4.43A shows the observed mean number of type I branches per axon in function of P_b. The same way as for real axons, we consider as type I branches those with length higher than 10 μm. During the rest of this chapter, we will consider only first order branches. As previously mentioned, second order branches are not only fewer, but also shorter (95 % $\leq 10\mu m$). In addition, the second order branch length distributions results very similar to that for type II first order branches, which allows to consider they are mostly type II.

As mentioned in the beginning of this section, more branches mean higher proportion of occupied volume in the lobe and thus more mechanical constraints for every axon. Also as mentioned, this constitutes a paradox, as real axons present branches and however reach the stopping region.

Strikingly, an increasing number of branches per axon actually increases the percentage of elongated axons in the simulations, rather than decrease it. As observed in Fig. 4.43B, up to around $4\frac{b}{ax.}$ result positive for axon elongation, comparing to zero branch (for which the percentage of non elongated axons is of 15 %, Fig. 4.31) and despite the increase of occupied volume. In particular, the presence of 1.5 to 2.5 $\frac{b}{ax.}$ enhances the overall probability that axons reach the stopping region, reducing the non-elongated axon percentage from 15 % (with no branch, Fig. 4.31) to 9%. Fig. 4.43 C shows the travelled distances for the simulated γ population with $1.5\frac{b}{ax.}$, where 91% of the axons are elongated.

These results suggest that axon branching may be a mechanism promoting axon growth in a population context *in vivo*. As we observed in Subsection 4.5.1, a redundant mechanism seems to take place in the development of γ axons, where branches grow the same way as main axons, trying to reach the stopping region as well. This redundancy gains a new sense if type I branch formation is considered as a strategy to guarantee that *at least* one tip of the axon reaches the stopping region to generate an adult elongated axon, allowing further connectivity. This new hypothesis can only be held considering axon development as a collective space-embedded phenomena, where the neurons interact mechanically with each other in limited space and time. A similar idea is expressed in Francisco et al. (2007): *Increased axon branching (. . .) may allow neurons to extend more branches and thus have a greater probability of establishing an axon.*

Nevertheless, the presence of more than $\sim 4\frac{b}{ax.}$ appeared prejudicial to the growth of the population (Fig. 4.43B), reflecting a dramatic increase in neuronal density and in the probability of encountering mechanical obstacles. Thus, forming type I branches is a good strategy for individual elongation in a population context until a certain limit, from which the excessive fraction of occupied volume wins over the benefits of branching.

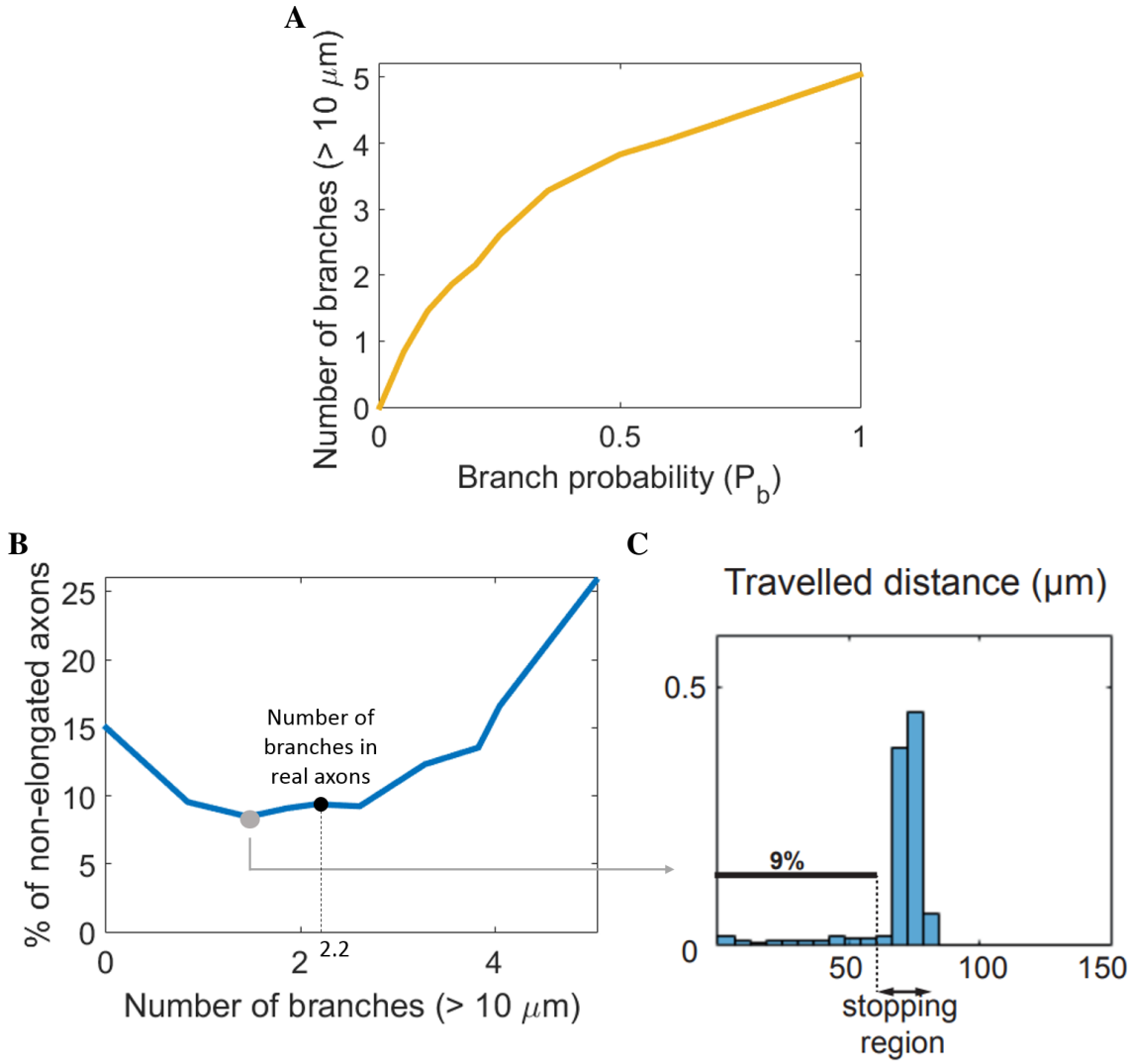

Fig. 4.43 γ population with random branching occurrence. (A) Mean number of branches (>10 μm) per axon in function of the branching probability P_b. (B) Percentage of axons not reaching the stopping region (definition in Fig. 4.2 D) in function of the mean number of branches (>10 μm) per axon. (C) Distribution of travelled distances for the γ population with $1.5\frac{b}{ax.}$ ($P_b = 0.1$), where 91% of the axons are elongated.

Complete model: Branching upon interaction

Following the hypothesis that branching is a strategy to improve axon elongation *in vivo*, the question that emerges is: do branches actually emerge randomly? Or is there any *better* strategy for branch occurrence?

Fig. 4.44 displays a spatial density heat-map of the medial lobe from the registered axons in data. This analysis shows that neuronal density is not homogeneous, peaking in the most central part of the medial lobe and towards the mid-line. More importantly, it reveals that spatial variation in axonal density (number of axonal segments) correlates very well with the density profile of type I branching point (type I branching point number per axon, Fig. 4.44, bottom). The phenomenon is also observed in 2D (Fig. 4.45). As it can be seen, the general branch point density is higher in the central area of the lobe (Fig. 4.45A), and while type II branching points are distributed everywhere in the lobe, including the dorsal and ventral areas (Fig. 4.45B and D), type I are concentrated mostly in the central area (Fig. 4.45C and D), coinciding with the general ML density raise (see Fig. 4.44). The fact that real axons present more (type I) branches where the density is higher, suggests that type I branches may not form with a uniform probability, but rather arise upon encountering a mechanical obstacle and thus improve axonal elongation. Davenport et al. (1999) has already observed, even though *in vitro*, that interstitial branches can be formed after mechanical constrains encountered by the growth cone. Also, Hayashi et al. (2014) showed that actin polymerization regulators (elements relevant for branch formation) are recruited to axon-axon contact sites during the collective axon extension of mice amygdala neurons. They find out that this mechanism facilitates the growth of the implicated growth cones, and suggest that it may also have other functions. Our hypothesis additionally takes into account observations showing that the formation of interstitial branches is decided at the level of the growth cone and occurs after it stops its trajectory. We hypothesize that the axon stops when it finds a mechanical barrier in its way and reorganizes itself to leave a branching point, from where another neurite in parallel (type I branch) will try to overcome the difficulty. The branch will then behave the same way as the main axon, trying to reach the same goal (Kalil et al., 2000), but following a different path.

To test this new branching occurrence hypothesis, we actualized our model such that axons encountering a mechanical constraint would not only try alternative directions until they find an available space, but also create a type I branch (if its distance to the previous branch is more or equal to a random Poisson number with its parameter estimated from data). Thus, the new *CONDITION* (Fig. 4.41) is that the axon grew less than n_{max} steps during t_j (or equivalently, that it encountered two consecutive mechanical constraints, Fig. 4.21 in Section 4.4). The new branch is then placed at the axon tip, as shown in Fig. 4.46.

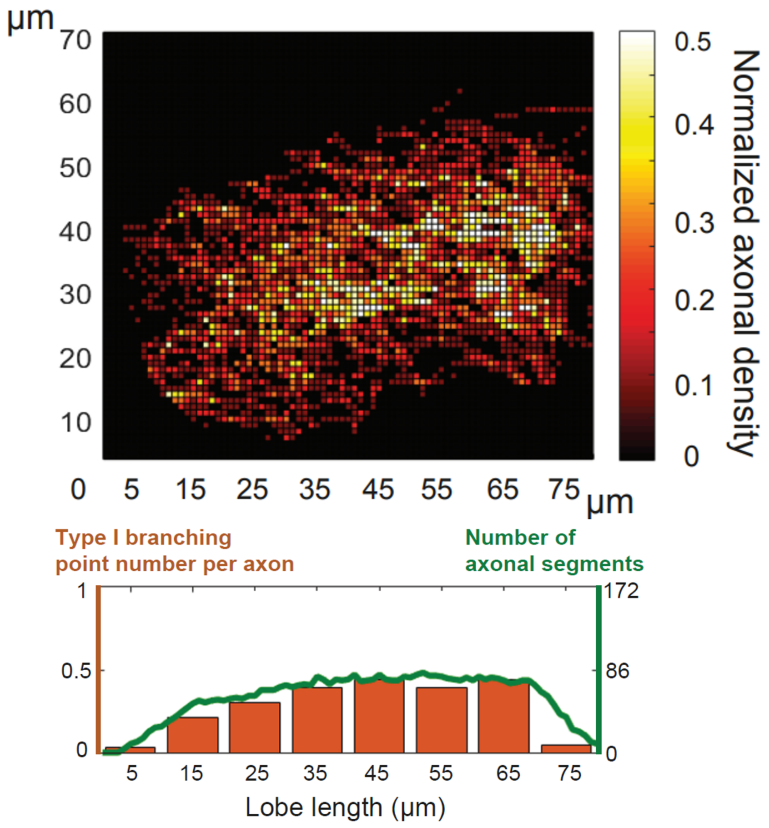

Fig. 4.44 Medial lobe density and branch point distribution. Upper panel: neurite 2D spatial density measured from data (number of neurites per μm^2). The values are normalized to 1, and values higher than 0.5 are shown in the same colour for clarity. Lower panel: Correlation between the spatial distribution of type I branching point number and axonal density. The orange bars represent the number of type I branching points per axon along the lobe axis (see Fig. S6 C for a representation of how the lobe was divided). The green curve represents the number of axonal segments (main axons and axonal branches) found in each lobe region.

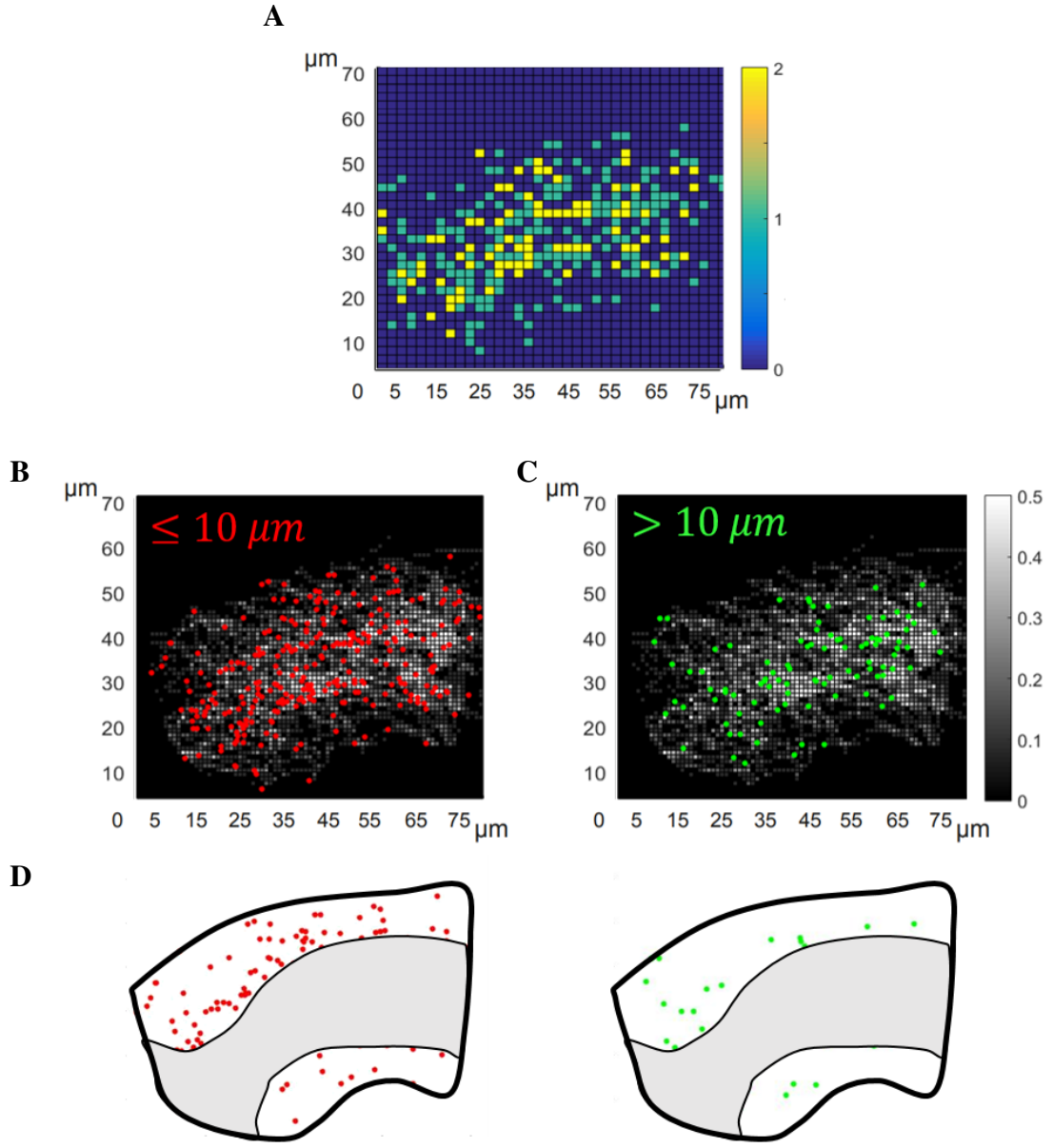

Fig. 4.45 Medial lobe branch point distribution. (A) 2D branch point spatial density measured from data (number of branching points per $2x2\mu m^2$). (B) Type II branching points from data drawn over the ML neurite density map (same as in Fig. 4.44A) for visual correlation. Each red point represents the birth point of a single branch with length $\leq 10\mu m$. (C) Type I branching points from data drawn over the ML neurite density map (same as in Fig. 4.44) for visual correlation. Each green point represents the birth point of a single branch with length $> 10\mu m$. (D) Same type I (right) and II (left) branching points distribution on the ML, where the central area is masked, for a better visual appreciation of the difference between them.

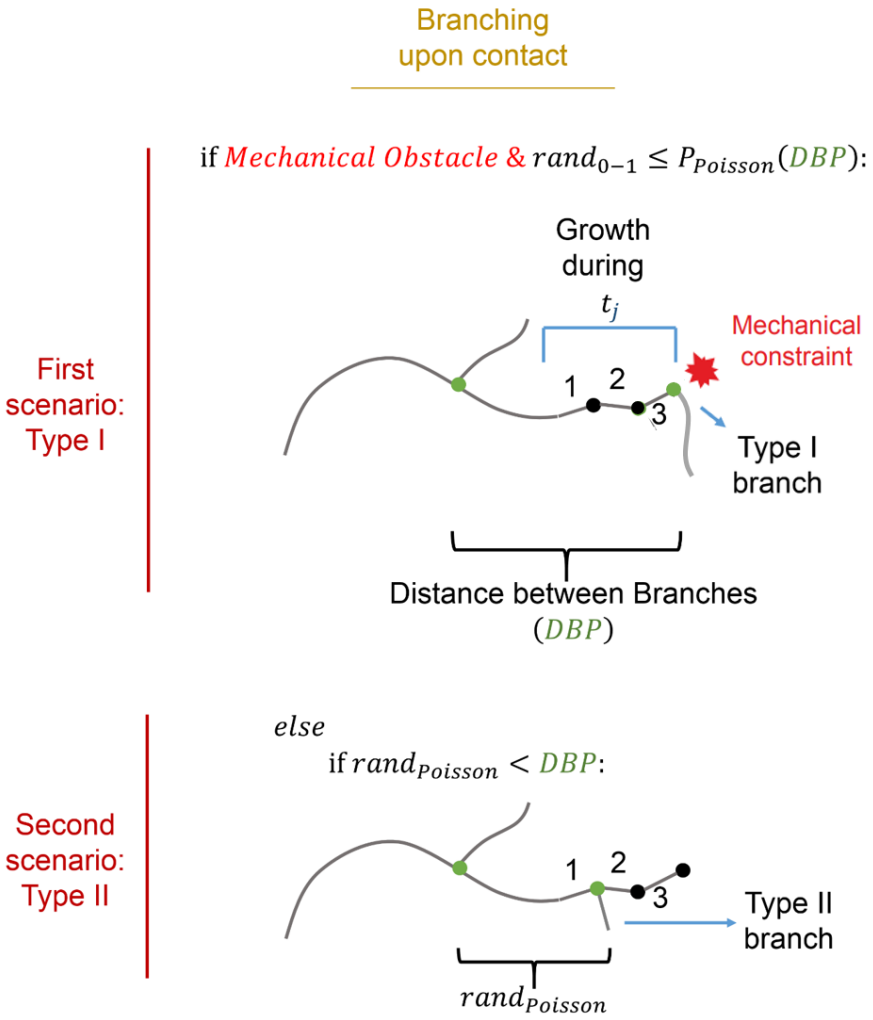

Fig. 4.46 Modelling of the branch process - Branching upon contact. Schematic representations illustrating branch occurrence. At the end of each time point and after elongation, a type I branch is created if the axon has encountered a mechanical obstacle, and a random number from zero to one is smaller or equal to the Poisson probability regarding the distance to the previous branch (DBP). The branching point is placed at the neurite tip. If a type I branch is not created, then a type II may appear if a random number from the estimated Poisson distribution ($rand_{Poisson}$) is smaller or equal than DBP. The new branch is thus placed at a distance $rand_{Poisson}$ from the previous branch.

As shown in Fig. 4.47, axons simulated according to this model have on average 2.0 type I branches along the main axon, a number very close to that of real axons (2.2). Strikingly, the percentage of non-elongated axons is lower than any one obtained with random type I branches, as 95% of axons reached the stopping region when mechanical branches where implemented. This reveals that branching exactly when needed enhances axonal elongation in a population context. This result is not trivial, as it proposes that the percentage of non elongated axons depends not only on the fraction of occupied volume (and other model parameters), but also on when exactly the branches are created.

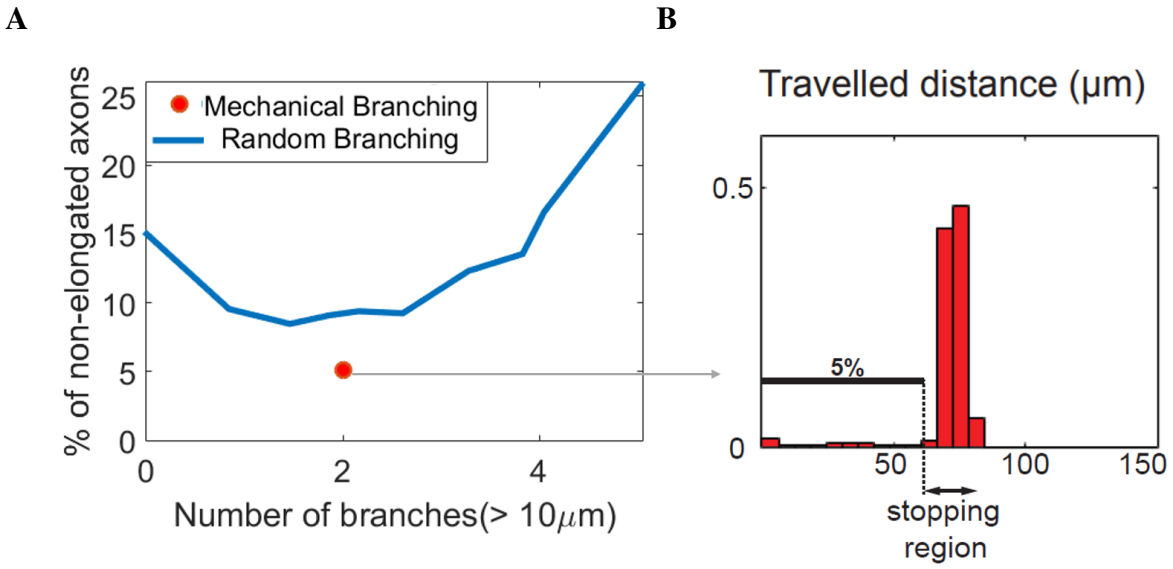

Fig. 4.47 γ population with mechanical branching occurrence. (A) Percentage of axons not reaching the stopping region (definition in Fig. 4.2 D) in function of the mean number of branches (>10 μm) per axon. The sole value for mechanical branching, is represented in red. The blue curve represent the case of random branches as in Fig. 4.43B. (B) Distribution of travelled distances for the γ population simulated with mechanic branching, where 95% of the axons are elongated.

An important question is to know if the simulated axons resemble real ones morphologically. Fig. 4.48A shows examples of axons reconstructed from data and simulated ones, from which the general similarity can be appreciated. Remarkably, the model also generated axons with a wide-range of growth and branching patterns, mimicking the morphological diversity observed for γ axons grown *in vivo*. A deeper analysis on intra-group variability is performed in Section 4.7. To quantify the similarity between simulated and real axons, we used the distance between trees (ESA distance) developed previously in the team (Mottini et al., 2015), and based on the previous work of Srivastava et al. (2005). This measurement takes into account the length, the shape, and the directionality of axons, as well as branching character-

istics. It allows to assign a scalar value to measure the similarity of two tree-like structures. However, one-by-one comparison does not make sense, as we simulated a general population and not mimicked a particular axon. To overcome this difficulty, we then performed the comparison between each axon from the real data and the rest of them, and between each axon of the real data and 43 of the simulated ones (same group size). Fig. 4.48B shows the results of these comparisons in the shape of box-plots. As it can be observed, both cases look very similar, suggesting there is no significant difference between comparing real axons with real axons than real axons with simulated ones. However, when applying the Kruskal Wallis test, the difference comes out significant. Even though this is in part due to the big size of the samples (combinatory of 43), to further understand this difference we observed the distributions in detail (Fig. 4.48C). Nicely, both distributions look very much alike. The sole difference is that when comparing real axons to themselves, the distribution adopts a bimodal behaviour, with a peak in smaller values which is absent when comparing real axons with simulated. This study allows to conclude that simulated axons present realistic axon morphologies, very close to real γ axons (more morphologies in Figs. B.8, B.9 and B.10 in Appendix B).

Fig.4.49A shows the branch length distribution of an entire population of simulated γ axons. Even though the number of type I branches ($> 10\ \mu m$) is very close to that of real axons ($2.0\frac{b}{ax.}$for simulated and $2.2\frac{b}{ax.}$ for real), the number of type II is slightly higher ($8.0\frac{b}{ax.}$ and $6.6\frac{b}{ax.}$). We suggest that this difference could be originated in the later pruning of small branches that γ axons undergo after metamorphosis, thus not included in the model. Tessier and Broadie (2008) observed that axons lose $\sim 19\%$ of their branches in this process, eliminating small ones ($<5\mu m$). This percentage is in accordance with the difference between real and simulated data. Another observation is that type I branches are longer in simulated axons that in real. This suggests that, in reality, there might be other branch termination programs than those we propose in our model.

Fig. 4.49B presents the linear density (*i.e.* neurite number per distance inside the lobe) of the reconstructed lobe from data and the simulated one. Nicely, both cases present similar maximum and minimum values, and corresponding shapes. However, the simulated lobe is denser in the beginning of the lobe, while the real one presents its maximum towards the mid-line. We analyse the origin of this difference in Section 4.8.

