Cell Polarity Under Extreme Morphological Conditions

Georg R. Walther

A thesis submitted to the University of East Anglia
for the degree of Doctor of Philosophy

John Innes Centre

Norwich, July 2014

Page left blank intentionally.

Abstract

Cell polarity is an important phenomenon in a multitude of cellular and developmental processes. The cellular contexts that polarity occurs in include a wide array of morphological properties such as size, shape, and growth.

An important, conserved system of cell polarity depends on the intracellular localisation of proteins that act as diffusive molecular switches. Since the localisation of these proteins depends on their reactive and diffusive properties, cell size and growth may alter polarity induced by localisation.

My work contributes extensive analyses of an established protein localisation model under extreme morphological conditions such as extremely small and rapidly growing cells. My work also uncovers non-trivial, biologically relevant behaviour caused by the inclusion of these morphological properties and further discusses the mechanisms underlying the observed behaviour.

In addition, I contribute and discuss a novel computational tool that can continue to aid the research community in understanding cell polarity under extreme morphological conditions.

Contents

4

List of Figures

List of Tables

Publications

The manuscripts below include material from this thesis and have been published.

Chapter 2

Walther, G. R., Maree, A. F. M., Edelstein-Keshet, L., and Grieneisen, V. A. (2012). Deterministic versus stochastic cell polarisation through wave-pinning. Bull. Math. Biol., 74, 2570–99.

Chapter 4

Walther, G. R., Hartley, M., and Mincheva, M. (2014). GraTeLPy: graph-theoretic linear stability analysis. BMC Syst. Biol., 8, 22.

Chapter 1

Introduction and Motivation

In this thesis I discuss cell polarity which is an essential phenomenon in a plethora of biological processes. A large body of both theoretical and experimental literature exists on the molecular basis of cell polarity. While known molecular networks, which underlie cell polarity, capture many related biological phenomena there is also growing evidence that cell morphology can play crucial roles in the establishment and maintenance of cell polarity. With the work presented in this thesis I contribute to the literature on the interplay between cell morphology and cell polarity: I present results that touch upon the influence exerted by both cell morphology and changes in morphology on cell polarity, and I show novel behaviour that emerges when cell morphology and cell polarity are linked through a feedback loop. I also present a novel computational tool that can aid in further unravelling the interplay between cell morphology and cell polarity.

1.1 Cell Polarity

Cell polarity is essential in a multitude of cellular and developmental processes (Jilkine & Edelstein-Keshet, 2011; Johnson, 1999), Figure 1.1: A few examples include plant reproduction, bud formation that precedes reproduction in *Saccharomyces cerevisiae*, directed migration of keratocytes, and shape development in early plant and animal morphogenesis.
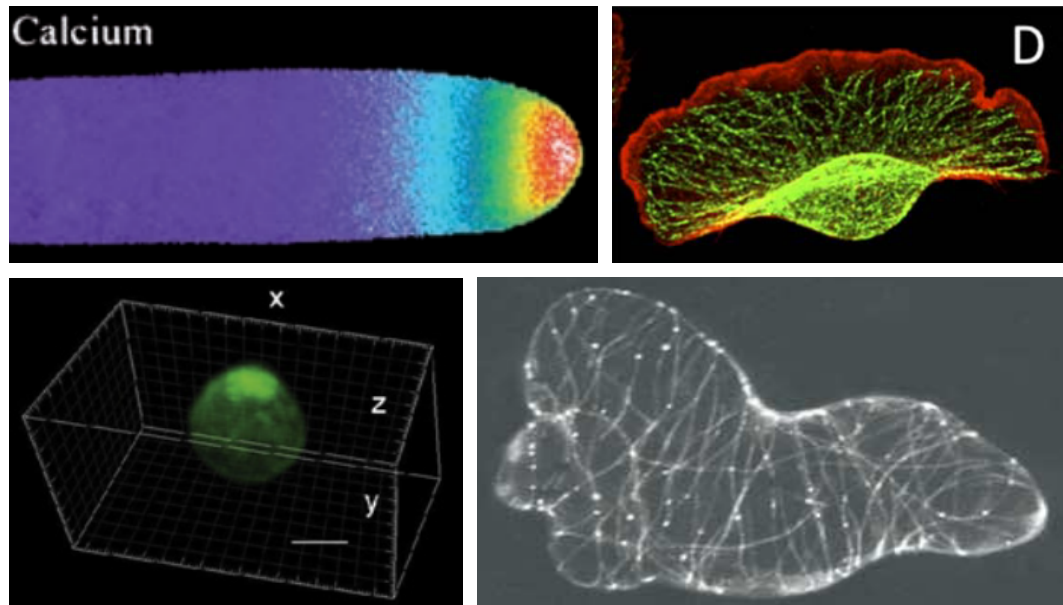
Figure 1.1: **An Overview of Cell Polarity in Different Biological Systems.** (top left) A pollen tube with tip polarity indicated by an observed gradient in calcium concentrations. (top right) A keratocyte cell with F-actin (red) localised at the leading edge which is defined by an underlying polarity process. (bottom left) A yeast cell, *Saccharomyces cerevisiae*, with a bud forming in a process that is dependent on Cdc42-induced polarity (Cdc42 is a homologue of Rho). The scale bar represents 2 μm. (bottom right) A pavement cell of a plant leaf with an effector of a Rho-homologue highlighted as white spots – this Rho-homologue induces polar growth in multiple directions. Pollen tube reproduced from Hepler et al. (2006). Keratocyte reproduced from Chen et al. (2013). Budding yeast reproduced from Slaughter et al. (2009). Pavement cell reproduced from Fu et al. (2005).

Cell polarity is established in most of these processes by spatially localised activation of specific cellular processes which is usually achieved through redistribution and localisation of multiple proteins and lipids (Johnson, 1999; Jilkine & Edelstein-Keshet, 2011). Eukaryotic cells are generally endowed with a system for gradient sensing that permits them to polarise in response to external environmental and developmental cues (Jilkine & Edelstein-Keshet, 2011).

While the molecular details of gradient sensing and polarity differ between cells and conditions, there are a number of universal features of cell polarity that are common to most scenarios (Jilkine & Edelstein-Keshet, 2011): Cells are able to sense both steep and shallow gradients while ignoring the absolute levels of the signal (adaptation), many cells can respond to changing gradients and maintain polarity even when the gradient is removed, and many cells can polarise spontaneously and usually maintain exactly one axis of polarity. Pavement cells, Figure 1.1 (bottom right), and other exceptions exist that maintain multiple axes of polarity.

1.2 Mathematical Models of Cell Polarity

In light of these exceptions, such as multiaxial polarity, it is unreasonable to assume that one class of mathematical model could capture all or most cell polarity phenomena. To this end, there are a number of models that capture cell polarity in certain phenomena better than other models (Jilkine & Edelstein-Keshet, 2011).

Local excitation global inhibition (LEGI) models usually include three chemical species (Jilkine & Edelstein-Keshet, 2011; Levchenko & Iglesias, 2002): A local activator A, a global inhibitor I, and a local response-generating species that is present in an inactive form R and an active form R^*. Production of the activator A and inhibitor I is regulated by an external signal $S(x, t)$ that upregulates both A and I in fixed proportions (Levchenko & Iglesias, 2002). Oftentimes, the activator A is a membrane-bound receptor, the inhibitor I a cytoplasmic enzyme, and the response R^* a second messenger that either diffuses along the plasma membrane (e.g. phospholipids) or in the cytoplasm (e.g. cAMP) (Postma & van Haastert, 2001). A key aspect of cell polarity that these models address is adaptation of the mechanism: Adaptation describes polarity systems that respond to gradients in the signal $S(x, t)$ rather than the absolute value of it (Levchenko & Iglesias, 2002). While these models can show perfect adaptation they are usually limited to sensing and amplifying gradients as they do not sustain polarity once the signal $S(x, t)$ has been removed (Jilkine & Edelstein-Keshet, 2011).

Another major family of cell polarity models is based on spatial instability driven by a discrepancy in diffusivity (Turing, 1952). The most basic cell polarity models with Turing-type mechanism describe two chemical species, a local self-enhancing activator and a global inhibitor (Turing, 1952; Gierer & Meinhardt, 1972). The polarity patterns predicted by these models often share the following characteristics (Jilkine & Edelstein-Keshet, 2011): They are induced spontaneously, the unpolarised state of the cell is unstable, and the concentration profile of the activator tends to show multiple maxima unless the domain size and diffusion coefficients are regulated carefully. In some models of this type, the unpolarised state of the cell is kept stable and exposure to critical levels of a uniform chemoattractant pushes the system beyond a Turing bifurcation point where it then polarises spontaneously (Narang,

2006). Another property that most models in this family predict is that cell polarity does not reorient in response to changing external signal gradients. Alterations of these models exist, however, in which the inhibitor possesses a long half life, triggering oscillatory repolarisation which permits cell polarity to reorient (Meinhardt, 1999).

A more recent family of models with a "wave-pinning" mechanism also describes the interaction between a local self-enhancer and a global inhibition process (Mori et al., 2008). Specifically, Mori et al. (2008) described the dynamics of Rho GTPases that exist on the membrane in an active form, and both on the membrane and in the cytosol in an inactive form. The phenomenological mini-model proposed by Mori et al. (2008) describes two chemical species: the active, slowly diffusing form of Rho, and the inactive membrane-bound and cytosolic forms of Rho summarised as one fast diffusing species. Cell polarity arising from this mechanism relies on the following core assumptions: The active form of Rho diffuses slowly and exerts positive feedback on its activation, and the inactive form diffuses fast and is depleted by Rho activation where depletion is usually achieved through mass conservation (Jilkine & Edelstein-Keshet, 2011). Unlike Turing-type mechanisms, the wave-pinning mechanism allows for stable rest states of the cell and the obtained wave front can be reversed in response to sufficiently strong stimuli. In further contrast to Turing-type mechanisms, while wave-pinning models can be stimulated to form multiple maxima initially these are usually resolved into one single maximum after a sufficient amount of time (Jilkine & Edelstein-Keshet, 2011). The wave-pinning model has been applied successfully to the study of multiple phenomena that depend on Rho-induced cell polarity, including polarity in human cells (Holmes et al., 2012b), F-actin dynamics (Holmes et al., 2012a), keratocyte migration (Marée et al., 2012, 2006), and neutrophil migration (Houk et al., 2012).

Since Rho GTPases are often found as a central module of polarity (Yang, 2008) I focus my work in this thesis on the localisation or polarisation of homologues of Rho. I further centre my work on the wave-pinning mini-model described by Mori et al. (2008) since it is both tractable and inherently suited for describing Rho dynamics.

As a side note, in line with existing literature I will use the phrases *Rho GTPases*

and *small G proteins* interchangeably to denote members and homologues of this family of proteins.

1.3 Cell Morphology and Changes in Cell Morphology

As one of the smallest units of life, cells occur in many contexts ranging from unicellular organisms, via individual specialised cells of larger organisms, to cells embedded in complex tissues. A few examples are shown in Figure 1.1 where pollen tubes (top left) are specialised plant cells important in plant reproduction, keratocytes (top right) are specialised cells of the connective tissue important in wound healing, budding yeasts (bottom left) are unicellular organisms, and pavement cells (bottom right) are multiaxial epidermal cells in plant leaves. In many biological contexts cells maintain a wide range of morphologies that have the potential to convey vastly different intracellular properties to the biochemical processes occurring within them.

For instance, lamellipodia situated at the leading edge of motile keratocytes are usually many microns wide but only about $0.1-0.2\ \mu$m high (Mogilner & Keren, 2009). The morphology of lamellipodia stands in stark contrast with the more spherical shape of the cell body of keratocytes that are typically $10\ \mu$m in diameter. Hence in keratocytes, biochemical processes that may encounter a relatively small effective volume in the lamellipodium will encounter a far bigger volume in the trailing cell body. This raises the question of how biochemical processes that occur in keratocytes react to potentially vastly different copy numbers of their involved proteins when comparing lamellipodia with the cell body.

Saccharomyces cerevisiae (budding yeast) divides by forming a bud and undergoing mitosis that pinches the developed bud off as a daughter cell. Reproducing budding yeast cells undergo dramatic changes in size as they double their volume in just under two hours before undergoing mitosis (Cross, 1988). It is therefore an interesting question to consider how biochemical processes that underlie polarity in the budding yeast can cope with rapid size doubling.

Pollen tubes are highly specialised, elongated cells that grow up to 1 mm/h and, in some cases, increase their effective volume by a factor of twenty over their lifes-

pan (Nishikawa et al., 2005; Rounds & Bezanilla, 2013). Further, considerable differences in the effective volume of the cytoplasm can be observed for different pollen tubes of the same species grown under identical conditions (Qin et al., 2012). Therefore it is likely that conserved biochemical processes important in establishing polarity in the pollen tube encounter vastly differing volumetric conditions in different pollen tube species and different tubes of the same species.

The morphology of the cytoplasm can also have effects on biochemical reactions: The phenomenon of macromolecular crowding has been argued by multiple authors to have substantial effects on both the diffusivity and reactivity of proteins inside cells (Ellis, 2001). In many cell types this phenomenon arises from the presence of a mix of macromolecules in the cytoplasm that excludes between 5% and 40% of the total volume accessible to biochemical reactions (Schnell & Turner, 2004). This phenomenon can confine biochemical reactions to one-dimensional channels with restricted effective volume and therefore reduced copy numbers of the involved biochemical species (Schnell & Turner, 2004).

As this summary of differences in cell morphology within and between cells shows, conserved biochemical processes are likely to occur in a multitude of different effective reaction volumes. Further, as cells grow and divide these same biochemical processes become subject to very dynamic volumetric properties.

The polarisation of Rho homologues entails biochemical processes and transport phenomena that can be affected by both the volumetric properties of the system and the copy number of the involved biochemical species. It is therefore important to understand the influence of cell morphological differences and changes on the processes that underlie cell polarity. Further it is interesting to ask whether potential feedback between morphology, biochemical processes, and cell polarity has merit in explaining observed biological phenomena.

1.4   Thesis Outline

In this thesis I study the interplay between cell morphology and Rho GTPase-dependent cell polarity modelled with the wave-pinning mini-model presented by Mori et al. (2008).

The outline of this thesis is as follows: In Chapters 2 and 3 I study the wave-pinning mini-model in cellular contexts at opposite extremes of the cell size spectrum.

In Chapter 2 I focus on the behaviour of the mini-model in small effective reaction volumes which either arise in cells of minute size or cells that are subject to macromolecular crowding. This restriction to small reaction volumes gives rise to conditions where only a few hundred or thousand protein copies interact in the reactions of the mini-model. Here I elucidate the effect of small copy numbers on Rho-dependent polarity and some of the mechanisms through which stochastic effects may challenge polarity.

In Chapter 3 I focus on the effects of large reaction volumes and dynamically increasing reaction volumes on cell polarity modelled with the Rho mini-model. Here I also close the loop between cell morphology and cell polarity by linking the rate of morphology change to the state of the polarity system. My work in Chapter 3 reveals that this closed feedback loop enriches the dynamic behaviour of the mini-model dramatically.

While in Chapters 2 and 3 I study the existing wave-pinning mini-model, in Chapter 4 I make the case for unfurling this phenomenological model. As I discuss in Chapter 4, biological studies make it clear that molecular details of the biochemical processes underlying cell polarity can be crucial in illuminating the ties between morphology and polarity. The tool I develop in Chapter 4 can be used for rapid model disambiguation in the search for these molecular details and is therefore a useful tool for further studies of the interplay between cell morphology and cell polarity.

Chapter 2

Deterministic Versus Stochastic Cell Polarisation Through Wave-Pinning

Abstract

Cell polarisation is an important part of the response of eukaryotic cells to stimuli, and forms a primary step in cell motility, differentiation, and many cellular functions. Among the important biochemical players implicated in the onset of intracellular asymmetries that constitute the early phases of polarisation are the Rho GTPases, such as Cdc42, Rac, and Rho that present high active concentration levels in a spatially localised manner. Rho GTPases exhibit positive feedback-driven interconversion between distinct active and inactive forms, the former residing on the cell membrane, and the latter present predominantly in the cytosol. A deterministic model of the dynamics of a single Rho GTPase described earlier by Mori *et al.* exhibits sustained polarisation by a wave-pinning mechanism. It remained, however, unclear how such polarisation behaves at typically low protein copy numbers, as stochasticity could significantly affect the dynamics. We therefore study the low copy number dynamics of this model, using a stochastic kinetics framework based on the Gillespie algorithm, and propose statistical and analytic techniques that help us analyse the equilibrium behaviour of our stochastic system. We use local perturbation analysis to predict parameter regimes for initiation of polarity and wave-pinning in our deterministic system, and compare these predictions with deterministic and stochastic spatial simulations. Comparing the behaviour of the stochastic

with the deterministic system we determine the threshold number of molecules required for robust polarisation in a given effective reaction volume. We show that when the molecule number is lowered wave-pinning behaviour is lost due to an increasingly large transition zone and fluctuations in the pinning position that cause the collapse of the wave, while the high and low equilibrium levels are affected far less.

2.1   Introduction

Many eukaryotic cell types undergo directed movement in a variety of scenarios. Such motility is important in embryogenesis (Charest & Firtel, 2007), wound healing, immune surveillance (Ridley et al., 2003), and cancer metastasis (Ridley et al., 2003).

As a first step in this process, cells polarise, forming a distinct front and rear distinguished by biochemical profiles of signalling molecules that regulate lamellipodial extension (Ridley, 2006). An important part of that internal polarising biochemistry is based on the activity and distribution of Rho GTPases.

These switch-like signalling proteins exhibit a distinct active (GTP, membrane-bound) form and an inactive (GDP) form that is largely cytosolic. Only the active, GTP-form is able to interact with downstream effectors to exert its biological function. Interchange between these two forms is mediated by GTPase-activating proteins (GAPs), which augment inactivation, and guanine nucleotide exchange factors (GEFs), which facilitate activation. It has been established that the active form increases its own rate of activation via various self-recruitment mechanisms (Raftopoulou & Hall, 2004; Li et al., 2003). While the active form binds the plasma membrane, the inactive form can be both in the membrane or released to the cytoplasm, a process that is positively regulated by binding to guanine nucleotide dissociation inhibitors (GDIs).

When a cell is stimulated, some Rho GTPase activity (notably, Cdc42 and Rac1) is focused at the leading edge (Ridley et al., 2003), inducing localised actin polymerisation that generates protrusive forces propelling the cell (Raftopoulou & Hall, 2004).

Here we are concerned about the onset of polarity and its maintenance, thus focusing only on the polarisation of the Rho pattern, and not on the downstream remodelling of the cytoskeleton (or possible feedback that this might generate). In a later chapter of this thesis we present and study a model that includes feedback between Rho localisation and effective downstream remodelling of the cytoskeleton.

Based on general interactions between Cdc42, Rac and Rho, and taking into account known parameters for the kinetics and diffusion of the active and inactive forms, it has been shown that sustained polarisation within a cell is possible while the homogeneous steady state (the rest state of the cell) remains stable (Marée et al., 2006; Jilkine et al., 2007). Later, Mori et al. (2008) determined the mathematical essence of the mechanism by studying a reduced deterministic model of cell polarisation, coining it "wave-pinning". We refer to this reduced model as a mini-model for Rho-dependent cell polarisation.

It remains, however, unclear to what extent stochasticity at low molecule numbers can influence the potential of the mechanism to initiate and sustain polarity within the cell. We therefore compare and contrast the deterministic and the stochastic version of the mini-model for wave pinning.

A simple 1D geometry in which this generic Rho GTPase can be studied is shown in Figure 2.1, where organelles and the nucleus are omitted, L is a cell diameter, and the chemical system is modelled by a two-component reaction with distinct rates of diffusion $D_a \ll D_b$ across L, since proteins diffuse much more slowly in the lipid membrane than in the cytosol. Here A is the active and B the inactive small GTPase (with concentrations $a(x, t)$ and $b(x, t)$). The height H and width W of the compartment are assumed to be reasonably small, so gradients are described in the x direction for $x \in [0, L]$, $t \geq 0$.

The system of reaction-diffusion equations of the deterministic GTPase model is (Mori et al., 2008):

$$\frac{\partial a}{\partial t} = D_a \frac{\partial^2 a}{\partial x^2} + f(a, b), \tag{2.1a}$$

$$\frac{\partial b}{\partial t} = D_b \frac{\partial^2 b}{\partial x^2} - f(a, b). \tag{2.1b}$$

Here $f(a, b)$ is the rate of GTPase interconversion. Equations (2.1) are taken with no-flux boundary conditions: $a_x(0) = a_x(L) = 0, b_x(0) = b_x(L) = 0$ and the interconversion rate is modelled to include auto-activation of A (through a positive feedback of A onto its own production):

$$f(a, b) = k_0 b + \frac{\gamma a^2}{K^2 + a^2} b - \delta a, \tag{2.2}$$

where k_0 and δ denote the basal rates of activation and inactivation of A respectively, γ is the rate of maximal feedback strength and K is the concentration of A leading to a half-maximal feedback level.

We first briefly describe the deterministic aspects of this model, and introduce a local perturbation analysis that leads to insights on how the initiation of polarisation depends on the parameters and on the total amount of molecules. We then explore how the polarisation mechanism reacts when only a limited number of molecules is available and stochasticity starts to impact on the polarisation state of the cell. To do so, we describe and analyse an analogous stochastic version (low copy number regime) of the mini-model. We confirm correctness of our stochastic implementation by showing that simulations with large molecule and lattice numbers approach the thermodynamic limit. To analyse the equilibrium behaviour of the stochastic system we introduce statistical tools that provide us with intriguing insights regarding the dynamics in the low copy number regime. Using these tools we discover that in the low copy number regime the system is dominated by spatial fluctuations of the transition zone rather than temporal fluctuations in the activity levels. We further show that loss of polarity is due to the region of high activity reaching a broadness that is unsustainable, causing the sudden collapse of the whole wave. Bifurcation analysis of a simplified model of a pinned wave provides an explanation for the behaviour of the stochastic system close to the point where the wave is lost due to stochastic fluctuations.
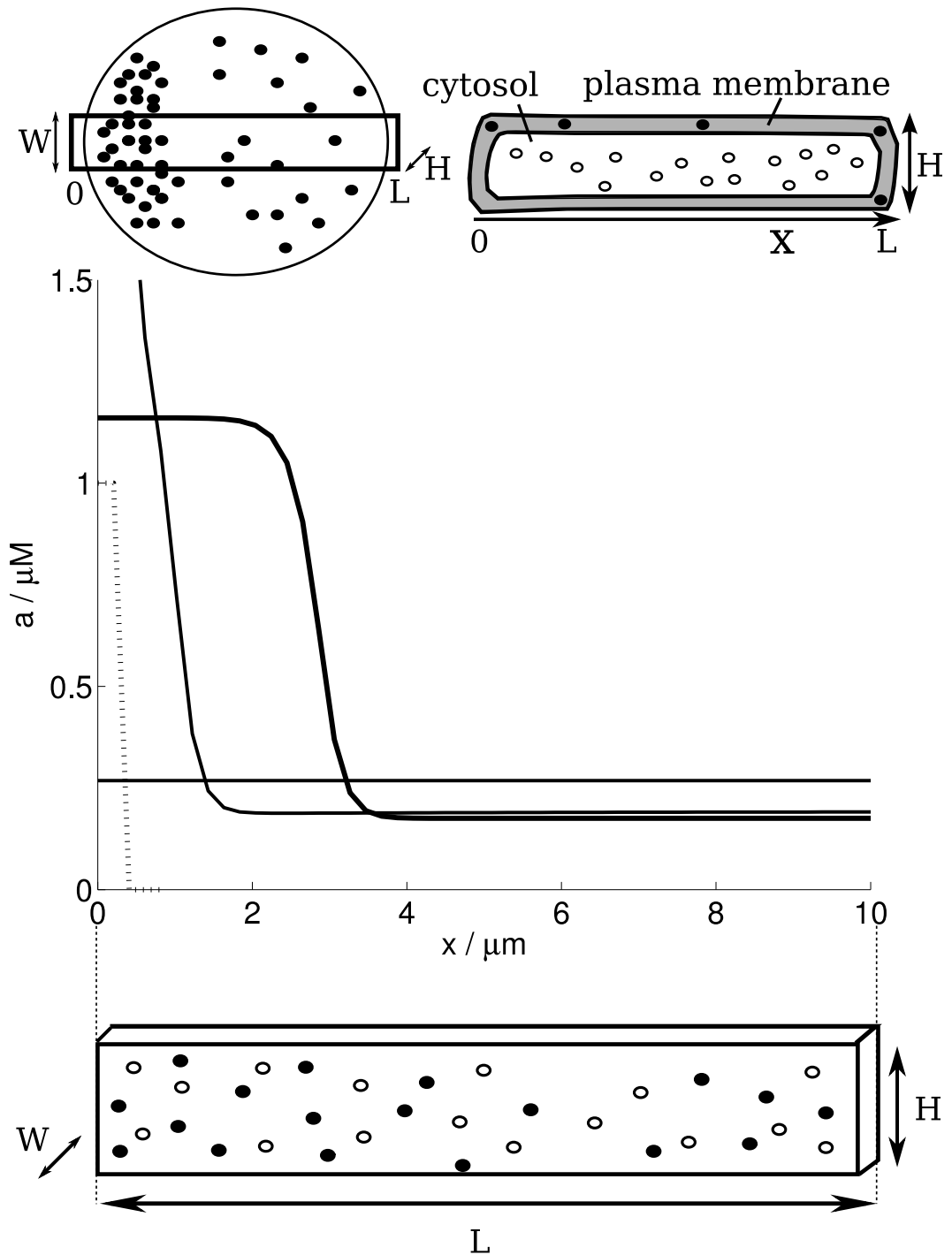
Figure 2.1: **The Modelled Cell**. (top): Schematic diagrams of a cell showing the "slab" of length L, height H, and width W in a top-down and side view. The model distinguishes membranous (solid circles; called A in this manuscript) and cytoplasmic (open circles; called B) proteins only by their distinct rates of diffusion. A typical "polarisation" state is shown in grey/white in the top-down view. (bottom): In the deterministic polarisation model (2.1) proposed by Mori et al. (2008), a small stimulus (dashed line, not to scale) produces a pinned wave (solid black line).