Finally, Figs. 4.50 and 4.51 show the different branch orientation distributions, as done previously with real axons (Figs. 4.38 and 4.39). Nicely, in every case, simulated axons present equivalent behaviours as real ones. Whilst for the angle distributions in Fig. 4.50 regarding the entire branch orientation the results are a direct consequence of the implementation of the model (we imposed type I branches to behave the same way as

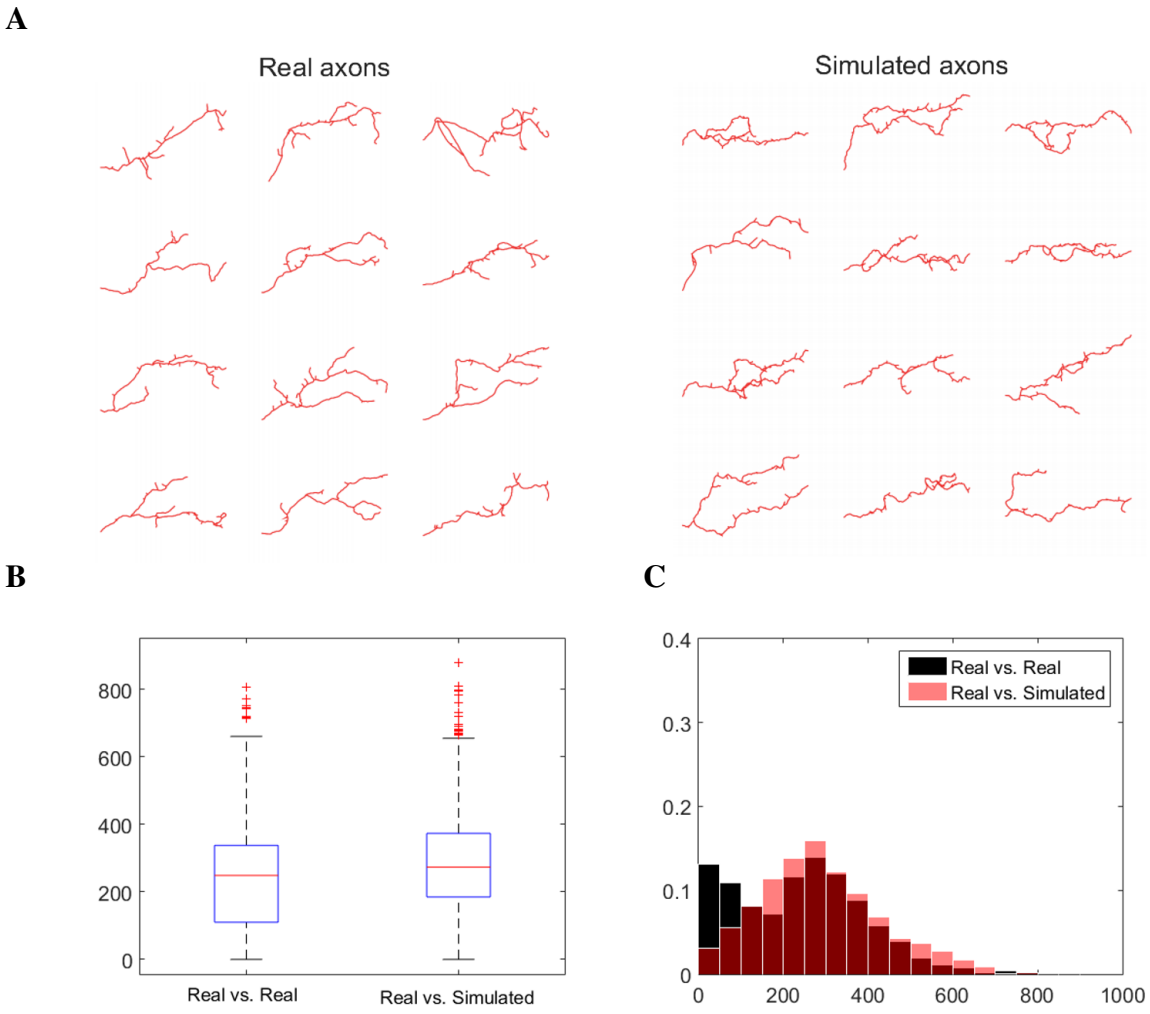

Fig. 4.48 Real and simulated axons are morphologically alike. (A) Examples of real (WT) and simulated axons. (B) Box-plot representation of the ESA distance between trees for real axons vs. real axons and real axons vs. simulated ones. (C) Distribution of the ESA distance between trees for real axons vs. real axons and real axons vs. simulated ones.

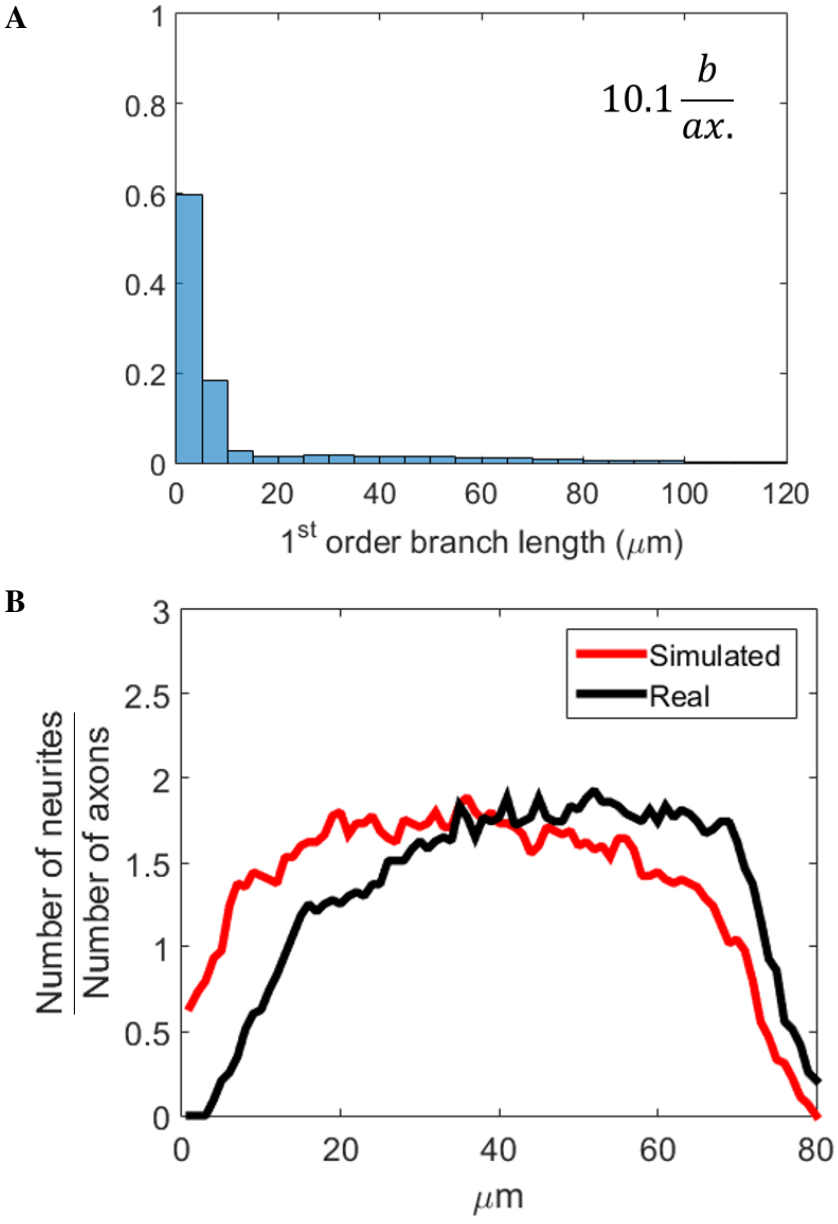

Fig. 4.49 Branch length distribution and linear density of simulated γ axons with mechanical branching. (A) Branch length distribution. In average, axons present 10.1 first order branches (of both types) per axon ($\frac{b}{ax.}$). (B) Linear number of neurites (both type of branches considered) over total number of axons in function of the ML depth for simulated vs. real axons.

the original main axons), the results in Fig. 4.51 on the orientation of the initial branch segments are not trivial. Even though branches are generated with uniform random angles, we observed that while type II branches are mostly orthogonal to their main axons and uniformly distributed in the lobe (Fig. 4.51, middle and top), type I branches had more chances of surviving (*i.e.* find enough free place to be established and continue elongation) when they were born originally oriented as their main axons as well as the external field (Fig. 4.51, middle and bottom). This suggests that the orientations regarding the initial segment observed for real axons (Fig. 4.39) may be a consequence of the mechanical environment and the collective growth, and not pre-defined.

Together, integrating axon-axon interactions in our model generated populations of γ axons with a realistic range of morphologies. Furthermore, it revealed that axon branching impacts on the efficiency of axon growth at the population level, and suggested that branching in response to physical constraints increases the chance that axons successfully reach their final destination.

4.6 Mutant phenotype predictions

Altered branching

Until now, we described a mathematical model and a simulation framework that allows the space-embedded simulation of a whole population of γ axons growing collectively in the ML during metamorphosis. We first analysed the growth of single axons, and observed the impact of mechanical interactions on general axon elongation. Further we added branches and, surprisingly, they did not aggravate the percentage of non-elongated axons due to the natural increase on neurites in the lobe and thus occupied volume, but significantly ameliorated it. This suggested branching as an elongation strategy. Finally, we proposed a branching mechanism that generates long branches after mechanical collisions between axons, with which we successfully simulated complete populations of fully elongated γ axons, morphologically very similar to real ones.

In this section we go further in the analysis and study hypothetical scenarios, where mechanical competition, the branching strategy and/or the Markov chain parameter values are challenged. We first wondered what would happen if the properties of only a single neuron would be altered in a context where the rest of the population grew according to our model. To address this question, we simulated 45 entire γ populations where a single axon is unable to generate type I branches, while the rest of them are normally capable of branching. Fig. 4.52 shows the distributions of main axon lengths and travelled distances recovered in this condition. Not only half of the population failed to fully elongate, but also a

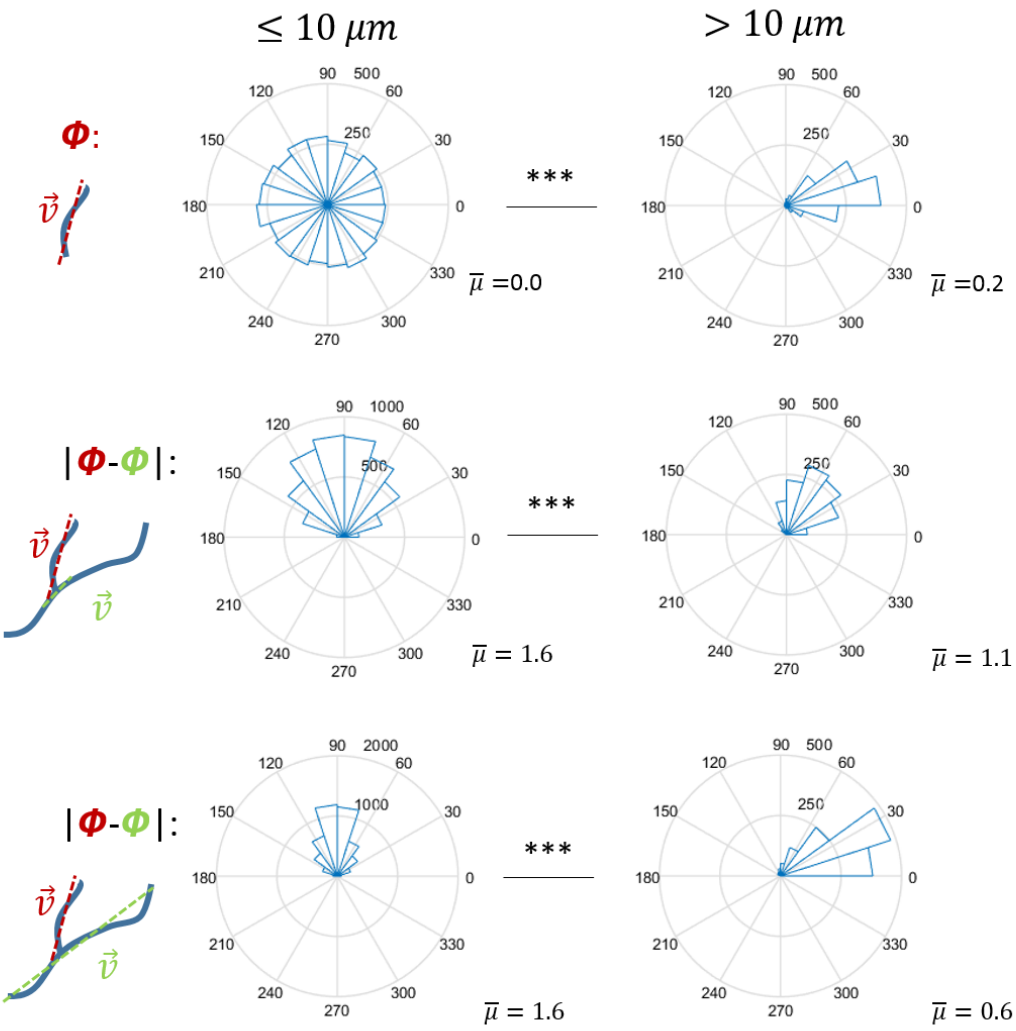

Fig. 4.50 Relevant branch angle distributions from simulations. Comparison between angles for branches $\leq 10\mu m$ and $> 10\mu m$. (top) Whole branch angle relative to the X axis in the SML (Fig. 4.1). (middle) Angle between the entire branch and the main axon locally at the branch point. (bottom) Angle between the whole branch and the entire main axon. $***$ accounts for $p < 0.001$ from the Kruskal Wallis test.

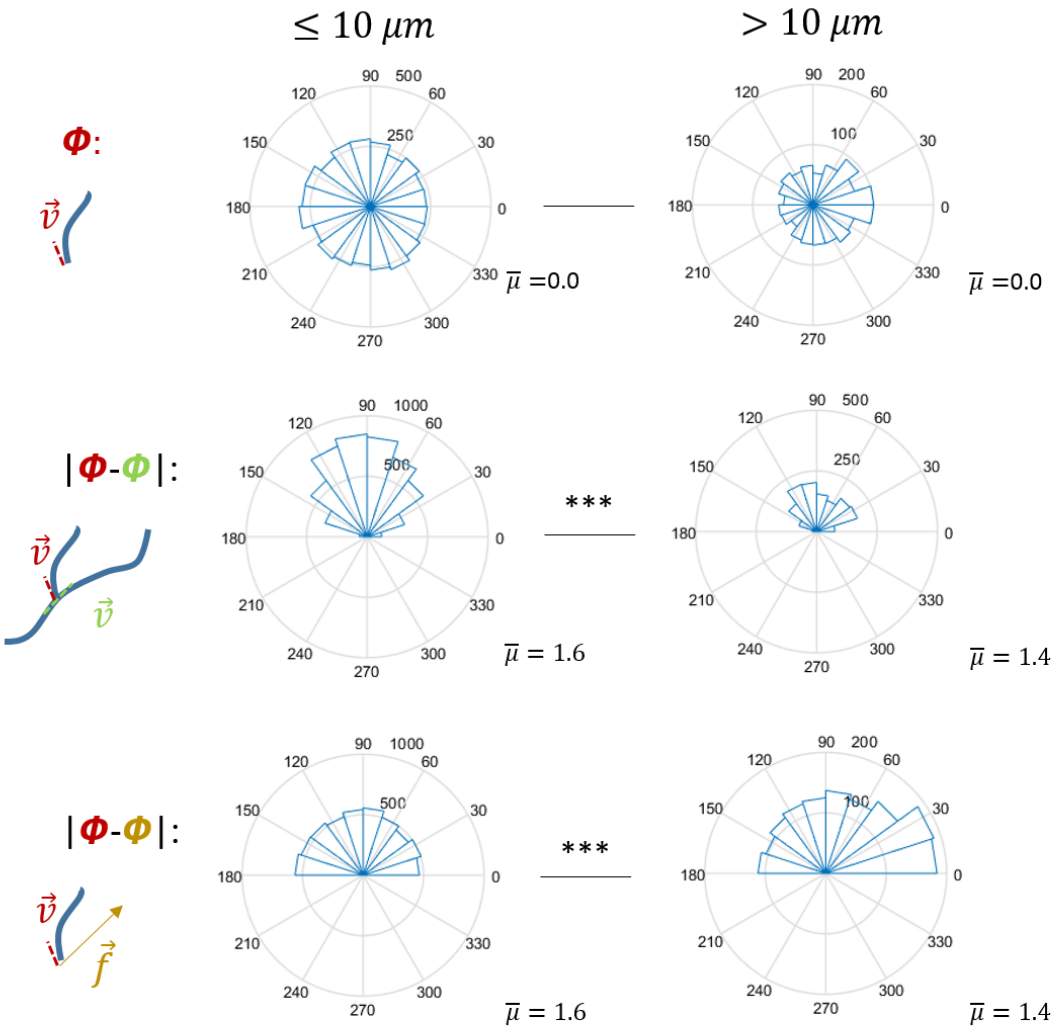

Fig. 4.51 Relevant branch angle distributions from simulations - bis. Comparison between angles for branches $\leq 10\mu m$ and $> 10\mu m$. (top) Initial branch angle relative to the X axis in the SML (Fig. 4.1). (middle) Angle between the initial branch segment and the main axon locally at the branch point. (bottom) Angle between the initial branch segment and the direction of the external field (Section 4.4). $***$ accounts for $p < 0.001$ from the Kruskal Wallis test.

bimodal behaviour emerged in both histograms. This situation is very different to the case in Fig.4.31 where no axon presented branches, as in that case the 15 % of non-elongated axons were distributed uniformly in the histograms of length and travelled distance, and not creating a sub-population. This result showcases the importance of considering uneven competition, where a mutated axon in the context of wild-type individuals reveals a new behaviour, different than when the entire population is mutated.

Thus, a prediction of our model is that single neurons mutant for genes essential for axon branching would fail to efficiently elongate their axons *in vivo*. To validate this prediction, we analysed the properties of Imp axons in our data. As we largely discussed in Chapter 3, the inactivation of *imp* was shown to significantly decrease adult γ axon branching, in particular long branches, and to result in defective axon growth (Medioni et al., 2014; Razetti et al., 2017). We thus studied the distributions of lengths and travelled distances of the Imp axons in our data, normalized to the SML (Section 4.1). As shown in Fig. 4.53, the bimodal distribution profiles of *imp* mutant axons were very similar to that displayed in Fig. 4.52, with about half of the mutant individual axons failing to reach the extremity of the medial lobe. Furthermore, the morphology of reconstructed *imp* mutant axons is very similar to that of individual simulated axons (Fig. 4.53 and 4.52, bottom). Regarding branch length, while for type II branches the distribution and number (5.4 and 6.6 $\frac{b}{ax.}$) is very close to that of wild-type axons, *imp* mutant axons present fewer (0.8 vs. 2.2 $\frac{b}{ax.}$) and shorter type I branches (distributions in Fig. B.11A in Appendix B).

These results thus suggest that branching deficiency may be the primary defect induced by *imp* inactivation, and axon growth defects are consequently due to competition with surrounding wild-type neurons. Once again, the results showcase the importance of considering axon developing as a space-embedded collective phenomenon. Regarding the Markov chain parameter values used in these simulations, as shown in Section 4.3.2, Imp axons present estimated values of α and β that are slightly different than those for WT and Imp Rescue (see Fig. 4.15). Nevertheless, the results in Fig. 4.52 were done considering the parameter values estimated for WT for every axon, included the mutant ones. This choice was motivated by two reasons. The first one is that previous simulations considering different parameter values (*i.e.* wild-type axons with the parameters estimated from the WT population, and *imp* mutants with the parameters estimated from Imp axons) showed no difference in the results regarding the double phenotype. The second one is that the fact of using the same parameter values showcases that this double phenotype can be exclusively explained by the branch shortage, with no need of other aberrant characteristics.

We then performed similar experiments but with single axons that can create mechanical type I branches with different probabilities $P_{bp} < 1$, in a normally branching environment.

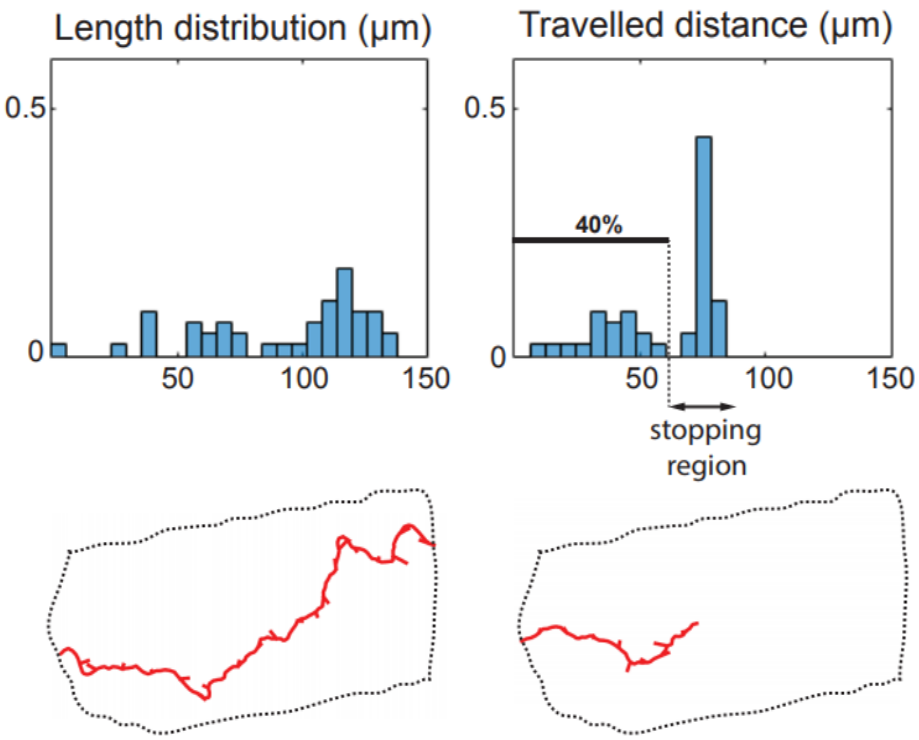

Fig. 4.52 Single non-branching axon in a normal population. Upper panels: frequency distributions of axonal length and travelled distance of single simulated non-branching axons grown in the context of otherwise branching axons. 40% of axons fail to reach the lobe extremity in this condition (n=45). Lower panels: examples of fully elongated (left) and non-elongated (right) simulated single axons.

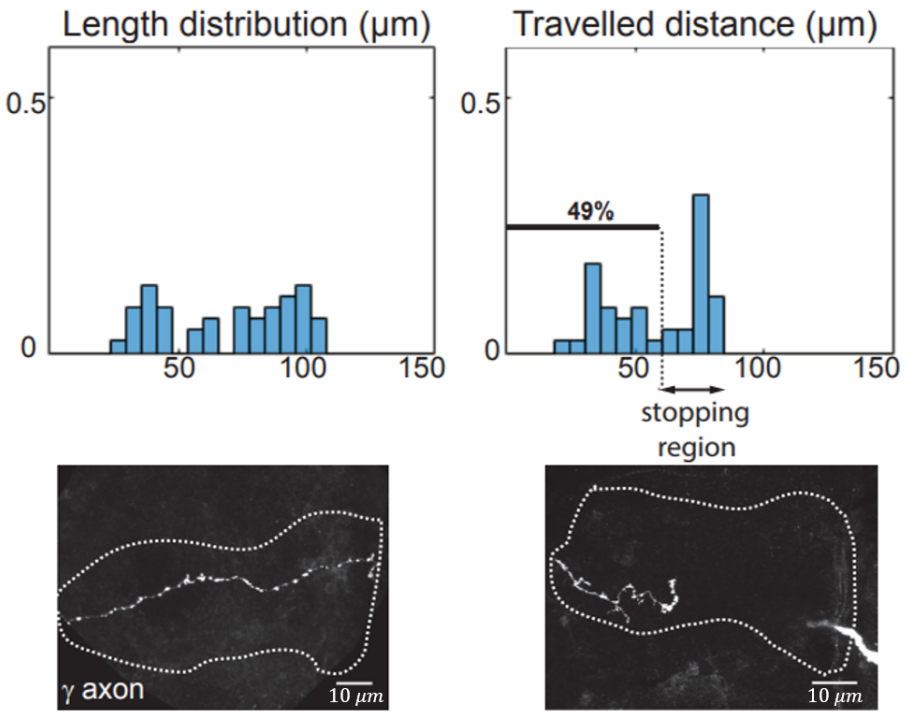

Fig. 4.53 *imp* mutant γ axons. Upper panels: frequency distributions of axonal length and travelled distance of single *imp* mutant axons grown in the context of a wild-type population. A bimodal behaviour similar to that generated by simulations is observed (n=45). Lower panels: example of fully elongated (left) and growth-defective (right) *imp* single axons from our data-base. Scale bar $10\mu m$.

Fig. B.11 B shows the mean number of branches in function of different values of P_{bp}. Naturally, increasing values of P_{bp} diminish the percentage of non-elongated axons. Fig. 4.54 (red curve) shows this percentage in function of the observed number of type I branches. Interestingly, this behaviour is equally observed in real axons. When *imp* mutant axons are rescued with Imp re-expression, the percentage of short axons diminishes to 36 % and when rescued with Profilin to 33 % and, at the same time, these phenotypes display also an increasing number of type I branches (analysis obtained from the corresponding images in our database and showed in Fig. 4.54, black curve). This observations reinforce the hypothesis of branching as an axonal elongation boost, and highlights the predictive aspect of our model.

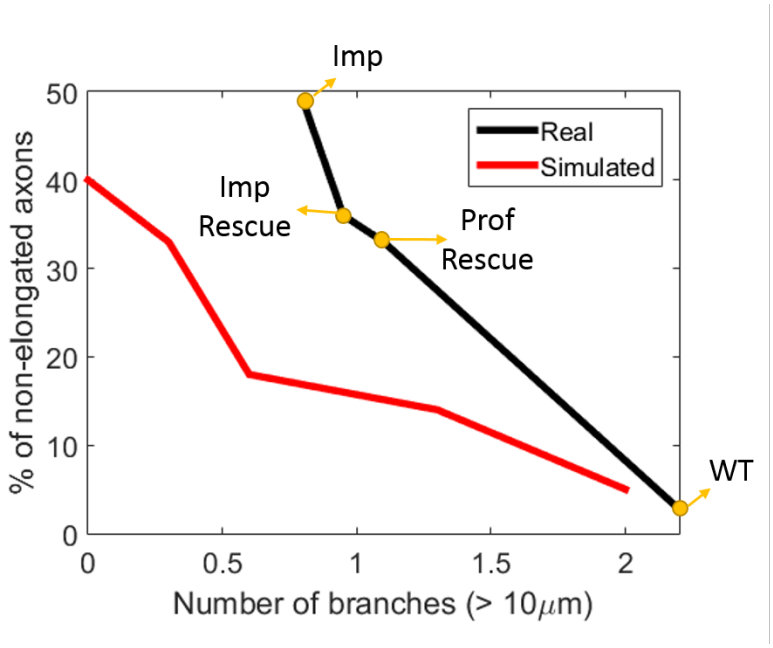

Fig. 4.54 Type I branch proportion influences axonal elongation. Percentage of non-elongated axons in function of the observed mean number of branches (>10 μm) per axon (in red for simulated axons and in black for real ones). Each data point of the real axon curve is labelled with the corresponding group of axons from real data.

Altered branching and guidance

We finally explored the consequences of changing the Markov chain parameter values. For that purpose we considered the *unf* mutant phenotype and *unf* phenotype rescued by TOR over-expression (a pathway known to control growth in general) published in Yaniv et al. (2012). The authors proposed that Unf promotes axonal regrowth *via* at least two pathways, the TOR pathway related to axon elongation and an unknown pathway regulating axon

guidance. They did experiments where 25 % of the γ population (one out of four neuroblasts) is mutated for *unf*, and the rest of the axons are wild-type. The results show that the mutated axons fail to elongate, stopping in the first half of the lobe. In parallel, they present a similar experiment but with the *unf* axons rescued by over-expresion of TOR. In this case the elongation was partially rescued and they observed that the axons occupied only the ventral side of the lobe or the ventral and dorsal, but not central areas. Fig. 4.55 presents examples of both cases, done at Besse Team at IBV (France). We also observe axons growing only ventrally and dorsally and ventrally, but not in the central area. To recreate these experiments, we changed the characteristics of 160 axons out of the 650 γ axons in the ML. We first perturbed guidance by applying a low value of β ($\beta = 0.1$) and leaving the rest of the growth aspects unchanged, to simulate *unf* rescued by TOR. Fig. 4.56A shows time points from a movie of the simulated *unf* mutant rescued by TOR growing axons. Nicely, we observe the same phenotype than described in Yaniv et al. (2012), as axons occupy mainly the ventral and dorsal areas of the ML rather than the centre. In addition, as shown in Fig. 4.33, simulated wild-type axons occupy firstly the central region of the ML and then the sides, in the exact opposite way as the observed phenotype. Then we analysed the possibility that the TOR pathway promotes axonal elongation partly by regulating the branching mechanism. This idea comes from the observation of *Tsc*1 mutants, resulting in the TOR pathway over-activation, that presented an important increase in the number of type I branches ($3.9\frac{b}{ax.}$) compared to wild-types ($2.2\frac{b}{ax.}$). Thus, we simulated populations of γ axons with 160 individuals with $\beta = 0.1$ (*unf* mutant) and different probabilities of (mechanical) branch permission (P_{bp}), to simulate different possible levels of growth rescued by TOR over-expression. Therefore, $P_{bp} = 0$ means *unf* mutant without rescue, where the axons have guidance deficiency and cannot branch. Fig. 4.56B (left) shows this case where, as observed for real axons in Yaniv et al. (2012), the axons stopped growing in the first half of the lobe. When $P_{bp} > 0$, the simulated rescued axons elongate more, but maintaining the phenotype regarding the lobe occupancy. These results suggest, on the one hand, that the lobe occupancy phenotype can be explained by a misguidance and mechanical competition with wild-type axons that occupy the central part of the lobe first. On the other hand, it proposes that the TOR pathway may be directly implied in branching, which results in an elongation mechanism which has not been described yet.

Nicely, the explored cases showcase the biological relevance of our mathematical model, and its capacity to predict the *in vivo* behaviour of mutant axons in a population context and a realistic mechanical environment.

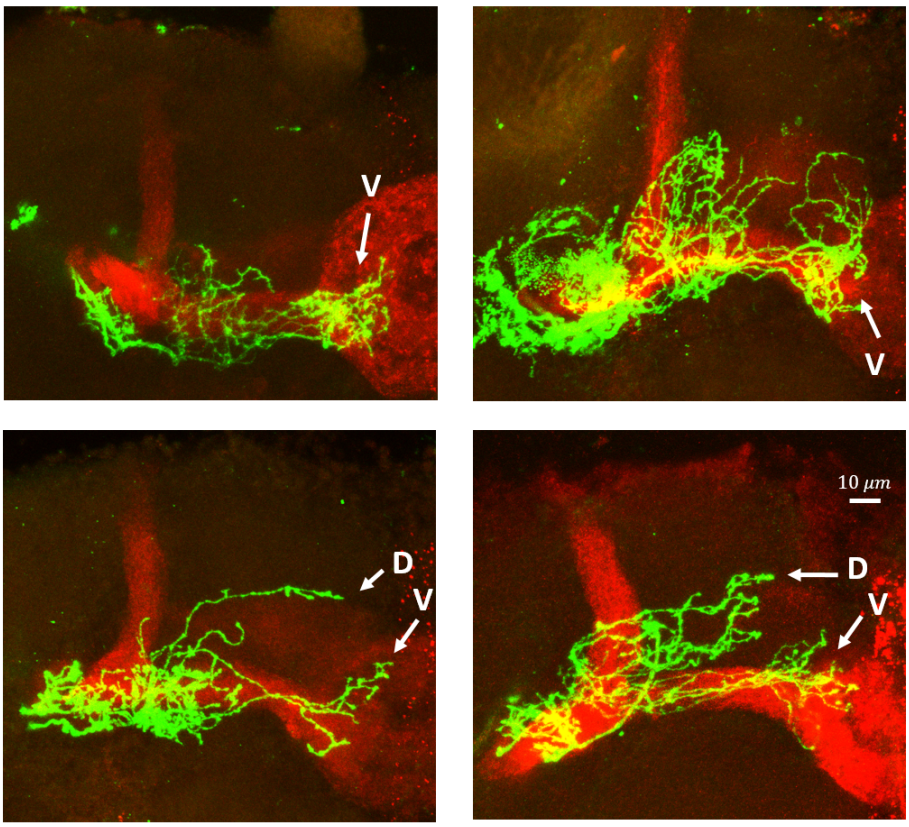

Fig. 4.55 *unf* mutant recued by TOR over-expression. Maximum intensity projections of confocal images depicting *unf* mutants rescued by TOR over-expression that developed in an otherwise wild-type environment (green) and the dorsal and medial lobes of the MB (red). (top) Examples of axons growing only ventrally, and (bottom) exaples of axons growing dorsally and ventrally. Scale bar: 10 μm.

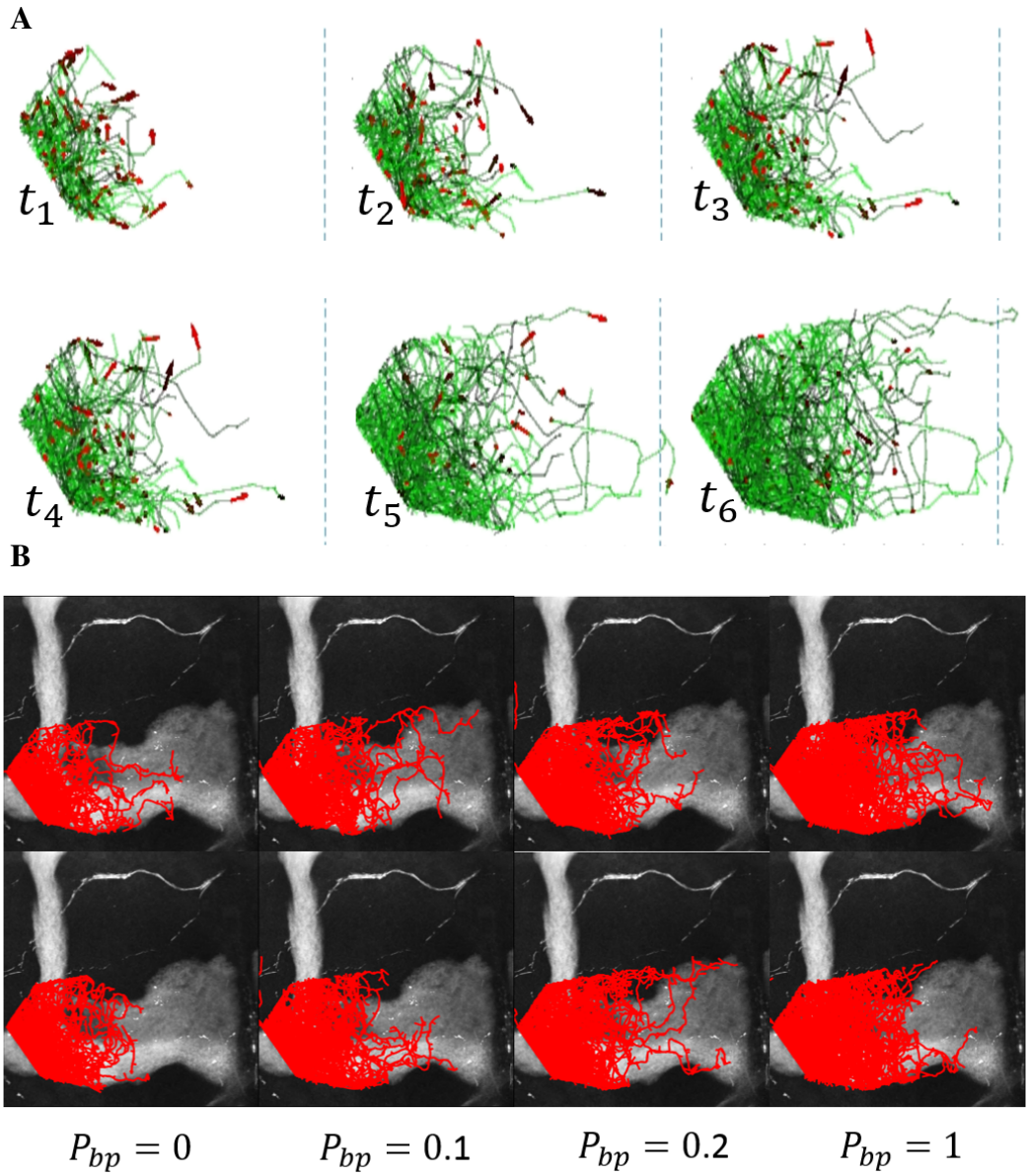

Fig. 4.56 Misguided axons in a normal population. (A) Time points from a movie of 160 simulated axons with $\beta = 0.1$ growing in a wild-type context. Unlike wild-type axons, they occupy preferentially the dorsal and ventral areas of the ML. (B) Results of simulations where 160 of the γ axons are misguided ($\beta = 0.1$) and present different (mechanical) branch occurrence probabilities. $P_{bp} = 0$ simulates *unf* mutant, and the increasing values represent different degrees of phenotype rescue by TOR. Two examples per condition are shown (one in each row).

4.7 Parameter value significance and other considerations

α and β

Equations 4.4 and 4.30 give the expressions of the variance of θ_i conditioned by θ_{i-1} (σ_0^2), and of θ_i (σ_θ^2) in function of the model parameters, α and β. These variances represent the noise (or energy) of the system, and have thus an important influence on the elongation rate of the axons. Trajectories with zero variance will continue straight, following their initial directionality. If axons started ordered and parallel to each other, they would continue that way without interacting.

Axons with noisy steps (high σ_0^2) generate more axon-axon contact during the growth, thus more mechanical obstacles and elongation difficulty. On the other hand, for any value of σ_0^2, if the variance of the chain is high the axonal trajectory tends to expand, generating more collisions with the geometrical limits of the growth cavity (ML). Fig. 4.57 shows how the values of σ_0^2 and σ_θ^2 change for different values of parameters($\alpha \in [10^{-5}, 10^3], \beta \in [10^{-2}, 10^3]$). σ_0^2 reaches high values (disorder region) when α and β are both small. On the other hand, σ_θ^2 is clearly more sensible to changes in β, and presents high values (disorder region) for small values of β, mostly independently from α. In both cases, the values estimated from data lie in the ordered region, near the disorder-order transition (*i.e.* when the noise of the system suddenly falls).

To study in detail how the noise of the system (or equivalently, the parameter values) affects axonal elongation, we first analysed the efficiency of population axon growth when varying the value of a single parameter while leaving the other one fixed, for axons simulated with mechanical branching and without branching. Fig. 4.58 presents the results of this analysis. The variance of the chain goes to infinity and zero when β tends to zero and infinity respectively. As the results show, $\sigma_\theta^2 \to \infty$ means a level of noise that generates total chaos (lack of elongated axons), with any axon reaching the end of the lobe, with or without branches. In both cases (branch and no branch) there is an optimal interval starting in our estimated value of β and until $\beta \sim 10$. For every value of β, axons simulated with mechanical branches elongated more efficiently than those without branching. When varying α, this is not true any more. Strikingly, to generate branches becomes the best strategy only from values of α near to that estimated from data which, in addition, is located in the minima of the curve. These results directly link the axonal shape -thus estimated parameter values- with theoretical elongation strategies such as creating branches and optimal noise of the system. In particular, the value of σ_θ^2 is bounded for variations in α.

To further investigate the influence of the noise of the system in axon elongation regarding axon-axon and axon-geometrical limits interactions, we performed simulations varying α

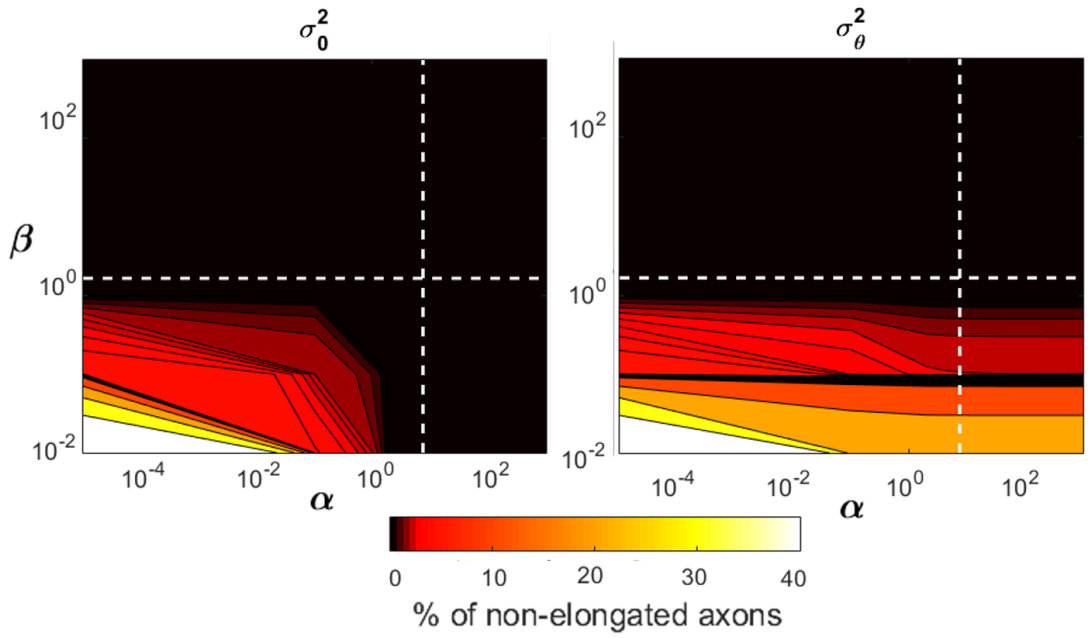

Fig. 4.57 Variance of θ_i conditioned by θ_{i-1} (σ_0^2) and of θ_i (σ_θ^2), in function of α and β (Eqs. 4.4 and 4.33). The dotted lines represent the values estimated from data $\alpha = 7.5$ and $\beta = 1.7$.

and β, with and without branches and within the medial lobe and in a tube (Fig. 4.59). To simulate within a tube we replaced the medial lobe by a tube of transversal area equal to the maximum area in the medial lobe, and a uniform field parallel to X axis (Fig. 4.22). For the case of *in vivo* γ axons within the medial lobe and with mechanical branches (LOBE BRANCH), we observe a closed region of optimal elongation (percentage of non-elongated axons $\leq 5\%$). Interestingly, as observed also in Fig. 4.58, the parameter values estimated from data lie within this region, in its inferior limit. For the case within the lobe but without branches (LOBE NO BRANCH), this region is smaller and located towards slightly smaller values of α. In general, small values of β seriously disturbed axon elongation, in accordance with the observation that the overall energy of the system is high in this condition (σ_θ^2 small, Fig. 4.57). However, for small values of both parameters but $\beta > \alpha$, the elongation in the population context is partially recovered. Thus, for very noisy trajectories, $\beta > \alpha$ induces order in the system by following the external force direction. However, the elongation does not increase constantly with the value of β, but presents an optimal region around $\beta = 10$ (for $\alpha < 10$) and then decays. When the population is simulated within a tube (TUBE BRANCH and TUBE NO BRANCH), this does not happen and from around $\beta = 1.7$ the elongation is total. This means that the geometry of the lobe also plays a role in the collective elongation.

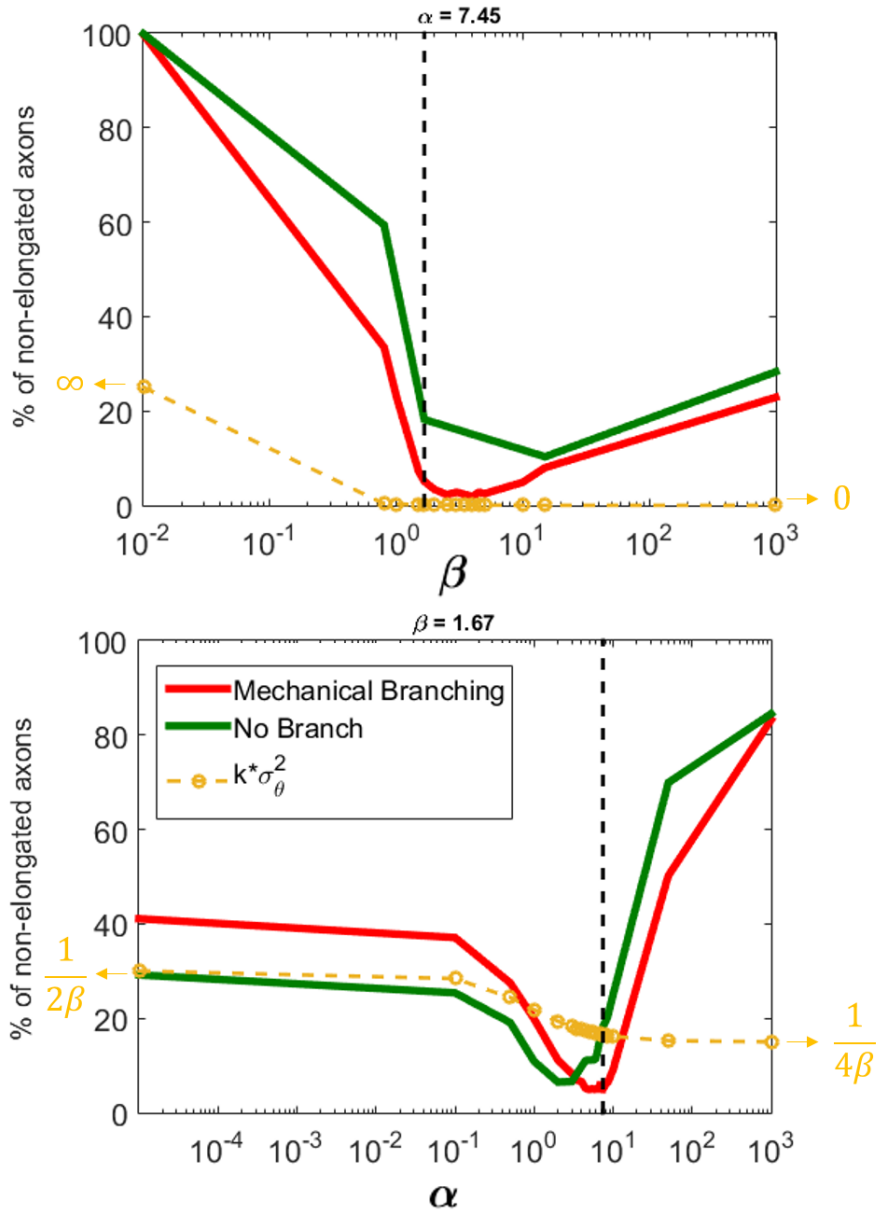

Fig. 4.58 Axon elongation in function of varying parameter values. (top) Percentage of non-elongated axons in function of β and with α fixed. (bottom) Percentage of non-elongated axons in function of α and with β fixed. In both cases, the analysis was done considering mechanical branching and no branching. The shape of the theoretical variance of the chain σ_θ^2 is also shown. $k = 1$ and $k = 100$ for the graphs in the top and bottom respectively. The limits of σ_θ^2 for $\alpha, \beta \to 0$ and $\alpha, \beta \to \infty$ are also indicated.

In fact, high values of β generate very ordered straight trajectories that stop axons from being able to overcome constraints imposed by a non-uniform external geometry. In the case of a tube, axons with low noise follow the straight trajectories imposed by the uniform external field thus avoiding collisions and fully elongating.

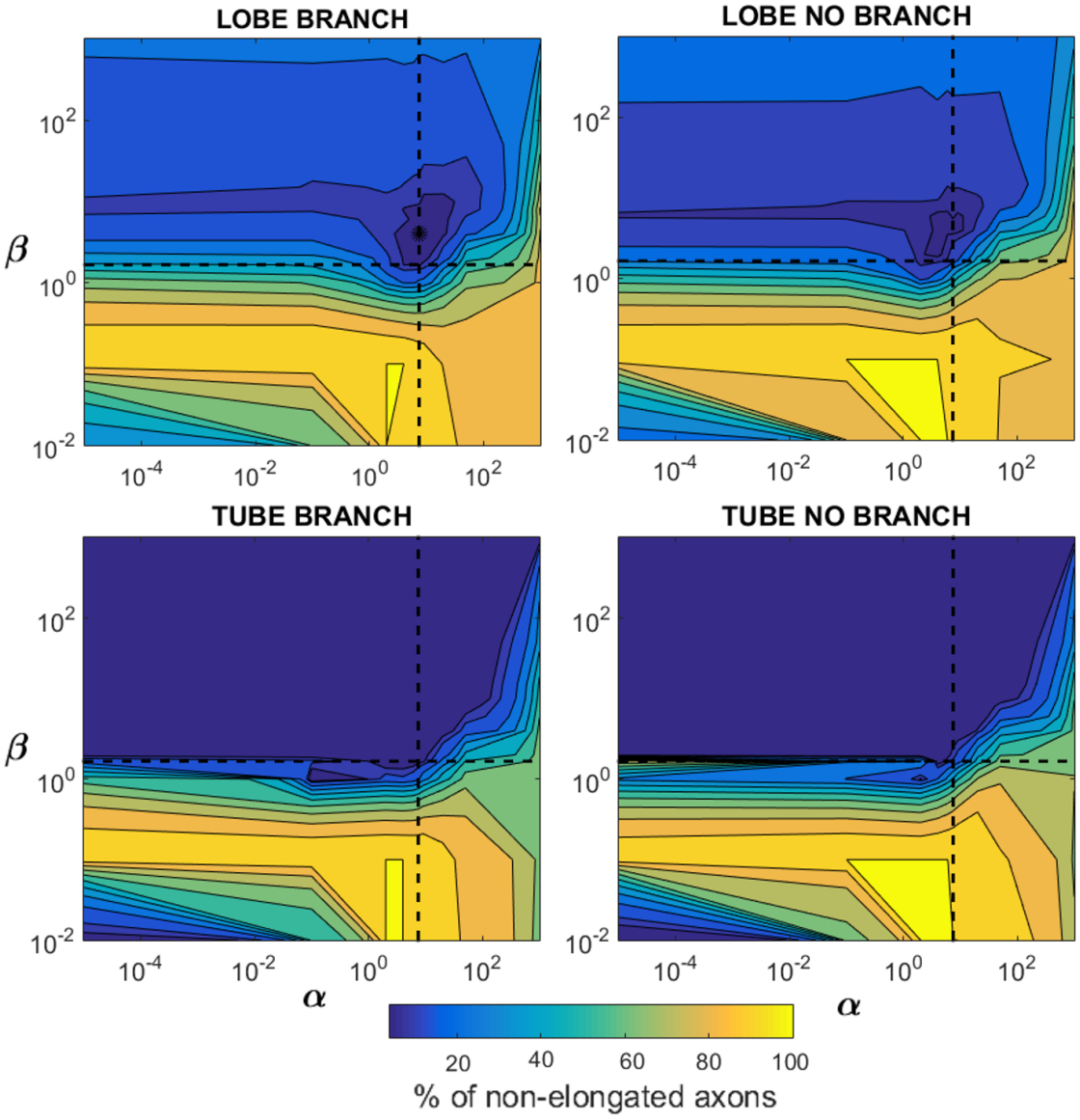

Fig. 4.59 Axon elongation in function of varying parameter values in the medial lobe and in a tube, with mechanical branches and without branches. The dotted lines represent the values estimated from data $\alpha = 7.5$ and $\beta = 1.7$, and the black disc the parameters for which elongation is maximal $(\alpha_{bp}, \beta_{bp})$.

This analysis shows that the parameter values estimated from data assure axonal trajectories with a level of order that avoids, on the one hand, too many axon-axon encounters and on the other hand enough noise to be able to adapt to non-uniform external limits. However, regarding the elongation percentage, they are not the optimal couple. In fact, the lower percentage of non-elongated axons corresponds to $\alpha = 7.445$ (the estimated value) and $\beta = 4$, which we called $(\alpha_{bp}, \beta_{bp})$.

Fig. 4.60 shows a morphological analysis of the γ axons obtained with $(\alpha_{bp}, \beta_{bp})$. Interestingly, we noticed that the $(\alpha_{bp}, \beta_{bp})$ combination, while optimal for axon growth efficiency, generated axons with a reduced complexity, and a lower number of type I branches (1.6 *vs.* 2.2 for real axons), as it can be observed in the examples of Fig. 4.60A (more examples in Fig. B.12 in Appendix B). Furthermore, the population of axons simulated with these values was more homogeneous than the real axons, failing to reproduce the diversity of axon morphologies observed *in vivo*. To quantify intra-population variability, we measured within each group inter-axon similarity (similarity between every possible pair of axons in the same population *i.e.* real, simulated and simulated with best parameters), using the distance between trees developed by Mottini et al. (2015), and based on the previous work of Srivastava et al. (2005). This measurement takes into account the length, the shape, and the directionality of axons, as well as branching characteristics (as done in Fig. 4.48 to compare between simulated axons with mechanical branches and real ones). As shown in Fig. 4.60B, neurons simulated with these parameters were much more similar to each other than were real axons, or axons simulated with (α, β) estimated from data. Fig. 4.60C shows the detail of the distribution of the distances within the populations. While the histograms for real axons and simulated with estimated parameters are practically the same, the one for simulated axons with $(\alpha_{bp}, \beta_{bp})$ presents higher frequencies of smaller values, meaning that they look more alike. Finally, one interesting observation is that these distributions are bimodal, meaning that (for real axons and simulated with estimated parameters) half of the axons look alike and other half is specially different. This suggests the existence two morphological groups within the populations that is nicely reproduced by the model without being particularly imposed.

Together, our results indicate that the *in vivo* stiffness of individual γ axons, as well as their sensitivity to the attractive gradient, are not only well described by our model, but also represent an optimal combination ensuring both morphological diversity and complexity, and efficient population axon growth.