2.2 Deterministic Behaviour

2.2.1 Wave-Pinning

Given appropriate conditions and using Equation (2.1), a stimulus-pulse of GTPase located within an otherwise homogeneous domain leads to the formation of a travelling wave of activation that slows down and stalls, delimiting a spatial region of activation, thus creating a robustly polarised cell (Mori et al., 2008).

In our simulations, this initial pulse is captured by a first-order reaction converting b to a within a small domain of the cell. This is achieved through the term $k_s b(x, t)$, which is both added to Equation (2.1a) and subtracted from Equation (2.1b).

The wave-pinning regime depends on the relative rates of diffusion and the total amount $T = \int_0^L a(x, t) + b(x, t) dx$ in the system. The mechanism of wave-pinning can be attributed to the following: relatively rapid diffusion of b ($D_b \gg D_a$) leads to a relatively constant level of b over the cell, while the existence of three roots of $f(a, b)$ for the fixed well-mixed equilibrium b level allows for a sufficiently large local perturbation in a to locally reach a distinct activity level (a process that we have coined "Δ-perturbability", see below). Mass conservation ensures that, while this peak of a levels expands similarly to a propagating wave, the relatively constant homogeneous level of b drops. This global decrease of b slows down and eventually limits the spatial propagation of the wave, pinning it in an equilibrium position (Mori et al., 2008).

Even though wave-pinning requires the existence of three roots of $f(a, b)$ for fixed b, it is important to realise that it is not a consequence of bistability and subsequent front propagation between two stable states. (Note that in reaction-diffusion systems, the terminology bistability is used to denote cases in which the corresponding well-mixed system has two distinct stable steady states.) The well-mixed ODE system has only one equilibrium, and in the PDE this equilibrium is stable against both homogeneous and small non-homogeneous perturbations. Nevertheless, in the spatial setting a sufficiently large local perturbation can trigger the travelling wave, which subsequently stalls, giving rise to sustained polarity.

2.2.2 Wave-Pinning Versus Propagation Failure

Because we compartmentalise space in this study to perform stochastic simulations, it is relevant to introduce yet another mechanism, coined propagation failure (Britton, 1985; Keener, 1987). As a possible source of confusion, propagation failure has previously also been referred to as "pinning of waves" (Fáth, 1998), thus evoking the need to emphasise its clear distinction from "wave pinning" as defined in Mori et al. (2008).

Propagation failure describes a specific phenomenon that can be observed in bistable systems in which travelling waves fail to propagate when space is discrete. This may occur when both the wave velocity is low and the discretisation of the space is coarse (relative to the diffusion coefficient) (Keener, 1987; Fáth, 1998). Under such conditions, propagation failure can manifest itself if, in the location of the wave front, the diffusive flux from one subdomain into the next becomes insufficient to bring the levels of that subdomain above the threshold required for the amplification and subsequent propagation of the wave. In contrast, the phenomenon of wave-pinning does not require discretised space. Instead, when the triggered wave spreads over the domain the velocity of the wave decreases because of the drop in the available inactive form that is used up by being converted into the active form.

Nevertheless, we here find that both phenomena become coupled to one another when space is discretised. Due to slowing of the wave as a result of wave-pinning the velocity of the wave eventually becomes sufficiently low that propagation failure will occur in coarse grids. Consequently, when we discretise space in this study, which we do in both numerical PDE simulations and in Gillespie simulations, propagation failure occurs for coarse grids as well as low diffusion rates.

Given that subdivision of space into grid points is a computational method but does not represent a biological property of the cell, we will ensure below that propagation failure does not play a role in the dynamics presented in this manuscript nor influences the biological insights we derive here.

2.2.3 Analysis of Polarity Initiation

The full bifurcation analysis of any system of partial differential equations (PDEs) is a challenging undertaking. While Mori et al. (2011) focused on the requirements of the travelling wave to stop, we will here discuss an analysis regarding the potential to initiate polarity and a travelling wave, in which we probe the homogeneous state of the cell with a local perturbation.

In short, we ask what happens if a local perturbation is introduced into a resting cell (being at a uniform steady state), by observing whether such a perturbation will diverge to a distinct local equilibrium (eventually causing polarisation through wave-pinning), or alternatively dampen out, returning to the rest state corresponding to the globally homogeneous state of the cell. This analysis provides a straightforward test whether a (sufficiently large) perturbation can "invade" the initially uniform steady state solution.

We refer to this reduced model as the "local perturbation analysis" (LPA) model, or system, as it allows us to study invasion criteria for a local perturbation of any given amplitude. Note that such Δ-perturbability does not directly imply sustained polarisation through wave-pinning, see below.

To address the onset of polarisation without having to deal with the full complexity of the PDEs, we break down the spatial system into two effective compartments, one corresponding to the levels of the active and inactive form at the site of the local perturbation, and another corresponding to the global values over the remainder of the cell. The simplified representation of the deterministic system is given in Equations (2.4), and is used to predict the total amount of small GTPase T for which to expect initiation of polarisation and wave-pinning. This analysis allows us to compare results from the deterministic and the stochastic version of the model later on. For the local perturbation analysis, we introduce the following assumptions and approximations into the PDE model:

- We ask whether the value of A at a site of the localised pulse $a_L(0)$ will diverge from the uniform global concentration of active GTPase $a_G(t)$. Since this active form has a very low rate of diffusion, we consider the limit $D_a \approx 0$ and treat $a_L(t)$ as a purely local variable, that can vary independently from $a_G(t)$. This

is equivalent to assuming that any perturbation in A will be spatially confined to the site of the perturbation and will initially evolve independently of the rest of the domain.

- Since the inactive GTPase B has a relatively fast rate of diffusion, we take the limit at which $D_b \approx \infty$ and consider $b(t)$ to be a purely global variable ($b_L(0) = b_G(t) \equiv b(t)$). Restated, any local perturbation in B caused by the local perturbation in A will be instantly adjusted to the global, homogeneous concentration profile.

This leads to the following LPA model:

$$\frac{da_L}{dt} = f(a_L, b), \quad \frac{da_G}{dt} = f(a_G, b), \quad \frac{db}{dt} = -f(a_G, b). \tag{2.3}$$

Furthermore, given that we consider only a narrow initial pulse of activation that hardly affects the overall cell levels, it is reasonable to approximate $T(t) \approx \int_0^L a_G(t) + b(t) dx \approx$ constant, so that $b(t) \approx (T/L) - a_G(t)$. Eliminating $b(t)$ by conservation leads to a system of two ODEs:

$$\frac{da_L}{dt} = f(a_L, (T/L) - a_G(t)), \tag{2.4a}$$

$$\frac{da_G}{dt} = f(a_G, (T/L) - a_G(t)). \tag{2.4b}$$

The approximations required for the polarity-invasion analysis are depicted in Figure 2.2. Note that since the activating pulse is confined to a sufficiently small subsection of the domain its variation will not affect the total amount of inactive form, and therefore the dynamics of B solely depends on the global level of A in the extended domain.

We can use such approximations to address the following questions: under what circumstances would a localised pulse of activation grow in magnitude compared with the surrounding levels? How large should the amplitude of the stimulus be to trigger a new state (e.g. depicting an initial polarisation)? And, if a new, bounded state exists, what values do we expect it to have? These answers depend on the parameters and on the total number of molecules (i.e. amount of small GTPase) in

the system. Note that we do not restrict attention to small amplitude perturbations as is common in linear stability analysis of reaction-diffusion equations.

The LPA system, Equation (2.4), can be studied in several ways, e.g. via the $a_G a_L$-phase plane and ODE bifurcation analysis. Performing such analysis reveals five qualitatively distinct regions of polarity behaviour, depicted in Figure 2.3.

The behaviour in each region can be captured with a sequence of corresponding $a_G a_L$-phase plane diagrams. In the bifurcation diagram, we plot the steady state value a_L versus the bifurcation parameter T, representing the total amount of the small G-protein. The five distinct regimes of behaviour (labelled I–V) are separated by two saddle-node and two transcritical bifurcations. The shape of the curve traversing the diagram from lower left to upper right is shared with the bifurcation curve that would be obtained when the PDE system is taken to be well-mixed ($D_a = D_b = \infty$). It represents a steady state where $a_G = a_L$ (unpolarized cell, "rest state"). The two actually coincide, meaning that the equilibrium level in the local patch is neither lower nor higher than the uniform global background level of a_G. Note that for any equilibrium found in the well-mixed case, there should exist a corresponding equilibrium $a_G = a_L$ in the LPA model. The stability of the equilibrium, however, can change (between stable and unstable) in a transcritical bifurcation. For example, while the equilibrium in the well-mixed case is always stable, the portion of that curve in Region III of the LPA model is unstable. We indicate the line $a_G = a_L$ in the $a_G a_L$ planes, using a dashed grey line, representing the absence of any local perturbation.

We now explain how to interpret the diagram and its implications for polarity behaviour. (I) In region I the total amount of molecules is low ($T < 19.09$) and a regime is found in which only one steady state value $a_L = a_G < 0.2$ exists. That state of low activation is unresponsive to stimuli, and no pulse can "invade". The time-space plot of $a(x, t)$ stays at, or rapidly returns to, a low uniform level (solid curve) no matter how large the amplitude of an applied stimulating pulse. That is, a cell will not polarise when stimulated and it will remain in a rest state. (II) This is the region corresponding to the deterministic regime of cell polarisation (wave-pinning). Here, for an intermediate level of substance $19.09 < T < 23.0$, there are three coexisting equilibria of Equation (2.4), the outer two of which are stable. A value of a_G

corresponding to the lower branch can be "invaded" by a local pulse, provided its amplitude is large enough to surpass a threshold depicted by a dashed curve in region II. The lower branch corresponds to the $a_G = a_L$ equilibrium, i.e. to a cell that is in a homogeneous rest state. Thus, the rest state is stable against small perturbations, but sufficiently large perturbations can polarise the cell. Note that as the total amount increases, the required amplitude to trigger polarisation decreases sharply, so that close to but just short of $T{=}23.0$ a pulse of very small size can lead to polarisation. We see that the full PDE solutions (top panels in Figure 2.3) show the invasion of such a pulse in this regime, which becomes established at a finite amplitude over some fraction of the domain. (III) Other patterned states (e.g. with one or more patches of active GTPase) occur in this region. For $23.0{<}T{<}25.99$, the global steady state $a_G{=}a_L$ is unstable to any perturbation. Here small amplitude noise or a pulse of small magnitude will disrupt the global state leading to other patterned states. This kind of behaviour is typical of a Turing instability. Indeed, the full PDE solution (with random noise initial conditions such that the total amount falls in this range) produces patterns with multiple peaks. (IV) For even higher values, $25.99{<}T{<}35.58$, the total amount of GTPase is so high that the global level of activation is at an elevated steady state level (highest solid branch of the diagram). Here an invading "pulse" has to locally deactivate a region in order to "invade" (i.e. the pulse is a dip below the uniform global level). The amplitude of that "dip" must cross the threshold (dashed portion of curve) to trigger the polarisation, as otherwise it decays back to the uniform activation level. As shown in the solution of the PDEs, a dip of sufficiently large amplitude leads to a stable local patch of depressed activity in an otherwise high global level of activity. (V) Finally, above $T{>}35.58$ the potential of polarisation is lost again. That is, no pulse nor dip can invade, and the uniform global state is one of high activity everywhere in the domain.

The LPA does not address the question in what position the wave will be pinned, but rather if a wave can be triggered and how tall it will become. The next question therefore is in what position along the cell length the wave stalls. We denote the wave position by L_0, and the equilibrium value of L_0 where the wave stalls by $L_0{}^*$. In Mori et al. (2008), the wave-pinning position has been derived mathematically for the limiting case of an infinite difference in diffusion rate between the active

and inactive form (i.e. using a sharp front approximation). In the bottom panel of Figure 2.3 we show the steady state value $L_0{}^*$ as a function of T. We elaborate on our numerical estimation of the sharp-front approximation in Chapter 3. Regarding the wave-pinning itself, three regions of qualitatively different behaviour can be discriminated. In regions *i* and *iii* no stable polarity can be found, because any wave would completely retract or expand over the whole domain, respectively. In contrast, in region *ii* we find the possibility for stable coexistence of a high and a low state. Note that the pinning position $L_0{}^*$ depends on the value of T. Importantly, the figure shows that the interval of sustained polarity is smaller than the interval of regions II–IV. It illustrates that even when a wave can be triggered, it does not always follow that it can also be sustained.

Note that both the results of the LPA and of the wave-pinning position act as an approximation of the behaviour of the full PDE, where actual rates of diffusion are finite and initial conditions can affect whether initially a single peak or several peaks emerge. However, they capture correctly the basic boundaries that determine the potential to polarise, the minimum perturbation amplitude required to do so, the expected values to be reached in the local perturbation, and the position where the wave pins. We will show how this contributes valuable insight when interpreting the role of stochasticity in cell polarisation through wave-pinning.
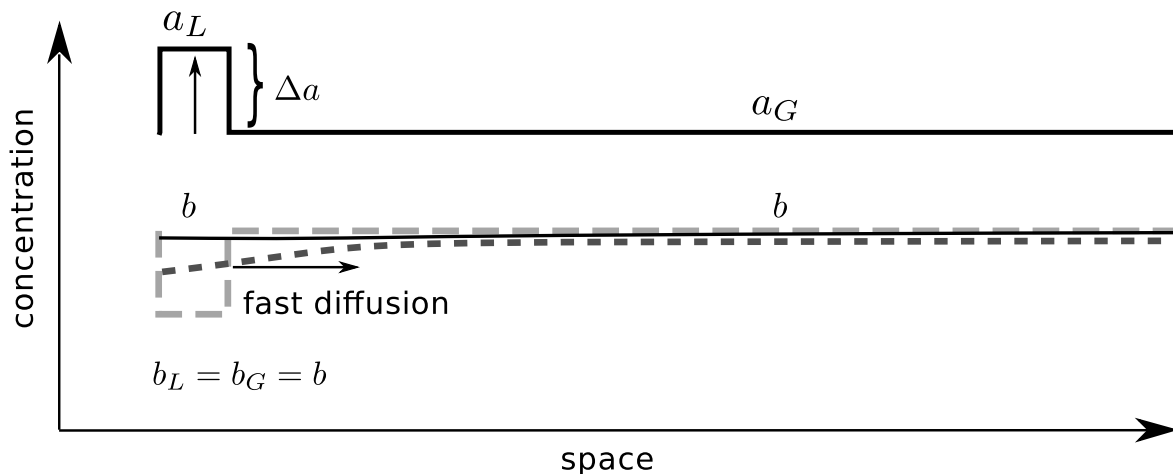
Figure 2.2: **Local Perturbation Analysis.** Schematic representation of the deterministic local perturbation analysis (LPA). In our reduction, we introduce a local perturbation Δa, and consider the evolution of the local variable a_L, the inactive protein (represented by the dotted lines, evolves to the straight black line, such that $b_L = b_G = b$), and the active protein in the rest of the cell, a_G.
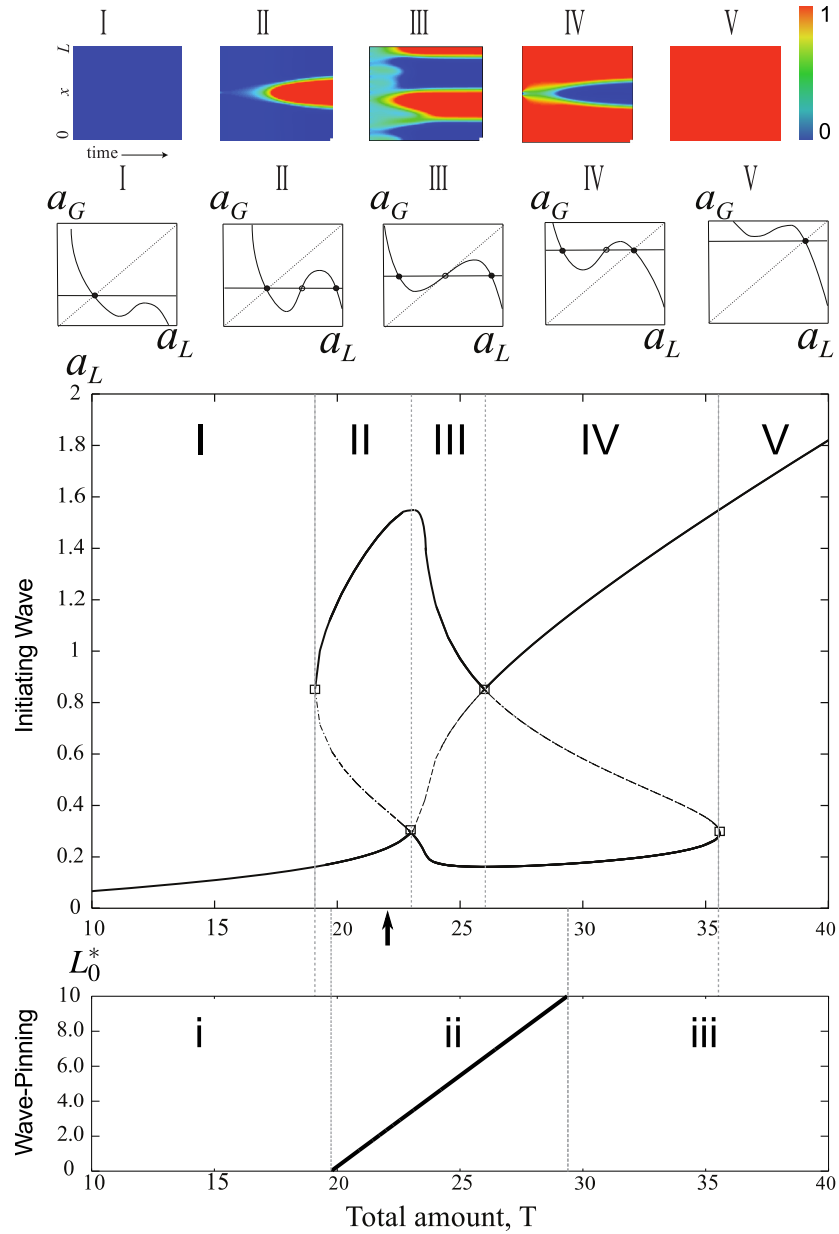
Figure 2.3: **Regions of Polarisation Behaviour in the Deterministic Model.** Unless otherwise indicated, all model parameter values are chosen as in Table 2.1. Middle graph: Bifurcation diagram of the reduced deterministic model system, Equation (2.4), showing steady-state local activated form, a_L, versus the total amount of material in the domain (T). Five distinct regimes of behaviour are found, as explained in the text. Top row: colour plots of the solutions of the PDEs, starting close to the uniform steady state in the given region. In these space-time plots, the time axis is horizontal and the space axis is vertical. The activity level is depicted using a colour gradient, with red indicating the highest and blue the lowest activity levels. Middle row: phase plane behaviour in the $a_G a_L$ planes, showing the number and stability of steady states of the reduced system in each of the regions. In regions I and V only a uniform level of global activity is stable, and no pulse or stimulus can grow. In the intermediate regions only a sufficiently large pulse (II), or a sufficiently low dip (IV) can grow. In region III, the homogeneous state is unstable to any perturbation, and a variety of patterns can form, depending on initial conditions. Bottom graph: Wave-pinning position as a function of T. In regions i and iii no sustained polarised state is possible, while in region ii wave-pinning occurs, with the position of pinning, L_0^*, monotonically increasing with T.

2.3 Stochastic Version

Next we ask how the same polarisation mechanism would behave in the low copy number regime. We ask under what conditions a stochastic equivalent of the deterministic model still presents wave-pinning, i.e. after triggering the formation of sustained regions within the cell with respectively low and high levels of the active form, and if our approach predicts biologically relevant conditions under which wave-pinning may be unsustainable in live, noisy cells.

We resort to the stochastic formulation of chemical kinetics (McQuarrie, 1967), and use the stochastic simulation algorithm (SSA) due to Gillespie (1976) for our stochastic simulations of Equation (2.1). For the discrete nature of stochastic simulations, we subdivide the domain of length L into N compartments of equal width $h=L/N$, see Figure 2.4.
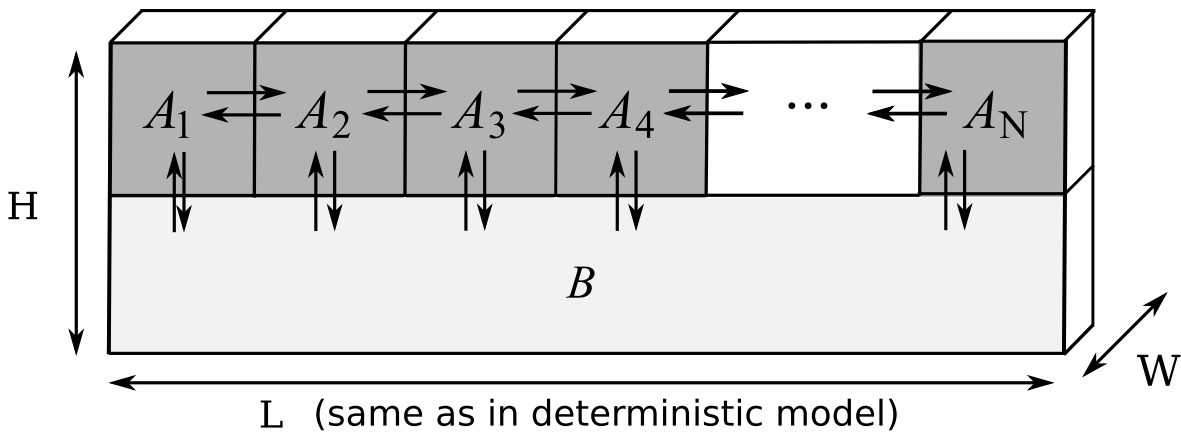
Figure 2.4: **Spatial Set-Up of the Stochastic Simulations.** The membrane is divided into N equally wide well-mixed subdomains. Diffusion within the membrane is modelled as first-order reactions while we assume that B diffuses fast enough to warrant modelling it as occupying a well-mixed cytoplasmic pool. Dimensions H and W associated with the domain of length L are necessary for conversion of concentrations to numbers of molecules but do not play any other role: all stochastic simulations are one-dimensional.

In the deterministic case, we choose experimentally-supported (Postma & van Haastert, 2001; Postma et al., 2004) relative diffusion values, $D_b=100D_a$, that effectively render B homogeneous. In our stochastic simulations we approximate B as a homogeneous cytoplasmic pool ($D_b\rightarrow\infty$) to reduce the computational cost, while maintaining a full computation for the heterogeneous A distribution. Diffusion of A in the membrane is treated as a series of first order reactions (spatial SSA (Erban et al., 2007)), each with propensity

$$\frac{D_a}{h^2} A_i, \quad i = 1, \ldots, N,$$

where A_i denotes the current number of molecules of A in membrane lattice point i. The above propensity equals zero for diffusion to the left when $i = 1$, and diffusion to the right when $i = N$, respectively (no-flux boundary conditions). Similarly, we choose the remaining propensities according to Gillespie (1976):

$$\begin{aligned}
\text{Background activation into membrane lattice point } i : \quad & k_0 \frac{B}{N}, \\
\text{autoactivation of } A \text{ in membrane lattice point } i : \quad & \frac{\gamma A_i^2}{K_N^2 + A_i^2} \frac{B}{N}, \\
\text{background inactivation out of membrane lattice point } i : \quad & \delta A_i, \\
\text{pulse activation into membrane lattice point } i : \quad & k_s \frac{B}{N}.
\end{aligned}$$

We note that B, the current number of molecules of the inactive species in the well-mixed cytoplasmic pool, is rescaled with $1/N$ because each membrane lattice point only senses this fraction of the available total number of molecules of B. For each propensity p, the probability that the corresponding event occurs within the next dt units of time equals $p \cdot dt + o(dt)$, where $o(\cdot)$ denotes terms that converge to zero quicker than its argument (little o notation, $\lim_{dt \downarrow 0} \frac{o(dt)}{dt} = 0$).

In the above propensity expressions, most kinetic constants are equal to those used in the deterministic system, Equation 2.1, since they are independent of the units of $a(x,t)$ and $b(x,t)$, or A_i and B correspondingly. However, the Michaelis-Menten constant K of Equation 2.2 has the same units as $a(x,t)$ and needs to be rescaled for our stochastic simulations:

$$K_N = 10^{-21} N_A V_a K,$$

where N_A is Avogadro's constant, V_a denotes the volume associated with each membrane lattice point, $V_a = \frac{L}{N} W \frac{H}{2}$, and the factor 10^{-21} is required to rescale units (V_a has units of μm^3, and K units of μM). Homogeneous initial concentrations of A and B, $a(x,0) = a_0$ and $b(x,0) = b_0$, are rescaled to numbers of molecules equivalently:

$$A_i(t = 0) = 10^{-21} N_A V_a a_0 ,$$

$$B(t = 0) = 10^{-21} N_A V_b b_0 ,$$

where $V_b = LW\frac{H}{2}$.