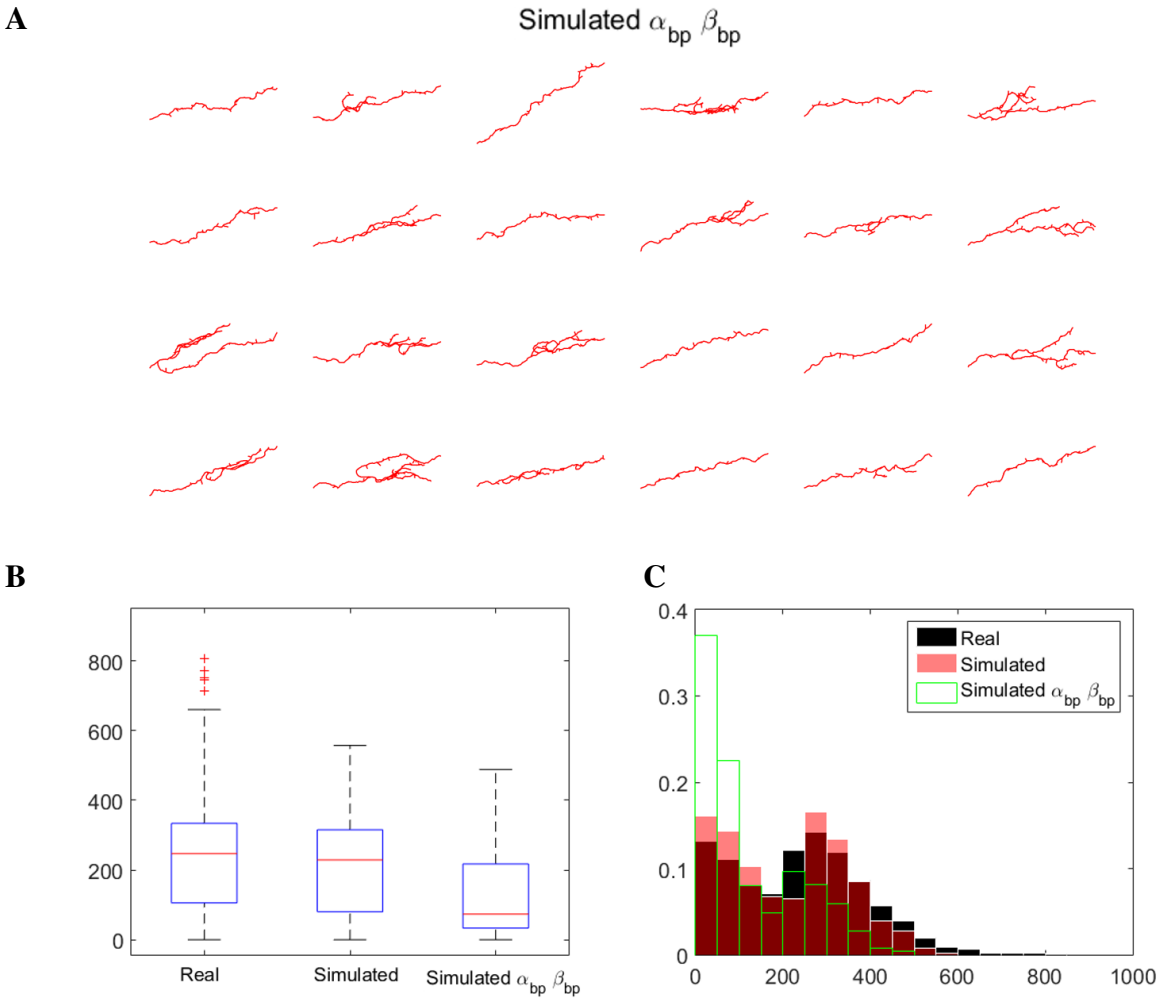

Fig. 4.60 Morphologies of axons simulated with parameters optimal for elongation. (A) Examples of axons simulated with the optimal parameters regarding elongation, $(\alpha_{bp}, \beta_{bp})$. (B) Box-plot representation of the ESA distances within the populations of real axons, simulated with the parameters estimated from data and with $(\alpha_{bp}, \beta_{bp})$. (C) Distribution of of the ESA distance within the populations of real axons, simulated with the parameters estimated from data and with $(\alpha_{bp}, \beta_{bp})$.

Branching mechanisms

Fig. 4.47A shows the elongation rate at the population level with branches that appear randomly at any time point t_j with some probability P_b (random branches) and with branches that appear upon interaction. One relevant observation is that, as a consequence of the design of the growth algorithm, random branches can be placed also after the encounter of a mechanical obstacle (*i.e.* at the end of each t_j, random branches emerge with probability P_b independently from the fact that the axon may or not have encountered an obstacle). Moreover, if a neurite is actually blocked in a certain region (*e.g.* a particularly occupied one), not only during t_j, but also during t_{j+1}, t_{j+2}, etc., the probability of a random branch to appear in that place naturally increases (even though its emergence mechanisms is originally independent of mechanical encounters). To analyse the influence of this effect on axon elongation, we simulated populations of γ axons with random branches with different probabilities P_b but that can appear *only* when the neurite did not encounter mechanical constraints, representing an extreme case (or equivalently, only when it advanced n_{max} steps in t_j, CASE A in Fig. 4.21). For didactic reasons, we call this branching mechanism Anti-mechanic, contrary to Mechanic (branching upon contact), and we rename as Random mix when random branches can appear at any time point (which is equivalent to random branches in Fig. 4.47). Fig. 4.61 shows the percentage of non-elongated axons for the three cases, in function of the obtained number of type I branches per axon. For the Mechanic case, we completed the left area of the graph by considering only mechanical branches but with branching permission probabilities (P_{bp}) smaller than 1 (upon interaction, the branch is created with probability P_{bp}, dotted red line in Fig. 4.61). Naturally, with this condition we obtained less branches than for the case with $P_{bp} = 1$. Similarly, we completed the right side of the graph with mechanic branches with $P_{bp} = 1$ plus random branches at any t_j with different probabilities P_b (branches emerge upon interaction and in any other t_j with probability P_{bp}, discrete red lines in Fig. 4.61). The number of branches in function of P_{bp} and in function of P_b are shown in Fig. B.13.

For the Anti-mechanic case, the lowest percentage of non-elongated axons is of 10%, with $1\frac{b}{ax.}$. For the Mechanic case, the curve shows that the most effective strategy is to place mechanic branches with $P_{bp} = 1$ rather than with smaller probabilities. However, when adding random branches, the percentage of non elongated axons starts once again to increase. One can notice, also, that for an increased number of branches ($> 4\frac{b}{ax.}$) the curve for Random mix and Mechanic get closer, as the mechanically originated branches in Mechanic are overcame by those placed randomly (right side of the graph is completed with mechanic branches plus random with increasing probabilities), in consequence of high values of P_{bp}.

In summary, the three considered cases resulted in curves that are clearly apart from each other, highlighting the importance of the branching strategy for collective elongation.

Interestingly, this analysis shows a difference between the curves of Anti-mechanic and Random mix, stressing the advantage of placing branches upon mechanical obstacles beyond anywhere else. In addition, even though artificial, the Anti-mechanic case serves as a counter situation of the Mechanic one. The observation that for the same number of branches the best strategy is always mechanical branches, followed by mix and in last Anti-mechanic (never mechanic), helps to support the idea that the system is sensible not only to the number of branches (fraction of occupied volume) but also to the emergence strategy, which is a non-trivial result.

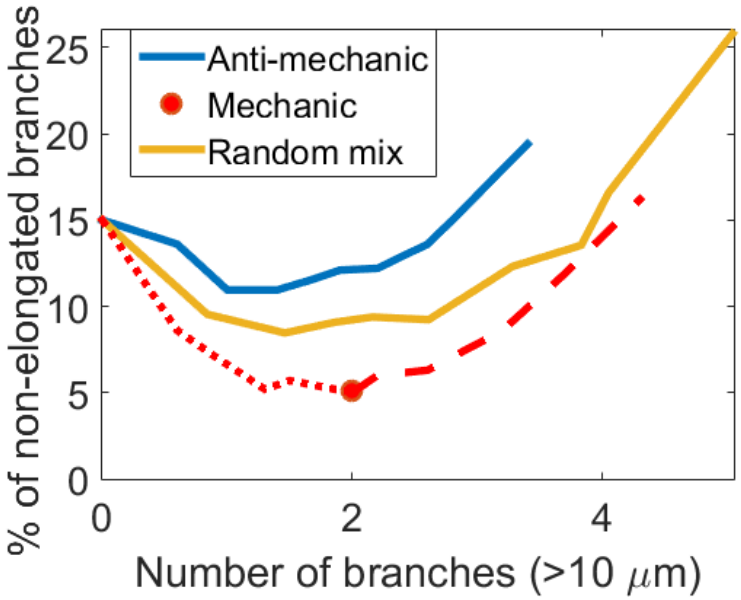

Fig. 4.61 Anti-mechanic branching. Comparison of three different branching mechanisms. Random mix is equivalent to random in Fig. 4.47, knowing that some random branches are also placed upon interaction as a consequence of the design of the algorithm. Thus, Anti-mechanic means random branches that are only placed at the end of t_j when the tip encountered no mechanical obstacle (for reference see Fig. 4.21, CASE A). The red dot represents mechanic branches (the same as in Fig. 4.47), while the dotted lines to the left show the results of simulations with mechanic branches with different permission probabilities, resulting in less branches but always mechanic. The discrete lines to the right of the dot mean mechanic branches with permission probability equal to 1 as in Fig. 4.47, plus random branches with different probabilities after t_j with no interaction, which results in more branches per axon.

Random pauses

Another possible consideration is that the neurite makes random pauses during its growth. In our algorithm, this means that with a certain probability P_{Pause} the axon tip will not grow in t_j.

Fig. 4.62 shows the percentage of not elongated axons for γ populations that pause without making branches (blue full line) and that pause and independently branches upon interaction (red full line). Even though pausing has a negative influence in elongation for both cases, the effects are intensified when axons do not branch ($\sim 10\%$ of difference for small P_{Pause}, and $\sim 15\%$ of difference for higher values of P_{Pause}). One may also consider that random pauses trigger branching, as pausing is classically described as necessary to develop a branch (Szebenyi et al., 1998). Adding branch occurrence when pausing clearly improves global elongation compared to the non-branching cases (Fig. 4.62, blue discrete). However, for populations that already performed branches upon interaction the addition of branches upon pausing has a negative influence on global elongation (Fig. 4.62, red discrete).

To conclude, when axons perform random pauses the optimal strategy for elongation is also to branch only upon interactions. Nevertheless, when pauses are not very frequent branch occurrence also upon pausing does not include significant changes (Fig. 4.62, discrete and full red curves are very similar for small values of P_{Pause}). Thus, mechanical branching as main branch occurrence mechanism plus eventually branch after random pauses is also plausible.

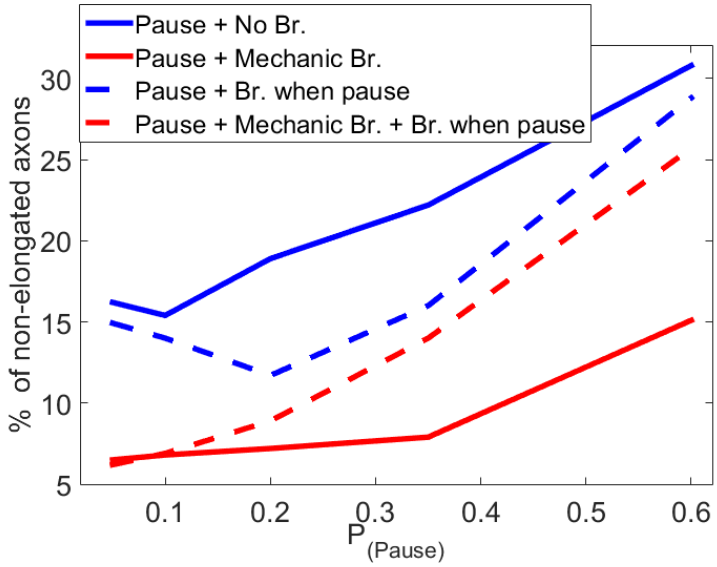

Fig. 4.62 Implementation of random pauses during the neurite growth. Percentage of non-elongated axons in function of the pausing probability P_{Pause} for cases when the axons do not branch (blue full line), branch only upon mechanical obstacles (red full line), branch only upon pausing (blue discrete lines) and upon mechanical obstacles plus when pausing (red discrete lines).

Dynamic parameters: *counter* and time

Fig. 4.63 presents the elongation rates of γ populations under different neurite growth stop mechanisms, for axons that do not branch, that branch upon interactions and do not branch and embedded in a branching population (Imp-like). First, we investigated the case of varying the *counter* value. As it can be seen, the largest differences between the three studied cases appear for rather small values of *counter*. However, for larger values the difference is still present, showing that some neurites get actually mechanically blocked, and would not elongate even with a very high number of trials. The second study considers that neurite growth stop not by a *counter*, but by a fixed number of time-points. Because type I branches follow the exact same rules as main axons, they start their growth with $counter = 0$ (Section 4.5.2). This gives a special advantage to neurons that branch respect to those that do not, as they do not only have more growing tips but also more trials. Interestingly, the time analysis in Fig. 4.63 shows that this advantage is not necessary to make branching an efficient elongation strategy. Even though for long time periods no branching has the same elongation rate in spite of being in a wild-type environment (*i.e.* Imp) or not, a difference is still visible between these two cases ($\sim 4\%$ of non-elongated axons) and mechanical branching ($\sim 1\%$ of non-elongated axons). *In vivo*, axons have limited time to develop completely which, in our model, is represented by low values of time or *counter*. However, the lack of precise temporal information to calibrate our model represents a limit to this study.

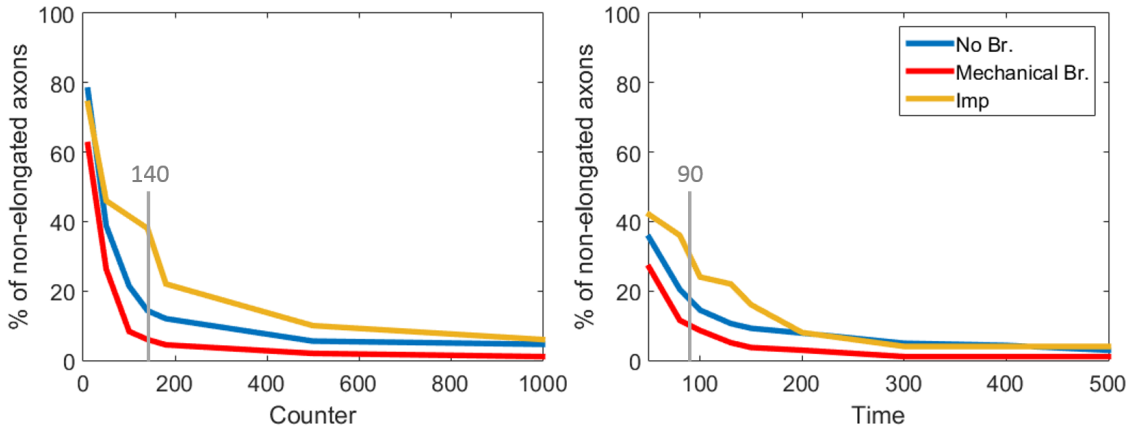

Fig. 4.63 Percentage of not elongated axons for different strategies of growth stop. Three relevant cases are studied: Imp (*imp* mutant *i.e.*, single axons that do not branch in an otherwise branching environment), non-branching axons and axons that branch upon interaction. (left) For different values of the *counter*, where the grey bar indicates the value of the *counter* used in this work, and (right) considering the same time limit for every neurite, where the grey bar indicates the time-point in which typically the 95% of the neurites stop growing in our simulations, considering $counter = 140$ (see Fig. B.14 in Appendix B).

Emergent growth speed

Finally, we analysed the spatial distribution of the growth speed ($\frac{n_{steps}}{\Delta t}$) of the axon tips in the lobe. Fig. 4.64 shows the mean speeds corresponding to the entire population of simulated γ axons that branch upon interaction, that do not branch and that branch upon interaction but representing only the neurites that are totally elongated. Even though it is visually more evident for the entire mechanical branching population, in every of the three cases a *slow* region shows up towards the centre of the ML, and faster regions dorsally, ventrally and distally. Slower regions mean they are also more dense, showing thus that axons initially occupy the centre of the lobe, augmenting the density and thus making growth more difficult, forcing neurites to expand dorsally and ventrally. Also, the end (distal area) of the lobe appears less dense, which makes sense as neurites *in vivo* stop within this region, diminishing progressively its density. An interesting observation is that in the studied mutant phenotypes (Imp, Imp Rescue, Prof. Rescue and Prof.), the axons that stop their growth before the stopping region (non-elongated axons), stop in the *slow* region. The arrows in Fig. 4.64 indicate the non-elongated stopping region and the normally-elongated stopping region, with the "gap" observed in the *imp* mutant phenotype which is also interestingly conserved in the other ones. This result suggests that axons under a disadvantageous condition find it difficult to trespass this dense area, increasing their probability to stop their growth there. However, the lucky ones that trespass it, find no other relevant difficulties (they enter a faster zone) and always reach the stopping region, as wild-type axons do. The comparison with the distribution of growth speed of the elongated segments reinforces the idea that many unlucky neurites get blocked in this area.

4.8 Discussion and contributions

In this chapter, we developed a theoretical and practical framework to simulate the growth of an entire population of space-embedded growing axons, and applied it to simulate the γ neuron population regrowth during metamorphosis. The major novelty in this work is that we considered the mechanical interactions between the growing axons and between them and their confined environment, *via* volume exclusion. Further, an important strength is that we relied close to biological data from *in vivo* and *live* images for the design, calibration, parameter estimation and validation of our model.

The first needed action was to stop considering the single axons in our *in vivo* images as isolated units, but as part of a population. This concerned mostly the normalization of the axon skeletons to the same medial lobe, to be able to fairly compare lengths and elongation as well to analyse, even though artificially, how the lobe is covered (density map) and the

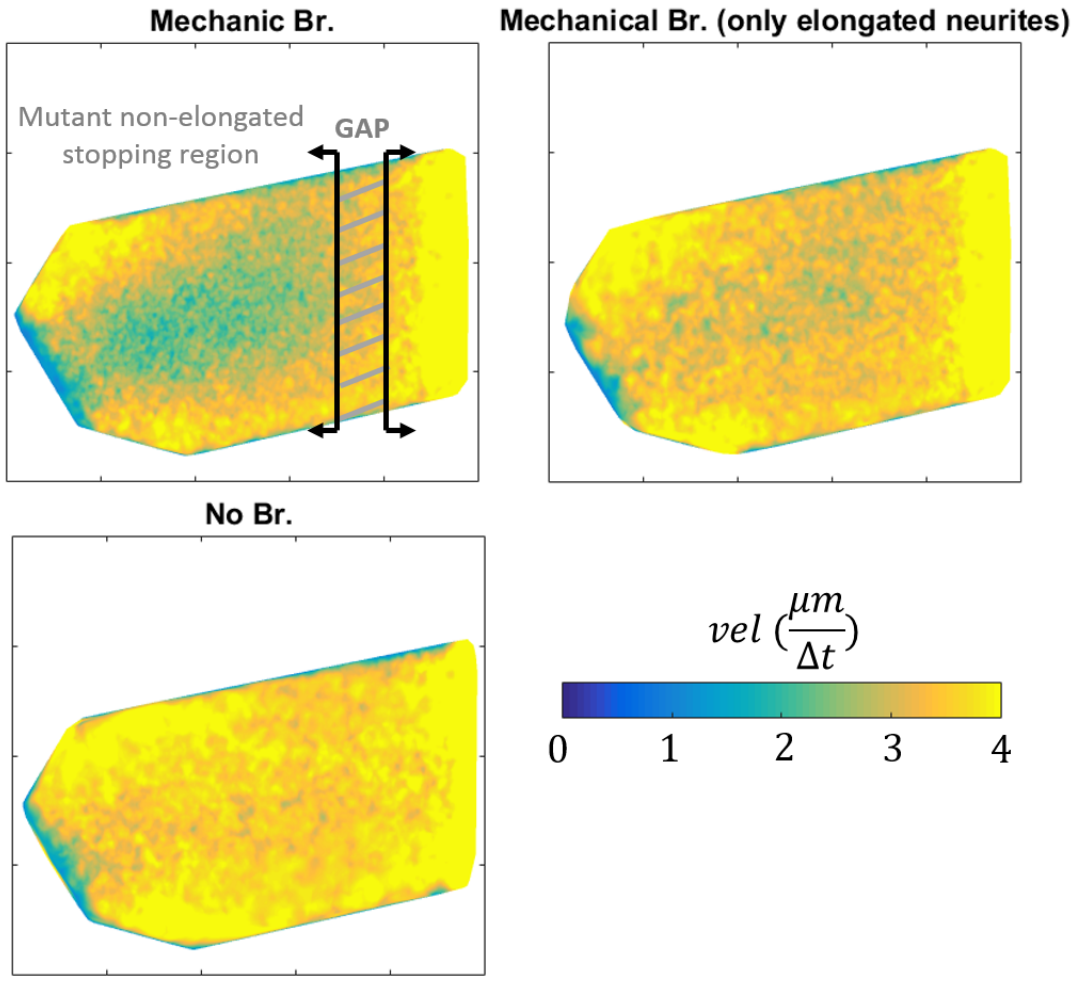

Fig. 4.64 Neurite tip growth speed distribution in the medial lobe. Mean growth speed in function of 2D space for all the neurites in a population of γ axons that branch upon interaction (Mechanic Br.), that do not branch (No Br.) and that branch upon interaction but only depicting the neurites that are totally elongated. The speed is represented in $\frac{\mu m}{\Delta t}$ (one step is one μm in the model). The arrows separate the regions where non-elongated mutant axons stop, and the normal stopping region (as observed in *imp* mutants and in the other studies phenotypes). A gap is present separating both regions.

directionality of the axons within it. This constitutes a first contribution of this chapter, as it allowed for the first time to understand and quantify the stopping region of the axons in the medial lobe, its density, the bimodal elongation distribution of certain mutants and overall axonal directionality.

We then designed an axon growth model where axonal trajectories are considered as the sequential addition of segments following a 3D Markovian model on their angles. This model can be easily embedded into a realistic environment including spatial constraints imposed by confinement, as well as physical interactions between neighbouring axons of the growing population. An important advantage of this model is that it relies on two single parameters (regarding the axon path) that are not only directly linked to the physiological and mechanical properties of individual axons, but can also be estimated from biological samples. Indeed, the Gaussian Markov Chain governing the directionality of each new growth step embeds a first term reflecting the axon stiffness (α), and a second term modelling the attraction towards the external field (β). Classically, the parameters of such models can be estimated from the data with second order statistics. However, because we treat with real biological data presenting noise and jumps in the model, the estimation acquires a higher level of complexity. Thus, instead of considering that the model does not fit the data, we assumed that it actually does but with special considerations. We then proposed that the model is broken when axons growing in a population context hit surrounding cells or other axons. We developed different estimation techniques to overcome this difficulty, and succeeded in estimating the parameter values, which allowed to simulate realistic axonal trajectories.

Another major strength of our model is that it is predictive, and can provide mechanistic interpretations of neuron behaviour. For example, our model explains the wide-range of morphologies observed for MB γ neurons grown *in vivo*. These neurons exhibit axonal trees of variable complexity and architecture, and project side branches of diverse length and tortuosity. Such a diversity contrasts with the very stereotypic axonal projection pattern observed for companion neurons innervating distinct layers of the Mushroom Bodies and reproducibly forming two main straight dorsal and medial branches (Lee et al., 1999). The γ axon variability suggests their growth to be the result of a stochastic process, and thus that each observed adult morphology is one of the many possible outcomes of some aleatory function. This randomness in the morphology generation and intrinsic intra-population variability is probably related to posterior synaptic plasticity and connectivity needs, linked to the complex and dynamic γ neuron role in olfactory memory formation. Another relevant aspect of γ axons is that even though they all reach the end of the medial lobe, some of them do it with only one tip, others with two and up to four. This becomes intriguing as, while for some of them just one is enough, why would others present more tips reaching the end?

This aspect suggests a redundant mechanism respect the arrival to the mid-line. In addition, a general observation done by experts is that the determination of the γ *main* axons is not trivial, and usually subjective to a certain level. For example, in Fig. 2.5, the second axon of the lower row presents two elongated neurites which, a priori, look alike in the sense of hierarchy. Which one of them is the main axon and which one a branch? This question can also be asked for almost all the axons in data. Again from Fig. 2.5, in the first image of the same row, which of the two smaller branches that reach the lobe end is part of the main axon and which one a branch? In general, experts finish up by taking the longest path, and eventually other characteristics like directionality and sense. This selection is generally done visually, and in this work we propose an automatic algorithm to do the task. A final observation is that, sometimes, the selected main axon seems actually as a part of the neurite that originally entered the lobe plus an elongated branch. A nice example is given by the third axon in the bottom row of Fig. 2.5, where even though nobody would doubt to choose as main axon the longest neurite reaching the end of the ML, one gets the impression that, actually, the neurite reaching the end is a branch of the one that initially entered the ML which, on its side, stops its growth somewhere in the middle of the ML.

These aspects inspired us with the idea that the hierarchy of the tree and its connectivity is established after the growth process, during which the initial single axon generates branches that elongate, sometimes reaching the end of the ML. This initial axon -the one that enters the lobe- may even stop before its end and not represent the entire main axon a posteriori. This idea of branches being generated and elongating to reach the end also can explain the difference in number of tips reaching the end of the ML between axons. However, the question rising is: why? Or, more formally, what adversity could have pushed γ neurons to develop such complex an redundant growth mechanism?

This question leads us to the central point of this chapter, which is the consideration of mechanical interactions between axons and their confined space. Even though recent bibliography already started to address the question of space-embedded axon growth, little is known on the relevance of local mechanical interactions on the dynamic growth process and final axon morphology. We thus simulated the entire γ population taking into account these considerations, and analysed their impact. To start, one very first result is that reaching the end of the ML is not trivial. While some *lucky* axons may encounter few difficulties in their way, others may collide frequently, loosing time respect to their neighbours while the ML volume continues to get filled and thus reduce even more their probability of finding a free way. These neurites can get either completely blocked, or partially, meaning that even if there is some free place around it, the time that it would take to found it would exceed the developmental period. This matches very well the observation mentioned before, when the

axon entering the lobe seems to have stopped near its entrance and a branch reaches its end instead. So, this led us to our second important consideration: branches can help axons to reach the end of the ML with at least one tip. If an axon formed (for example) just one branch, this would then double the probability of the axon to have at least one neurite -original axon or the branch- reaching the end of the lobe. Even though this may look trivial, one must remember that, on the other hand, neurites grow in populations. Thus, more neurites means more general occupied volume in the lobe and therefore more mechanical constraints and difficulty to traverse the lobe. To test this idea we included in our simulations the possibility of randomly creating branches, that continue then growing following the same rules as their main axons. We observed that, effectively, branching increased overall elongation at some level. In particular, while having more than zero and until $\sim 2.5\frac{b}{ax.}$ increases axonal elongation at the population level, more branches become prejudicial for the same purpose. This is, naturally, due to the increase of occupied volume that creates mechanical obstacles.

What if forming branches, in addition to help create connectivity, was also an elongation strategy in such a complex developmental environment? If we consider so, may they appear not randomly but when the neurite finds mechanical constraints during their growth? This mechanism may be due, for example, to the fact that neurites are forced to stop their growth upon mechanical obstacles and this pause triggers branching.

The application of this principle in our simulations led to interesting results. Indeed, branching upon interaction actually increments axonal elongation at the population level, and also generates axons with the same average number of branches than real ones. Even though this does not prove the mechanism, it shows that, at least theoretically, it is mechanically plausible. The fact that, in addition, the simulated axons resemble very much to real ones, could be a happy coincidence as well as have more serious implications; but this depends on a subjective appreciation. When we estimated the model parameters from real data, we reached the conclusion that they present jumps in their trajectories, probably due to the interaction with their environments during growth. The hypothesis of branching upon interaction plus the idea that branches are essential for elongation and target innervation (thus that main axons are frequently composed by the original main axon and a branch), finally explains these jump points in the trajectories, and allows to reproduce them with our model. Other kind of interactions may be considered in future work to further explain these jumps.

One major contribution of this work is that our model can predict and suggest mechanistic explanations of mutant phenotypes. In biology, the morphological characterization of phenotypes upon mutations is generally performed by describing a set of relevant measurements -i.e. less or shorter branches than wild-type, existence of non-elongated axons- and, reasonably, unless other experiments are available no more assumptions are done on *why* the particular

phenotype is observed. In this point our work is novel: we propose a framework where with no more information than the axon adult morphology we are able to reproduce and propose an explanation about its origin.

In particular, our analysis allows to suggest, for the first time, a functional link of causality between the shortage in branching and elongation deficiency. To understand the relationship between both, one must necessarily take into account the mechanical interactions between axons in the developing population. In particular, these phenotypes put in evidence the competition for space and elongation between axons. Even though this competition is always present, in wild-type situations it becomes less palpable as all axons finally adopt a functional and correctly elongated morphology. However, in cases of uneven competition like single mutants in a wild-type environment, the existence of this competition becomes clearer as it may leave evidence on the adult final morphology. An interesting example of competition during development in mouse is seen in a model of angiogenesis using embryonic stem cell lines forming embryonic bodies, where endothelial cells compete dynamically for the tip position (Jakobsson et al., 2010). In this model, authors describe the constant and stochastic re-evaluation of the implicated molecular circuit as the cell meets new neighbours, and cell position dynamic shuffling, promoting the formation of a robust network. Interestingly, they studied the impact of a certain mutation on the probability of taking the leadership. Their results showed that even though completely mutated embryoid bodies showed normal vascular generation, when mixing mutated with wild-type cells, mutated ones drastically reduce their probability of occupying the leading position, highlighting the effects of uneven competition. Another similar example of developmental roles between cells being decided dynamically and stochastically (and not pre-defined) is the tracheal branching morphogenesis in *Drosophila* (Ghabrial and Krasnow, 2006). In this model, two different roles exist (lead and subsidiary positions) which are not pre-determined, but established upon competition during development.

Fig. 4.65 summarizes the impact on elongation at the population level of different studied scenarios. The best found strategy for correct elongation is branching upon interactions with a 95% of elongation rate. This 5% of non-elongated axons can fairly be considered as the error of the model, regarding all the approximations, assumptions and simplifications that were naturally done. Another option, however, is to give this (small) percentage of non-elongated axons biological significance. In particular, one possibility is that, *in vivo*, some axons actually do not enter the lobe and stay bocked in the peduncle. If this is the case, we would not notice this non-elongated axons when creating the database and thus do not actually be aware of their existence. Futher studies should be done to investigate this possibility. With almost double the error (9% instead of 5%) follows the case where every

axon produce branches randomly. Similarly, no branching at all triplicates the error of the best strategy,up to 15%. However, the worst scenarios regarding elongation are, naturally, those with uneven competition, where a single axon cannot branch (40% non-elongated axons) or the 25% of the population cannot branch and have guidance deficiency ($\beta \sim 0$), resulting in a total failure in elongation.

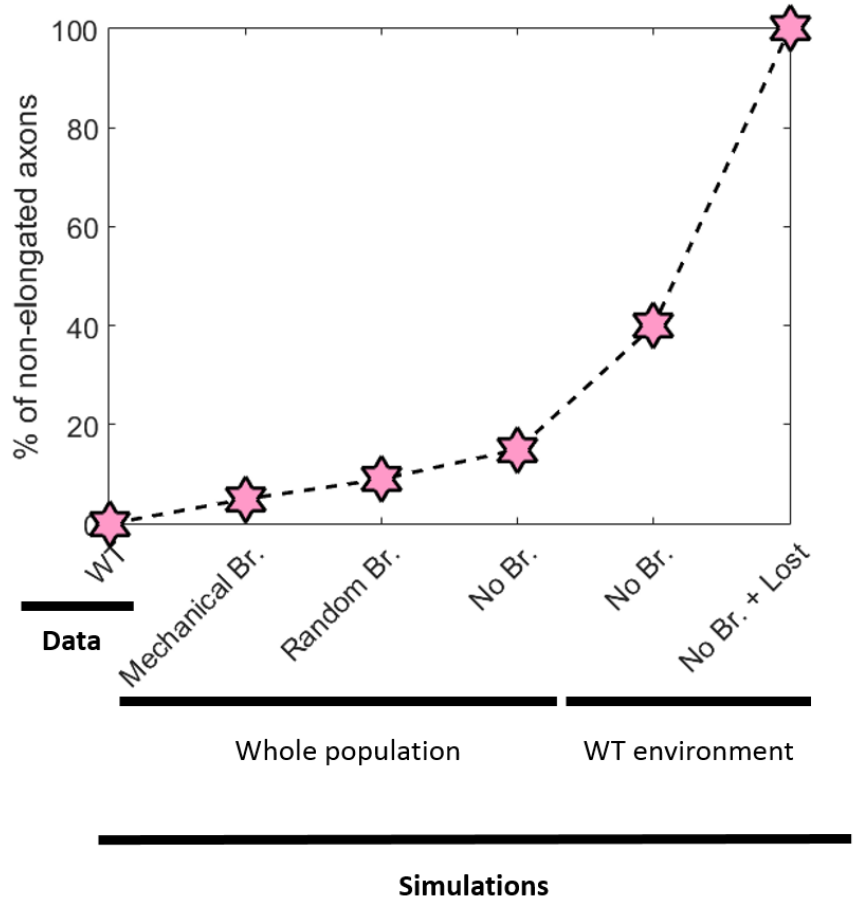

Fig. 4.65 Non-elongated axon percentage for different studied cases. From left to right: real wild-type axons, simulated axons: with mechanical branching, with random branching, with no branching, single non-branching axons in a wild-type environment and 25% non-branching and with guidance deficiency ($\beta \sim 0$) axons in a wild-type environment.

Nevertheless, one may think that branches are created where they are, not by a random process, but by a previously defined specific program. The same can be considered then about the exact place where each neurite (branch or initial axon) stops, as well as its orientation and path. Even though we do not have the tools to fully argument against this idea, we propose a process where randomness operates creating a large variety of morphologies, and

that connections between the created neurites as well as between them and other populations are created a posteriori. The proposed philosophy does not pretend that branches have as sole objective to guarantee axon elongation, but they are needed for connectivity as well and the role that each neurite will fulfil in adulthood is decided after development, depending on the result of the generative process (*i.e.* if the neurite reached the end of the medial lobe, or not, exactly where it ends, etc.).

Previous work has already established connections between axonal branching and mechanical properties of their environments. Flanagan et al. (2002) observed mouse spinal cord primary neuronal cells developing *in vitro* in substrates of varying deformability. In their experiments, neurons grown on softer substrates formed significantly more branches as those grown on stiffer gels. These results show that the mechanical properties of the substrate specifically direct the formation of branches. Koser et al. (2016) also observed *in vitro* and *in vivo* that axons branch more in softer tissues. In addition, they describe that, when growing in less dense environments, axons navigate more and inter-cross themselves more. This can be related with γ neurons, as in the peduncle they are tight and grow parallel to each other until they reach the medial lobe, where they have more available space and generate a large variation of morphologies in a heavily inter-crossed network. However, both works consider the density of the environment where they are embedded in, and not the density generated by the axons themselves. This is why we relate these studies with γ axons regarding the difference between the peduncle and the medial lobe, but not within the lobe. In addition, this would contradict our hypothesis that branches occur after mechanical interactions, as they both observed that axons branch more in less dense environments.

Interestingly, the diversity in γ axonal trees observed *in vivo* is well reproduced by our model. Part of the variability observed after simulation is explained by intrinsic factors, including the stochasticity of the axon growth process modulated by the mechanical and chemical properties of the axons, which is traduced by the model parameters values. Increasing axon sensitivity to the attractive field, indeed, generated populations of neurons with reduced morphological heterogeneity, suggesting that differences in intrinsic properties may partly explain the various degrees of morphological variance observed *in vivo* in different populations. Interactions with the surrounding environment also largely contribute to the variability in axon morphology observed in our model. In particular, mechanistic constraints imposed by other γ axons growing synchronously and competing for space define final axon paths. In addition, neuron-specific contacts established stochastically within the population impact on the formation of side branches and the final morphology of axonal trees, generating further variability. Remarkably, previous work has shown that, like mammalian brain structures, MBs exhibit a unique degree of flexibility in their organization, with neurons

establishing plastic synapses and receiving unstructured rather than stereotyped inputs (Caron et al., 2013; Cassenaer and Laurent, 2007; Murthy et al., 2008). Thus, establishing a dense network of non-stereotypic axonal branches may be an optimal strategy for MB γ neurons to perform their described integration function, and in particular contextualize novel sensory experiences to provide adapted output behaviour (Heisenberg, 2003; Hige, 2017).

Looking closer to the exact value of the model parameters estimated from data, we found out that they are very near to optimal theoretical values. Furthermore, as mentioned, they generate populations of axons morphologically very similar to axons grown *in vivo*, highlighting the biological relevance of the model. Interestingly, our study shows that the exact values of the parameters estimated from data provide the axonal trajectories with an energy sufficient to recreate variate morphologies, as well as to be able to find alternative paths when they are blocked. In addition, they insure enough collective order to allow the correct elongation of the entire population.

For simplicity (and lack of extra information), we considered the parameters to be constant through the population, as well as in space and time. However, this assumption may not be true and other possibilities should be further explored. To analyse possible variabilities between axons within the the same population, we performed simulations of entire γ populations where the parameter values of each axon where stochastic and thrown from different distributions (Bayesian approach). Some of the proposed distributions went from simple independent uniform distributions for each parameter, independent Gaussian distributions, bivariate Gaussian distributions, etc., all of them adapted from the parameter estimated surfaces in Fig. 4.15B-D. However, we could not observe any evident difference with respect to fixed values, neither in elongation, branch number or general axonal morphology. Further, we performed similar experiments but including a single axon that cannot branch (*imp* phenotype) with the intention of finding a correlation between the short phenotypes and the associated value of parameter of each mutated axon. However, once again, we did not observe such a correlation and thus no evidence of the impact of the parameter variation. Even though the estimation of the parameters axon by axon results in different values for each of them, this can be explain by the nature of the model itself (the axons lengths and the presence of jump points) and does not necessarily mean that axons present intrinsically such difference. Thus, even though from data we cannot directly answer if there exist parameter value variation within the population, simulations suggest that even if it existed (until some point), it would not impact correct development.

Other possible variations are on space/time. Fig. 4.66A shows the parameter values estimated (using Eq. 4.34 for the whole population) in function of X (the axon position along the lobe) with a sliding window of $5\mu m$. The curves are quite noisy and the exact values

cannot be completely trusted as they are obtained from a very small sample. However, the curves suggest a decrease of the parameter values along the X axis (*i.e.* along the ML). Fig. 4.66B shows the parameter estimation surfaces (Section 4.3) for WT, Imp and Imp Sh from data. In addition, we calculated the estimation surface for only the first halve of WT and Imp L. While Imp Sh axons present a value of α slightly smaller than considering Imp altogether and WT -suggesting that parameters may be smaller in the beginning of the ML- the first half of WT and Imp L axons present an elevated value of both parameters. From the analysis from WT data we observe that parameters seem to be higher in the entrance of the ML and lower afterwards, in accordance to Fig. 4.66A. Another apparent evidence of the decrease of the parameters values along the lobe, relies in the neurite density distribution of the ML. Fig. 4.66C (top) shows the density heat map obtained from real axons, (middle) for simulated axons with fixed parameters and (bottom) simulated axons with β decreasing along the X axis. While the density increase towards the centre of the ML is well respected by simulated axons with fixed parameters, real axons also present a peak distally that is not reproduced by these simulations. However, when decreasing β along the X axis, this behaviour is reproduced. This is due to the fact that *lost* (*i.e.* with guidance deficiency) neurites wander more generating more tortuous and longer trajectories before reaching the end of the ML. In consequence, they generate more local density and thus, following our model, more branches. This distal *lost* behaviour is visually evident in some real axons, such as that one in Fig. 4.34A where the main axon continues growing parallel to the mid-line. In addition to this observation, real axons also present lower densities than simulated ones in the region near the entrance to the ML. These observations may indicate that real axons may enter the lobe with higher values of parameters -thus lower energies and more ordered/parallel trajectories- which progressively descend through the ML (or in time, which is similar).

These observations suggest the interest of analysing the energy of the system in function of space/time. We thus measured the entropy of the axons paths in function of their evolution in the X axis. We considered the angles (in the plane XY) of all the segments $\Delta\rho_i$, ϕ_i in intervals of $\Delta X = 5\mu m$. We approximated for ΔX the distributions of the ensemble of ϕ_i observed using an Epanechnikov kernel and the functions $fitdist$ and pdf provided by Matlab. Fig. 4.67 presents the obtained distributions.

We the calculated the entropy $S(X,\phi)$ by

$$S(X,\phi) = -\int_{-\pi}^{\pi} P(\phi) log(P(\phi)) d\phi. \tag{4.44}$$

As we can observe in Fig. 4.68B, even though simulated and real axons present the same values of $\mu_S = 10$ and $\sigma_S = 4$, its behaviour in function of the evolution of the system differs. In the first case it remains quite constant through X, while for real axons the entropy augments

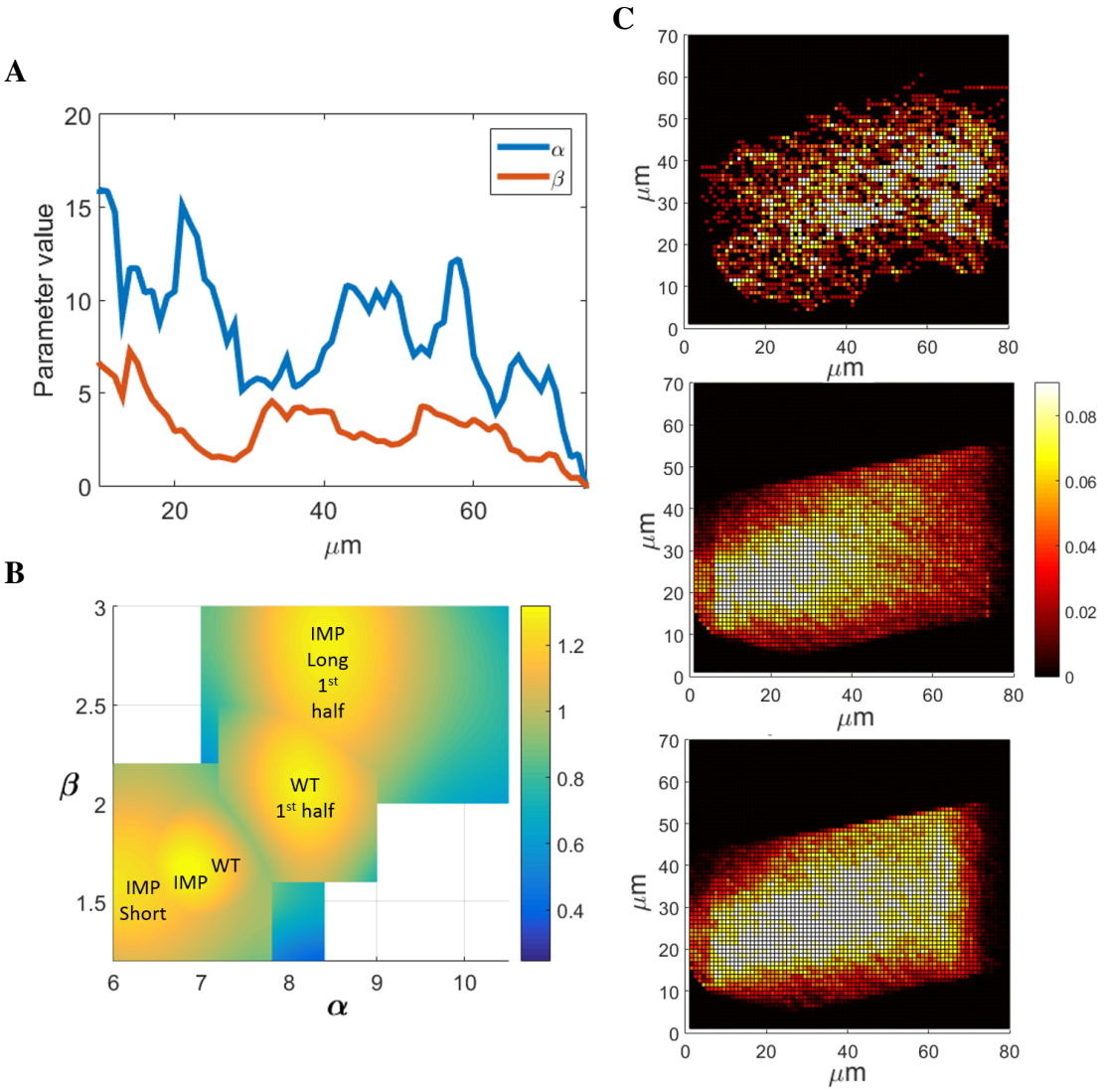

Fig. 4.66 Parameter variation analysis. (A) Parameters calculated from wild-type real axons considering a sliding window of size $5\mu m$ along the X axis of the lobe. (B) Parameter estimation surfaces (equivalent to those in Fig. 4.15B-D) for WT, the entire *imp* population (IMP), *imp* only with short phenotype (IMP Short), as well as estimated from only the first halve of wild-type (WT 1^{st} half) and the first halve of *imp* only with long phenotype (IMP Long 1^{st} half). (C) Density heat maps (normalized by number of axons in the sample, 43 for real axons and 650 for simulated) for wild-type real axons (top), simulated axons with mechanical branching and fixed values of parameters (middle) and simulated axons with mechanical branching and β decreasing with the spatial localization along the X axis (bottom).

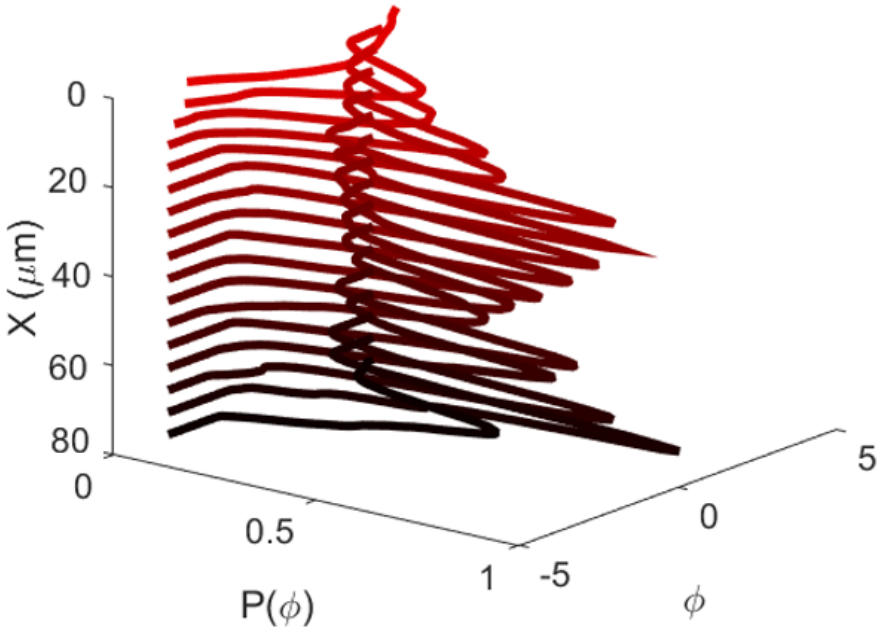

Fig. 4.67 Axon entropy calculation. Distribution of the observed angles ϕ_i corresponding to the axons paths segments every $\Delta X = 5\mu m$ in the domain $[-\pi, \pi]$ approximated using the functions $fitdist$ and pdf provided by Matlab.

until $\sim X = 25\mu m$ (first peak) and then descends until $X \sim 60\mu m$ when it starts augmenting again until the end of the ML (second peak). This suggests that the estimated parameter values nicely capture the general energy of the system, but they may not be actually constant through its evolution. In particular, the shape of $S(X, \phi)$ can be compared with the density of broken points regarding our model and studied in Section 4.3 (Fig. 4.68A). Regarding real axons, the model is broken mostly when the entropy of the system is high. This can be due in part to the generation of type I branches that *a posteriori* form part of the main axons and are originally born with random angles (thus detected as broken points), as simulated axons also present a rise in broken points at $X \sim 30\mu m$ and they are generated with constant parameters. However, simulated axons do not present a second peak on broken points distribution by the end of the ML, and real axons do; suggesting that this behaviour cannot be due exclusively to branch generation. Finally, the observed first peak on the entropy and breakdown points coincide with area where the *imp* short axons stopped, both in real data and in simulations. This is particularly interesting as, initially, they are unrelated measurements from different databases. However, when interpreting the γ population as a collective system that evolves in space and time, with impact in the agents (axons) energy, we can link these phenomena.

γ axons enter the lobe from the peduncle where they are placed altogether in an ordered way (parallel trajectories) as pictured in Fig. 4.22B. Then, they enter the ML where they adopt random and less coherent (disordered) trajectories, which impacts their evolution in several ways. The most studied one in this work is the fact that independent tubular-like shapes growing altogether with disordered trajectories in a constraint space will collide and become eventually blocked, in a transient or permanent way. In our model, this phenomenon alone may explain: • the observed ascyncronicity within the population, • neurites stopping their growth before reaching the end of the ML, • the need of type I branches, • the presence of jump points and • the double phenotype observed in *imp* knock-down (single non-branching axons in a wild-type environment). Nonetheless, this may not be the only emergent phenomena in our system going from order to dis-order. For example, in addition to already exploited observations of parameter value variation, ML neurite density, energy, etc., real *imp* axons are divided half and half between short and long phenotypes, while our model predicts only 40% of short. An interesting perspective is thus to better study this collective phenomena, to understand if the variations in entropy are due to intrinsic characteristics (variation of parameters values, *i.e.* changes in elasticity or more probable in the external field attractivity) or to an emergent phenomenon consequence of other inter-axon interactions that were not considered in this study.

As we mentioned in the introduction of this work, the bibliography on axon growth models is vast, covering many different aspects of this complex process. To end this discussion, we establish a parallelism between our model and the model published in Nguyen et al. (2016) and discussed in Chapter 2. We chose this one as its mathematical design is similar to ours, and because they perform detailed validations *in vitro*. They are interested at analysing how the axons react and bend towards a certain source of attractive cues. They simulate axons that are born with a certain angle orthogonal to the direction of the external attraction field, and observe how the turning angles (angle of the segment connecting the initial and the last points of the axon) evolve. Their model includes a term on the tendency to grow straight (similar to our stiffness), on the effect of the chemotactic gradient (similar to our external field attraction) plus some noise. Their noise parameter is independent from the previous two, in contrast to our case. An important contributor in their model are the anchor points, which determine the preferred direction of the next step regarding axon stiffness. In Fig. 4.69A, the anchor point is the circle placed in the beginning of the axon, and $\phi(t)$ the preferred direction for each step $\Delta\theta(t)$ regarding the stiffness. Their three parameters (stiffness, external attraction and noise) are tuned by hand, and the resulting trajectories are closely compared to axons from peripheral nervous system growing on a laminin substrate. Their study shows that most probably, *in vitro*, the only anchor points are located where the

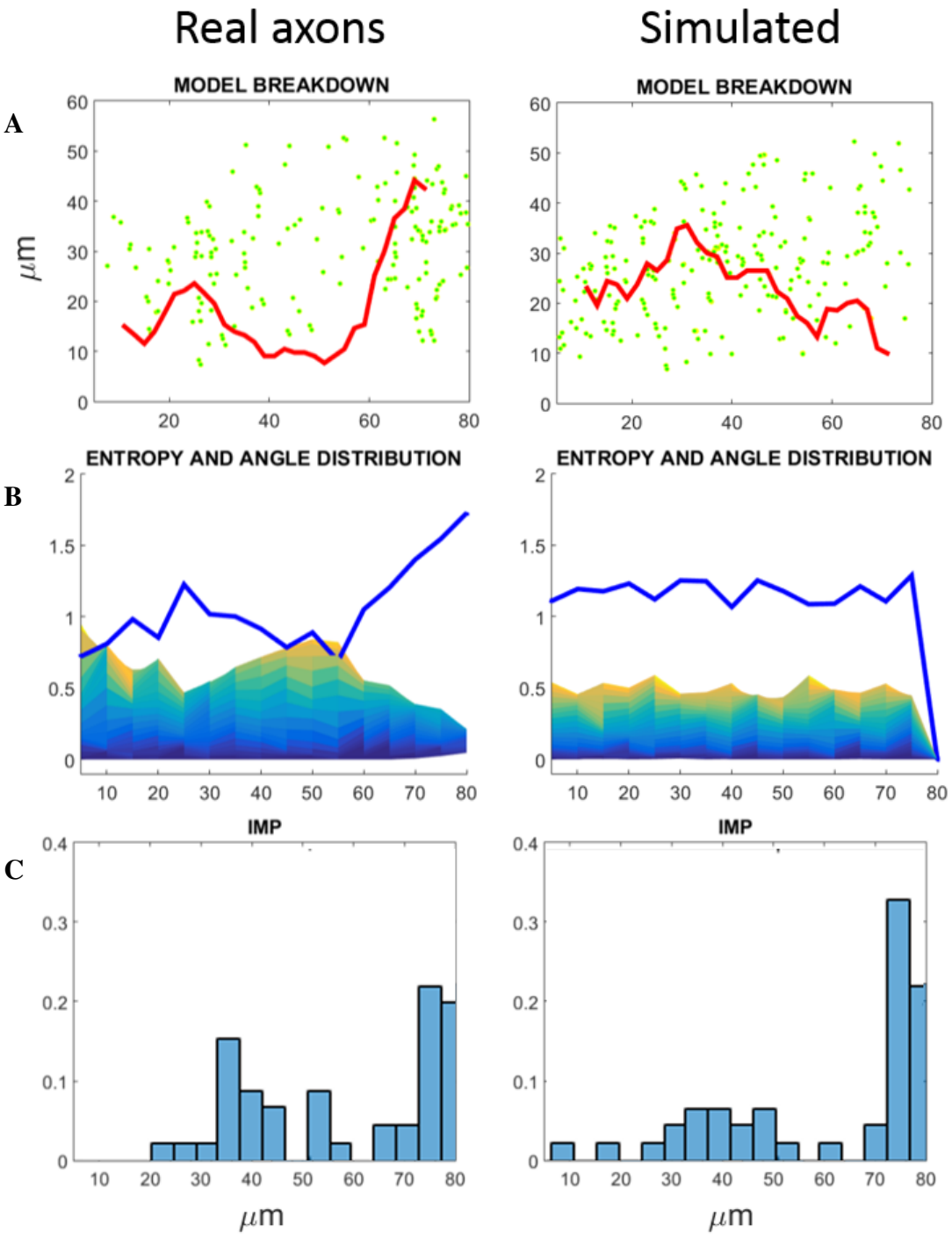

Fig. 4.68 Breakdown and energetic analysis for real vs. simulated axons. (A) Points where the Markov model is broken using the statistical framework in Section 4.3 with fixed $T = 2$ and the null hypothesis of $\beta = 0$ (green). The red line represents the shape of the distribution of broken points along the X axis. (B) The surface represents the distribution of angles ϕ_i in function of X (Fig. 4.67) and the blue line the obtained entropy $S(X, \phi)$ using Eq. 4.44 and multiplied by a factor of $k = 0.1$. In both cases $\mu_S = 10$ and $\sigma_S = 4$. (C) Distribution of travelled distances for *imp* axons.

axon emerges from the soma or at branching points. However, they propose that *in vivo* the environment is more complex and axons may establish anchor points at several sites with their substrate. They investigated how the turning angles depend on the increasing number of anchor points, with r representing the frequency of anchor points respect to normal steps (Fig. 4.69B (left)). Because their model is well calibrated with observed data, we wanted to test if the trajectories proposed by our model matched theirs. We thus generated axons with an initial angle of $\pi/2$ and an external field constant direction of 0. Because they simulated 150 steps, for each value of r (anchor point frequency) the number of steps of our simulations was $150r$ (each step with fixed size of $1\mu m$). The results are very similar as those using their model, and are shown in Fig. 4.69B (right). This means that our model using the parameters estimated from *in vivo* data has equivalent behaviours as their model, considering that each step is originated in an anchor point. These results do not only cross-validate our model and estimation with a previously published model, but also suggests a relevant difference between axons *in vitro* (which have been well studied) and *in vivo* (mostly unknown). In particular, our model and estimated parameter values follows the idea that anchor points are established frequently, in our case every $1\mu m$ (size of the step). They report that more anchor points increase the fidelity of turning but with noisier trajectories, and conclude that the frequency of anchor points plays a key role in the axonal turning response to a gradient. Finally, they propose that *in vivo*, axons should present many anchor points and/or additional cues such as mechanical ones or axon-axon interactions. We thus extend their results by proposing that anchor points may be very frequent *in vivo*, at $\sim 1\mu m$. Interestingly, type II branches in the *live* data are observed roughly at this distance, suggesting further interest on studying their function.