Note that in our stochastic simulations we need to associate volumes with each lattice point (both for the membrane and for the cytoplasm), since we reformulate a concentration-based model, Equation 2.1, as a molecule-based stochastic model. Even though we specify a volume for each lattice point in this conversion, the stochastic simulations are effectively one-dimensional, as are our deterministic simulations. That is, in both the deterministic and the stochastic system we focus on radial polarisation along the diameter of a cell.

2.4 Results

For our simulations and comparison between the stochastic and deterministic case, we use parameter values based upon Mori et al. (2008), and summarised in Table 2.1: $k_0=0.067\,s^{-1}$, $\gamma=1\,s^{-1}$, $K=1\,\mu M$, $\delta=1\,s^{-1}$, and diffusion coefficients $D_a=0.1\,\mu m^2/s$, $D_b=10\,\mu m^2/s$ (the latter in the deterministic case only). For the initial stimulus we choose $k_s=10\,s^{-1}$ for $50\,s\leq t\leq70\,s$ and $0\,\mu m\leq x\leq0.4\,\mu m$, and $k_s=0$ otherwise. For biologically reasonable concentrations (Marée et al., 2006), we set $b(x,0)=b_0=2\,\mu M$ and use $a(x,0)=a_0$ such that $f(a_0,b_0)=0$ and (a_0,b_0) is linearly stable ($a_0=0.2683\,\mu M$). We fix the slab height $H=0.2\,\mu m$ and length $L=10\,\mu m$ (see Figure 2.1) and simulate our stochastic system for varying numbers of molecules by varying the width W of our cell slab. Varying the width W (Figure 2.4) allows us to change the number of molecules without changing the initial concentrations of A and B.

D_a	$0.1\,\mu m^2/s$	diffusion of the active form A in the membrane
D_b	$10\,\mu m^2/s$	diffusion of the inactive form B in the cytoplasm (deterministic system only)
k_0	$0.067\,s^{-1}$	rate of background activation
γ	$1\,s^{-1}$	maximal rate of auto-activation of A
δ	$1\,s^{-1}$	rate of background inactivation
K	$1\,\mu M$	concentration of A resulting in half-maximal rate of auto-activation (deterministic system)
K_N	$\cdots$	as constant K but rescaled depending on current volume of the system ($K_N \propto N_A HLW \cdot K$, where N_A is Avogadro's number)
k_s	$10\,s^{-1}$	rate of activation due to transient pulse
L	$10\,\mu m$	length of the domain
N	50	number of membrane lattice points

Table 2.1: Summary of the parameter values used in our simulations.

2.5 Propagation Failure Does not Affect the Deterministic Simulations

We first determine that discretisation of space does not cause propagation failure near the parameter values used for our analysis on stochasticity. To do so we make use of the fact that the initial perturbation that triggers a wave does not necessarily have to be small in width. A wide (but not too wide) perturbation of sufficient amplitude can also trigger a wave. Perturbations that are wider than the final pinning position, however, trigger waves with a negative velocity, i.e. waves that retract until they come to halt at the pinning position. If indeed propagation failure plays a role, both the extending and the retracting wave are expected to halt before their velocities would have become zero in the continuous case. Thus, propagation failure would cause extending and retracting waves to halt in distinct positions.

We therefore devise the following numerical experiment: We use a total amount of $T = 22.68$ which lies in the wave-pinning regime, in which a Δ-perturbation is required to trigger a wave that gives rise to sustained polarity, see black arrow in Figure 2.3. We then start simulations, for varying values of the active form diffusion coefficient D_a and number of lattice sites N, with two different wave-shaped initial conditions (see Figure 2.5): one narrower and higher, one broader and lower, both indicated by dashed lines in the figure, where the former lies to the left and the latter

to the right of the equilibrium wave-pinning position.
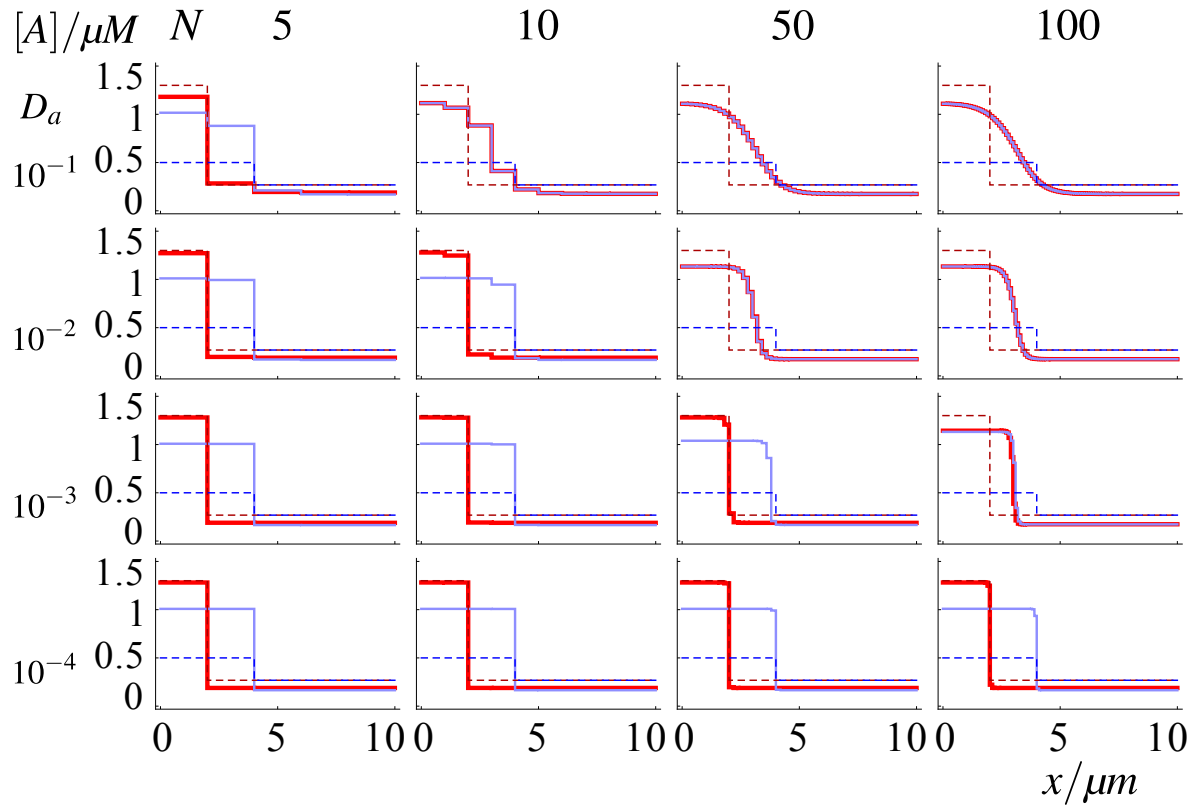
Figure 2.5: **Propagation Failure in the Deterministic System as a Function of Grid Coarseness and Diffusion Coefficient of the Active Form.** For each column a different number of compartments, N, has been used, and for each row a different diffusion coefficient, D_a. The values of both are indicated in the figure, with diffusion coefficient D_a in units $\mu m^2/s$. Dashed lines denote initial conditions, solid lines are equilibrium profiles (at $t = 20,000\,\mathrm{s}$), thick red lines denote narrow, high initial waves, thin blue lines denote broad, shallow initial waves. Discrepancies between the red and blue equilibrium profiles for slow diffusion coefficients and coarse grids can be attributed to propagation failure. At our default setting of $D_a = 0.1\,\mu m^2/s$ and $N = 50$ propagation failure is not observed.

Figure 2.5 shows the equilibrium wave profiles for the narrowly initiated waves (thick lines) and broadly initiated waves (thin lines). Sufficiently large box sizes (low N) and sufficiently low diffusion coefficients show a discrepancy in the final position of the pinned wave between the retracting and the extending waves. This illustrates that sufficiently low diffusion combined with sufficiently coarse subcompartmentalisation leads to propagation failure. For the default values used in this study ($N = 50$; $D_a = 10^{-1}\,\mu m^2/s$), however, no propagation failure can be observed. A 10-fold coarser grid or 100-fold slower diffusion would be needed for propagation failure to occur.

2.6 Equilibrium Behaviour of the Stochastic System

Individual SSA runs (Figure 2.6, top) resemble the deterministic behaviour for sufficiently large numbers of molecules ($\approx$6800), and the averaged concentration profile clearly shows wave-pinning after a stimulus is applied (Figure 2.6, bottom). With fewer molecules ($\approx$700), individual SSA runs become more erratic and the averaged profile is homogeneous across the domain, resulting in loss of wave-pinning (Figure 2.6, bottom, inset). As we increase the number of molecules by increasing W, we observe convergence of the stochastic system to the deterministic one, indicated by a decreasing observed variance and smoothing of the black solid curve.

To study the distribution of pinning positions, we fit, well after wave-pinning, the concentration profile of the active species to the symmetric Richards model (Richards, 1959), which captures well sigmoidal profiles and contains a parameter, c, that defines the position of the inflection point, which we consider to be the pinning position of the wave:

$$a_{\text{fit}}(x) = M \left(1 - \frac{1}{1 + \exp(-(x - c)/h)} \right) + m \, ,$$

where M and m are the concentrations in the high and low plateau respectively, h determines the slope in the inflection point, and c is the position of the inflection point. The insets in Figure 2.6, top, show the distribution in pinning position for the indicated number of molecules, based upon concentration profiles from 100 simulation runs collected between 150 s and 200 s and recorded every 0.1 s (i.e. based upon $50,100$ concentration profiles). We fit this observed distribution to a normal distribution, indicated by the black line. At great molecule numbers the distribution closely follows a Gaussian distribution, with the mean corresponding to the predicted value from the PDE analysis (indicated by the dashed line). At small molecule numbers, the pinning position becomes biased to smaller pinning positions $L_0{}^*$ (i.e. to the left), due to the fact that stochasticity creates an effectively lower level of auto-activation, see discussion below. At high molecule number, fluctuations in the pinning position become small and closely spaced around the position within the deterministic model (indicated by the dashed line). At lower molecule numbers the profile broadens.

To quantify convergence of the stochastic case, $A_i(t)$, to the deterministic case, $a_i(t)$, at time t and in lattice point i, we use the normalised Euclidean distance

$$d(t) = \sqrt{\sum_{i=1}^{N} [A_i(t) - a_i(t)]^2} \Big/ \sqrt{\sum_{i=1}^{N} a_i(t)^2} \,. \tag{2.5}$$

We compute this distance for $J=100$ SSA runs, $d(t)^{(j)}$ $(j = 1, \ldots, J)$, and report mean observed distances $\overline{d(t)} = (1/J) \sum_{j=1}^{J} d(t)^{(j)}$.

The observed mean Euclidean distance in Figure 2.7 shows the expected convergence of the stochastic case to the deterministic case for increasing copy number. In Figure 2.7, disappearance of the wave profile becomes apparent by the increasing value of $\overline{d(200)}$ for fewer molecules. These results show the stochastic limit to the deterministic behaviour, confirming the qualitative intuition of a stochastic lower limit to the wave pinning mechanism. What is not clear *a priori* however, is the dynamics by which wave-pinning is lost at low molecule numbers and how this is affected by other parameter changes.
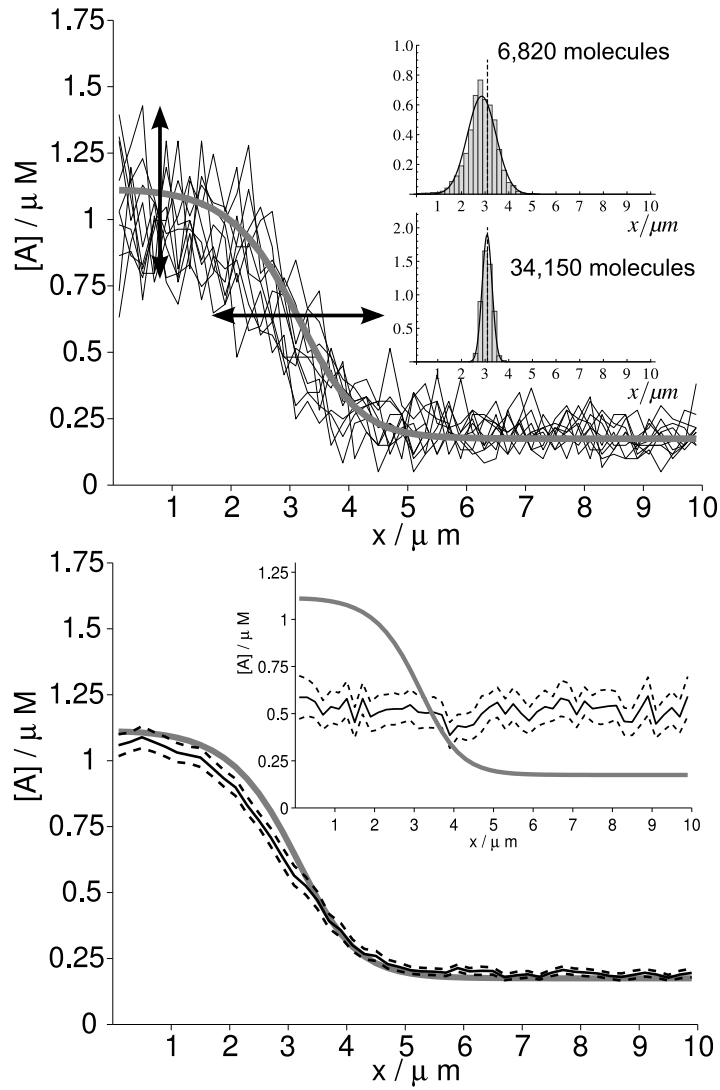
Figure 2.6: **Stochasticity in Wave-Pinning.** (top) Black solid lines: ten individual runs of the stochastic model using the SSA at $t=200\,s$. Vertical and horizontal noise of the jagged lines are indicated by the corresponding arrows. Grey line: behaviour of the deterministic system. (Parameter values $k_0=0.067\,s^{-1}$, $\gamma=1\,s^{-1}$, $K=1\,\mu M$, $\delta=1\,s^{-1}$, $D_a=0.1\,\mu m^2/s$, $D_b=10\,\mu m^2/s$; diffusion of B in deterministic model only). Total amount $T=22.68$. $W=2.5\,\mu m$, number of molecules equals 6,820. (insets) Histograms of observed pinning positions when number of molecules is $6,820$ and $34,150$, respectively. Concentration profiles of 100 simulation runs between 150 and $200\,s$ were fitted to symmetric Richards model and inflection point used as pinning position (bottom) Black solid line: observed mean of the stochastic system over 100 runs, shown together with a grey line representing the deterministic system. Black dashed lines enclose the 95 % c.i. of the sample mean. (bottom inset) Lines as in bottom image, and parameters as in top image except for width, $W=0.25\,\mu m$, resulting in a reduction of the number of molecules to 702. The top image shows that individual SSA runs closely reproduce the deterministic behaviour for appropriate parameters. The corresponding mean behaviour in the bottom image confirms this. The bottom inset depicts loss of wave-pinning from the mean behaviour when too few molecules are present in the system.
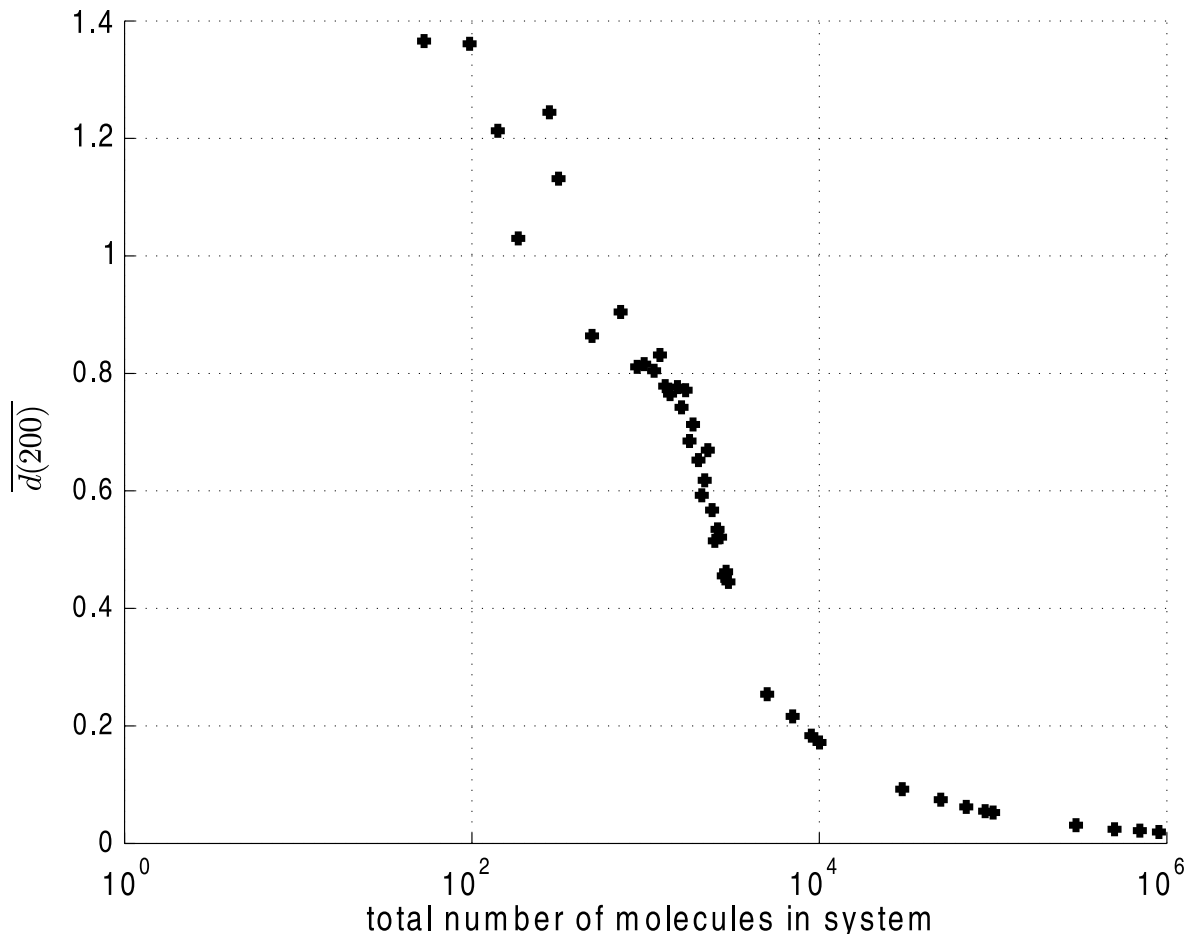
Figure 2.7: **Convergence of Stochastic to Deterministic System.** Mean Euclidean distance over $J = 100$ SSA runs, $\overline{d(200)}$ at time $t{=}200\,s$, after the stimulus. Simulation parameters as in Figure 2.6, except for W. W is varied to obtain different total copy numbers. This plot shows convergence of the stochastic system to macroscopic predictions as we increase the number of molecules.

2.7 Stochastic Simulations Are not Affected by Propagation Failure

Before we discuss in detail the loss of wave-pinning at low molecule numbers, we revisit the phenomenon of propagation failure. We repeat the procedure described above to determine potential propagation failure within the stochastic description, but now for varying molecule numbers rather than diffusion rates (Figure 2.8). As before, when the grid becomes coarser, secondary effects of propagation failure manifest themselves and affect the pinning positions through this independent mechanism. Again however, as we showed before in Figure 2.5, those effects are not noticeable in the simulation regimes we focus on, and only to a very small extent under extreme conditions of a very coarse grid.
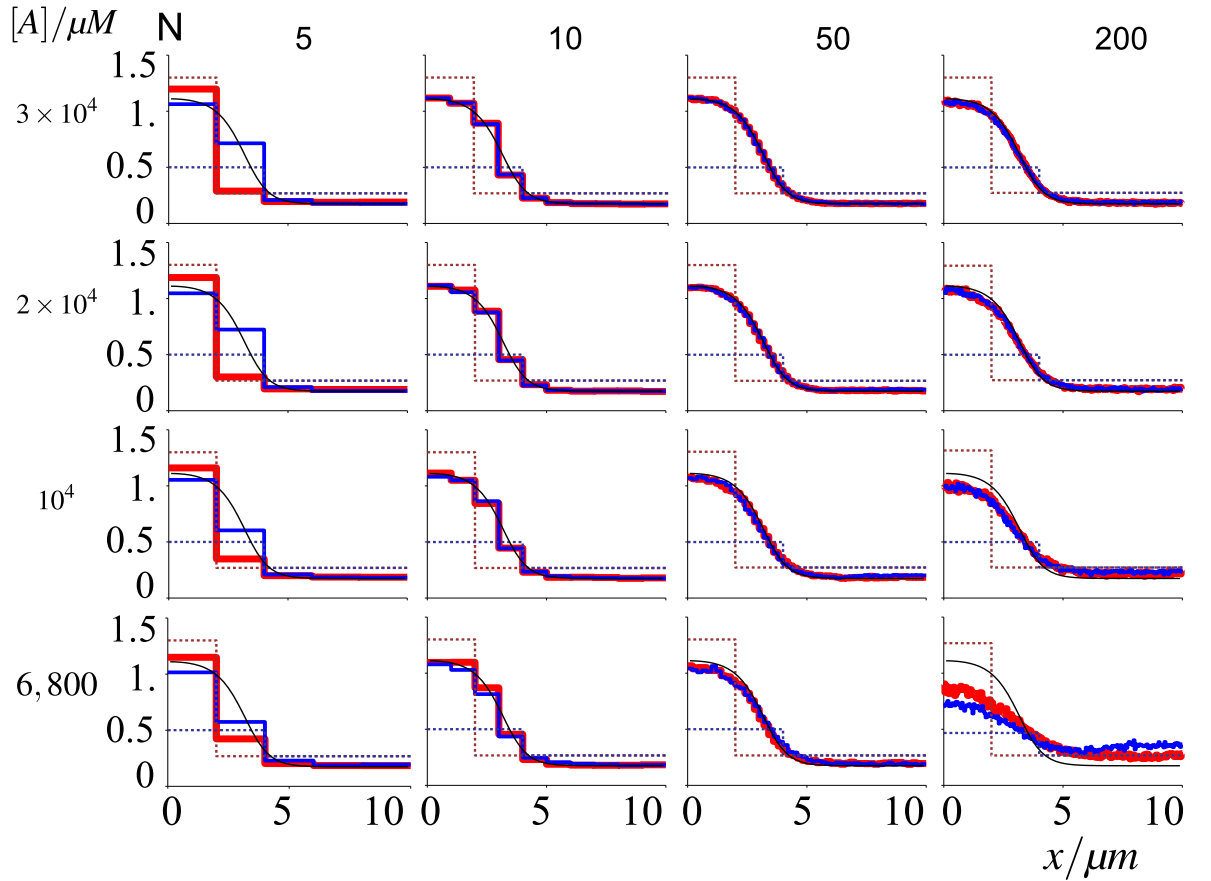
Figure 2.8: **Propagation Failure Within Stochastic Simulations as a Function of Grid Coarseness and Molecule Number.** Dashed profiles indicate initial conditions, correspondingly coloured solid lines depict mean equilibrium profiles at $t = 200\,$s, averaged over 100 runs. Confidence intervals are omitted for reasons of clarity, but standard errors of mean are similar to Figure 2.6. $D_a = 0.1\,\mu m^2/s$; the number of compartments and molecules are indicated in the figure. Propagation failure does not affect 50 lattice point simulations. At small molecule numbers, however, larger numbers of lattice points present deviations that are unlike propagation failure and we discuss this effect in the main text.

We find that stochasticity reduces rather than increases the parameter regime for which propagation failure can be observed (compare the column corresponding to $N = 5$ in Figure 2.8 with the upper left graph in Figure 2.5). This can be understood by realising that stochasticity can help overcome the threshold to propagate the wave. This means that for coarse grids and low diffusion rates the stochasticity at low molecule numbers can even increase the precision of the pinned wave. Besides deviations at low box numbers, we also observe deviations from the stationary solution of the deterministic model (as indicated in each frame with a black line) for a combination of low molecule and high box numbers (see lower right graph in Figure 2.8). This is an artifact attributed to an effective change in the auto-activation function when the number of molecules in a box becomes very small, which will be

41

discussed further below.

2.8 Comparison of Deterministic and Stochastic Predictions

We test our predictions from the LPA system, Equation (2.4), by classifying individual SSA runs (using the same parameter values as in Figure 2.6) as either homogeneous in A (i.e. uniform in A well after the stimulus, e.g. at $t{=}200\,s$) or inhomogeneous in A (i.e. where a local pulse invaded the global concentration profile of A, creating at least one high plateau or peak in A).

Comparing the predictions of the deterministic version of the model with the SSA simulations we find interesting differences. With parameter values as in Figure 2.6, specifically with $D_a{=}0.1\,\mu m^2/s$ (central inverted cup in Figure 2.9), we observe a somewhat more stringent condition in the stochastic wave-pinning regime, $22{\leq}T{\leq}27$, than predicted by Figure 2.3 (arrow indicates default value $T = 22.68$). This discrepancy may be explained by the relatively fast diffusion of A, compared with the limiting rates used in the theoretical treatment, which destabilises local perturbations in A (through high curvature in the concentration profile of A), making it harder for perturbations to stabilise and invade the remaining profile of A. Indeed, as we decrease D_a by orders of magnitude (Figure 2.9, broadening inverted cups), increasingly wider ranges of T allow for stabilisation of perturbations in A and their invasion of the global homogeneous A level. Our observations for decreasing D_a, and certainly for $D_a{=}0\,\mu m^2/s$ (Figure 2.9, dashed line), indicate increasing agreement between our stochastic simulations and the predicted wave-pinning criteria.
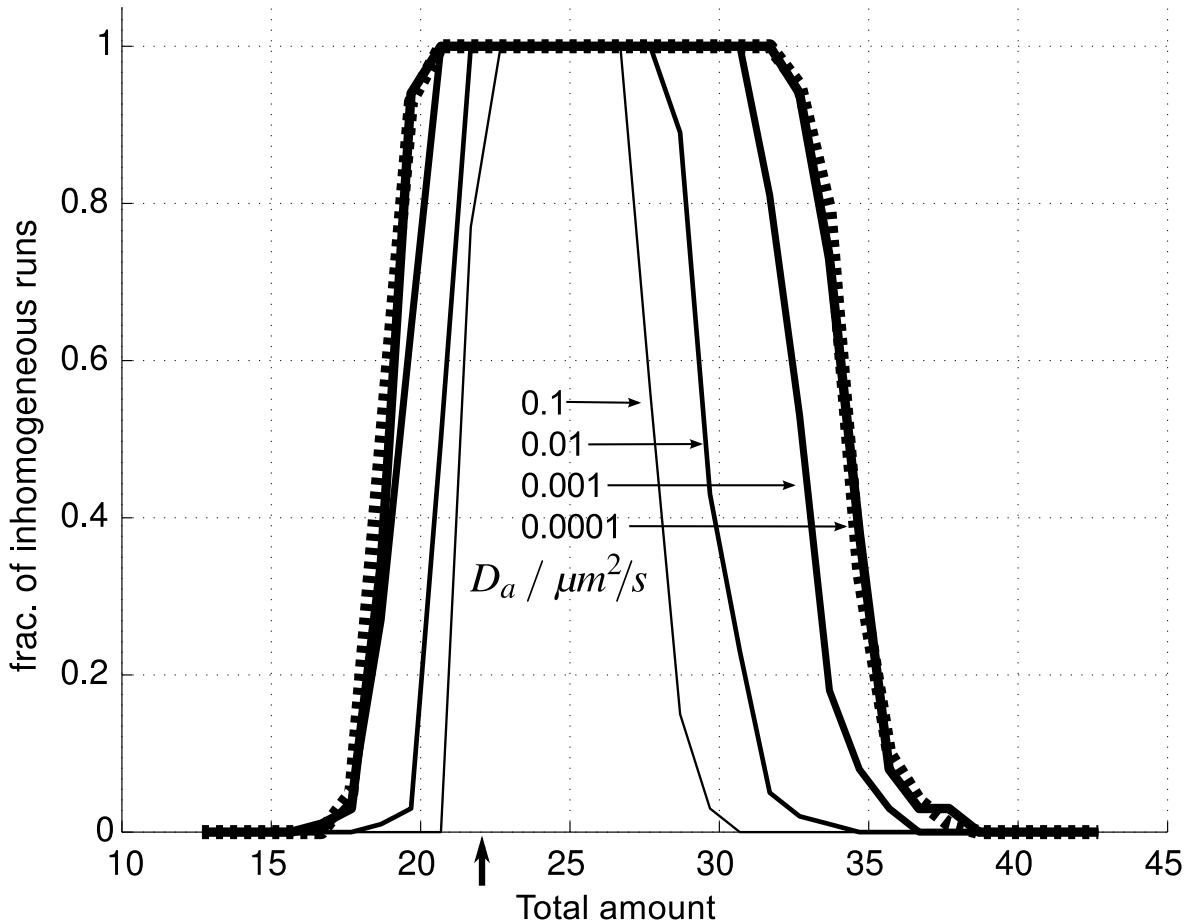
Figure 2.9: **Dependence of Wave-Pinning Stability on Diffusion and Total Amount.** Fraction of SSA runs (out of 100 runs) observed at $t=200\,s$ to be stably inhomogeneous (wave-pinned, or with multiple peaks in A). System parameters as in Figure 2.6 with varying D_a. Each inverted-cup-shaped curve corresponds to a different D_a value, decreasing from $D_a=0.1\,\mu m^2/s$ to $D_a=10^{-4}\,\mu m^2/s$ from the innermost to the outermost cup, using order of magnitude changes in the parameter value (as indicated by the arrows). Dashed line: $D_a=0\,\mu m^2/s$. This plot shows for what parameter values (changing total amounts T and diffusion constant D_a) the stochastic system shows stabilisation of perturbations in the uniform concentration profile of A (high fraction of inhomogeneous runs). Stabilisation of perturbations is difficult for high diffusivity of A ($D_a=0.1\,\mu m^2/s$ innermost, narrowest cup) since spikes in the concentration profile of A are more readily smoothed out before they grow to sufficiently high levels. For decreasing values of D_a our observations match progressively better with the predictions made using the LPA in the main text. That is because the LPA assumes zero diffusivity for A and therefore ignores any effects due to the diffusion of the active form.

2.9 Loss of Wave-Pinning in Small Number Regimes

Loss of wave-pinning for fewer molecules may either result from increased randomness along the concentration axis (vertical noise, Figure 2.6), or greater random fluctuations in the pinning position of a travelling wave (horizontal noise, Figure 2.6): With a low copy number, individual runs may show wave-pinning but the

steepest point of the concentration profile (pinning position) may fluctuate along the horizontal axis due to inherent stochasticity.

To distinguish the effects caused by these distinct types of noise, we quantify the horizontal noise in lattice point i, using a sample of $J=100$ SSA runs, with a function similar to autocorrelation, here denoted "spatial autocorrelation":

$$RS_k(i) = \frac{\mathbb{E}\left[\left(A_i(t) - \overline{\mathbb{E}[A(t)]}\right)\left(A_i(t+k) - \overline{\mathbb{E}[A(t+k)]}\right)\right]}{\sqrt{\overline{\mathrm{Var}[A_i]}}}, \tag{2.6}$$

where k is the time lag, $\overline{\mathbb{E}[A(t)]}$ is the spatial mean of A at time t, and $\overline{\mathrm{Var}[A_i]}$ is a spatial variance. (We use bars to denote statistics related to spatial autocorrelation.) Statistic $\overline{\mathbb{E}[A(t)]}$ measures the average concentration of A at time t across the domain, and the spatial variance, $\overline{\mathrm{Var}[A_i]}$, gauges the spread of the concentration of A in lattice point i, $A_i(t)$, about the spatial mean. Observing the concentration of A in M equidistant time points ($A_i(m)$, $m=1,\dots,M$), we estimate these statistics as follows:

$$\overline{\mathbb{E}[A(m)]} \approx \hat{\bar{\mu}}(m) = \frac{1}{N}\sum_{i=1}^{N} A_i(m), \tag{2.7a}$$

$$\overline{\mathrm{Var}[A_i]} \approx \hat{\bar{\sigma}}_i^2 = \frac{1}{M-1}\sum_{m=1}^{M}(A_i(m) - \hat{\bar{\mu}}(m))^2. \tag{2.7b}$$

Our estimator of spatial autocorrelation then becomes:

$$\widehat{RS}_k(i) = \frac{\frac{1}{M-k}\sum_{m=1}^{M-k}[A_i(m) - \hat{\bar{\mu}}(m)][A_i(m+k) - \hat{\bar{\mu}}(m+k)]}{\hat{\bar{\sigma}}_i}. \tag{2.8}$$

This function centres observations $A_i(m)$ about the spatial mean and gauges the correlation between observations k time steps apart. We also normalise our estimator ($\widehat{RS}_k(i) \in (-1,1)$) by dividing by the observed standard deviation about the spatial mean. A high-value $\widehat{RS}_k(i)$ denotes a lattice point i where A_i is stably far away from the spatial mean, while a low value suggests that A_i repeatedly comes near the spatial mean.

We report estimates of autocorrelation (spatial, and temporal further on) as sample means $\overline{\widehat{RS}_k}(i) = (1/J)\sum_{j=1}^{J} \widehat{RS}_k(i)^{(j)}$ of the corresponding estimates for J SSA

runs, $\widehat{RS_k}(i)^{(j)}$. We further compute these estimates in time periods when we expect our system to be stable, long after the application of a stimulus and wave-pinning (stimulus applied between $50\,s$ and $70\,s$ and observations between $500\,s$ and $1500\,s$ used for computations).

In a regime where the mean behaviour presents wave-pinning we observe a dip in spatial autocorrelation in a section of the domain that includes the pinning position (Figure 2.10, top). The width of this dip (shaded area in Figure 2.10, top) increases with decreasing numbers of molecules in the system (Figure 2.10, bottom), and spans the entire domain ($10\,\mu m$) when wave-pinning is lost.

For quantifying vertical noise, we use an autocorrelation function denoted "temporal autocorrelation" for clarity:

$$RT_k(i) = \frac{\mathbb{E}\left[(A_i(t) - \mathbb{E}[A_i])\,(A_i(t+k) - \mathbb{E}[A_i])\right]}{\sqrt{\text{Var}\,[A_i]}}, \tag{2.9}$$

where $\mathbb{E}[A_i]$ is the expected value and $\text{Var}[A_i]$ is the variance of concentration A_i. We use the common sample mean and sample variance as estimators of these statistics:

$$\mathbb{E}[A_i] \approx \hat{\mu}_i = \frac{1}{M} \sum_{m=1}^{M} A_i(m), \tag{2.10a}$$

$$\text{Var}[A_i] \approx \hat{\sigma}_i^2 = \frac{1}{M-1} \sum_{m=1}^{M} (A_i(m) - \hat{\mu}_i)^2. \tag{2.10b}$$

Using these, our estimator of temporal autocorrelation becomes:

$$\widehat{RT_k}(i) = \frac{\frac{1}{M-k} \sum_{m=1}^{M-k} [A_i(m) - \hat{\mu}_i][A_i(m+k) - \hat{\mu}_i]}{\hat{\sigma}_i}. \tag{2.11}$$

Temporal autocorrelation measures the randomness of the time evolution of A, which is governed by a continuous-time Markov process. (If we know the distribution of $A(t)$ at present, the future distribution of $A(s)$, $s > t$, only depends on $A(t)$ and is independent of observations of A before t.) Given the Markovian character of $A(t)$, we expect the temporal autocorrelation to be greatly dependent on the lag k: a small lag ($k{=}1\,s$) results in relatively high temporal autocorrelation (black area, Figure

2.11 top right and bottom right), while a bigger $k=20\,s$ yields small temporal auto-correlation (light grey area, Figure 2.11 top right and bottom right). We also observe that the Markovian character of $A(t)$ is comparable for small and large numbers of molecules since the magnitude of temporal autocorrelation does not change significantly when altering the number of molecules (Figure 2.11 right panels). This means that in both small and large copy regimes, temporal autocorrelation decreases comparably fast.

While the vertical noise does not seem to decrease when increasing the number of molecules, wave-pinning still shows up as a marked peak in temporal autocorrelation in the transition zone of the wave (Figure 2.11, bottom right). Since overall vertical noise is relatively constant, this peak seems to be caused by the wave fluctuating about its pinning position: concentration $A_i(t)$ in the transition zone fluctuates between high and low values (high and low plateau of wave) causing $A_i(t)$ to be far away from its average repeatedly (high temporal autocorrelation).

We find spatial autocorrelation to behave markedly different from temporal autocorrelation for increasing numbers of molecules (Figure 2.11, left panels). While spatial autocorrelation is comparable in magnitude to temporal autocorrelation in a small copy number regime (Figure 2.11 top panels), we observe much greater spatial autocorrelation than temporal autocorrelation for various lag k values in a large copy number regime (Figure 2.11 right panels). The stably high spatial autocorrelation, even for large lags k, far away from the pinning position suggests that the high and low plateau of the wave are persistent in wave-pinning regimes (Figure 2.11 bottom right). The random fluctuation of the pinning position is highlighted by the decreasing spatial autocorrelation for increasing lags k (dark to light grey area in Figure 2.11 bottom left) in this part of the domain.

To elucidate the source of the horizontal noise and the impact it has on the sustainability of a polarised cell, we make use of another spatially simplified representation of the deterministic system. Without loss of generality, once a wave has been formed, the concentration profile of a can be split into a plateau of length L_0 and approximate concentration a_L on the left, and a plateau of length $L - L_0$ and

approximate level a_R on the right (Figure 2.12 left) where

$$a(0 \leq x < L_0, 0) = a_L(0) = a_1, \ a(L_0 \leq x \leq L, 0) = a_R(0) = a_2 .$$

Here $a_1 > a_2$ are two roots of $f(a, b_{wp}) = 0$, where b_{wp} is the uniform concentration of b for this wave-shaped profile, while it is assumed that $D_b = \infty$ and $D_a = 0$. Then the total amount T of protein in the domain can be approximated as $T_{wp} = Lb_{wp} + L_0a_1 + (L - L_0)a_2$. In the limit $D_a \to 0$ (no direct communication between plateaus), $a_L(t)$ and $a_R(t)$ evolve independently. Eliminating $b(t)$ using mass conservation leads to:

$$
\begin{aligned}
\frac{da_L}{dt} &= f(a_L, b), \quad \frac{da_R}{dt} = f(a_R, b), \\
b(t) &= (T_{wp}/L) - a_L(t)(L_0/L) - a_R(t)(L - L_0)/L,
\end{aligned}
\tag{2.12}
$$

which also implies that the position of the wavefront is

$$L_0 = \frac{T_{wp} - L[b(t) + a_R(t)]}{a_L(t) - a_R(t)} . \tag{2.13}$$

System (2.12) is a second deterministic reduction that leads to a way of comparing the stochastic and deterministic model versions. In the deterministic case we expect that the wave-pinned configuration, Figure 2.12, left panel, is stable over time when $L_0 = L_0^*$ (pinning position). As Equation (2.13) indicates, noise in the concentration variables will propagate to the width of the high plateau of the pinned wave. We expect that this propagation of noise is qualitatively the same in the full spatial system and conjecture that noise in the pinning position (horizontal noise) is the result of noise in the concentration levels (vertical noise).

Given that the total amount is fixed at T_{wp}, a_L and a_R adjust to varying widths of L_0: increasing L_0 will typically decrease a_L and increase a_R accordingly, and vice versa. For system (2.12), we plot the steady-state concentration of a_L as a function of L_0 (Figure 2.12, right panel) and observe two critical values for L_0, $L_0^{(1)} < L_0^{(2)}$, at which saddle-node bifurcations occur. The bifurcation plot in Figure 2.12 shows that if we initialise our simplified system in a wave-pinned configuration (high plateau on left, low plateau on right, as in Figure 2.12, left panel, with $L_0 = L_0^* < L_0^{(2)}$),

sufficiently large horizontal noise may drive the effective plateau width, L_0, away from L_0^* and past the critical point $(L_0 > L_0^{(2)})$ which would then cause the wave to collapse. Due to hysteresis in the bifurcation plot, L_0 may fluctuate back to the left of $L_0^{(2)}$ after collapse of the wave without triggering restoration of the wave.

We expect that the saddle-node bifurcation at $L_0^{(2)}$ and hysteresis explain the sharp increase in dip width (Figure 2.10, bottom) observed in the full spatial system: decreasing the number of molecules in the system increases vertical noise which propagates into horizontal noise. As horizontal noise increases, the likelihood that the pinning position randomly overshoots the critical value increases and we are more likely to observe wave collapse (dip width approaching $10\,\mu m$, Figure 2.10, bottom). This prediction is further supported by our observation that under conditions equivalent to those of Figure 2.12, and with $6,820$ molecules, we do not observe any pinning positions greater than $6\,\mu m$ (Figure 2.6, top inset).

Figure 2.12 is closely linked to Figure 2.3. While the latter shows the impact of a Δ-perturbation of infinitesimally small width, corresponding to an $L_0 = 0\,\mu m$, as a function of T, the former shows the impact of perturbations of varying width, for a fixed value of T. The link between the analysis on the loss of wave-pinning with system 2.12 and the analysis on polarity initiation (LPA, system 2.4) is very direct. By allowing $L_0 \downarrow 0$, system 2.12 becomes equivalent to system 2.4: a_L in Equation 2.12 describes the active level on a vanishingly small domain and therefore ceases to affect the inactive species b (i.e. a_L – active left – becomes the local perturbation a_L – active local – in Equation 2.4), while a_R starts occupying the entire domain of length L and therefore only the presence of a_R affects b (i.e. a_R – active right – becomes the global active form a_G – active global – in Equation 2.4). Figure 2.13 brings both pieces of information together, showing a two-parameter bifurcation plot, with L_0 along the x-axis, and T along the y-axis. The right panel of Figure 2.12 corresponds to a horizontal cross-section at $T = 22.68$ (Figure 2.13, zone VI), while the bifurcation diagram of Figure 2.3 corresponds to a vertical cross-section through this bifurcation diagram at $L_0 = 0$, along which line we have indicated the two fold bifurcations (FB) and two transcritical bifurcations (TB) that can be seen in Figure 2.3, indeed reproduced at exactly the same parameter values.

Increasing L_0 away from zero, the critical total amounts of both fold bifurcations