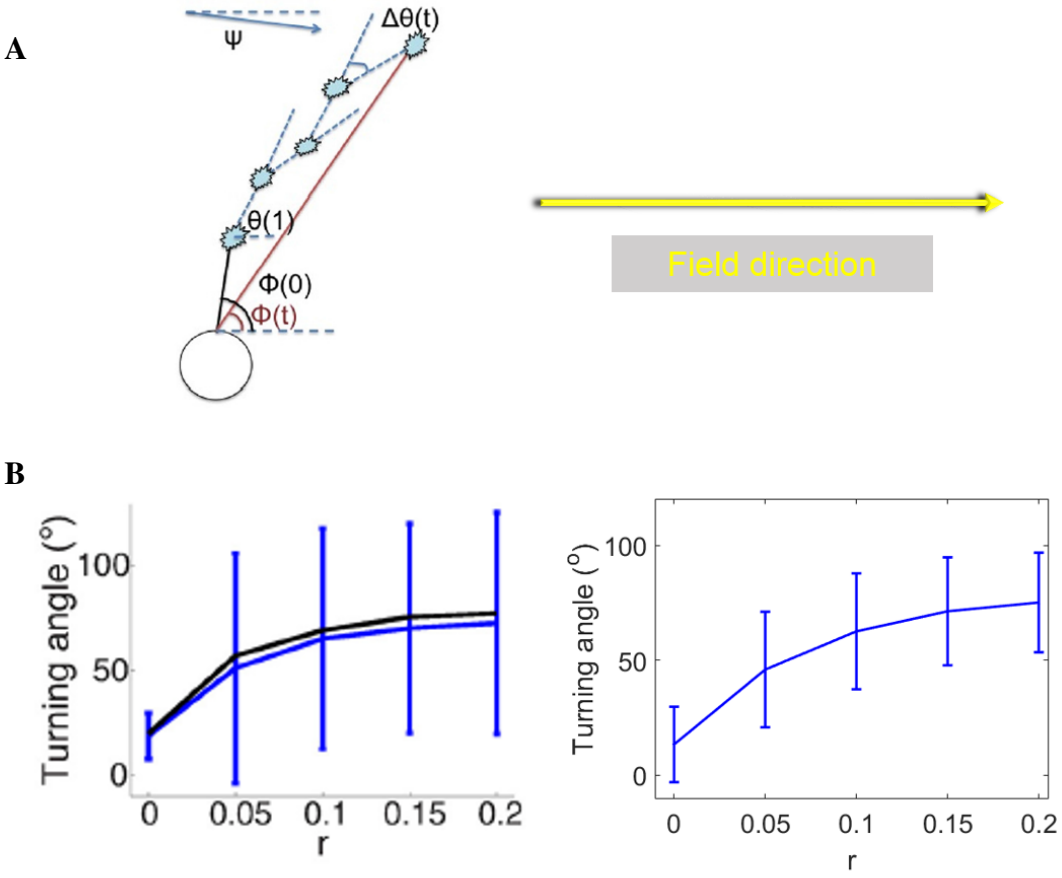

Fig. 4.69 Comparison with another published axon growth model. (A) Schema of the model proposed in Nguyen et al. (2016) for axon elongation *in vitro* and present in Fig. 1 A. (B) (left) Fig. 14 G from Nguyen et al. (2016), depicting the mean and standard deviation of the turning angles $\phi(t)$ (defined as the angle between the origin and the tip of the axon) in function of the frequency of anchor points (r). (right) Mean and standard deviation of turning angles for 1000 axons simulated using our model ($\alpha = 7.5$ and $\beta = 1.7$) and number of steps equal to $150xr$.

Chapter 5

General conclusion, contributions and perspectives

Todas las teorías son legítimas y ninguna tiene importancia. Lo que importa es lo que se hace con ellas.

Jorge Luis Borges

This manuscript has two main parts where the original contributions are described, Chapter 3: Morphological Study and Chapter 4: Space-Embedded Axon Growth Model. Both of them present their own discussion and contributions section. We now perform a general conclusion, and point out the main contributions of this thesis. Finally, we suggest possible improvements and further development related to this work.

5.1 General conclusion

The aim of our work was the design of a methodological framework to unveil new aspects of axonal dynamic growth *in vivo* (*i.e.* in the brain), counting with a collection of confocal 3D images of adult axons as main data-set, representing the final result of the process. The central challenges have been: • to integrate previous knowledge on axonal development from the literature, where most experiments were carried out *in vitro*, with new original hypothesis on the impact of a realistic environment on axonal growth; and to • use this to design mathematical models and algorithms allowing simulations (generative property of the model), in order to validate the initial assumptions, as well as to better understand the impact on axonal elongation of different aspects of the process (predictive property of the model).

First, we needed to deeply describe and understand the axonal morphology, as well as the differences and similarities between wild-types and different considered phenotypes,

where mainly branching and elongation were affected. To do so, we designed a stochastic framework to compare neuron groups based on their morphology. We proposed probabilistic models to describe the behaviour of the main morphological features, estimated from data the associated parameters and performed a likelihood analysis combined with statistical tests. The most relevant and novel biological conclusion of this study is that overall axonal elongation correlates with the number and length of branches.

This morphological study suggested an interesting correlation between branching and elongation, which had not been yet studied in the literature. In addition, the diversity of morphologies between genetically identical axons inspired us with the idea that axonal growth should be the result of a stochastic process and not of a pre-defined mechanism. We proposed a mathematical model of axonal trajectories where the directionality of each new step follows a Gaussian Markov chain with two parameters, which can be estimated from data: accounting for the axonal rigidity and the attraction to an external field. We then combined these model with some simulation rules on branch formation and dynamics, to finally recreate the growth of an entire population of axons developing altogether, as it occurs *in vivo*. The parameter estimation from real data resulted to be not trivial, and the approaches we developed to overcome this inconvenient proposed that axons can actually be described by Gaussian Markov chains with constant parameters, but piecewisely between jump points. These places where the chains are broken may indicate obstacles that the axons encountered during their growth *in vivo*. Interestingly, the parameters estimated from data create a system that is quite noisy, and where the mechanical shocks between growing axons not only must exist, but may even stop an important percentage of them from attaining full elongation in the limited developmental time. This implies that other mechanisms must exist to allow full elongation of axons in a dense population context. Because of the previous observation on the correlation between elongation and branching, we turned our attention to the branching characteristics in the axons of our database to be able to include their development in our model. We concluded that there exist two types of branches, small and dynamic and longer ones, which grow following the same rules as main axons regarding path and elongation. This implies that the axonal hierarchy is actually set at the end of the process (*i.e.* the longest path may be composed by the initial axon and a branch), and not pre-defined, as it is classically assumed. We have also shown that even though an increment in branching elevates the general density thus the difficulty of each axon to elongate, axons as a population reach better growth rates when they create branches than when they do not. Strikingly, we showed that for the same mean number of branches, the elongation strategy is even more efficient when branches are born after mechanical collisions, rather than uniformly randomly. Thanks to the data processing which allowed to register all the axons to the same spatial

frame, we observed that the axonal density correlates with the number of branch points, suggesting that this mechanism is actually plausible. We also succeeded in reproducing the phenotypes of the studied mutants (which presented phenotypic branch size, number and overall axonal lengths), which reinforced the idea that the shortage of long branches results in a consequent deficiency in axonal elongation, thus proposing that the main deficiency caused by this mutations is branching, and elongation is a consequence of this. Finally, an interesting observation is that the model parameters estimated from data are not the most efficient in terms of growth at the population level, but they insure a compromise between elongation and morphological diversity, which is essential to guarantee accurate connectivity.

5.2 Perspectives

We propose the following perspectives for this work:

Data analysis

- The registration technique described in Section 4.1 and used to map all the axons to the same spatial frame does not take into account possible rotations on the X or Y axis. In further studies, this could be taken into account to ameliorate the axonal morphological reconstructions.
- The segmentation and analysis tools developed in the team allowed to obtain quantitative information on the length and number of the small and dynamic branches, which was used in the study. However, further analysis and automatic tracking techniques could be used to obtain dynamic information on the axonal tip elongation and pathfinding, specifically from videos as those shown in Fig. 4.33. This could refine the dynamic aspect of the simulations, as well as suggest other ways of limiting time rather than implementing an artificial counter.

Model

- In Section 4.2 we describe the mathematical formulation of the model for spatial trajectories evolution, that we use to simulate the axonal paths. One advantage of this model is that it allows renormalization in space. This implies that one can segment the trajectories in a convenient scale to maximize the signal/noise ratio, estimate the parameters of the model, and then obtain the equivalent parameters for different scales, including smaller ones. One perspective would be to take advantage of this, to compare trajectories from datasets of different characteristics (*i.e.* different image resolution, or

axon types of different sizes). Another use could be to simulate smaller step sizes, to also obtain more continue-like trajectories.

- An interesting (and ambitious) perspective would be to design a model which inherently takes into account the interactions between trajectories, instead of reproducing them by simulations. Even better, if the interaction parameter can be estimated from data, one could obtain interesting information on the behaviour of the system just by analysing independent trajectories.

Parameter estimation

- In Section 4.3.2 we describe a framework to detect the broken points from otherwise homogeneous Gaussian Markov chains and thus estimate the hidden parameters. As presented, neither the statistical approach nor the bayesian one with a Bernoulli prior, take into account the density of broken points (or the distance between two continuous jump points). In consequence, they may assign several neighbour points as jump points. This is of course useful to detect every point that does not follow the Markov model. However, further improvements may consider a prior following a Poisson distribution or an accumulation process to impose constraints in the distance between consecutive jump points.

- Here, we developed the jump detection and estimation algorithm with the purpose of estimating the parameters from the axons in our database. However, it would be interesting to apply the methods to study trajectories of other nature, as well as any kind of process susceptible of being described as a piecewisely homogeneous Gaussian Markov chain (for example in the domains of finance or audio).

Simulations

- As studied in Section 4.8, the values of the parameters of the model seem to vary in space and time. It would be interesting to further integrate these variations in the model, to see the impact on axonal development and analysis the feasibility of the hypothesis. Even though we have not observed significant impact on variations of parameters between axons, this could have important implications and thus may be further studied. Variations in parameter could be at the origin of the entropy peaks discussed in the same section.

- Based on the estimation results and hypothesis considered for the jump point detection framework (Section 4.3.2), the simulation rules can be modified such that after

encountering mechanical obstacles the axon is allowed to break the Markov model. This would imply, also, higher parameter values for the rest of the shaft (as observed when estimating the parameters with this method). Also, it would be interesting to analyse in the simulated axons present the same percentage of jump points as real ones with the same jump hypothesis and estimated parameters. This could have interesting implications, and help to validate other estimated parameters such as the axonal diameter (as it suggests that the number of collisions encountered by simulated axons is the same as real ones).

- In this work, we considered mechanical interactions between axons, *via* volume exclusion. Further work could include interactions of other nature, even attractive ones, as well as interactions with other components of the axonal environment such as the extracellular matrix or glial cells.

Biological validation

- Here we used wild-type γ axons to design and calibrate the model, and a series of mutant ones and predicted their phenotypes. It would be interesting to generate databases with other mutants, for example where the axonal rigidity may be changed (which has not been tested in this work).

- Finally, the most interesting perspective of this work is to apply ou model to test the impact of mechanical interactions in axonal development in other neuron types. This may allow to further understand the phenomenon as well as to, maybe, unveil different elongation mechanisms.

  In particular, it would be interesting to apply our model to populations of axons presenting varied morphologies, and that form dense interconnected structures in the vertebrate brain. For example, cerebellum-like structures, sometimes described as the vertebrate counterparts of Mushroom Bodies, also consist of large collections of equivalent cells that project axon-like processes into densely-packed parallel fibres (Bell et al., 2008; Farris, 2011).

References

Aeschlimann, M. and Tettoni, L. (2001). Biophysical model of axonal pathfinding. *Neurocomputing*, 38:87–92.

Armijo-Weingart, L. and Gallo, G. (2017). It takes a village to raise a branch: Cellular mechanisms of the initiation of axon collateral branches. *Molecular and Cellular Neuroscience*.

Ascoli, G. A. and Krichmar, J. L. (2000). L-neuron: a modeling tool for the efficient generation and parsimonious description of dendritic morphology. *Neurocomputing*, 32:1003–1011.

Aso, Y., Grübel, K., Busch, S., Friedrich, A. B., Siwanowicz, I., and Tanimoto, H. (2009). The mushroom body of adult drosophila characterized by gal4 drivers. *Journal of neurogenetics*, 23(1-2):156–172.

Aso, Y., Hattori, D., Yu, Y., Johnston, R. M., Iyer, N. A., Ngo, T.-T. B., Dionne, H., Abbott, L. F., Axel, R., Tanimoto, H., and Rubin, G. M. (2014). The neuronal architecture of the mushroom body provides a logic for associative learning. *Elife*, 3.

Bell, C. C., Han, V., and Sawtell, N. B. (2008). Cerebellum-like structures and their implications for cerebellar function. *Annu. Rev. Neurosci.*, 31:1–24.

Bixby, J. L. and Harris, W. A. (1991). Molecular mechanisms of axon growth and guidance. *Annu Rev Cell Biol*, 7:117–59.

Caron, S. J., Ruta, V., Abbott, L., and Axel, R. (2013). Random convergence of olfactory inputs in the drosophila mushroom body. *Nature*, 497(7447):113–117.

Cassenaer, S. and Laurent, G. (2007). Hebbian stdp in mushroom bodies facilitates the synchronous flow of olfactory information in locusts. *Nature*, 448(7154):709–713.

Cuntz, H., Forstner, F., Borst, A., and Häusser, M. (2010). One rule to grow them all: a general theory of neuronal branching and its practical application. *PLoS computational biology*, 6(8):e1000877.

Davenport, R. W., Thies, E., and Cohen, M. L. (1999). Neuronal growth cone collapse triggers lateral extensions along trailing axons. *Nat Neurosci*, 2(3):254–9. Erratum in Nat Neurosci 1999 May;2(5):485.

Davis, B. M., Frank, E., Johnson, F. A., and Scott, S. A. (1989). Development of central projections of lumbosacral sensory neurons in the chick. *Journal of Comparative Neurology*, 279(4):556–566.

De Rooij, R., Miller, K., and Kuhl, E. (2017). Modeling molecular mechanisms in the axon. *Computational Mechanics*, 59(3):523–537.

Demyanenko, G. P. and Maness, P. F. (2003). The l1 cell adhesion molecule is essential for topographic mapping of retinal axons. *Journal of Neuroscience*, 23(2):530–538.

Dent, E. W., Gupton, S. L., and Gertler, F. B. (2011). The growth cone cytoskeleton in axon outgrowth and guidance. *Cold Spring Harbor perspectives in biology*, 3(3):a001800.

Deppmann, C. D., Mihalas, S., Sharma, N., Lonze, B. E., Niebur, E., and Ginty, D. D. (2008). A model for neuronal competition during development. *Science*, 320(5874):369–373.

Dickson, B. J. (2002). Molecular mechanisms of axon guidance. *Science*, 298(5600):1959–1964.

Eberhard, J. P., Wanner, A., and Wittum, G. (2006). Neugen: a tool for the generation of realistic morphology of cortical neurons and neural networks in 3d. *Neurocomputing*, 70(1):327–342.

Engle, E. C. (2010). Human genetic disorders of axon guidance. *Cold Spring Harbor perspectives in biology*, 2(3):a001784.

Farris, S. M. (2011). Are mushroom bodies cerebellum-like structures? *Arthropod structure & development*, 40(4):368–379.

Flanagan, L. A., Ju, Y.-E., Marg, B., Osterfield, M., and Janmey, P. A. (2002). Neurite branching on deformable substrates. *Neuroreport*, 13(18):2411.

Francisco, H., Yellen, B. B., Halverson, D. S., Friedman, G., and Gallo, G. (2007). Regulation of axon guidance and extension by three-dimensional constraints. *Biomaterials*, 28(23):3398–407.

Franze, K. (2013). The mechanical control of nervous system development. *Development*, 140(15):3069–3077.

Friedl, P., Locker, J., Sahai, E., and Segall, J. E. (2012). Classifying collective cancer cell invasion. *Nature cell biology*, 14(8):777.

Gallo, G. (2011a). The cytoskeletal and signaling mechanisms of axon collateral branching. *Developmental neurobiology*, 71(3):201–220.

Gallo, G. (2011b). The cytoskeletal and signaling mechanisms of axon collateral branching. *Dev Neurobiol*, 71(3):201–20.

Ganguly, S., Liang, X., Grace, M., Lee, D., and Howard, J. (2016). Mathematical relationships between neuron morphology and neurite growth dynamics in drosophila melanogaster larva class iv sensory neurons. In *APS Meeting Abstracts*.

Ghabrial, A. S. and Krasnow, M. A. (2006). Social interactions among epithelial cells during tracheal branching morphogenesis. *Nature*, 441(7094):746.

Gibson, D. A. and Ma, L. (2011). Developmental regulation of axon branching in the vertebrate nervous system. *Development*, 138(2):183–195.

Goodhill, G. J. (2007). Contributions of theoretical modeling to the understanding of neural map development. *Neuron*, 56(2):301–311.

Goodhill, G. J. and Richards, L. J. (1999). Retinotectal maps: molecules, models and misplaced data. *Trends in neurosciences*, 22(12):529–534.

Graham, B. P. and Van Ooyen, A. (2001). Compartmental models of growing neurites. *Neurocomputing*, 38:31–36.

Guan, K.-L. and Rao, Y. (2003). Signalling mechanisms mediating neuronal responses to guidance cues. *Nature Reviews Neuroscience*, 4(12):941–956.

Guerra, L., McGarry, L. M., Robles, V., Bielza, C., Larrañaga, P., and Yuste, R. (2011). Comparison between supervised and unsupervised classifications of neuronal cell types: a case study. *Developmental neurobiology*, 71(1):71–82.

Haeger, A., Wolf, K., Zegers, M. M., and Friedl, P. (2015). Collective cell migration: guidance principles and hierarchies. *Trends Cell Biol*, 25(9):556–66.

Hakim, Y., Yaniv, S. P., and Schuldiner, O. (2014). Astrocytes play a key role in drosophila mushroom body axon pruning. *PLoS One*, 9(1):e86178.

Hannezo, E., Scheele, C. L., Moad, M., Drogo, N., Heer, R., Sampogna, R. V., van Rheenen, J., and Simons, B. D. (2017). A unifying theory of branching morphogenesis. *Cell*, 171(1):242–255.

Hayashi, S., Inoue, Y., Kiyonari, H., Abe, T., Misaki, K., Moriguchi, H., Tanaka, Y., and Takeichi, M. (2014). Protocadherin-17 mediates collective axon extension by recruiting actin regulator complexes to interaxonal contacts. *Dev Cell*, 30(6):673–87.

Heisenberg, M. (2003). Mushroom body memoir: from maps to models. *Nature Reviews Neuroscience*, 4(4):266–275.

Hely, T. A., Graham, B., and Van Ooyen, A. (2001). A computational model of dendrite elongation and branching based on map2 phosphorylation. *Journal of theoretical biology*, 210(3):375–384.

Hige, T. (2017). What can tiny mushrooms in fruit flies tell us about learning and memory? *Neuroscience Research*.

Hong, W. and Luo, L. (2014). Genetic control of wiring specificity in the fly olfactory system. *Genetics*, 196(1):17–29.

Hua, J. Y., Smear, M. C., Baier, H., and Smith, S. J. (2005). Regulation of axon growth in vivo by activity-base competition. *Nature*, 434(7036):1022.

Jakobsson, L., Franco, C. A., Bentley, K., Collins, R. T., Ponsioen, B., Aspalter, I. M., Rosewell, I., Busse, M., Thurston, G., Medvinsky, A., et al. (2010). Endothelial cells dynamically compete for the tip cell position during angiogenic sprouting. *Nature cell biology*, 12(10):943.

Kalil, K. and Dent, E. W. (2014). Branch management: mechanisms of axon branching in the developing vertebrate cns. *Nature Reviews Neuroscience*, 15(1):7.

Kalil, K., Szebenyi, G., and Dent, E. W. (2000). Common mechanisms underlying growth cone guidance and axon branching. *Journal of neurobiology*, 44(2):145–158.

Kerstein, P. C., Nichol, R. H. I., Nichol RH, I., and Gomez, T. M. (2015). Mechanochemical regulation of growth cone motility. *Front Cell Neurosci*, 9:244.

Kobayashi, T., Terajima, K., Nozumi, M., Igarashi, M., and Akazawa, K. (2010). A stochastic model of neuronal growth cone guidance regulated by multiple sensors. *J Theor Biol*, 266(4):712–22.

Koene, R. A., Tijms, B., van Hees, P., Postma, F., de Ridder, A., Ramakers, G. J. A., Ramakers, G., van Pelt, J., and van Ooyen, A. (2009). Netmorph: a framework for the stochastic generation of large scale neuronal networks with realistic neuron morphologies. *Neuroinformatics*, 7(3):195–210.

Kolodkin, A. L. and Tessier-Lavigne, M. (2011). Mechanisms and molecules of neuronal wiring: a primer. *Cold Spring Harbor Perspectives in Biology*, 3(6):a001727.

Kong, J.-H., Fish, D. R., Rockhill, R. L., and Masland, R. H. (2005). Diversity of ganglion cells in the mouse retina: unsupervised morphological classification and its limits. *Journal of Comparative Neurology*, 489(3):293–310.

Koser, D. E., Thompson, A. J., Foster, S. K., Dwivedy, A., Pillai, E. K., Sheridan, G. K., Svoboda, H., Viana, M., da F Costa, L., Costa, L., Guck, J., Holt, C. E., and Franze, K. (2016). Mechanosensing is critical for axon growth in the developing brain. *Nat Neurosci*, 19(12):1592–1598.

Lee, T., Lee, A., and Luo, L. (1999). Development of the drosophila mushroom bodies: sequential generation of three distinct types of neurons from a neuroblast. *Development*, 126(18):4065–4076.

Leondaritis, G. and Eickholt, B. J. (2015). Short lives with long-lasting effects: filopodia protrusions in neuronal branching morphogenesis. *PLoS biology*, 13(9):e1002241.

López-Cruz, P. L., Larrañaga, P., DeFelipe, J., and Bielza, C. (2014). Bayesian network modeling of the consensus between experts: An application to neuron classification. *International Journal of Approximate Reasoning*, 55(1):3–22.

Lowery, L. A. and Van Vactor, D. (2009). The trip of the tip: understanding the growth cone machinery. *Nature reviews. Molecular cell biology*, 10(5):332.

Luczak, A. (2006). Spatial embedding of neuronal trees modeled by diffusive growth. *Journal of neuroscience methods*, 157(1):132–141.

Luo, L. (2002). Actin cytoskeleton regulation in neuronal morphogenesis and structural plasticity. *Annual review of cell and developmental biology*, 18(1):601–635.

Luo, L. and Flanagan, J. G. (2007). Development of continuous and discrete neural maps. *Neuron*, 56(2):284–300.

McLean, D. R., van Ooyen, A., and Graham, B. P. (2004). Continuum model for tubulin-driven neurite elongation. *Neurocomputing*, 58:511–516.

Medioni, C., Besse, F., Descombes, X., and Malandain, G. (2015a). Motion compensation in two-photon microscopy temporal series. In *International Symposium on Biomedical Imaging (ISBI)*, page 4, Brooklyn, United States.

Medioni, C., Ephrussi, A., and Besse, F. (2015b). Live imaging of axonal transport in drosophila pupal brain explants. *Nat Protoc*, 10(4):574–84.

Medioni, C., Ramialison, M., Ephrussi, A., and Besse, F. (2014). Imp promotes axonal remodeling by regulating profilin mrna during brain development. *Curr Biol*, 24(7):793–800.

Mehes, E. and Vicsek, T. (2014). Collective motion of cells: from experiments to models. *Integr Biol (Camb)*, 6(9):831–54.

Mitchison, T. and Kirschner, M. (1988). Cytoskeletal dynamics and nerve growth. *Neuron*, 1(9):761–772.

Mortimer, D., Feldner, J., Vaughan, T., Vetter, I., Pujic, Z., Rosoff, W. J., Burrage, K., Dayan, P., Richards, L. J., and Goodhill, G. J. (2009). A bayesian model predicts the response of axons to molecular gradients. *Proceedings of the National Academy of Sciences*, 106(25):10296–10301.

Mortimer, D., Fothergill, T., Pujic, Z., Richards, L. J., and Goodhill, G. J. (2008). Growth cone chemotaxis. *Trends Neurosci*, 31(2):90–8.

Mottini, A., Descombes, X., and Besse, F. (2013). Tree-like shapes distance using the elastic shape analysis framework. In *BMVC-British Machine Vision Conference*.

Mottini, A., Descombes, X., and Besse, F. (2015). From curves to trees: A tree-like shapes distance using the elastic shape analysis framework. *Neuroinformatics*, 13(2):175–191.

Murthy, M., Fiete, I., and Laurent, G. (2008). Testing odor response stereotypy in the drosophila mushroom body. *Neuron*, 59(6):1009–1023.

Myatt, D. R., Hadlington, T., Ascoli, G. A., and Nasuto, S. J. (2012). Neuromantic - from semi-manual to semi-automatic reconstruction of neuron morphology. *Front Neuroinform*, 6:4.

Nguyen, H., Dayan, P., Pujic, Z., Cooper-White, J., and Goodhill, G. J. (2016). A mathematical model explains saturating axon guidance responses to molecular gradients. *Elife*, 5.

Pawlik, M. and Augsten, N. (2011). Rted: a robust algorithm for the tree edit distance. *Proceedings of the VLDB Endowment*, 5(4):334–345.

Petrovic, M. and Schmucker, D. (2015). Axonal wiring in neural development: Target-independent mechanisms help to establish precision and complexity. *BioEssays*, 37(9):996–1004.

Plachez, C. and Richards, L. J. (2005). Mechanisms of axon guidance in the developing nervous system. *Current topics in developmental biology*, 69:267–346.

Rabinovich, D., Yaniv, S. P., Alyagor, I., and Schuldiner, O. (2016). Nitric oxide as a switching mechanism between axon degeneration and regrowth during developmental remodeling. *Cell*, 164(1):170–182.

Razetti, A., Descombes, X., Medioni, C., and Besse, F. (2016). Statistical characterization, modelling and classification of morphological changes in imp mutant drosophila gamma neurons. In *BIOSTEC 2016-The 9h International Joint Conference on Biomedical Engineering Systems and Technologies*.

Razetti, A., Descombes, X., Medioni, C., and Besse, F. (2017). A stochastic framework for neuronal morphological comparison: Application to the study of imp knockdown effects in drosophila gamma neurons. pages 145–166.

Rørth, P. (2012). Fellow travellers: emergent properties of collective cell migration. *EMBO reports*, 13(11):984–991.

Sagasti, A., Guido, M. R., Raible, D. W., and Schier, A. F. (2005). Repulsive interactions shape the morphologies and functional arrangement of zebrafish peripheral sensory arbors. *Current biology*, 15(9):804–814.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. *Nature methods*, 9(7):676–682.

Schlüter, K., Jockusch, B. M., and Rothkegel, M. (1997). Profilins as regulators of actin dynamics. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, 1359(2):97–109.

Schmidt, H. and Rathjen, F. G. (2010). Signalling mechanisms regulating axonal branching in vivo. *Bioessays*, 32(11):977–985.

Schmidt, H., Stonkute, A., Jüttner, R., Schäffer, S., Buttgereit, J., Feil, R., Hofmann, F., and Rathjen, F. G. (2007). The receptor guanylyl cyclase npr2 is essential for sensory axon bifurcation within the spinal cord. *J Cell Biol*, 179(2):331–340.

Schumacher, L. J., Kulesa, P. M., McLennan, R., Baker, R. E., and Maini, P. K. (2016). Multidisciplinary approaches to understanding collective cell migration in developmental biology. *Open Biol*, 6(6).

Segev, R. and Ben-Jacob, E. (2000). Generic modeling of chemotactic based self-wiring of neural networks. *Neural Networks*, 13(2):185–199.

Simpson, H. D., Mortimer, D., and Goodhill, G. J. (2009). Theoretical models of neural circuit development. *Current topics in developmental biology*, 87:1–51.

Song, H.-j. and Poo, M.-m. (2001). The cell biology of neuronal navigation. *Nature cell biology*, 3(3):E81–E88.

Srivastava, A., Joshi, S. H., Mio, W., and Liu, X. (2005). Statistical shape analysis: Clustering, learning, and testing. *IEEE Transactions on pattern analysis and machine intelligence*, 27(4):590–602.

Suleymanov, Y., Gafarov, F., and Khusnutdinov, N. (2013). Modeling of interstitial branching of axonal networks. *J Integr Neurosci*, 12(1):103–16.

Szebenyi, G., Callaway, J. L., Dent, E. W., and Kalil, K. (1998). Interstitial branches develop from active regions of the axon demarcated by the primary growth cone during pausing behaviors. *J Neurosci*, 18(19):7930–40.

Tessier, C. R. and Broadie, K. (2008). Drosophila fragile x mental retardation protein developmentally regulates activity-dependent axon pruning. *Development*, 135(8):1547–1557.

Tessier-Lavigne, M., Goodman, C. S., et al. (1996). The molecular biology of axon guidance. *Science*, 274(5290):1123–1133.

Theveneau, E. and Mayor, R. (2012). Neural crest migration: interplay between chemorepellents, chemoattractants, contact inhibition, epithelial-mesenchymal transition, and collective cell migration. *Wiley Interdiscip Rev Dev Biol*, 1(3):435–45.

Torben-Nielsen, B. and De Schutter, E. (2014). Context-aware modeling of neuronal morphologies. *Front Neuroanat*, 8:92.

Triplett, J. W., Pfeiffenberger, C., Yamada, J., Stafford, B. K., Sweeney, N. T., Litke, A. M., Sher, A., Koulakov, A. A., and Feldheim, D. A. (2011). Competition is a driving force in topographic mapping. *Proceedings of the National Academy of Sciences*, 108(47):19060–19065.

Tsigankov, D. and Koulakov, A. (2009). Optimal axonal and dendritic branching strategies during the development of neural circuitry. *Frontiers in neural circuits*, 3.

Van Ooyen, A. (2011). Using theoretical models to analyse neural development. *Nature reviews. Neuroscience*, 12(6):311.

van Ooyen, A. and Ribchester, R. R. (2003). Competition in the development of nerve connections. *Modelling Neural Development*, pages 183–211.

van Pelt, J. and Schierwagen, A. (2004). Morphological analysis and modeling of neuronal dendrites. *Mathematical biosciences*, 188(1):147–155.

Vanherpe, L., Kanari, L., Atenekeng, G., Palacios, J., and Shillcock, J. (2016). Framework for efficient synthesis of spatially embedded morphologies. *Phys Rev E*, 94(2-1):023315.

Vedula, S. R. K., Vedula, S., Ravasio, A., Lim, C. T., and Ladoux, B. (2013). Collective cell migration: a mechanistic perspective. *Physiology (Bethesda)*, 28(6):370–9.

Verheyen, E. M. and Cooley, L. (1994). Profilin mutations disrupt multiple actin-dependent processes during drosophila development. *Development*, 120(4):717–728.

Wang, L. and Marquardt, T. (2013). What axons tell each other: axon–axon signaling in nerve and circuit assembly. *Current opinion in neurobiology*, 23(6):974–982.

Watts, R. J., Hoopfer, E. D., and Luo, L. (2003). Axon pruning during drosophila metamorphosis: evidence for local degeneration and requirement of the ubiquitin-proteasome system. *Neuron*, 38(6):871–885.

Watts, R. J., Schuldiner, O., Perrino, J., Larsen, C., and Luo, L. (2004). Glia engulf degenerating axons during developmental axon pruning. *Curr Biol*, 14(8):678–84.

Weijer, C. J. (2009). Collective cell migration in development. *Journal of cell science*, 122(18):3215–3223.

Woods, M. L., Carmona-Fontaine, C., Barnes, C. P., Couzin, I. D., Mayor, R., and Page, K. M. (2014). Directional collective cell migration emerges as a property of cell interactions. *PLoS One*, 9(9):e104969.

Wu, J. S. and Luo, L. (2006). A protocol for mosaic analysis with a repressible cell marker (marcm) in drosophila. *Nat Protoc*, 1(6):2583–9.

Yaniv, S. P., Issman-Zecharya, N., Oren-Suissa, M., Podbilewicz, B., and Schuldiner, O. (2012). Axon regrowth during development and regeneration following injury share molecular mechanisms. *Current Biology*, 22(19):1774–1782.

Yaniv, S. P. and Schuldiner, O. (2016). A fly's view of neuronal remodeling. *Wiley Interdisciplinary Reviews: Developmental Biology*, 5(5):618–635.

Zubler, F. and Douglas, R. (2009). A framework for modeling the growth and development of neurons and networks. *Frontiers in computational neuroscience*, 3.

Zubler, F., Hauri, A., Pfister, S., Bauer, R., Anderson, J. C., Whatley, A. M., and Douglas, R. J. (2013). Simulating cortical development as a self constructing process: a novel multi-scale approach combining molecular and physical aspects. *PLoS computational biology*, 9(8):e1003173.

Appendix A

Demonstrations

A.1 Bernoulli model

From equation 3.6 we can express $\mu_k(A+1,p)$ as

$$\mu_k(A+1,p) = \sum_k k \binom{k-1}{A+1} p^{A+2}(1-p)^{k-A-2}. \tag{A.1}$$

Using

$$\binom{k-1}{A} + \binom{k-1}{A+1} = \binom{k}{A+1}, \tag{A.2}$$

equation A.1 can be rewritten as

$$\mu_k(A+1,p) = \sum_k k \binom{k}{A+1} p^{A+2}(1-p)^{k-A-2} - \sum_k k \binom{k-1}{A} p^{A+2}(1-p)^{k-A-2}. \tag{A.3}$$

Taking out $\frac{p}{1-p}$ as a common factor from the second sum in A.3, we obtain

$$\sum_k k \binom{k-1}{A} p^{A+2}(1-p)^{k-A-2} = \frac{p}{1-p} \sum_k k \binom{k-1}{A} p^{A+1}(1-p)^{k-A-1} = \frac{p}{1-p}\mu_k(A,p). \tag{A.4}$$

Similarly, the first sum in A.3 can be worked out to obtain

$$\mu_k(A+1,p) = \frac{p}{1-p}\mu_k(A+1,p) - \frac{p}{1-p}. \tag{A.5}$$

From A.4 and A.5 we finally obtain

$$\mu_k(A,p) = \frac{A}{p}. \tag{A.6}$$

From equations 3.7 and A.6 we can express

$$\sigma_k^2(A,p) = \sum_k k^2 \binom{k-1}{A} p^{A+1}(1-p)^{k-A-1} - \frac{A^2}{p^2}, \tag{A.7}$$

from where

$$\sigma_k^2(A+1,p) = \sum_k k^2 \binom{k-1}{A+1} p^{A+2}(1-p)^{k-A-2} - \frac{(A+1)^2}{p^2}. \tag{A.8}$$

Equation A.8 can be rewritten using A.2 to get

$$\sigma_k^2(A+1,p) = \sum_k k^2 \binom{k}{A+1} p^{A+2}(1-p)^{k-A-2} - \sum_k k^2 \binom{k-1}{A} p^{A+2}(1-p)^{k-A-2} - \frac{(A+1)^2}{p^2}. \tag{A.9}$$

Working out A.9 similarly to A.3 we finally get

$$\sigma_k^2(A,p) = \frac{(1-p)A}{p^2}. \tag{A.10}$$

A.2 Parameter estimation from global variance. Part I

From

$$\theta_i = \frac{\alpha}{\alpha+\beta}\theta_{i-1} + \xi_i, \tag{A.11}$$

where

$$\xi_i \sim \mathcal{N}(0, \frac{1}{2(\alpha+\beta)}). \tag{A.12}$$

If we denote $\gamma = \frac{\alpha}{\alpha+\beta}$, we then have

$$\theta_{i+1} = \gamma\theta_i + \xi_{i+1} \tag{A.13}$$

$$\theta_{i+1} = \gamma[\gamma\theta_{i-1} + \xi_i] + \xi_{i+1} \tag{A.14}$$

$$\theta_{i+1} = \gamma^2 \theta_{i-1} + \gamma\xi_i + \xi_{i+1}, \tag{A.15}$$

where ξ_i and ξ_{i+1} are two independent random variables with the same Gaussian distribution (eq. A.11). Therefore, with $\sigma_0^2 = \frac{1}{2(\alpha+\beta)}$;

$$\gamma\xi_i = \sim \mathcal{N}(0, \gamma^2\sigma_0^2), \tag{A.16}$$

and

$$\gamma\xi_i + \xi_{i+1} = \mathcal{N}(0, (\gamma^2+1)\sigma_0^2). \tag{A.17}$$

Finally,

$$\theta_{i+1} = \gamma^2 \theta_{i-1} + \tau_{i+1}, \tag{A.18}$$

where $\tau_{i+1} \sim \mathcal{N}(0, (\gamma^2+1)\sigma_0^2)$.
Similarly,

$$\theta_{i+2} = \gamma\theta_{i+1} + \xi_{i+2}, \tag{A.19}$$

where $\xi_{i+2} \sim \mathcal{N}(0, \sigma_0^2)$.
Using A.18

$$\theta_{i+2} = \gamma[\gamma^2 \theta_{i-1} + \tau_{i+1}] + \xi_{i+2} \tag{A.20}$$

$$\theta_{i+2} = \gamma^3 \theta_{i-1} + \gamma\tau_{i+1} + \xi_{i+2}. \tag{A.21}$$

τ_{i+1} and ξ_{i+2} are independent, thus

$$\theta_{i+2} = \gamma^3 \theta_{i-1} + \rho_{i+2}, \tag{A.22}$$

with $\rho_{i+2} \sim \mathcal{N}(0, (\gamma^4+\gamma^2+1)\sigma_0^2)$.
By recursion we obtain the cumulated variance as the series

$$\sigma^2_{\theta_{i=1:M}} = \sigma_0^2 \sum_{i=0}^{M-1} \gamma^{2i}. \tag{A.23}$$

The series in eq. A.23 is geometrical, thus

$$\sigma_{\theta_i}^2 = \sigma_0^2 \frac{1-\gamma^{2M}}{1-\gamma^2}. \tag{A.24}$$

that converges to

$$\sigma_{\theta_\infty}^2 = \frac{\sigma_0^2}{1-\gamma^2}. \tag{A.25}$$

A.3 Parameter estimation from global variance. Part II

The difference $\Delta\theta = \theta_i - \theta_{i-1}$ is equal to

$$\Delta\theta_i = \gamma\theta_{i-1} + \xi_i - \theta_{i-1} = (\gamma-1)\theta_{i-1} + \xi_i. \tag{A.26}$$

The variance of $\Delta\theta = \theta_i - \theta_{i-1}$ is thus

$$\sigma_{\Delta\theta_i}^2 = (\gamma-1)^2\sigma_{\theta_{i-1}}^2 + \sigma_0^2 - 2E(\theta_{i-1}\xi_i). \tag{A.27}$$

θ_{i-1} and ξ_i are independent and centered in 0, thus, with eq. A.25

$$\sigma_{\Delta\theta_\infty}^2 = \sigma_0^2[\frac{(\gamma-1)^2}{1-\gamma^2} + 1] = \frac{2\sigma_0^2(1-\gamma)}{1-\gamma^2}. \tag{A.28}$$

Appendix B

Extra Figures

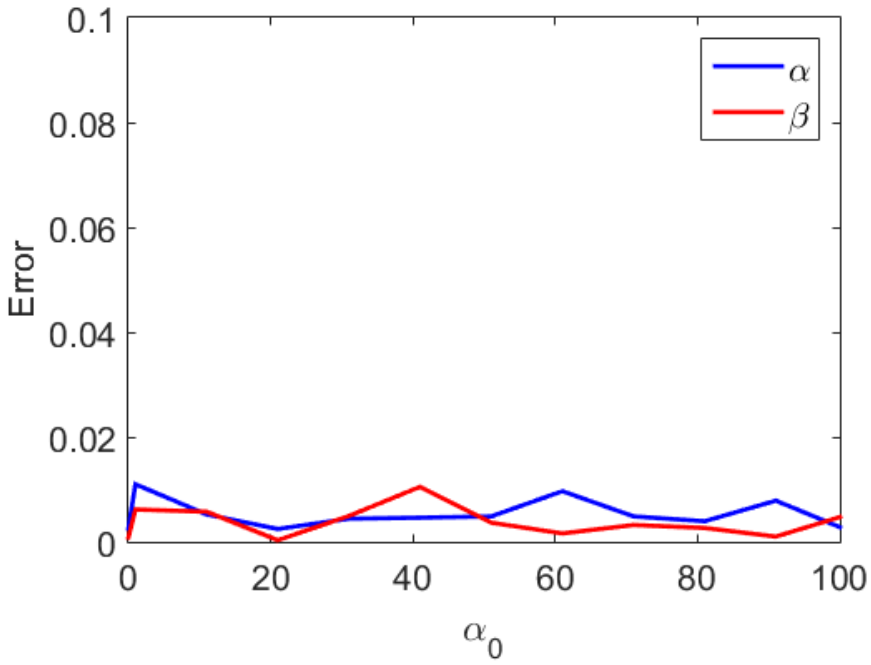

Fig. B.1 Iterative method algorithm: Initialisation. Absolute error ($\frac{|\hat{\alpha}_m - \alpha|}{\alpha}$) of the mean of the estimates ($\hat{\alpha}_m$) of 200 simulated chains of length 1000, with $\varepsilon = 0.00001$, $\alpha = \beta = 10$ and maximum number of iterations of 100 with varying values of initial condition α_0. The curves were fitted linearly with the Matlab Basic Fitting tool. The curve for α was fitted with $y = -6.6e-06 * x + 0.0056$, and for β with $y = -1.1e-05 * x + 0.0043$.

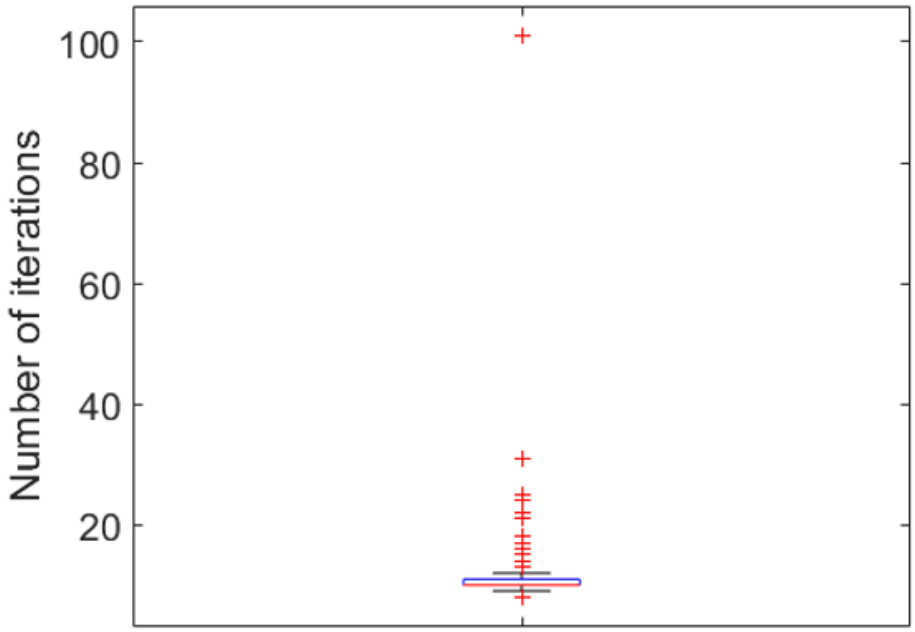

Fig. B.2 Iterative method algorithm: Application - Number or iterations. Distribution of the number of iterations corresponding to the estimation of parameters from 1000 simulated chains of length 100, with $\varepsilon = 0.00001$, $\alpha = \beta = 10$ and maximum number of iterations of 100.

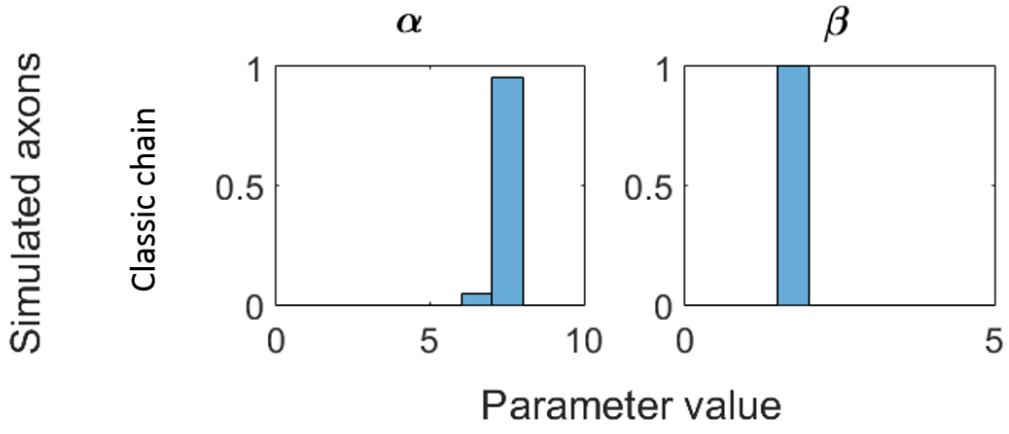

Fig. B.3 Axon by axon parameter estimation - Long simulated axons. Histograms of the parameter values estimated axon per axon from N=45 simulated chains using $\alpha = 7.5$, $\beta = 1.7$, and constant length $M = 10000$.

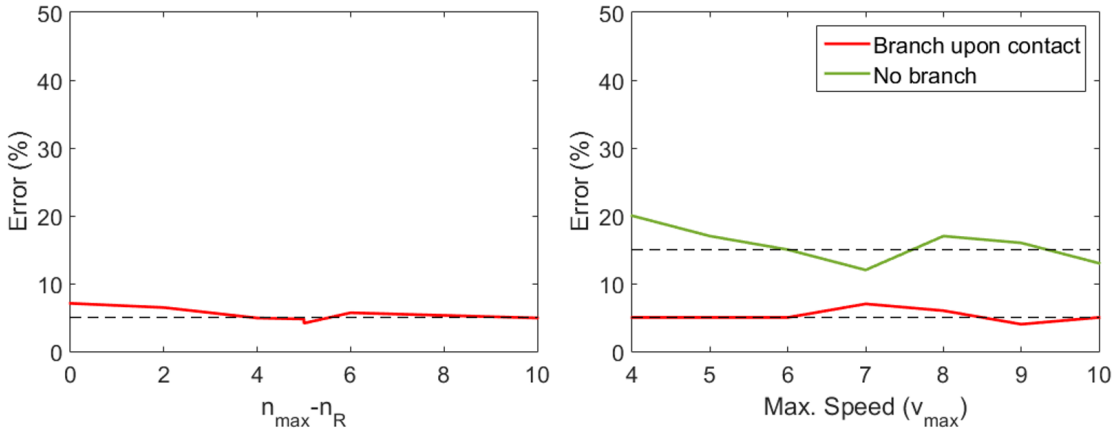

Fig. B.4 Error percentage in function of different n_{max} and n_R. (A) Percentage of non-elongated axons in function of the difference between maximum number of steps and retraction rate ($n_{max} - n_R$), for axons simulated with branching upon contact. The dotted line is placed at 5% to indicate the results with $v_{max} = 6$. Each point represents a single simulation. (B) Percentage of non-elongated axons in function of the maximum speed $v_{max} = \frac{n_{max}\Delta\rho}{\Delta t}$, for simulated with branching upon contact and without branching. The dotted lines are placed at 5 and 15% to indicate the results with $v_{max} = 6$, respectively. Each point represents a single simulation.

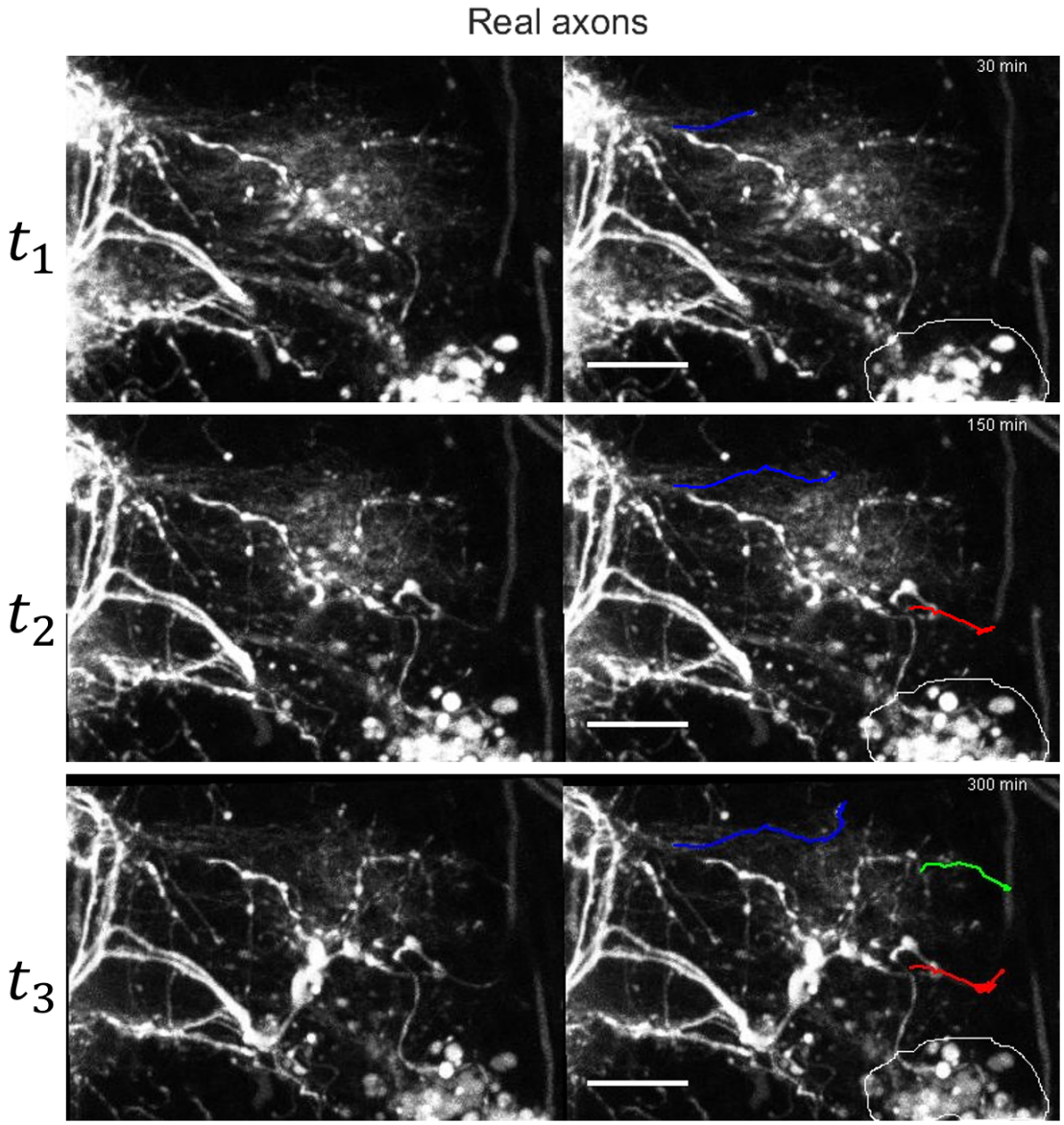

Fig. B.5 Emergent asynchronicity within the γ population (second example). Three time-points depicting an entire population of γ axons growing in the ML of a wild-type pupal brain (30h APF stage). Left: raw movie. Right: annotated movie where some examples of regrowing axons have been highlighted by colored tracks. Degenerating material found in the lower right corner of the field is delimited by a white line. Scale Bar:5μm.

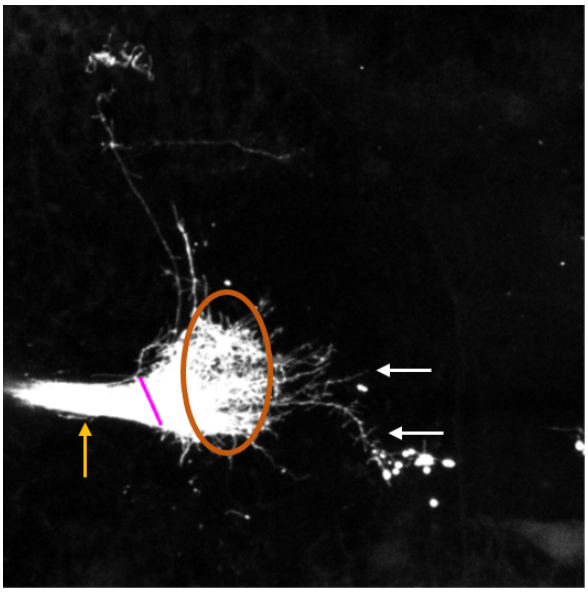

Fig. B.6 Emergent asynchronicity within the γ population (third example). Frame from a movie in our data set showing a whole neuroblast progeny at around 30h APF. White arrows show two axons that have already reached the stopping region. The oval highlights a region where we cannot easily spot the axon tips but the highly inhomogeneous image intensity suggests that they are not all equally grown. The yellow arrow indicates the peduncle, where axons grow tight and parallel to each other, and the purple line represents the peduncle transversal section, as in Fig. 4.22.