in Figure 2.3 change while the critical values of the transcritical bifurcations are unaffected (solid grey lines in Figure 2.13). This implies that at levels of T that are less favourable for sustained polarity, narrow Δ-perturbations are still able to trigger a wave, while broad ones can not do that any more. Moreover, this illustrates that when the well-mixed equilibrium is unstable against spatially inhomogeneous bifurcations (zones III–V), the width of the perturbation becomes irrelevant. We also observe four cusp bifurcation points (CP), highlighting where two fold bifurcation lines merge. These points imply the possible coexistence of waves with different amplitudes. Finally, there are two bifurcations where a fold bifurcation and a transcritical bifurcation collide (FT). These are linked to the symmetry of the two-parameter bifurcation plot about $L_0 = 5\,\mu m$, which is due to the lack of inherent bias for either a left-oriented or a right-oriented polarisation in system 2.12: when $L_0 < 5\,\mu m$, the pinned wave has its high plateau on the left, while for $L_0 > 5\mu m$ the high plateau is on the right. Due to this symmetry, we observe a bifurcation plot equivalent to Figure 2.3 when plotting a_R and setting $L_0 = 10\,\mu m$ (data not shown), confirming that initiation of a wave from the left is equivalent to initiation from the right. Together, Figure 2.13 reveals that there are seven qualitatively different zones of levels of T, each presenting diverse requirements on wave initiation and maintenance. It allows us to predict the potential to trigger (or sustain, when $L_0 = L_0{}^*$) a wave through a perturbation of any possible width and height, as well as the expected height that such a wave will reach while it travels, stalls, or stochastically fluctuates. Specifically, it sets the boundaries for horizontal fluctuations to trigger a collapse of the polarised cell state.
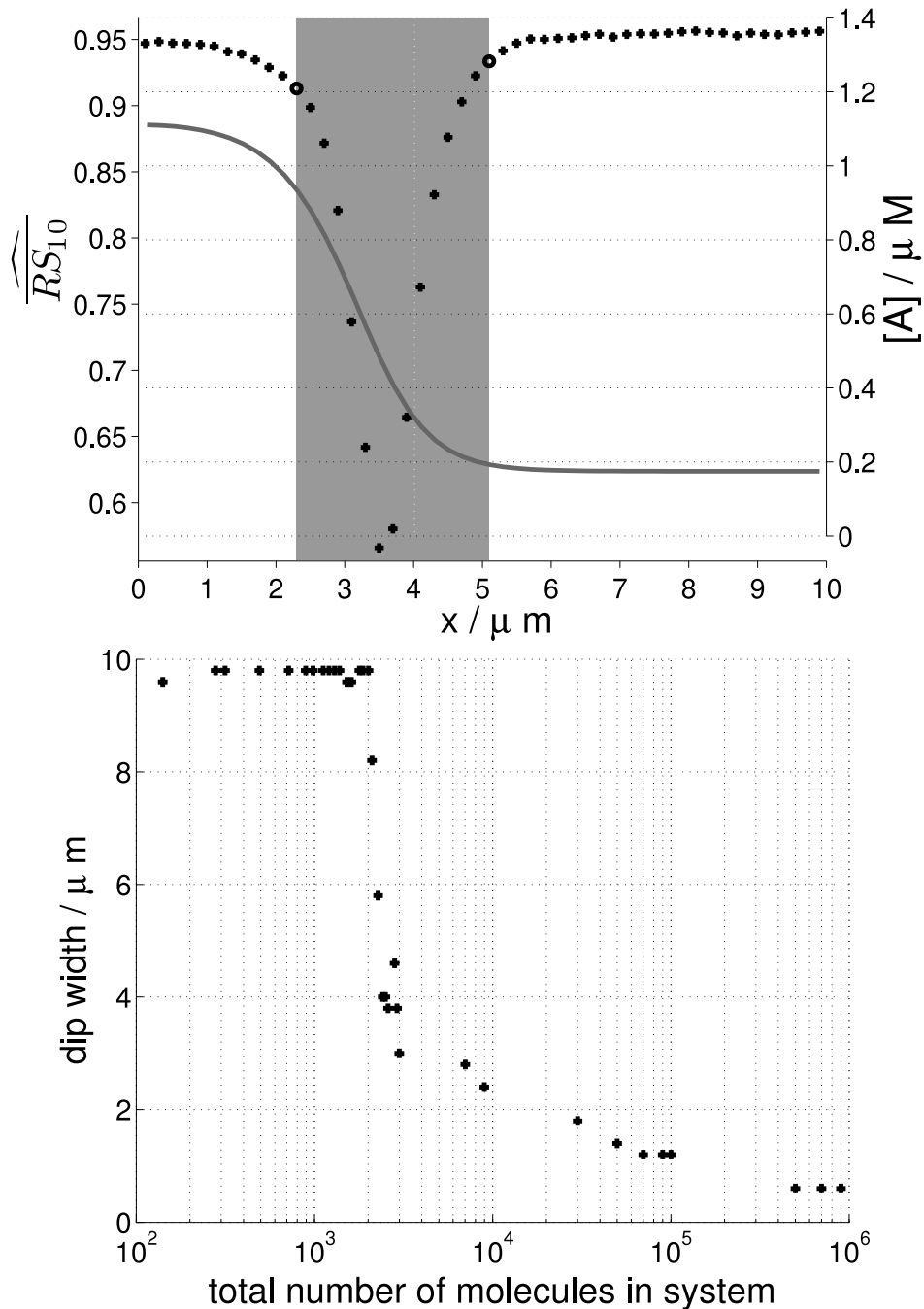
Figure 2.10: **Noisy Transition Zone.** (top) Averaged spatial autocorrelation estimate, Equation (2.8), lag $k=1\,s$, with deterministic solution as in Figure 2.6 for reference. Bold circles: delimiters of autocorrelation dip, highlighted by shaded area. This figure shows correspondence of a dip in spatial autocorrelation with the transition zone of the pinned wave. (bottom) As in top figure, mean autocorrelation dip width (width of grey rectangle in top figure) over 100 SSA runs, as a function of total number of molecules. When the dip width approaches the length of the domain of the cell ($10\,\mu m$), the pinning position stops conferring information that can be used by the cell. We note a sudden and drastic change in width at approximately $2,000$ molecules. Simulation parameters as in Figure 2.6, except for W which is varied.
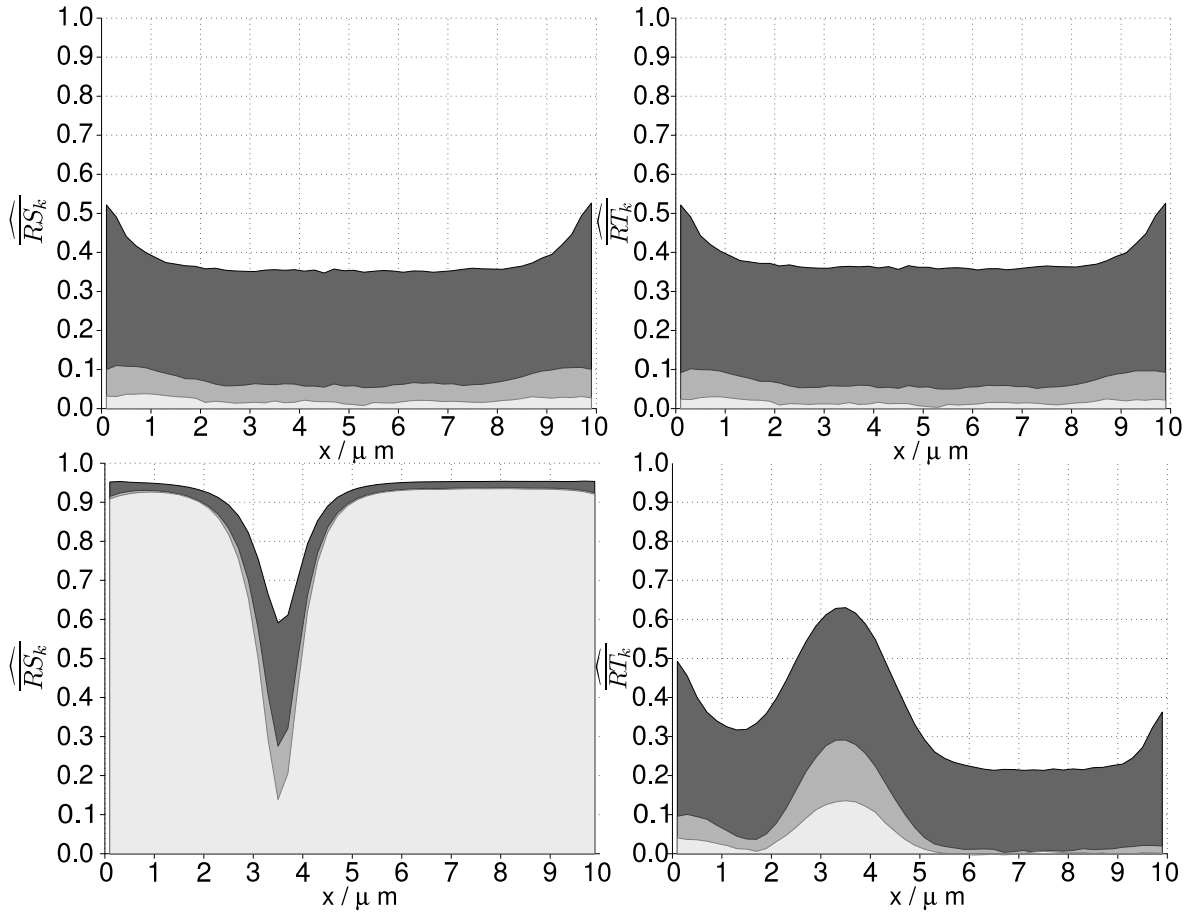
Figure 2.11: **Fluctuations in Space and Time.** Spatial / temporal autocorrelation sample mean (over $J{=}100$ SSA runs) with lag $k{=}1\,s$ (dark grey), $k{=}10\,s$ (grey), and $k{=}20\,s$ (light grey). (top row) 702 molecules ($W{=}0.25\,\mu m$) (top left) Spatial auto-correlation. (top right) Temporal autocorrelation. (bottom row) $6,820$ molecules ($W{=}2.5\,\mu m$) (bottom left) Spatial autocorrelation. (bottom right) Temporal autocorrelation. The plots in the top row show that for sufficiently few molecules both spatial and temporal autocorrelation are almost uniform and show no pattern for the equilibrium state (pinning position). In the bottom row and for a sufficiently high number of molecules, spatial autocorrelation reveals a clear pattern in the equilibrium state with the dip in autocorrelation robust across different lags k. In the same copy-number regime, temporal autocorrelation shows a pattern for small lags k which, however, vanishes for increasing values of k. Simulation parameters as in Figure 2.6, except for W.
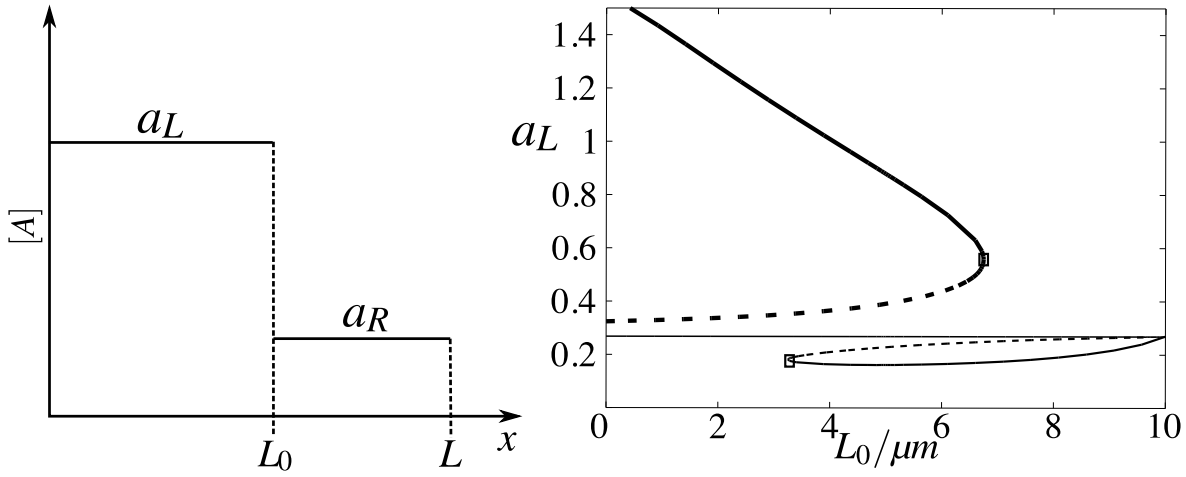
Figure 2.12: **Polarised Cell State Analysis.** (left) Schematic of the simplified deterministic wave-shaped state, consisting of a high plateau a_L of length L_0 on the left, and a low plateau a_R of length $L - L_0$ on the right. (right) Bifurcation plot of steady-state concentrations of a_L as a function of L_0, with parameters as in Figure 2.6 and with total amount $T{=}22.68$. Two saddle-node bifurcations occur at critical values $L_0^{(1)}{=}3.25\,\mu m$ and $L_0^{(2)}{=}6.75\,\mu m$. The corresponding bifurcation plot of a_R is symmetric to this one of a_L. In our stochastic simulations, for equivalent parameter values and number of molecules = 6,820, we did not observe pinning positions greater than $6\,\mu m$ (top inset, Figure 2.6), as is predicted by this analysis.
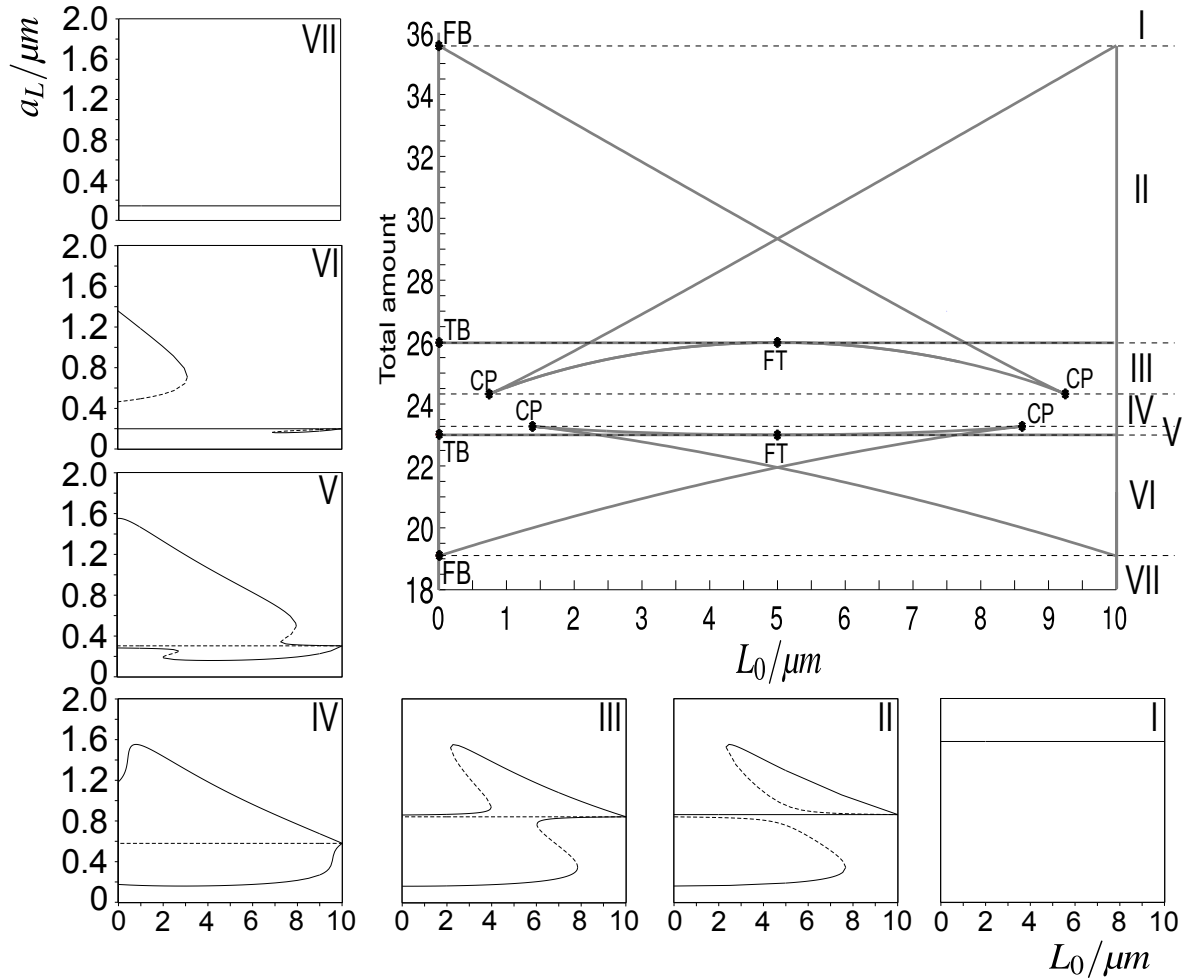
Figure 2.13: **Two-Parameter Bifurcation Plot of System 2.12.** This figure shows the connection between the LPA (Figure 2.3) and the loss-of-wave-pinning analysis (Figure 2.12). It also presents the requirements on both height and width of an initial Δ-perturbation to trigger cell polarity. (Top right) Two-parameter bifurcation plot, with total amount T along y-axis and wave position L_0 along x-axis. Grey lines starting at FB: fold bifurcation lines; grey lines starting at TB: transcritical bifurcation lines; CP: cusp bifurcation; FT: Collision of fold and transcritical bifurcation. (Surrounding subfigures) When varying T, seven distinct zones of qualitatively different behaviour are found, labelled I–VII. To understand the behaviour in each zone, bifurcation diagrams are plotted of the equilibrium value a_L (Equation 2.12) as a function of L_0 for a fixed value of T within that zone.

2.9.1 Validity of Stochastic Model at Large Compartment Numbers

Subdividing the domain into N compartments is a computational method that should not influence the biological insights derived here. Above we discussed that the coarseness of the lattice could introduce artifacts due to propagation failure, but that the simulations in this study are sufficiently fine-grained. Figure 2.8, however, also presented deviations from the expected wave profile when N was large. When testing the behaviour of the stochastic model as $N \to \infty$, we realised that this is due

to the fact that the effective auto-activation function starts to change as the number of molecules in a box becomes very small. Given that at high N only small variations between the mean values of neighbouring boxes are expected (as the flux between boxes goes to infinity), we tested if the change in auto-activation could be due to the expected stochastic variations within each box, assuming a Poisson distribution for the number of active molecules within each box. We therefore calculated the predicted mean activation rate as a function of the concentration a for different box sizes, by taking the kernel of the Poisson distribution with the auto-activation function itself. The insets of Figure 2.14 show the resulting auto-activation rates for different box sizes, to which subsequently the auto-activation function $\frac{\gamma a^2}{K^2 + a^2}$ was fitted, with both K and γ as the fitting parameters. The figure shows that when the number of active molecules becomes small the functional response effectively shifts to the right (i.e. the fitted value of K becomes larger), while the plateau hardly changes (i.e. γ remains more or less 1.0). This phenomenon is due to the plateau in the sigmoidal activation function: while the high spectrum of the Poisson distribution cannot further increase the activation rate, as it is capped, the low spectrum decreases it. Figure 2.14 repeats the analysis of Figure 2.12 for the estimated effective values of K. It shows that consequently at lower molecule numbers the wave becomes lower and is lost more easily due to horizontal fluctuations until, at very low molecule numbers per box, the wave cannot be sustained any longer. Also, the lower critical value of L_0, $L_0^{(2)}$, shifts to lower values (i.e. to the left). This shift to lower L_0 and drop in height of the wave closely corresponds to the observed changes in the stochastic simulations as can be seen in Figure 2.8, suggesting that the side-effects observed when increasing N are due to this modification of the auto-activation function only. This reduction in the number of molecules per box when increasing N can be overcome in our modelling set-up by increasing W. Thus, N can be made arbitrarily large, as long as, through modifications of W, the number of molecules per box is maintained above around 14 (below which value the effective K becomes too large to warrant sustainable wave-pinning).
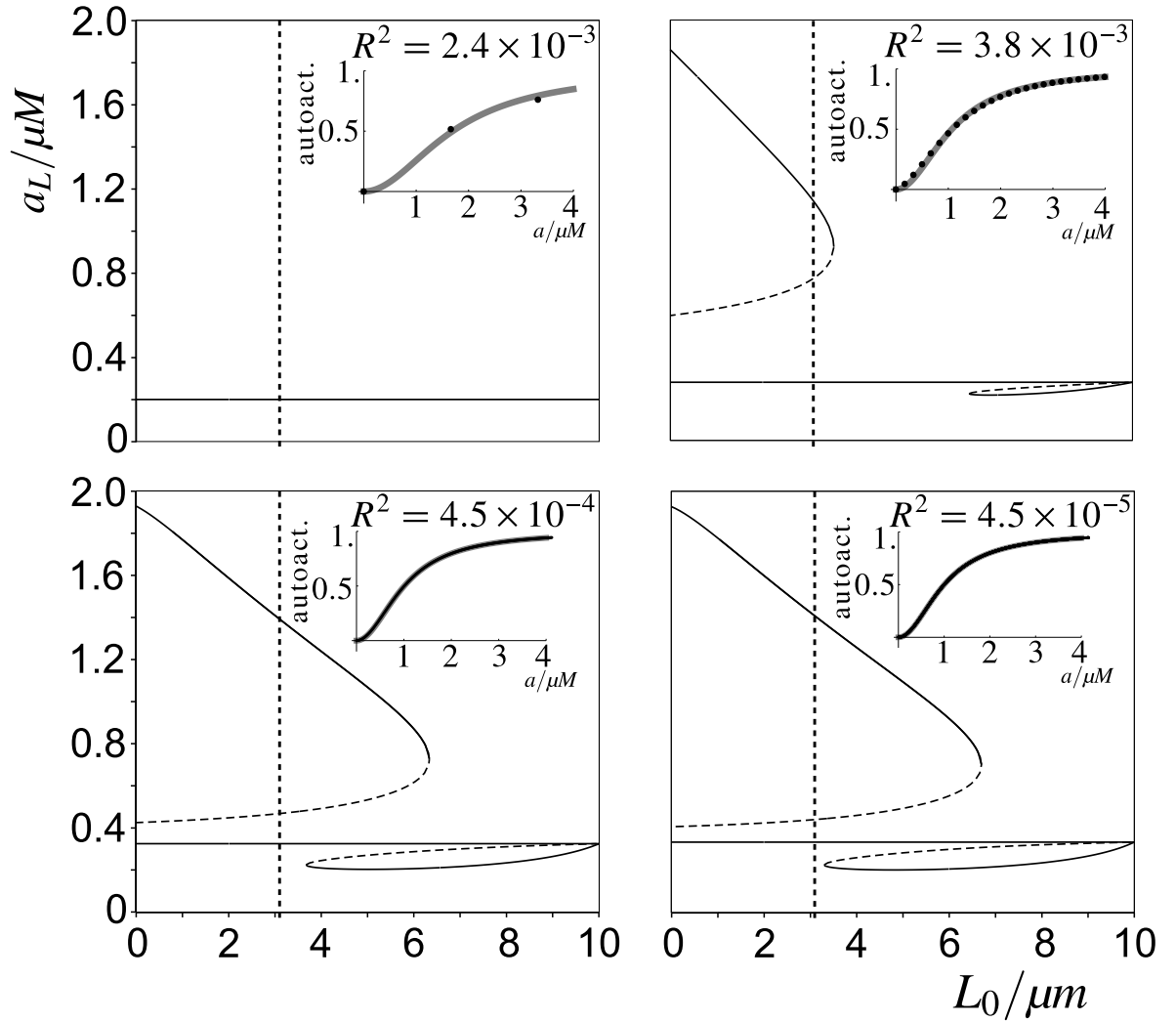
Figure 2.14: **Predicted Effect of Stochasticity on the Auto-Activation Term and its Consequences for Wave-Pinning.** Bifurcation diagrams to analyse the polarised cell state, as in Figure 2.12, for four different box volumes V_a (top left: $0.001\,\mu m^3$; top right: $0.01\,\mu m^3$; bottom left: $0.1\,\mu m^3$; bottom right: $1.0\,\mu m^3$). The different bifurcation diagrams are made by varying the parameter values γ and K of the Hill-type auto-activation $\gamma a^2 / \left(K^2 + a^2\right)$. The effective parameter values γ and K for small box sizes are determined as follows: For the mean number of molecules of active A being equal to $0, 1, 2, \ldots$ molecules per box (indicated by the black dots in the insets), a Poisson distribution for that number of molecules is assumed in order to calculate an expected rate of auto-activation. These observed values are then fitted to the auto-activation term itself using a least-squares fit. The fitted auto-activation functions are shown as grey lines in the insets, together with the residual sums of squares, R^2. We observe that while the fitted γ hardly differs from the originally used value ($\gamma \approx 1$), the fitted K increases as box volume decreases. The bifurcation diagrams show that this change in effective auto-activation lowers the higher plateau and narrows the range of permissible L_0 values. Note that the difference between the lower panels is virtually undetectable. Indeed, both are almost equivalent to the bifurcation diagram for the deterministic system, which was shown in Figure 2.12. It illustrates that a volume of $V_a = 0.1\,\mu m^3$ (bottom left) forms the upper bound for observing this stochastic effect. This corresponds to 137 molecules per box, given that the total concentration is $2.268\,\mu M$. The lower bound for wave-pinning to occur is reached when the effective K becomes 1.07, at a box size of $V_a = 0.009\,\mu m^3$, corresponding to 14 molecules per box.

2.10 Discussion

In this manuscript we compared deterministic and stochastic aspects of a model for cell polarisation of the wave-pinning class (Mori et al., 2008). This work was motivated by recent interest in the influence of stochastic noise in biological systems. In considering stochastic noise we account for possible effects due to low-copy numbers of signalling proteins, as has been done for instance by Isaacson et al. (2011).

Recent studies indicate that noise can have either constructive or detrimental effects in biological systems. For example, noting beneficial effects, Paulsson et al. (2000) observed that stochastic focusing increased sensitivity of cascades, Rao et al. (2002) found that noise-induced population heterogeneity improves fitness, Howard & Rutenberg (2003) argued that biologically relevant oscillations in a two-component dynamical system are more robust in the stochastic case than the deterministic one, and Gamba et al. (2005) showed that stochasticity could play a role in chemotactic responses to shallow gradients. On the other hand, detrimental effects were noted by, for example, McAdams & Arkin (1997) who showed that gene expression in a noisy regime resulted in bursts, rather than constant levels of gene expression. For our stochastic model of cell polarisation, we observe that at critically low molecule numbers stochastic noise has an impact on the behaviour of the system in a detrimental manner, eventually destroying polarisation.

As shown in Figure 2.6, we verify correctness of our stochastic implementation by comparing our stochastic simulation results in large copy number regimes with the deterministic system (approaching the macroscopic limit). The compartmentalised spatial Gillespie model used here, however, does present deviations when the lattice number N becomes very small or very large. Alternatively, an off-lattice Brownian dynamics model could have been developed (Andrews & Bray, 2004; van Zon & ten Wolde, 2005), to independently confirm the results based upon space discretisation presented here, given that both approaches present their own limitations (Erban & Chapman, 2007, 2009). However, to re-formulate the auto-activation in terms of mass-action kinetics only, which is required to use off-lattice alternatives, would require a replacement of the current parsimonious non-linear term by, for example, specific enzyme kinetics. Unfurling such a mechanism to specific mass-

action enzyme kinetics requires an iterative process where we propose a number of molecular species and reactions, test the resulting mechanism for cell polarity properties, and amend the mechanism if needed. Classically this process is not trivial, however in Chapter 4 of this thesis we develop a computational tool that alleviates some of the encountered complexity and we offer an outlook that promises to automate much of this iterative process.

Here we studied wave-pinning in a one-dimensional slab of cell material representing radial positional information available to a single cell. Our results suggest that for small copy numbers, radial information within a cell regarding its front and back that is available through the wave-pinning process decreases in quality because of fluctuations of the wave around its pinning position. If we extrapolate our results to a spherical $15\,\mu m$-diameter cell and assuming micromolar small G-protein concentration (corresponding to 10^6 molecules), stable polarisation should typically be observed. However, *in vivo* effective reaction compartment size is often restricted due to macromolecular crowding and resulting volume exclusion (Schnell & Turner, 2004; Grima, 2010): the volume encompassed by a cell is generally occupied by a range of macromolecules which do not participate in any of the relevant chemical reactions. As these macromolecules span the cell in a mesh-like fashion, individual effective reaction compartments may emerge which have a potentially small volume. Our results indicate that sufficiently small effective reaction compartments (those that hold 10^3 molecules, bottom Figure 2.10), will produce inaccurate positional information (fluctuation of the pinning position). Hence, a cell subject to great amounts of macromolecular crowding may integrate inaccurate positional information from its individual effective reaction compartments and therefore lose a global sense of directionality.

Instead of a gradual loss of wave-pinning we observe a threshold number of molecules (between 2,000 and 3,000 molecules) below which the wave is suddenly lost. Analysis of a limiting deterministic case shows that sudden loss of wave-pinning is due to saddle-node bifurcations with hysteresis. We conjecture that this also explains sudden loss of polarisation in our spatial stochastic system.

Altschuler et al. (2008) studied a related positive feedback model of Cdc42 with a homogeneous cytoplasmic pool of the inactive form. Their model includes self-

recruitment of the active form, but the conditions for polarisation and wave-pinning as defined through the LPA method described in this manuscript are not fulfilled, hence no stable polarisation can be observed. Nevertheless, by defining polarisation as a transient situation where 10% of the domain holds more than 50% of the molecules, they were able to show that within their model decreasing the numbers of molecules, which increases the random fluctuations, could trigger polarisation. This is opposite to what has been shown in our study, in which increasing fluctuations shut off polarity. For a fixed positive feedback strength, they found that 1,000 molecules yields the maximum probability (fraction of simulation runs) for polarisation. For 3,000 molecules or more, they observed that fewer than 50% of the runs would polarise. In contrast, our model predicts that polarisation fails below 2,000 molecules (Figure 2.10, bottom panel). Moreover, in our model, the behaviour of the stochastic system more closely resembles that of the deterministic system when the molecule number increases (bottom panel of Figure 2.6), i.e. unlike Altschuler et al. (2008) with more molecules the polarity becomes increasingly more robust.

Khain et al. (2011) recently discussed stochastic travelling waves in a spatial one-dimensional model of spruce budworm populations, in which the high plateau of the wave corresponded to parts of the environment with a great number of budworms (outbreak state) and the low plateau denoted few budworms (refuge state). Note that in their model wave-pinning does not occur. Nevertheless, they compared a deterministic version (thermodynamic limit) of their model with a stochastic version of it and observed differences in wave propagation velocity. They explained that these differences are caused by random jumps, possible within the stochastic system, from the high plateau to the low plateau and vice versa, similar to our explanation for the fluctuations observed in the pinning position. In their case, however, the stochasticity affects the velocity of the travelling wave, while in our study it affects the pinning position.

In future efforts it would be interesting to study the stochastic model for cell polarisation in higher dimensions, where effects of geometry are non-trivial (e.g. see Strychalski et al. (2010)), as well as in off-lattice Brownian dynamics models.

Chapter 3

Protein Localisation in a Growing System: A Simple Model of Pollen Tube Growth

Abstract

Pollen tubes are projections of pollen grains that carry sperm cells from the grain to the ovules in a growth process termed tip growth. Pollen tube growth is essential for plant reproduction and tip growth has been the focus of a number of theoretical studies. These existing studies either investigate the biophysical properties of tip growth and presume molecular detail, or they focus solely on molecular interactions with macromolecular growth as an implied outcome.

The model we present bridges this gap: We model the dynamics of a key protein, ROP1, that is known to localise to the apex of growing pollen tubes where it guides downstream processes that ultimately induce tip growth. The mathematical framework we use allows us to bridge the gap between a molecular description of ROP1 and macromolecular growth by introducing a biologically relevant phenomenological link.

Our model reproduces both oscillatory and constant growth modes that are observed commonly in growing pollen tubes. We further discuss an explanation of the existence of these growth modes that our model suggests.

3.1 Introduction

Fertilisation in flowering plants is effected by a specialised cell type called pollen tube. The pollen tube is an elongated cell that grows out of the pollen, carrying the two sperm cells that are intended to fertilise the ovules of the flower that the pollen lands on (Franklin-Tong, 1999).

Pollen tubes possess an interesting mode of growth referred to as tip growth: the cell wall near the apex is extended while the remainder of the tube, towards the pollen grain, remains unchanged (Franklin-Tong, 1999).

3.1.1 Distinct Biophysical and Ionic Properties Form the Basis of Tip Growth

Turgor pressure is the ultimate force that drives pollen tube tip growth (Palin & Geitmann, 2012). Since turgor pressure acts isometrically, tip growth requires tight regulation of the mechanical properties of the primary cell wall encasing the pollen tube (Geitmann, 2010): The cell wall needs to be extensible in the apex where the pollen tube expands and rigid away from the apex.

The mechanical properties, such as extensibility, of the cell wall are determined by its biochemical composition. An important part of the cell wall is a matrix composed primarily of pectin molecules that have the ability to interlink tightly under certain chemical and ionic conditions (Palin & Geitmann, 2012).

Pectin molecules are produced in the Golgi apparatus in a predominantly methyl-esterified form and inserted into the cell wall through exocytosis (Geitmann, 2010). In their methyl-esterified form, pectin molecules show low propensity to interlink thus resulting in a softer, more extensible cell wall at their site of exocytosis (Geitmann, 2010).

The enzyme pectin methyl esterase (PME) de-esterifies pectin molecules thus allowing them to interlink at their carboxyl groups in a calcium-dependent process (Palin & Geitmann, 2012). PME is highly active in a narrow band in the shoulder of the apical dome where its inhibitor is endocytosed locally thus permitting local rigidification of the cell wall in the shank of the tube (Palin & Geitmann, 2012).

As we already touched upon, calcium is known to play at least two important

roles in pollen tube tip growth. Firstly, localised exocytosis of pectin is thought to rely on a calcium gradient that peaks at the leading edge of the apex. This finding is analogous to other systems where calcium is an important regulator of secretion, such as neurotransmitter release and histamine secretion in mast cells (Messerli et al., 2000). Secondly, calcium is needed for the establishment of a rigid pectin-based matrix that prevents the tube from expanding at its shanks while permitting turgor-dependent expansion in the apex (Palin & Geitmann, 2012). Agents that reduce calcium uptake and treatments that disrupt the calcium gradient in the tip prevent pollen tube growth (Messerli et al., 2000; Holdaway-Clarke et al., 1997; Miller et al., 1992).

Tip-localised exocytosis and shank-localised interlinking of pectin provide a physical model of how isotropic turgor pressure may effect tip growth. However both of these localised event types are dependent on another system that localises them in the first place.

In the present study we are interested in a dynamic protein system that has the potential to self-localise and that may therefore act as the localisation agent of the pectin-dependent system required for directing the growth force of the turgor pressure to the tip. We therefore introduce the components of this protein system and their link to pollen tube growth in the following subsections.

3.1.2   Rho GTPases are Key Molecular Switches in Pollen Tube Growth

Rho GTPases are small G proteins that bind GTP and have weak hydrolase activity (Yalovsky et al., 2008). A key molecular player in pollen tube growth is a plant homologue of Rho called Rho-related GTPase from plants 1 (ROP1).

In their GTP-bound state, Rho GTPases are able to bind downstream effectors whereas in their GDP-bound state they do not communicate with their effectors. Owing to this discrepancy in downstream signalling, we will refer to the GTP-bound form of Rho GTPases as the active form and the GDP-bound form as the inactive form.

Inactivation (hydrolysis of bound GTP) and activation (exchange of a bound

GDP for a GTP) are regulated by two families of enzymes: GTPase-activating proteins (GAPs) bind the active form and catalyse hydrolysis of the bound GTP molecule while guanyl nucleotide exchange factors (GEFs) catalyse exchange of a GDP molecule for a GTP molecule (Yalovsky et al., 2008), Figure 3.1.

Rho GTPases are modified posttranslationally to include a hydrophobic lipid residue that allows them to attach to the plasma membrane. A class of proteins, guanyl nucleotide dissociation inhibitors (GDIs), inhibit GEF-catalysed GDP-GTP exchange and shield the lipid residue of Rho GTPases thus preventing rebinding to the membrane (Yalovsky et al., 2008). GDIs are observed commonly to have far greater affinity for the inactive than for the active form of the corresponding Rho GTPase, Figure 3.1.
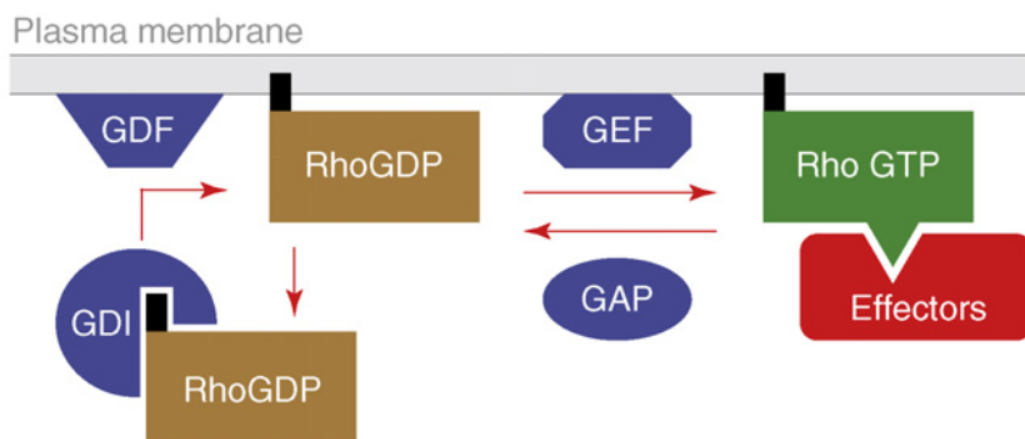
Figure 3.1: **Interchange Between the Active and Inactive Form of Rho GTPases.** The active, GTP-bound, form of Rho possesses high binding affinity for its effectors and low binding affinity for GDI. Exchange between the active and inactive form is facilitated by GEFs and GAPs. The inactive, GDP-bound, form of Rho has low binding affinity for its effectors but high binding affinity for the corresponding GDI. Thus, GDI sequesters the inactive form of Rho into the cytoplasm by shielding its lipid residue. Reproduced from Kost (2008).

3.1.3 Apical Polarity Directs Tip Growth

Pollen tube growth is facilitated by the alignment of cytoskeletal components and directed exocytosis of cell-wall material at the tip (Rounds & Bezanilla, 2013), Figure 3.2.

Active ROP1, localised at the apex of tip-growing pollen tubes, is necessary for the alignment of the cytoskeleton and, therefore, directed exocytosis (Lin & Yang, 1997; Gu et al., 2005; Hwang et al., 2005).

This model of how exocytosis is directed towards the tip is based on two observations. Firstly, experiments in which ROP1 was overexpressed indicated that ROP1 induces growth locally (Hwang et al., 2010). The key observation of these experiments was that a wider ROP1 peak in the apex produces a wider, ballooning pollen tube, Figure 3.3. Secondly, ROP1 has been observed to induce F-actin assembly locally via interaction with an effector called RIC4 (Gu et al., 2005). Tip-localised assembly and elongation of F-actin is important for tip growth (Rounds & Bezanilla, 2013).

Deposition of cell wall material, through directed exocytosis, has been shown to soften the cell wall (Rojas et al., 2011) thus hinting at a biophysical mechanism of pollen tube growth driven by an interplay between cell wall properties and turgor pressure. Despite these observations, the molecular and biophysical link between ROP1 and growth is not yet understood entirely. We therefore focus on the established functional link between ROP1 localisation and local growth (Hwang et al., 2010), Figure 3.3.
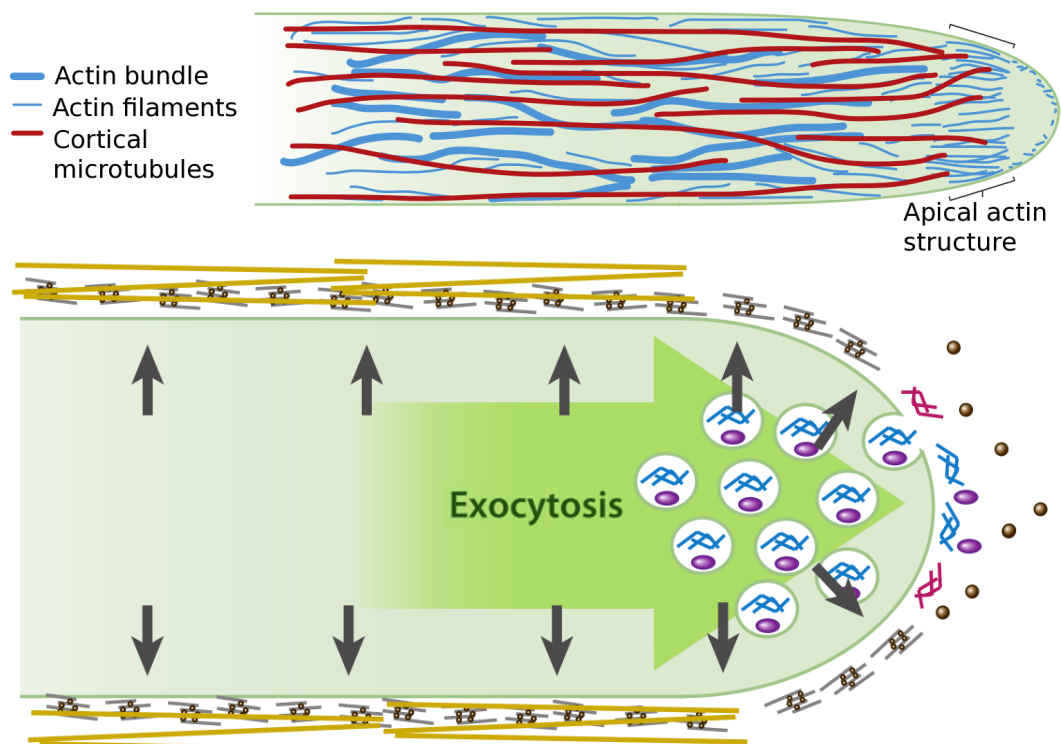
Figure 3.2: **Aligned Cytoskeleton and Resultant Directed Exocytosis.** The cytoskeleton of tip-growing pollen tubes is aligned longitudinally thus facilitating directed transport of vesicles towards the tip. At the tip, vesicles that carry cell-wall material and membrane components fuse with the plasma membrane in a process known as directed exocytosis. Adapted from Rounds & Bezanilla (2013).
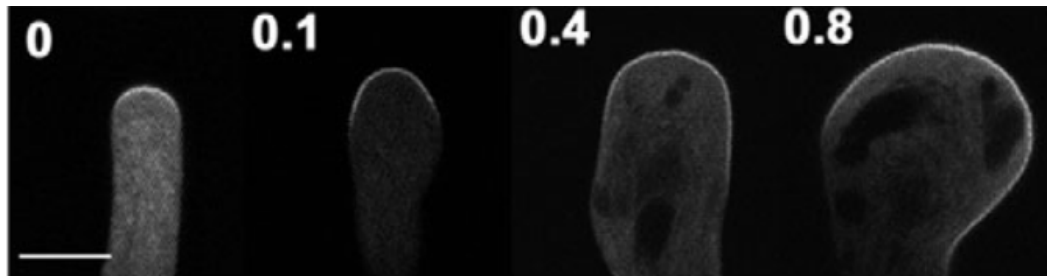
Figure 3.3: **ROP1 Induces Local Growth.** Bright lining of the apex is fluorescence from a GFP-RIC4 construct. RIC4 is an effector and direct binding partner of ROP1 and this construct has been shown to be a sensible probe for ROP1 localisation (Hwang et al., 2010). Different amounts (0. − 0.8 μg) of a ROP1 construct were expressed transiently and resulting pollen tube apex morphology is pictured. A wider spread of ROP1 causes the pollen tube to balloon at the apex, thus indicating that ROP1 induces local growth. The scale bar represents 10 μm. Reproduced from Hwang et al. (2010).

3.1.4 Pollen Tube Morphology Plays an Ambiguous Role in Apical Polarity

The functional role of ROP1 is dependent on its molecular abundance, Figure 3.3, and it is sensible to assume that the magnitude of this functional role is dependent on the concentration of ROP1. Since pollen tubes can grow up to 1 mm/h and increase their volume by a factor of twenty, the size of the cytoplasm and plasma membrane where ROP1 is located are likely important in determining the efficacy of ROP1 signalling (Nishikawa et al., 2005; Rounds & Bezanilla, 2013). Morphologically, the cytoplasm of the pollen tube is located near its apex and, in some cases, callose plugs separate the sperm cell-carrying cytoplasm from the remainder of the tube, Figure 3.4. Isolation of the cytoplasm by callose plugs is often interpreted as a mechanism employed to retain a manageable cytosolic volume (Nishikawa et al., 2005).

However, formation of callose plugs is observed to vary between pollen tubes of the same species: Nishikawa et al. (2005) reported that only about 81.5 % of *Arabidopsis* pollen tubes form callose plugs and that callose plugs are generally not observed to be necessary for pollen tube elongation. Further, Qin et al. (2012) reported the distance, measured from the grain, at which pollen tubes form their first callose plug varies significantly between different *Arabidopsis* ecotypes.

Nishikawa et al. (2005) have also demonstrated that mutant *Arabidopsis* pollen

tubes that do not form callose plugs still grow normally: These pollen tubes still elongated and grew to reach the flower while sustaining a twenty-fold increase in their volume (Nishikawa et al., 2005).

At another extreme, Zhang et al. (2008) described tomato pollen tubes that express antisense constructs to reduce the levels of LePRK2, a kinase that is important in pollen tube signalling. While wild-type tomato pollen tubes form callose plugs every 350 μm, antisense LePRK2 pollen tubes did so every 220 to 250 μm (Zhang et al., 2008). Despite a reduced size of the effective cytoplasm, antisense LePRK2 pollen tubes were still observed to grow albeit at a slower rate than wild type tubes (Zhang et al., 2008).

Since callose plugs have not been shown to be essential for pollen tube growth we do not consider them further as regulators of cytosolic volume.
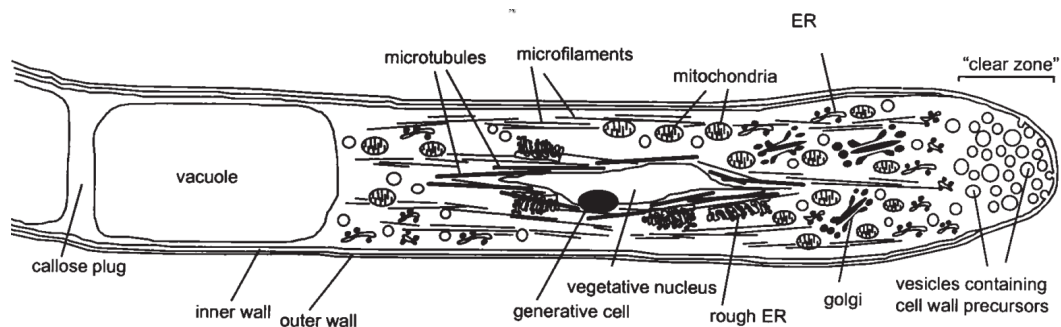
Figure 3.4: **Pollen Tube Morphology.** Notice the elongated shape of the pollen tube due to its characteristic tip growth. This sketch shows the cytoplasm with organelles and the sperm cell near the apex. The rear of the cytoplasm is closed off by callose plugs and resulting compartments become filled with vacuoles. The cytoplasm is often occupied by a central large vacuole (Li, 1999). Reproduced from Franklin-Tong (1999).

3.1.5   Pollen Tube Tip Growth is Both Steady and Oscillatory

Apical pollen tube growth is observed as both a steady and an oscillatory phenomenon where both the amplitude and period of oscillation are species-specific (Rounds & Bezanilla, 2013), Figure 3.5. In oscillatory growth regimes, growth velocity and apical ROP1 activity are generally observed to oscillate with similar periods but phase-shifted (Gu et al., 2005; Hwang et al., 2010).
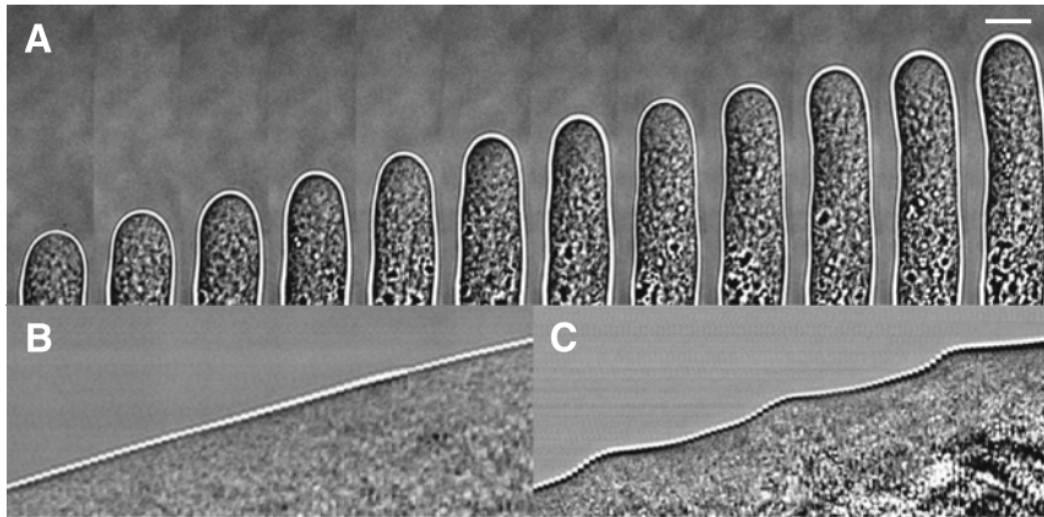
Figure 3.5: **Kymographs of Tip-Growing Pollen Tubes.** (Top) A pollen tube that displays tip growth. The scale bar corresponds to 10 μm and 6 s elapse between successive pictures. (Bottom) Kymographs of growing pollen tube tips indicated as white strips. Time runs horizontally, space vertically. Depending on growth conditions, pollen tubes can show steady growth (left) and oscillatory growth (right). The scale bar represents 10 μm. Reproduced from Rojas et al. (2011).

By studying the correlation between different oscillators in the pollen tube it has been hoped that a master timekeeper of oscillation would be discovered. However none has been described so far and some authors argue that oscillations are an emergent property of the system as a whole rather than being driven by one timekeeper (Holdaway-Clarke & Hepler, 2003; Rounds & Bezanilla, 2013).

3.1.6 Existing Models

Recently, Rojas et al. (2011) described a rheological model of pollen tube growth that captures both steady and oscillatory growth, Figure 3.5. The model of Rojas et al. (2011) is concerned with a balance between cell wall resilience and turgor pressure-induced growth where the properties of the cell wall are governed by deposition of new cell wall material. Deposition of cell wall material is assumed to occur in the tip region, and together with a negative feedback loop between pollen tube growth and deposition Rojas et al. (2011) were able to reproduce both oscillatory and constant growth.

While the model of Rojas et al. (2011) reproduces different modes and important geometric aspects of tip growth, it does not explain tip-directed exocytosis which lies at the root of tip-localised cell wall expansion.

In another recent biophysical model, Kroeger et al. (2011) applied the Lockhart equation that describes the rate of cell wall expansion as proportional to the ratio between turgor pressure and cell wall thickness: The pollen tube grows faster under greater turgor pressure and when its cell wall is thinner (and therefore more extensible). While Kroeger et al. (2011) assumed rapidly controlled, and therefore constant, turgor pressure they did assume dynamic cell wall thickness regulated by cytoplasmic calcium. Cytoplasmic calcium plays crucial roles in both exocytosis of cell wall building blocks and their eventual interlinking in forming the wall. The concentration of calcium is regulated by intracellular uptake and influx from the growth medium through stretch-activated ion channels (Kroeger et al., 2011).

The model described by Kroeger et al. (2011) relies on directed exocytosis of cell wall material and a direct relation between the expansion rate and the turgor pressure. The model of Kroeger et al. (2011) reproduces experimentally observed oscillatory growth, but their model fails to explain why exocytosis is directed and further relies on a direct relation between growth and turgor pressure that has not yet been proven experimentally.

The strength of the models of Rojas et al. (2011) and Kroeger et al. (2011) lies in explaining pollen tube tip growth from a biophysical point of view, however neither of these models establish how key processes such as exocytosis are localised to the tip. In our present work, we focus on establishment and maintenance of apical ROP1 localisation that is assumed to direct exocytosis to the tip. To this end, we summarise briefly two models of apical ROP1 localisation before we go into detail about the model that we propose to describe this process.

Apical Polarity Model

Both, existing models of apical ROP1 localisation and the model that we describe below are based on the *molecular switch* property of ROP1 described in Section 3.1.2.

Combined with this property, a number of experimental observations in tip-growing cells have been drawn together to infer a sensible mechanism of apical ROP1 localisation (Kost, 2008).

The key experimental observations that the model described by Kost (2008), Figure 3.6, is based on are the following. Fluorescently labelled GAP has been found

localised to the flanks of pollen tube tips and significant overexpression of GAP reduces pollen tube growth (Kost, 2008). Point-mutated ROP1 that has low association affinity with GDI was observed to accumulate at the flanks of pollen tube tips, whereas wild-type ROP1 localises to the apex (Kost, 2008). These observations have been interpreted to indicate that GAP inactivates ROP1 at the flanks, where GDI sequesters inactive ROP1 and recycles it to the apex. Further, GEF has been observed to localise to the site of Rho activity in fission yeast and animal cells and is therefore assumed to localise to the apex of growing pollen tubes (Kost, 2008).
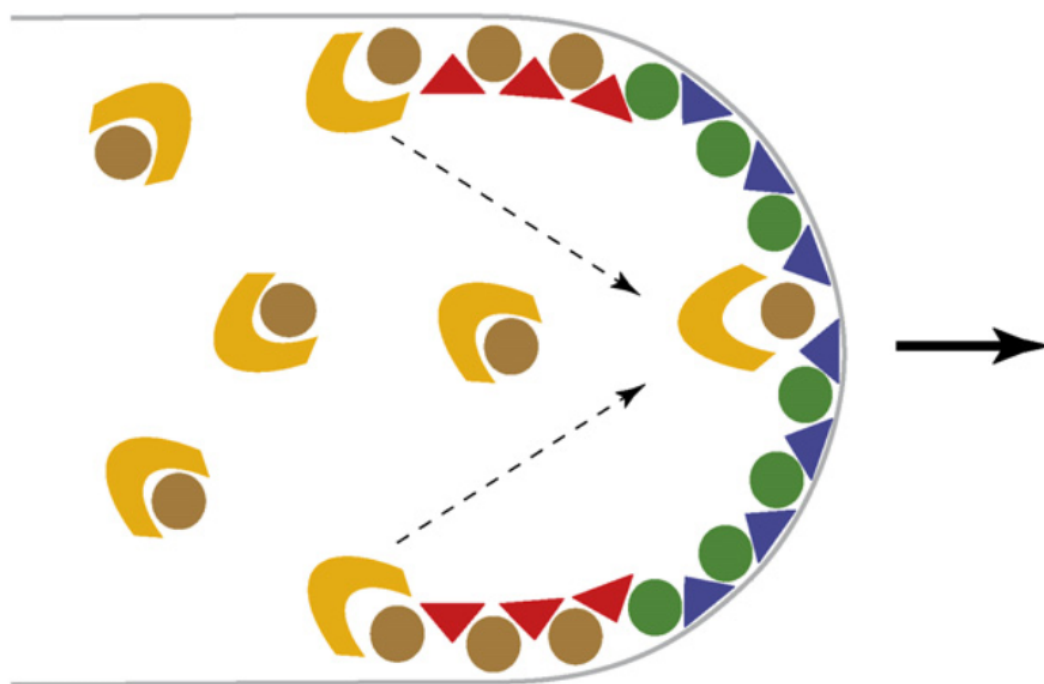
Figure 3.6: **Model of Apical ROP1 Localisation.** In this model, based on multiple experimental observations, there are three localised events: (i) Activation of inactive ROP1 (brown disks) at the apex by localised GEF (blue rectangles), (ii) inactivation of active ROP1 (green disks) at the flanks by localised GAP (red triangles), and (iii) sequestration of inactive ROP1 from the flanks by locally available GDI (yellow cups) and release of the inactive form at the apex. Directed transport of GDI bound to the inactive form follows the direction of vesicles transported directionally and exocytosed at the apex in tip growth. Reproduced from Kost (2008).

The model described in Figure 3.6 is sensible to the extent that it incorporates experimentally observed data and reproduces correctly apical ROP1 localisation. However, this model also leaves some basic questions unanswered: What processes localise GEF and GAP to their locations in the apex and flanks respectively? And if there are processes at play that "know" where the pollen tube tip is located in order to localise GEF and GAP to their respective positions, then why does the pollen tube still require apical ROP1 localisation for tip growth?

Oscillations Model

The preceding model, the Apical Polarity Model, predicts steady apical localisation of active ROP1 and therefore steady growth.

To model oscillation of apical ROP1 activity, Yan et al. (2009) devised mathematical models that include F-actin assembly (as a proxy of the growth rate), calcium concentration, and apical ROP1 activity. Their model is non-spatial so that all concentrations and activities are presumed to capture values near the apex of the growing pollen tube. Yan et al. (2009) incorporated two feedback loops, Figure 3.7: (i) active ROP1 upregulates its downstream effector RIC4 which in turn upregulates F-actin assembly, and (ii) active ROP1 also upregulates its downstream effector RIC3 which increases the concentration of calcium and in turn inhibits F-actin assembly via actin-binding proteins. As Yan et al. (2009) described, apical ROP1 activity can be observed to oscillate ahead of both growth rate and apical calcium concentration. Correlation analysis determined that ROP1 activity oscillates approximately 90 degrees ahead of growth rate while it oscillates 120 degrees ahead of calcium concentrations. This difference in delay between ROP1 activity and downstream observations was invoked by Yan et al. (2009) to introduce a delay, ν, between RIC3 and calcium (120 degrees) but none between RIC4 and F-actin (90 degrees), Figure 3.7.