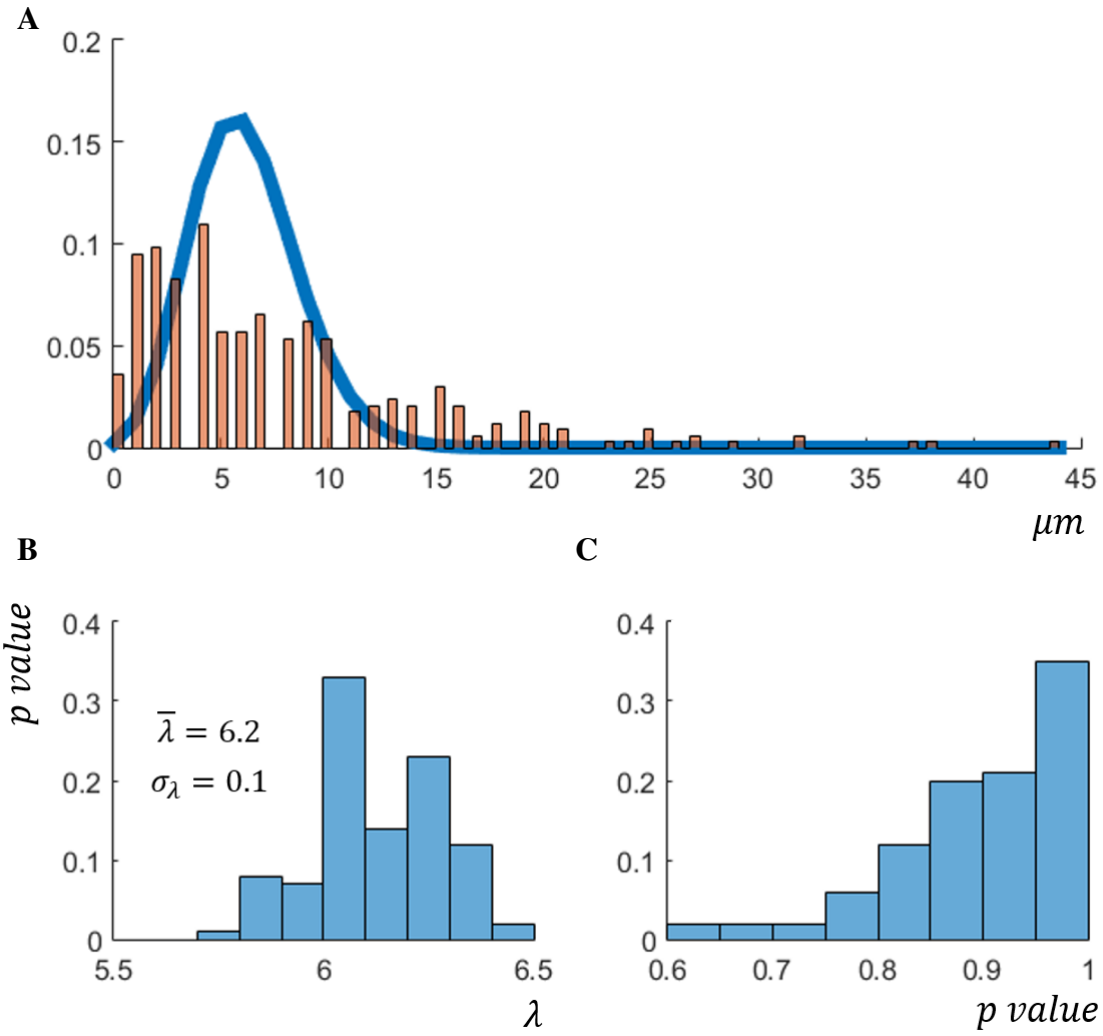

Fig. B.7 Analysis of the distances between branches in γ axons. (A) Branching point distances from γ axons in our data. To test its resemblance to a Poisson distribution, we threw samples (the same size as our data) from a Poisson distribution with $\lambda_b \in [1,20] \in \mathbb{N}$, and compared them to our data using the Kruskal-Wallis test. The highest p value corresponded to $\lambda_b = 6$. The orange bars represent the distance distributions from data, and the blue line represents the Poisson distribution with $\lambda_b = 6.2$ (tuned value). (B) To further tune the value of λ_b, we proceeded the same way as in (A) with $\lambda_b \in [5 : 0.1 : 7]$. We repeated this procedure one-hundred times and at each one took the value of λ with highest p value. The histogram represents the best λ_b value for each of the one-hundred iterations. The mean and standard deviation of the distribution are also shown. (C) Histogram of Kruskal-Wallis p values for the best λ_b at each iteration.

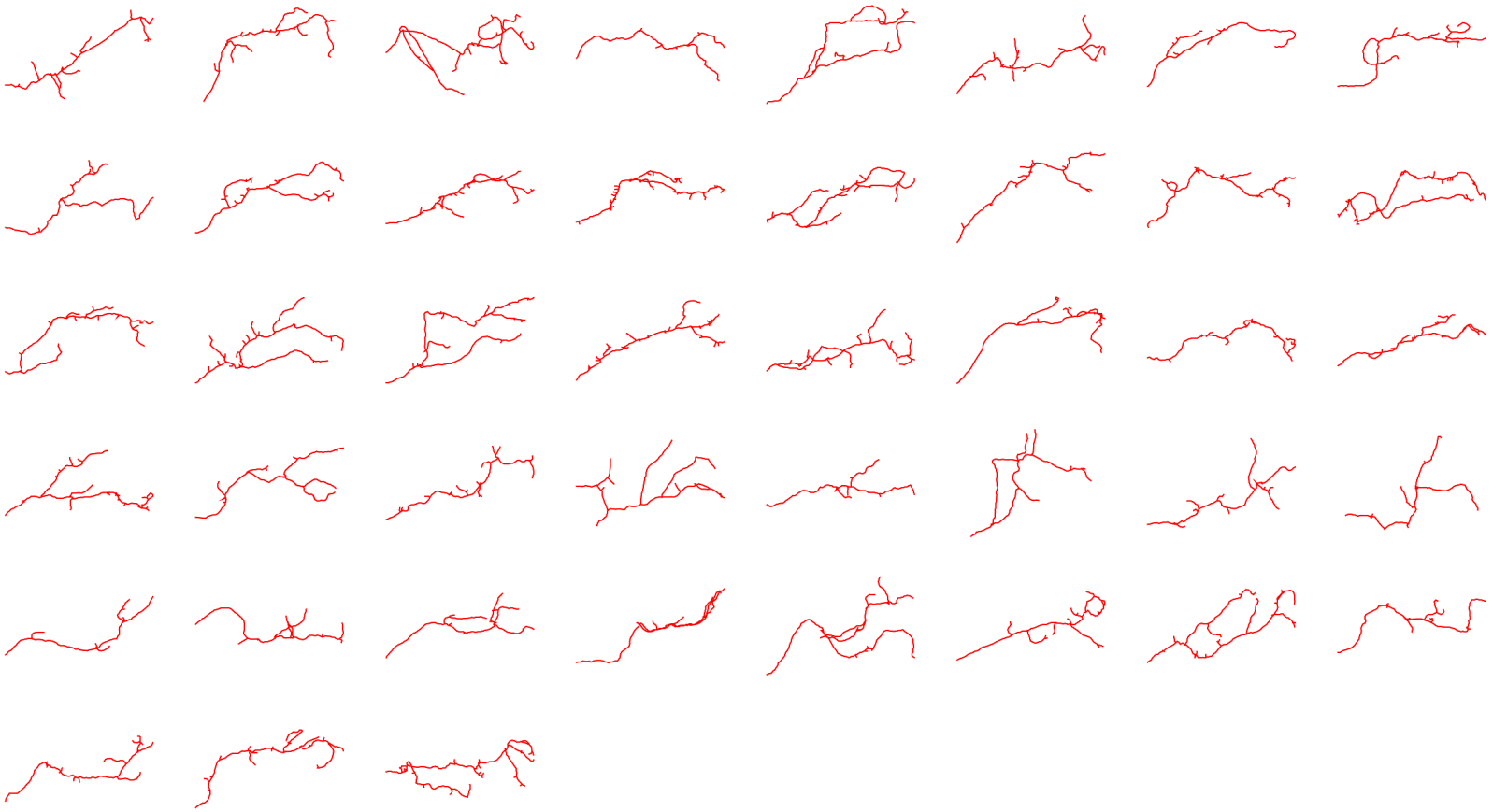

Fig. B.8 Wild type γ axons reconstructed from data.

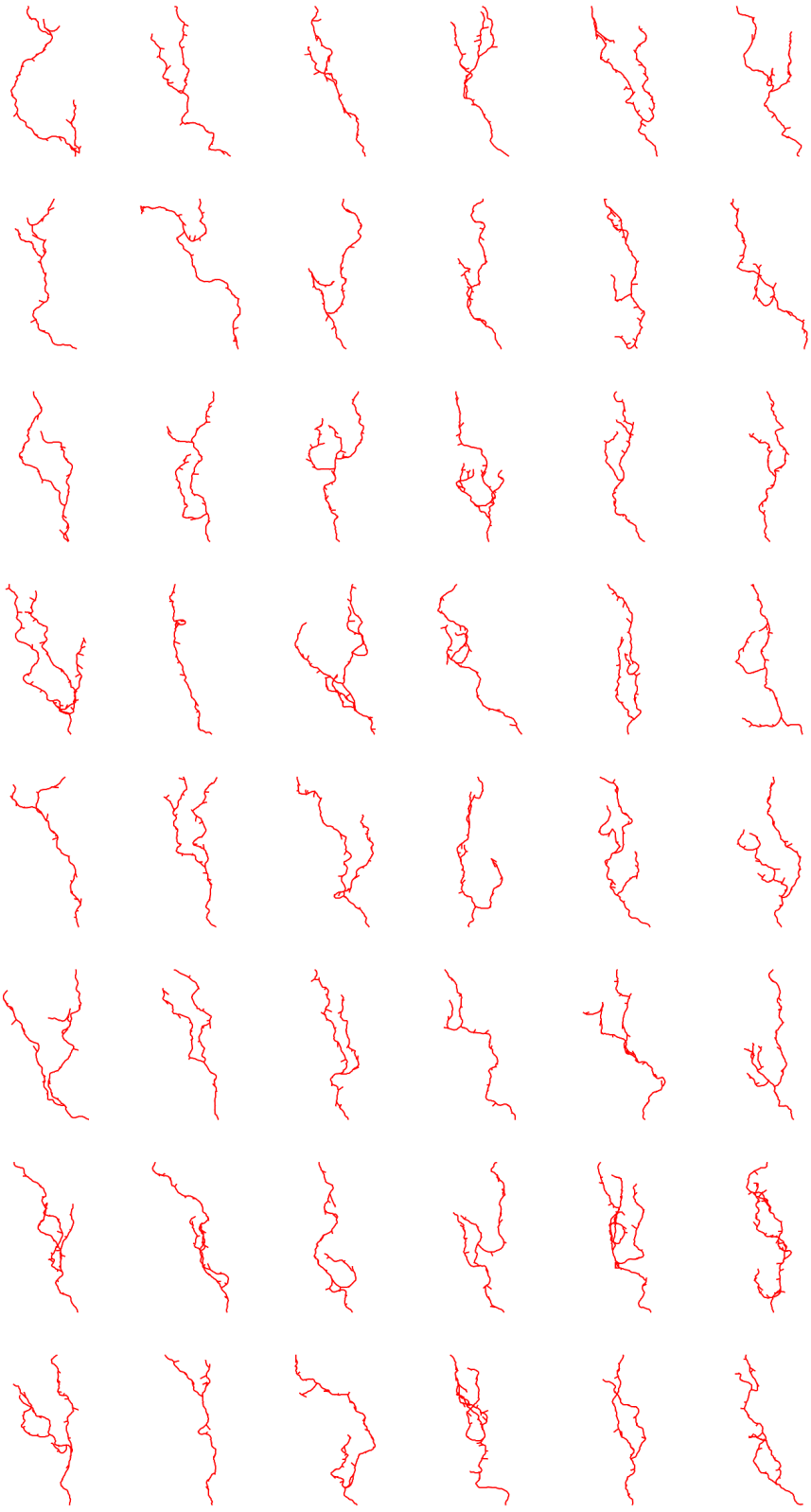

Fig. B.9 Simulated wild type γ axons.

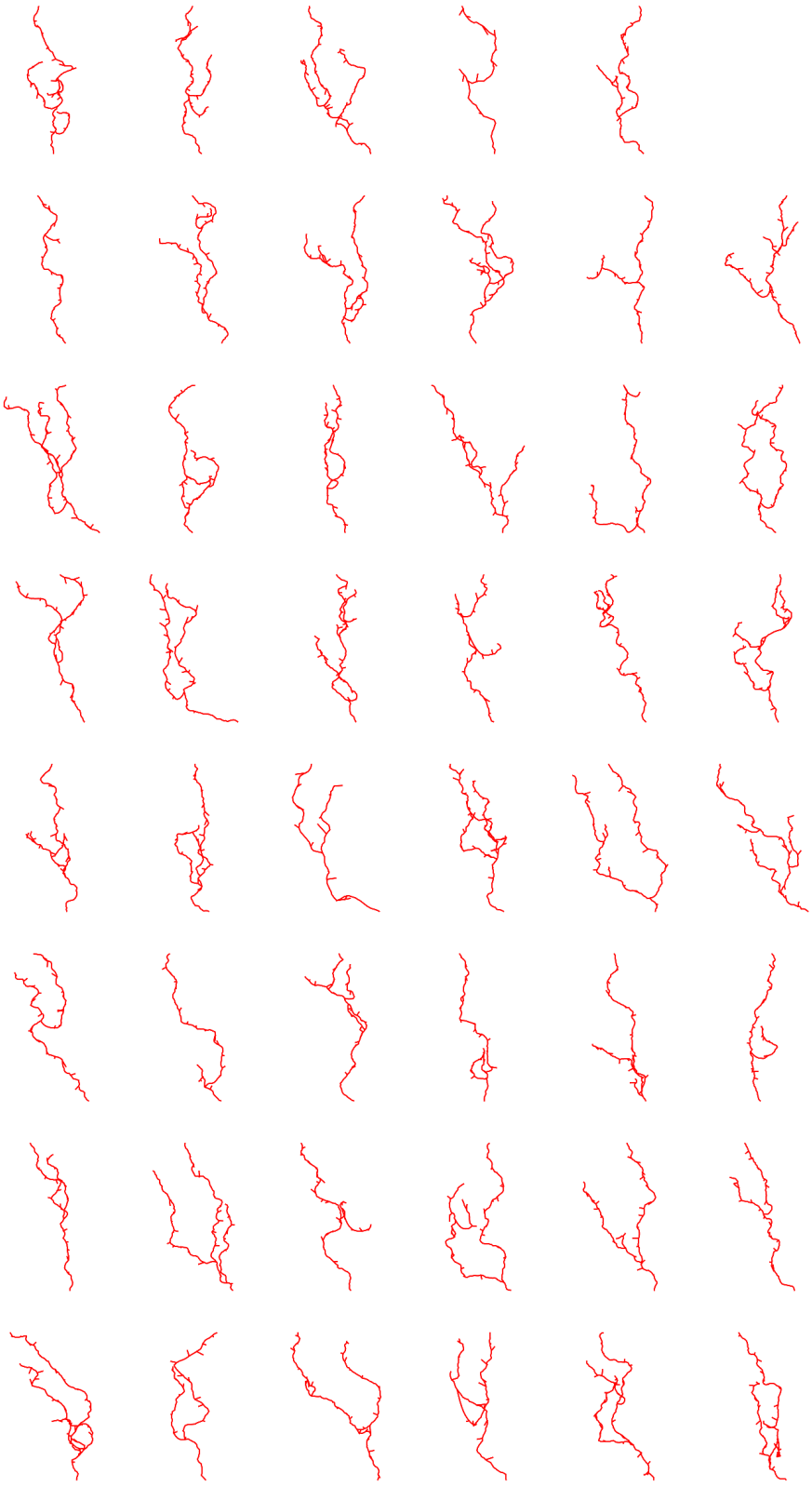

Fig. B.10 Simulated wild type γ axons - bis.

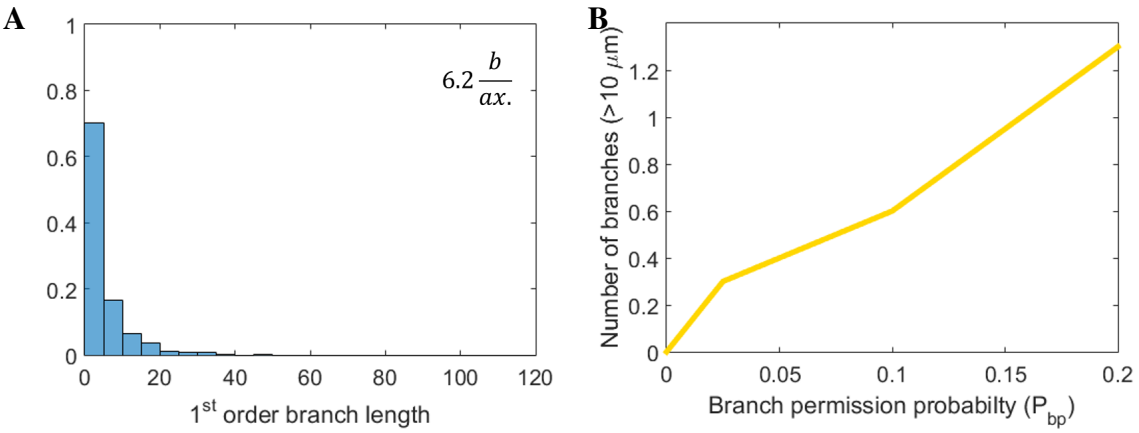

Fig. B.11 Branching deficiency phenotypes. (A) Branch length distribution from the Imp axons in our database. (B) Number of observed first order branches longer than 10 μm in simulated axons in function of the branch permission probability. Axons thus create a branch upon mechanical interactions with probability P_{bp}.

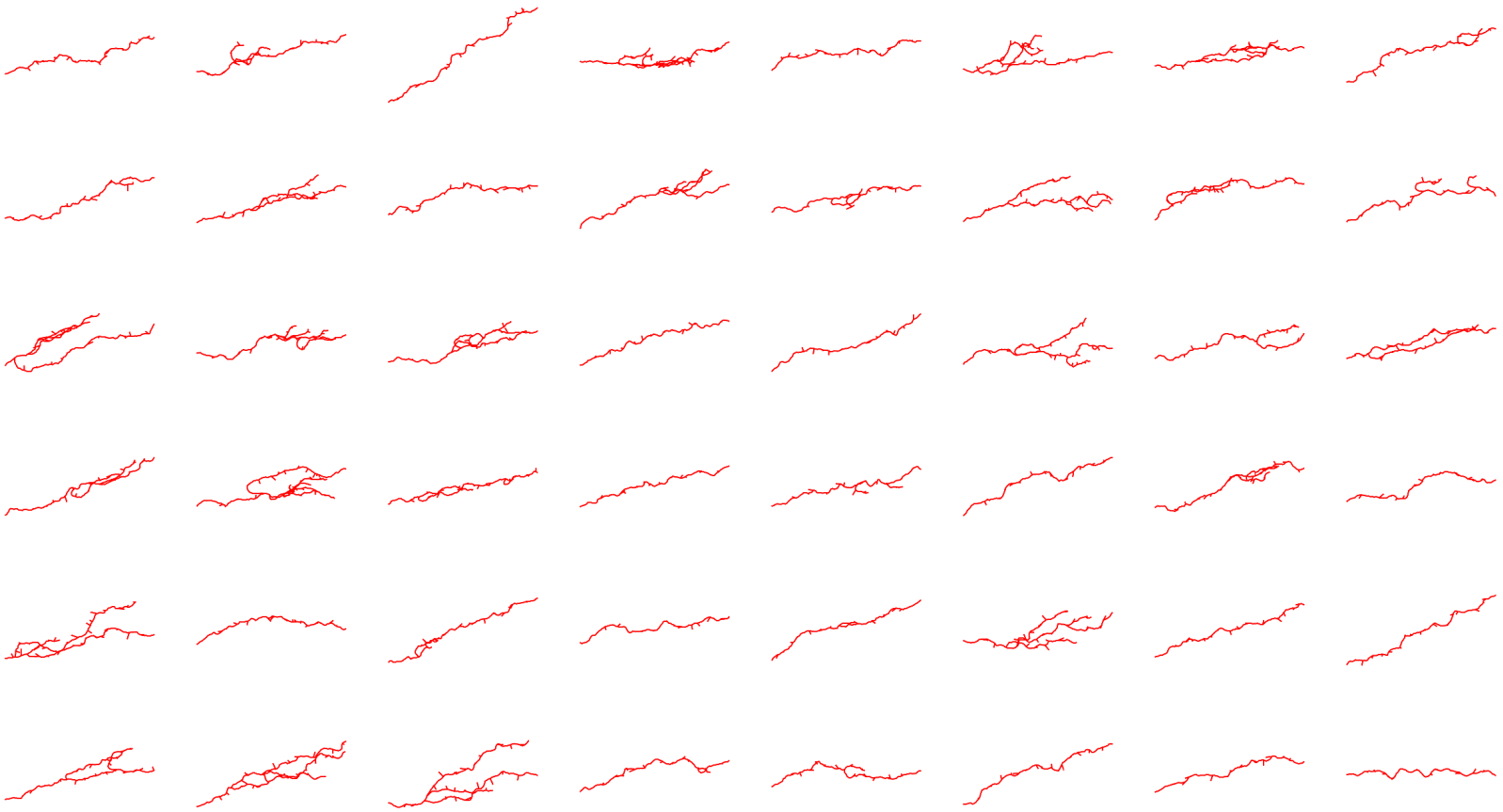

Fig. B.12 Simulated wild type γ axons with α_{bp} and β_{bp}.

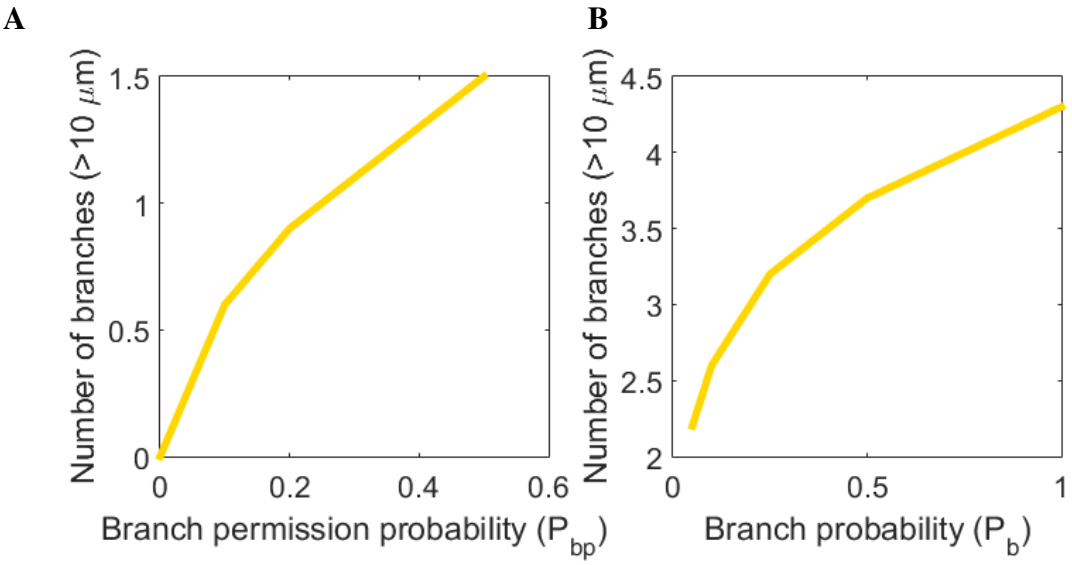

Fig. B.13 Number of branches in function of branching mechanisms. (A) Branches placed upon mechanical contact and with probability ${}_Pbp < 1$. (B) Branches placed upon mechanical contact and with probability ${}_Pbp = 1$ plus random branches, placed when no mechanical obstacles were encountered with probability P_b.

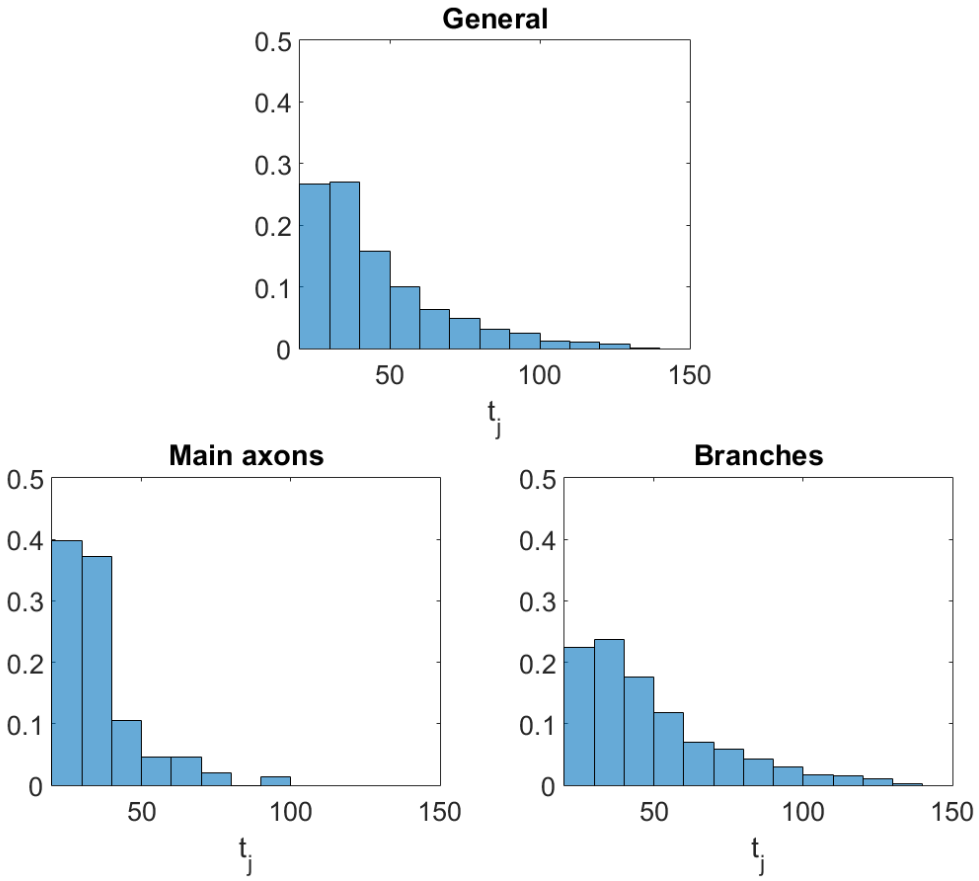

Fig. B.14 Distributions of elongation times for simulated γ axons. Three cases are presented: every neurite (General), only the Main axons and only Branches. Naturally, branches finished later than main axons as they are born afterwards.