For appropriate parameter values, Yan et al. (2009) find that their model reproduces the out-of-phase oscillations of apical ROP1 activity and apical calcium concentration that are commonly observed in growing pollen tubes.

While the model described in Figure 3.7 may have interesting dynamic and steady-state behaviour due to its two feedback loops, it is the introduction of time delay that permits this system to oscillate. To understand this, let us consider the following hypothetical model of apical ROP1 dynamics:

$$\frac{\mathrm{dROP1}}{\mathrm{dt}} = -\alpha \mathrm{ROP1}(t),$$

where we have chosen to model the time dynamics of apical ROP1 as a negative feedback loop – corresponding to the RIC3-dependent loop in Figure 3.7. For positive parameter values $\alpha > 0$, this model predicts that an initial concentration of

apical ROP1 decays as an exponential function described by the following expression:

$$\text{ROP1}(t) = \text{ROP1}_{\text{initial}} \exp{(-\alpha t)}.$$

Similar to Yan et al. (2009) we now amend this simplified negative feedback model with delay to obtain the following model:

$$\frac{d\text{ROP1}}{dt} = -\alpha \text{ROP1}(t - \nu), \tag{3.1}$$

where $\nu > 0$ is the delay parameter. As Smith (2011) described, it can be shown mathematically that this model produces oscillatory apical ROP1 concentrations for positive parameter values α and ν if $\alpha \nu > e^{-1}$. This inequality describes a condition for apical ROP1 oscillation based on the relative magnitude of the feedback α and delay ν.

To illustrate this, we choose $\nu = 1$ and $\alpha = 1/(e\,\nu) + 1$, and rename $\rho = \text{ROP1} + 5$ where we choose $\rho(t) = 5$ for $t \leq 0$. The resultant time behaviour of ρ shows oscillations, Figure 3.8.

What we show with this brief digression is that the model described by Yan et al. (2009) does not produce oscillatory apical ROP1 due to the double RIC3/RIC4 feedback loop but rather due to the introduction of delay. In fact, a model as simple as Equation (3.1) is already sufficient to produce oscillations as long as we introduce delay. To turn the damped oscillations displayed by our simplified model, Figure 3.8, into sustained oscillations as shown by the model presented by Yan et al. (2009) may require inclusion of feedback loops. However, the phenomenon of oscillation is achieved solely by introducing a delay.

Hence, a shortcoming of this model, Figure 3.7, is its limited explanatory value. This model reproduces both oscillatory and non-oscillatory growth as is observed for live pollen tubes, Figure 3.5, for different values of ν (Yan et al., 2009). However, the biological basis of the delay ν is not obvious and expanding this model to a mechanistic description of ν is non-trivial due to ambiguity.
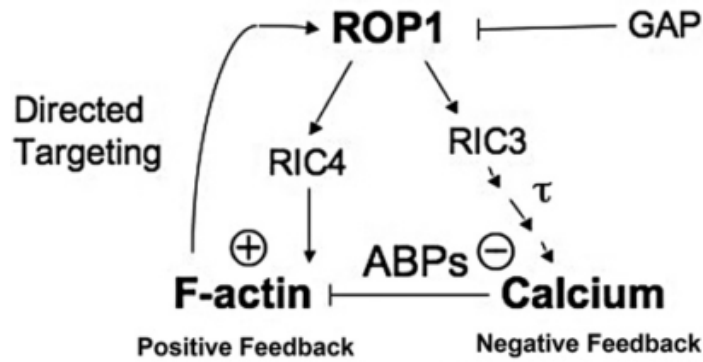
Figure 3.7: **Model of ROP1, F-Actin, and Calcium Oscillations.** In this model of ROP1 oscillations, Yan et al. (2009) included two feedback loops: (i) active ROP1 upregulates RIC4 which increases F-actin assembly, and (ii) active ROP1 upregulates RIC3 which downregulates F-actin assembly via calcium and actin-binding-proteins (ABPs). All concentrations and activities in this model are non-spatial and assumed to represent apical values. Further, F-actin assembly is understood as a proxy for pollen tube growth. Introduction of delay ν between RIC3 and calcium and not RIC4 and F-actin assembly was rationalised with a 120 degrees phase difference between the former and a 90 degrees phase difference between the latter pair of activities (Yan et al., 2009). Reproduced from Yan et al. (2009).
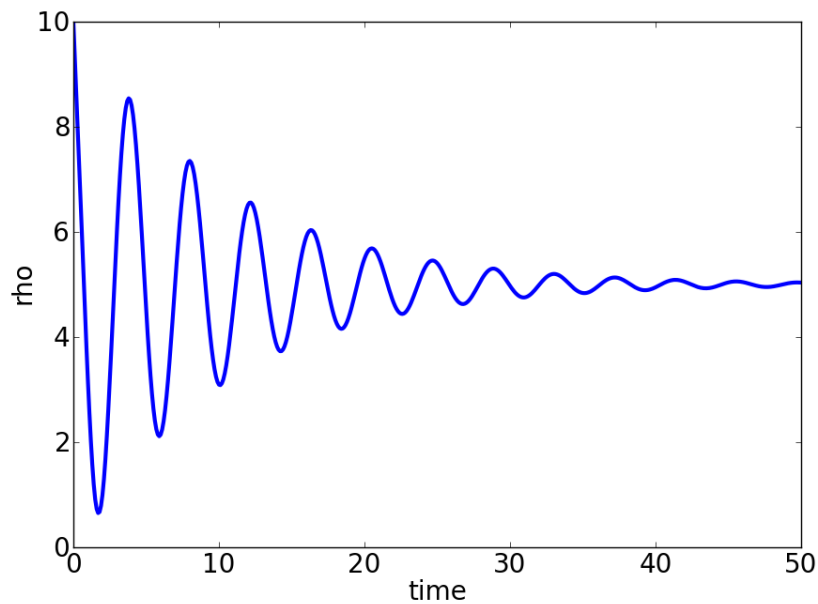
Figure 3.8: **Oscillatory Time Behaviour of ρ.** For appropriate parameter values, the time behaviour of Equation (3.1) shows oscillations. This illustrates that a model of ROP1 dynamics as simple as Equation (3.1) can show oscillatory time dynamics as long as we introduce a delay and choose appropriate parameter values. To turn these damped oscillations into sustained oscillations as discussed by Yan et al. (2009) may require adding more variables – however the phenomenon of oscillations is produced by introducing a time delay.

3.1.7 Pollen Tube Geometry and Line Concentrations

The pollen tube is modelled routinely as a cylinder and we apply the same geometric simplification (Holdaway-Clarke & Hepler, 2003). We simplify the pollen tube to a cylinder of length L and radius R. Then looking down the length of the cylinder, and assuming radial symmetry, we assume that all planes that run along the length of the tube are equivalent, Figure 3.9.
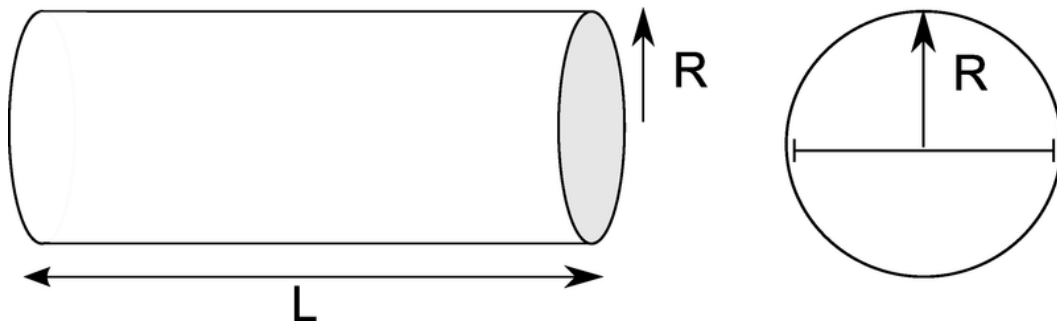
Figure 3.9: **Simplified Pollen Tube Geometry.** (left): The pollen tube simplified as a cylinder of length L and radius R. (right): A cross section of the pollen tube. The line denotes the top face of a plane that runs along the length of the tube. All planes of this kind are assumed equal (rotational symmetry).

Since all planes along the length of the pollen tube are assumed equal, we will model just one such plane, Figure 3.10.
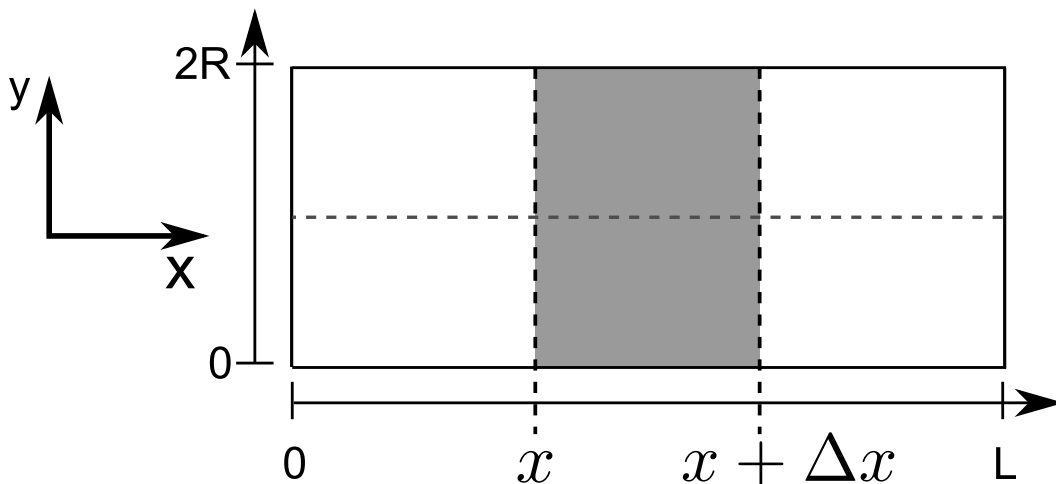
Figure 3.10: **Arbitrary Plane as Model of Pollen Tube.** We choose one arbitrary plane that runs along the length of the pollen tube as the basis for the domain of our dynamical system. Due to the dimensions of the cylinder that represents our pollen tube, this plane has length L and height $2R$.

We restrict our current model to one spatial domain and therefore we need to work with *line concentrations* of all involved chemical species. In the two-dimensional view of our pollen tube system, Figure 3.10, the plasma membrane is a line (top, or

equivalently, bottom horizontal edge of the plane) and therefore chemical species that live on the membrane are given in line concentrations by definition. The cytoplasm, however, is two-dimensional and we need to compute the effective line concentration of cytoplasmic species that interact with the plasma membrane.

Diffusivity of chemical species in the cytoplasm is orders of magnitude greater than that of species attached to the plasma membrane due to the difference in viscosity of these two phases (Postma et al., 2004). We assume that the radius of the pollen tube R is fixed and significantly smaller than its, potentially increasing, length L. Therefore we assume that any radial variation in a cytoplasmic species v vanishes and we take the line concentration of v to equal its mean radial concentration:

$$v(x,t) := \frac{1}{2R} \int_0^{2R} v(x,s,t)\mathrm{d}s. \tag{3.2}$$

Even though we assume radial symmetry in the cytoplasm, only a small band of the cytoplasmic species in the rectangle of area $2R \, \Delta \, x$, Figure 3.10 (grey area), interacts with membrane-bound species: Molecules of cytoplasmic species that are far away from the membrane will not interact with membrane-bound molecules and only cytoplasmic molecules sufficiently near the membrane can be expected to interact. We denote the width of this *biologically active* band of cytoplasmic concentration by r. Factors that contribute to r can be expected to include both the diffusion coefficient of the cytoplasmic chemical species and the timescale of their lifespan (protein turnover).

Given the above line concentration of a cytoplasmic chemical species v, the concentration of v, along a stretch of length Δx of the pollen tube, seen by a membrane-bound chemical species is given by

$$\frac{\Delta x r v(x,t)}{\Delta x 2R} = \omega v(x,t), \tag{3.3}$$

where we define $\omega = r/2R$ as the radial fraction of the cytoplasm that is visible chemically to membrane-bound chemical species. The ratio ω is a dimensionality correction factor that scales between line (membrane) and area (cytoplasm) concentrations.

3.1.8 Polarity Model

In this section we introduce a reaction-diffusion model of ROP1 dynamics that has been first described by Mori et al. (2008) for a generic Rho GTPase. This model differs from the ones we reviewed above in that it is a spatial model, and Rho localisation is emergent under certain conditions rather than prespecified.

The model of Rho GTPase regulation, Figure 3.1, introduced by Mori et al. (2008) makes the following assumptions: (i) The active form is bound to the membrane, and the inactive form is located mainly in the cytoplasm (hence, on average diffuses much faster than the active form), (ii) the active form upregulates its own production, presumably by regulating GEFs, and (iii) GAPs have a homogeneous distribution in space.

The reaction-diffusion equations that describe these assumptions are as follows (Mori et al., 2008):

$$
\begin{aligned}
\frac{\partial u}{\partial t} &= D_u \frac{\partial^2 u}{\partial x^2} + f(u,v), \\
\frac{\partial v}{\partial t} &= D_v \frac{\partial^2 v}{\partial x^2} - f(u,v),
\end{aligned}
\tag{3.4}
$$

where $u(x,t)$ and $v(x,t)$ denote the concentration of the active (low diffusivity) and inactive (high diffusivity) form respectively, and f is the reaction term. For simplicity we assume that concentrations are in *arbitrary units* (a.u.), since most quantitative data of ROP1 localisation stem from fluorescence measurements which are reported in arbitrary units relative to an appropriate baseline. The aforementioned discrepancy in diffusivity is expressed by $D_u \ll D_v$. The original reaction term f stated by Mori et al. (2008) did not include a dimensionality correction factor and we here modify f to include ω, Equation (3.3):

$$
f(u,v) = \omega v \left(k_0 + \frac{\gamma u^2}{K^2 + u^2} \right) - \delta u + k_s(x,t)v,
\tag{3.5}
$$

where the first two terms describe background activation and autoactivation respectively, the third term captures first-order inactivation, and the fourth term describes a spatially confined transient stimulus which we discuss further below (Mori et al.,

2008). Furthermore, zero flux is assumed on the boundaries (Mori et al., 2008):

$$\frac{\partial u}{\partial x} = 0, (x = 0, L),$$
$$\frac{\partial v}{\partial x} = 0, (x = 0, L).$$

(3.6)

Note that with these boundary conditions, no protein enters or escapes at the boundaries and the reaction term $f(\cdot)$ merely describes interconversion between the active and inactive form. A property of this system that we will discuss and use is the total concentration of protein

$$T = \frac{\tau}{L},$$

(3.7)

where L is the length of the system and τ is the total protein mass (or protein amount) and is expressed as

$$\tau = \int_0^L (u(x,t) + v(x,t))\mathrm{d}x.$$

(3.8)

Given appropriate parameter values, Table 3.1, Equations (3.4) can then be observed to evolve spontaneously from a homogeneous initial condition to a steady state with localised active ROP1, Figure 3.11. In our current work we are not interested in potential volume effects and therefore choose a dimensionality factor ω of value one, Table 3.1.

Spontaneous localisation of active ROP1, Figure 3.11, is important since spontaneous emergence of ROP1 localisation is likely involved in leading the pollen tube out of the grain in the first place.

For the remainder of our work we assume that there was an initial, spontaneous ROP1 localisation event that defined the pollen tube apex. The geometry of the apex conveys preference for localisation of active ROP1 to the tip: The reaction-diffusion system, Equation (3.4), minimises the length of the interface between localised u and dilute u. Generally, in a two-dimensional elliptical spatial domain we observe that the mass of localised u migrates as a whole and along the domain boundary towards either of the two end points of the major axis. We observe that the mass of

75

localised u only stops migrating once the boundary between localised and dilute u lies perpendicular to the major axis of the ellipse. We assume that the pollen tube apex is already formed and that the semi-elliptical shape of the apex exerts a pull on localised u due to this geometric effect. To account for this assumed geometry in our one-dimensional model, we introduce a stimulus term into Equation (3.4) which conveys preference of ROP1 localisation to the tip:

$$k_s = \begin{cases} 0 & : x < L - \varepsilon_s \\ k_{s,0} & : x \geq L - \varepsilon_s, \end{cases}$$

where ε_s denotes a small distance from the right-hand boundary, and $k_{s,0} > 0$ is a positive rate constant.

With an appropriate tip stimulus and our parameter values, Table 3.1, we can then observe tip-localised active ROP1, Figure 3.12.
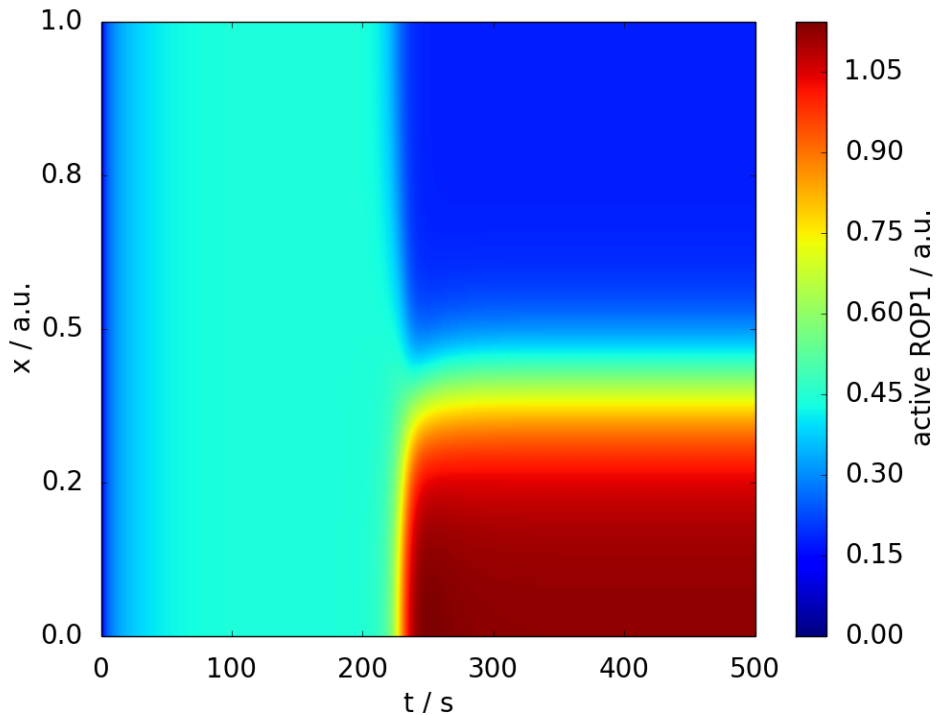
Figure 3.11: **The Active Form of ROP1 can Localise Spontaneously.** For appropriate parameter values, Table 3.1, the active form can localise spontaneously starting from a homogeneous initial concentration profile (Holmes, 2014; Walther et al., 2012).

Parameter	Value	Unit	Comment
L	1.0	a.u.	We assume that our pollen tube has length one initially.
D_v	0.1	a.u.^{2}/s	High diffusivity of the inactive, cytoplasmic form.
D_u	0.001	a.u.^{2}/s	Low diffusivity of the active, membrane-bound form. Postma et al. (2004) reported hundred-fold lower diffusivity for the active form compared with the inactive form.
k_0	0.067	1/s	Due to Mori et al. (2008).
γ	1.0	1/s	Due to Mori et al. (2008).
K	1.0	a.u.^{2}	Due to Mori et al. (2008).
δ	1.0	1/s	Spatially uniform GAP activity.
ε_s	–	a.u.	Grid-dependent: chosen so that stimulus occurs in right-most grid point.
$k_{s,0}$	0.0001	1/s	Choice of this value is arbitrary and determines the total protein concentration that suffice to trigger ROP1 localisation.
ω	1.0	–	For simplicity we choose a unit correction factor.
τ	2.36	a.u.	Value can be derived using the method described in Section 3.1.9.

Table 3.1: **Standard Parameter Values Used For Our Numerical Experiments.** These parameter values are based mostly on values presented by Mori et al. (2008) and Walther et al. (2012).
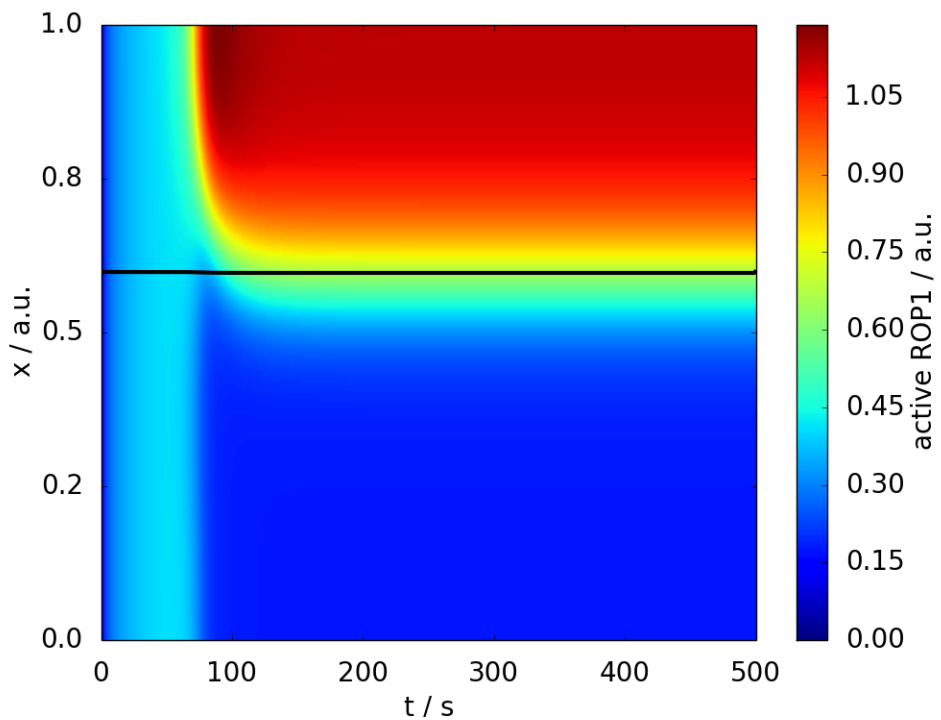
Figure 3.12: **Apical Localisation of Active ROP1.** Kymograph of the time behaviour of the active form u starting with a homogeneous initial condition and using the parameter values summarised in Table 3.1. Black line: position of the interface between high and low concentrations of active ROP1 as predicted with our method described in Section 3.1.9.

3.1.9 Analysing the Polarity Model

We now have an overview of the polarity model that we will study. While we can run simulations of this model for many different parameter values it will be helpful to analyse the model and understand what behaviour we expect for a given set of parameter values.

Perturbability: Local Perturbation Analysis

It is straightforward to see that our model, Equation (3.4), allows for at least one non-trivial homogeneous steady state if we relate concentrations u and v via $u+v=T$ (mass conservation, fixed domain length L) and set the stimulus to zero everywhere, $k_s=0$.

Suppose our system is in one of these homogeneous steady states where both the active and inactive form are uniform in space $u(x,t)=u^*$, $v(x,t)=v^*$ ($u^*>0$, $v^*>0$). To observe localisation of active ROP1, u, it needs to be possible to perturb our system out of the homogeneous steady state so that it can evolve to an inhomogeneous steady state.

As a side note, the homogeneous steady state $(u,v)=(u^*,v^*)$ of the reaction-diffusion system, Equation (3.4), is also a steady state of the kinetics term $f(u,v)$. However, the targeted inhomogeneous steady state of the reaction-diffusion system does generally not contain a steady state of $f(u,v)$.

Knowing the exact steady state that our system moves to after perturbation is hard, hence we will settle with answering two simpler questions: (i) Can we expect our system to move to a different steady state if we apply the "right" perturbation to it? (ii) If so, how big a perturbation (how much do we need to alter u and v) is sufficient so that our system does not just move back to (u^*,v^*)?

The Local Perturbation Analysis (LPA) answers these questions by determining the perturbable parameter regime of given reaction-diffusion equations (Walther et al., 2012; Holmes, 2014): In a perturbable regime, a sufficiently large perturbation in the active form can generate a stable interface between a patch of highly concentrated, localised, active form in a sea of dilute active form. Our system, Equation (3.4), is in a perturbable regime in Figure 3.12 where a stimulus on the far right is

sufficient to trigger such an interface.

Note that this concept of perturbability is different from classical analyses of pattern formation in reaction-diffusion systems. Classically, we would analyse the time behaviour of infinitely small perturbations away from the homogeneous steady state (u^*, v^*), (Tyson, 1975). Since the magnitude of these perturbations is infinitely small, we can usually approximate their initial time behaviour with the linear part of the Taylor series of the original dynamics term. The analytic solution of this approximated dynamical system is known and tells us directly if the applied perturbations will grow (the system leaves the homogeneous steady state) or vanish (the system falls back to the homogeneous steady state). The marked difference between this classical approach and LPA is that in the former we apply perturbations of infinitely small magnitude whereas in the latter we learn something about the magnitude of perturbation we need to apply to move the system away from the homogeneous steady state.

Both Holmes (2014) and Walther et al. (2012) describe the LPA in detail and here we will apply the LPA to our system, Equation (3.4). In the LPA, we consider an extreme case of our system where the active form has zero diffusivity and the inactive form has infinitely high diffusivity (Holmes, 2014; Walther et al., 2012). Applying these approximations to our system, Equation (3.4), we generate the following system of ordinary differential equations:

$$
\begin{aligned}
\dot{u}_G &= f(u_G, v_G), & (3.9) \\
\dot{u}_L &= f(u_L, v_G), \\
\dot{v}_G &= -f(u_G, v_G), \\
T &= u_G + v_G,
\end{aligned}
$$

where u_G is non-diffusing active form spread across the entire pollen tube, and u_L is a perturbation away from u_G with arbitrary magnitude that is located in an infinitely narrow interval on the membrane. Variable v_G denotes the spatially uniform concentration of the inactive form. Note that since u_L and u_G are assumed to be uniform concentration profiles and v_G is uniform due to our choice of diffusivity, the

diffusion operator vanishes for all three variables. The last relation of Equation (3.9) follows directly from Equation (3.7) for uniform concentrations u_G and v_G. Note that u_L is assumed not to contribute to the total protein concentration in the pollen tube since it lives in an infinitely narrow interval.

Using the same parameter values as in Table 3.1, and varying total concentration T we reproduce the bifurcation plot presented by Walther et al. (2012) in Figure 3.13.
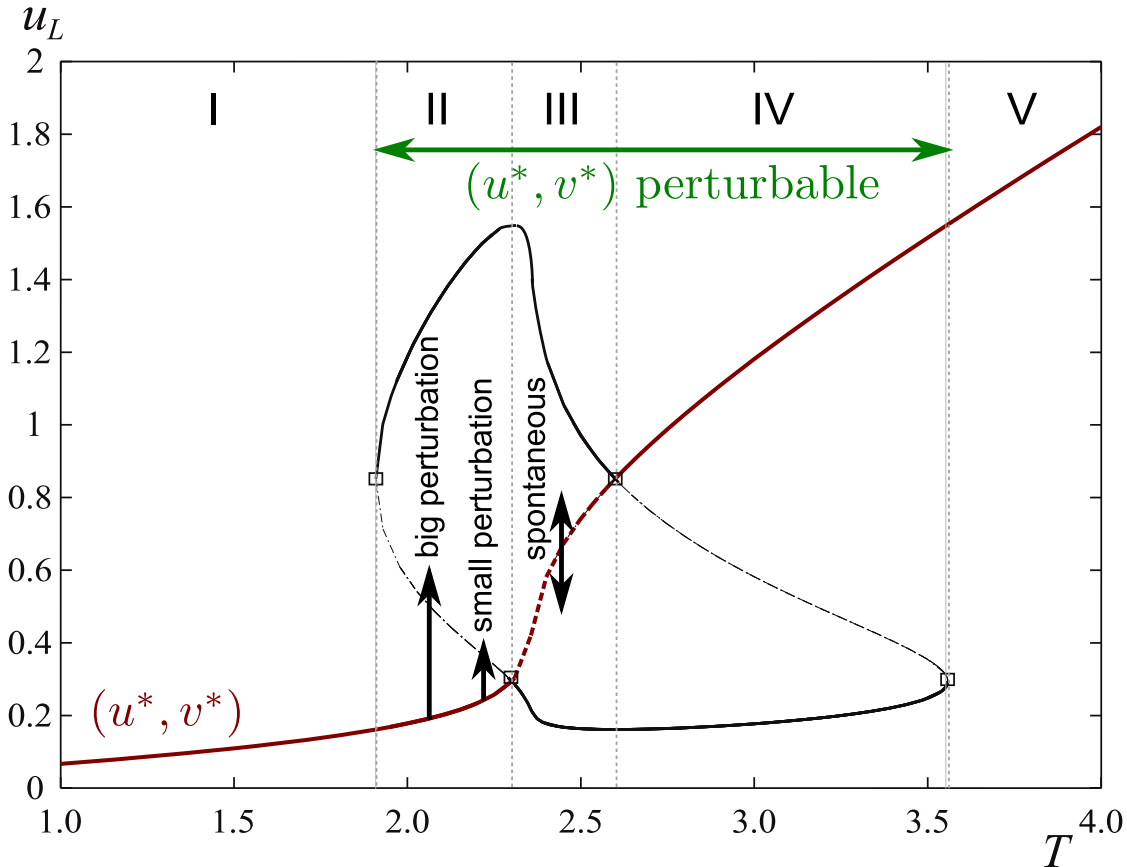
Figure 3.13: **Local Perturbation Analysis Bifurcation Plot.** Bifurcation plot of the perturbation in the active form in Equation (3.9). In zone II (IV) a sufficiently high (low) stimulus (suppression) can trigger a travelling interface between a patch of high (low) concentration of the active form in a sea of uniform low (high) concentration of the active form. In zone III an arbitrarily small perturbation will trigger such a travelling interface – here a travelling interface can be triggered spontaneously by stochastic events. In zones I and V the system is not perturbable. Adapted from Walther et al. (2012).

The LPA bifurcation plot, Figure 3.13, summarises the answers to our previous questions. We can perturb the homogeneous steady state to an inhomogeneous steady state in zones II - IV and the required size of the perturbation is visible in Figure 3.13. In zone II we need to elevate u_L above an unstable steady state, Figure 3.13 (black dashed line in zone II), and the size of the required perturbation depends on the total concentration T. In zone III the homogeneous steady state $(u_L, v_G) = (u^*, v^*)$

is unstable, Figure 3.13 (red dashed line), and therefore small random perturbations that either elevate or suppress u_L will drive the system to an inhomogeneous steady state. In zone IV we need to suppress u_L below an unstable steady state, Figure 3.13 (black dashed line in Zone IV), and the size of the required perturbation depends on the total concentration T. In the remainder of this study our system is located either in zone II or III and any applied perturbation, or stimulus, elevates u_L rather than suppress it.

Polarisability: The Sharp Front Approximation

The LPA guides us in choosing parameter values so that we can trigger a travelling interface in the active form of ROP1. An important question is now: Where will the interface stop? This is an important question to ask due to the functional link between localised ROP1 and local growth, Section 3.1.3: If, for instance, the interface stops near the base of the pollen tube then the tube would not grow apically and rather balloon to the sides.

Specific to our case, we need to ask how wide the plateau with high concentrations of active ROP1 will become, Figure 3.12 (red shades). Mathematically, if the width of the high plateau becomes too big, boundary effects will destroy the formed interface and our system evolves to a new stable rest state with the active species uniformly high in concentration. Biologically, an excessively wide high plateau has the aforementioned effect of pollen tube ballooning. Mathematically, if the width becomes too small then our system evolves back to the original stable rest state with the active species uniformly low in concentration. Biologically, an overly narrow high plateau may be insufficient to direct local growth.

To approximate the width of the high plateau we expand a numerical analysis technique introduced by Walther et al. (2012). This technique, similar to the LPA, assumes that diffusivity on the membrane is infinitely low and diffusivity in the cytoplasm is infinitely high:

$$D_u \rightarrow 0, \; D_v \rightarrow \infty.$$

In this limit an interface in the active form will have a very sharp boundary as

diffusion on the membrane tends to zero (Mori et al., 2011). As Mori et al. (2008) discussed, the velocity of the boundary of a sufficiently sharp interface on a finite domain such as ours can be approximated by the following expression

$$c_f \propto \int_{u_-}^{u_+} f(u, v(u)) \, \mathrm{d}u, \tag{3.10}$$

where f is the kinetic term of our system, Equation (3.5), and the concentration of v is a function of u due to mass conservation

$$v(u) = T - \frac{1}{L} \int_0^L u(x, t) \, \mathrm{d}x.$$

The front velocity, Equation (3.10), was formulated for a left-polarised interface so that

$$u(x, t) \to u_+ \quad , \quad x \to -\infty,$$
$$u(x, t) \to u_- \quad , \quad x \to \infty, \tag{3.11}$$

where $u_+ > u_-$ are both stable steady states of the kinetic term $f(\cdot)$, Equation (3.5). Hence, a positive velocity $c_f > 0$ indicates that the high plateau expands in size while a negative velocity $c_f < 0$ indicates that the high plateau retracts.

To adapt Equation (3.10) of the velocity to our right-polarised interface we make use of the following symmetry considerations: Left-polarised and right-polarised interfaces need to behave equally. Suppose our system displays a left-polarised interface whose high plateau has width w and travels with velocity c_f, Figure 3.14 (left). Then mirroring this concentration profile in the centre of the domain, we obtain a right-polarised interface whose high plateau has width w and travels with velocity c_f in the opposite direction, Figure 3.14 (right). This mirroring operation also reverses the velocity vector of our interface so that a left-polarised right-moving interface is turned into a right-polarised left-moving interface, Figure 3.14. Based on these considerations, we reinterpret the velocity c_f to approximate how fast width w of the high plateau expands or shrinks: If $c_f > 0$ then w expands, whereas when $c_f < 0$ then w shrinks.
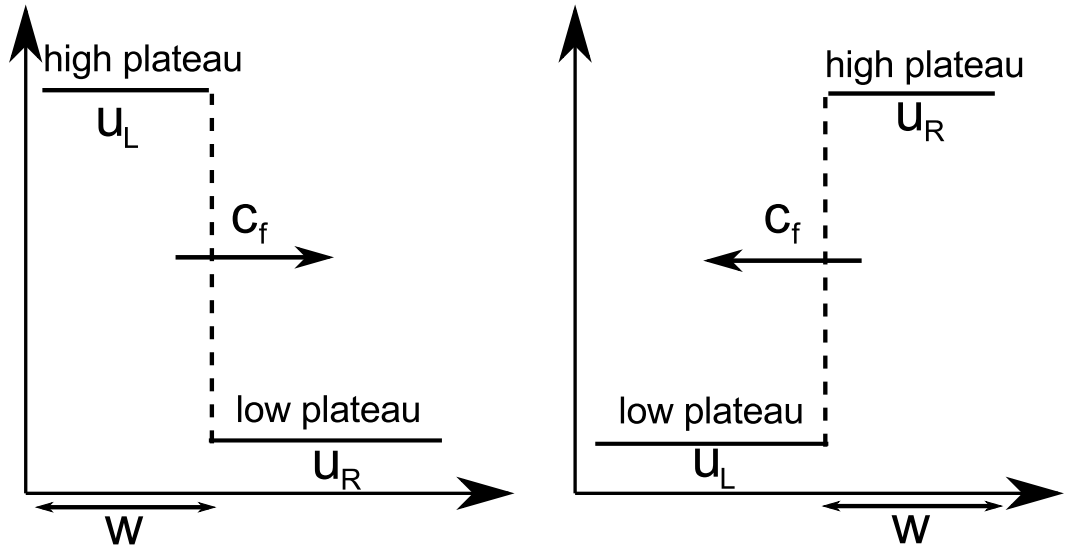
Figure 3.14: **Sharp Front Approximation.** We approximate the sharp front solution by representing it with two plateaus, u_L and u_R, and we track the width of the high plateau with variable w. We reinterpret the sharp front velocity c_f, Equation (3.10), to indicate the rate at which w expands or shrinks. For our purpose we assume that the proportionality constant in Equation (3.10) equals 1 since we are interested in where the interface stops, $c_f=0$, and not in the absolute velocity of the interface.

We approximate the behaviour of a sufficiently sharp interface with a three-variable system of ordinary differential equations, Figure 3.14:

$$
\begin{aligned}
\dot{u}_L &= f(u_L, v), \\
\dot{u}_R &= f(u_R, v), \\
\dot{w} &= \int_{u_-}^{u_+} f(u, v)\, \mathrm{d}u, \\
v &= T - wu_+ - (L - w)u_-,
\end{aligned}
\tag{3.12}
$$

where u_L and u_R denote the concentration levels in the left- and right-hand plateau respectively, and w denotes the width of the high plateau. The concentration in the high plateau, u_+, corresponds to u_L for a left-polarised interface, Figure 3.14 (left), and u_R for a right-polarised interface, Figure 3.14 (right). Note that we set the proportionality constant in Equation (3.10) to one since we are interested in the pinning position of the wave, $c_f=0$, and not its precise velocity.

For a left-polarised interface, $u_L \approx u_+$, $u_R \approx u_-$ hence $\dot{w}$ and v become

$$\dot{w} = \int_{u_R}^{u_L} f(u,v)\,\mathrm{d}u,$$

$$v = T - w \cdot u_L + (L - w) \cdot u_R, \tag{3.13}$$

whereas for a right-polarised interface, $u_L \approx u_-$, $u_R \approx u_+$ hence $\dot{w}$ and v become

$$\dot{w} = \int_{u_L}^{u_R} f(u,v)\,\mathrm{d}u,$$

$$v = T - w \cdot u_R + (L - w) \cdot u_L. \tag{3.14}$$

A stable steady state of the three-variable Equation (3.12) approximates the pinning position and plateau levels of an interface between low and high concentrations of the active form of ROP1. Hence, such a steady state approximates to what extent active ROP1 localises.

3.1.10 Modelling Growth

Since pollen tubes can grow fast and increase their volume by a factor of twenty, Section 3.1.4, we believe it is essential to study tip-localisation of ROP1 on a growing domain. To model growth of the pollen tube, we use a framework developed by Crampin E. J. (2000).

A central concept of this framework is that of local growth velocity. Suppose that we consider a volume element of the pollen tube located in position x then the local growth velocity of this element is given by:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = a(x,t). \tag{3.15}$$

The local growth velocity, $a(x,t)$, assigns a space- and time-dependent velocity to each volume element of the pollen tube. We note that our local velocity term $a(x,t)$ has one space component since we are interested exclusively in longitudinal growth and not any of the morphological deformations, such as ballooning, that are observed for pollen tubes under certain conditions, Figure 3.3.

Here we describe how our reaction-diffusion system, Equation (3.4), is altered by the introduction of $a(x, t)$ (Crampin E. J., 2000).

Derivation

For brevity we state the following derivation for the active form u only and introduce coordinates $\mathbf{r}$, and reaction term $R(\mathbf{r}, t)$. For an arbitrary volume element $V(t)$, Figure 3.15(a), we can write down a conservation equation for the protein mass contained within $V(t)$:

$$
\begin{aligned}
\frac{d}{dt} \int_{V(t)} u(\mathbf{r}, t) \mathrm{d}V &= \int_{\partial V(t)} -(\mathbf{n} \cdot \mathbf{j}) \mathrm{d}A + \int_{V(t)} R(\mathbf{r}, t) \mathrm{d}V \qquad (3.16) \\
&= \int_{V(t)} -(\nabla \cdot \mathbf{j}) \mathrm{d}V + \int_{V(t)} R(\mathbf{r}, t) \mathrm{d}V \\
&= \int_{V(t)} (D_u \nabla \cdot \nabla u + R(\mathbf{r}, t)) \mathrm{d}V,
\end{aligned}
$$

where u denotes the active form of ROP1, $\partial V(t)$ denotes the surface of volume element $V(t)$, $\mathbf{n}$ is the outward-pointing normal on $\partial V(t)$, $\mathbf{j}$ is some (initially unknown) flux of u on the boundary $\partial V(t)$, and $R(\mathbf{r}, t)$ is the reaction term (this term is a shorthand for $f(u(x, y, t), v(x, y, t))$). Note that we use shorthand notation for coordinates and volume elements: $\mathbf{r}$ stands for the two-dimensional coordinate (x, y), $\mathrm{d}V$ stands for $\mathrm{d}x\mathrm{d}y$, and $\mathrm{d}A$ corresponds to a line segment. In the derivation of the final form of Equation (3.16) we follow the standard steps of first applying the Divergence Theorem over the closed surface $\partial V(t)$ and then realising that the flux $\mathbf{j}$ is due to Fickian diffusion, hence $\mathbf{j} = -D_u \nabla u$ where D_u is the diffusion coefficient of u.

If the volume element $V(t)$ were constant in shape then we would be able to draw the $\mathrm{d}/\mathrm{d}t$ operator on the left-hand side under the integral and we would end up with a standard reaction-diffusion equation. However, since $V(t)$ will be distorted by growth, Figure 3.15(a), we need to take a few extra steps.

The time derivative on the left-hand side of Equation (3.16) is, by definition, equal to the following limit (if it exists):

$$
\frac{d}{dt} \int_{V(t)} u \mathrm{d}V = \lim_{\Delta t \to 0} \frac{\int_{V(t+\Delta t)} u(\mathbf{r}, t + \Delta t) \mathrm{d}V - \int_{V(t)} u(\mathbf{r}, t) \mathrm{d}V}{\Delta t}. \qquad (3.17)
$$

85

Let us expand the first integrand on the right-hand side in a Taylor series:

$$u(\mathbf{r}, t + \Delta t) = u(\mathbf{r}, t) + \Delta t \frac{\partial u}{\partial t} + O(\Delta t^2),$$

and assuming that Δt is sufficiently small we ignore higher order Δt-terms. We substitute this approximation of $u(\mathbf{r}, t + \Delta t)$ into Equation (3.17) and obtain

$$
\begin{aligned}
\frac{d}{dt} \int_{V(t)} u \mathrm{d}V &\approx \lim_{\Delta t \to 0} \frac{\int_{V(t+\Delta t)} \left(u(\mathbf{r}, t) + \Delta t \frac{\partial u}{\partial t} \right) \mathrm{d}V - \int_{V(t)} u(\mathbf{r}, t) \mathrm{d}V}{\Delta t} \quad (3.18) \\
&= \int_{V(t)} \frac{\partial u}{\partial t} \mathrm{d}V + \lim_{\Delta t \to 0} \frac{\int_{V(t+\Delta t)} u(\mathbf{r}, t) \mathrm{d}V - \int_{V(t)} u(\mathbf{r}, t) \mathrm{d}V}{\Delta t}.
\end{aligned}
$$

The Taylor approximation we introduce in Equation (3.18) allows us to view the difference between the two integrals on the right-hand side as the result of a volume element $V(t)$ that grows to a distorted volume element $V(t+\Delta t)$ in an unchanged field $u(\mathbf{r}, t)$, Figure 3.15(b). Assuming that Δt is sufficiently small, we can approximate this difference by

$$\int_{V(t+\Delta t)} u(\mathbf{r}, t) \mathrm{d}V - \int_{V(t)} u(\mathbf{r}, t) \mathrm{d}V \approx \int_{\partial V(t)} (\mathbf{n} \cdot u\mathbf{a}) \Delta t \mathrm{d}A,$$

and substituting this into Equation (3.18) we can drop the limit (since Δt cancels) and obtain

$$
\begin{aligned}
\frac{d}{dt} \int_{V(t)} u \mathrm{d}V &= \int_{V(t)} \frac{\partial u}{\partial t} \mathrm{d}V + \int_{\partial V(t)} u(\mathbf{a} \cdot \mathbf{n}) \mathrm{d}A \quad (3.19) \\
&= \int_{V(t)} \left(\frac{\partial u}{\partial t} + \nabla \cdot (u\mathbf{a}) \right) \mathrm{d}V,
\end{aligned}
$$

where we use the Divergence Theorem over the closed surface $\partial V(t)$ of the volume element $V(t)$ in Figure 3.15(b).

This concludes our derivation of how to draw the d/dt operator on the left-hand side of Equation (3.16) under the integral sign. Upon substituting Equation (3.19) into Equation (3.16), we obtain

$$\int_{V(t)} \left(\frac{\partial u}{\partial t} + \nabla \cdot (u\mathbf{a}) - D_u \nabla \cdot \nabla u - R(\mathbf{r}, t) \right) dV = 0,$$

and since $V(t)$ is an arbitrary choice, this integral can only equal zero if the integrand equals zero everywhere – thus giving us the more commonly seen differential form of this reaction-diffusion-advection system (we write $\nabla^2 u = \nabla \cdot \nabla u$ for brevity):

$$\frac{\partial u}{\partial t} + \nabla \cdot (u\mathbf{a}) = D_u \nabla^2 u + R(\mathbf{r}, t) \tag{3.20}$$

We are only interested in longitudinal growth and not any three-dimensional bending or ballooning that is observed for growing pollen tubes experimentally. To this end, we restate Equation (3.20) with our one-dimensional velocity $a(x, t)$:

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x} \cdot (u(x, t)a(x, t)) = D_u \frac{\partial^2 u}{\partial x^2} + R(x, t). \tag{3.21}$$

The resultant full set of equations with our kinetics term f of the one-dimensional concentration variables is:

$$
\begin{aligned}
\frac{\partial u}{\partial t} + \frac{\partial}{\partial x}(ua) &= D_u \frac{\partial^2}{\partial x^2} u + f(u, v), \\
\frac{\partial v}{\partial t} + \frac{\partial}{\partial x}(va) &= D_v \frac{\partial^2}{\partial x^2} v - f(u, v),
\end{aligned}
\tag{3.22}
$$

Compared with our system for a non-growing pollen tube, Equation (3.4), this system of equations that incorporates longitudinal growth possesses the following additional terms: dilution terms, $u\partial a/\partial x$ and $v\partial a/\partial x$, and advection terms, $a\partial u/\partial x$ and $a\partial u/\partial x$. The former terms describe dilution because $\partial a/\partial x$ is a local strain rate describing how fast points along the length of the pollen tube move apart. The latter terms are advection terms and describe directed transport of ROP1 due to growth.

Setup and derivation of Equation (3.22) may be a bit abstract, therefore to gain some intuition let us think of longitudinal growth as follows: In the initial pollen tube (solid, black pollen tube Figure 3.15(c)) we label all infinitesimally small volume elements with their initial position X. As the pollen tube grows, the position of element X describes a trajectory $G(X, t)$ (grey, dashed pollen tube, Figure 3.15(c))

which is equivalent to position x in Equation (3.22). Generally, the form of $G(X,t)$ is unknown: If we did know $G(X,t)$ from the start then we would know exactly how the pollen tube grows over time – specifying $G(X,t)$ would therefore impose a specific pollen tube length for every time point t. In the more general case, $G(X,t)$ is unknown and evolves over time as a function of our system of protein concentrations or other properties of the pollen tube. We impose, however, the following initial and boundary conditions on $G(X,t)$: (i) $G(X,0)=X$ which follows from the initial label of volume elements (dashed diagonal in Figure 3.15(c)), and (ii) $G(0,t)=0$ which serves to model elongation rather than translation. We assume that the pollen tube either grows or stalls and never shrinks, hence we expect that the plot of $G(X,t)$ never falls below the initial diagonal line $G(X,0)$, Figure 3.15(c).

For longitudinal growth, the connection between $G(X,t)$ and growth velocity $a(X,t)$ is:

$$G(X,t) = G(X,0) + \int_0^t a(X,s)\mathrm{d}s,$$

and using both the fundamental theorem of calculus and summarising the aforementioned initial and boundary conditions, we obtain

$$
\begin{aligned}
\frac{\partial G}{\partial t} &= a(X,t), & (3.23)\\
G(X,0) &= X,\\
G(0,t) &= 0.
\end{aligned}
$$

Equation (3.23) is a partial differential equation that we can solve alongside protein dynamics. This formulation also allows us to specify local growth velocities $a(X,t)$ as a function of protein concentrations or any other property of our system.

To study Equations (3.22) numerically, we make use of standard space and time discretisation and approximation techniques. The grid size of our spatial domain is altered by growth, potentially requiring intricate adaptive remeshing of our grid. Therefore we reformulate Equations (3.22) for our numerical experiments in terms of the initial positions X of the volume elements x in our domain (material coordi-

nates).

The volume elements x of our growing pollen tube are related to their initial positions X through the trajectory function G:

$$x = G(X, t),$$

which is a unique (one-to-one) mapping of an arbitrary X at time $t=0$ to a corresponding volume element x at time t. Using this mapping we can transform the spatial derivative operators in Equation (3.22) from x-space to X-space:

$$
\begin{aligned}
\frac{\partial}{\partial x} &= \frac{1}{G_X} \frac{\partial}{\partial X}, \\
\frac{\partial^2}{\partial x^2} &= \frac{1}{G_X^2} \frac{\partial^2}{\partial X^2} - \frac{G_{XX}}{G_X^3} \frac{\partial}{\partial X}.
\end{aligned}
\tag{3.24}
$$

Using these operator transformations we rewrite Equation (3.22) for X-space (material coordinates) and combine our result with our definition of $G(X, t)$, Equation (3.23):

$$
\begin{aligned}
\frac{\partial u}{\partial t} + u \frac{G_{Xt}}{G_X} + \frac{G_t}{G_X} \frac{\partial u}{\partial X} &= D_u \left(\frac{1}{G_X^2} \frac{\partial^2 u}{\partial X^2} - \frac{G_{XX}}{G_X^3} \frac{\partial u}{\partial X} \right) + f(u, v), \\
\frac{\partial v}{\partial t} + v \frac{G_{Xt}}{G_X} + \frac{G_t}{G_X} \frac{\partial v}{\partial X} &= D_v \left(\frac{1}{G_X^2} \frac{\partial^2 v}{\partial X^2} - \frac{G_{XX}}{G_X^3} \frac{\partial v}{\partial X} \right) - f(u, v), \\
\frac{\partial u}{\partial X} &= 0, \ (X = 0, 1), \\
\frac{\partial v}{\partial X} &= 0, \ (X = 0, 1), \\
\frac{\partial G}{\partial t} &= a(X, t), \\
G(0, t) &= 0.
\end{aligned}
\tag{3.25}
$$

All concentration variables in Equation (3.25) live in X-space, the initial configuration of the pollen tube, Figure 3.15 (a). So for instance, $u \equiv u(X, t)$. The only undefined entity in Equation (3.25) is the local growth velocity $a(X, t)$ which will be defined by our choice of growth model.
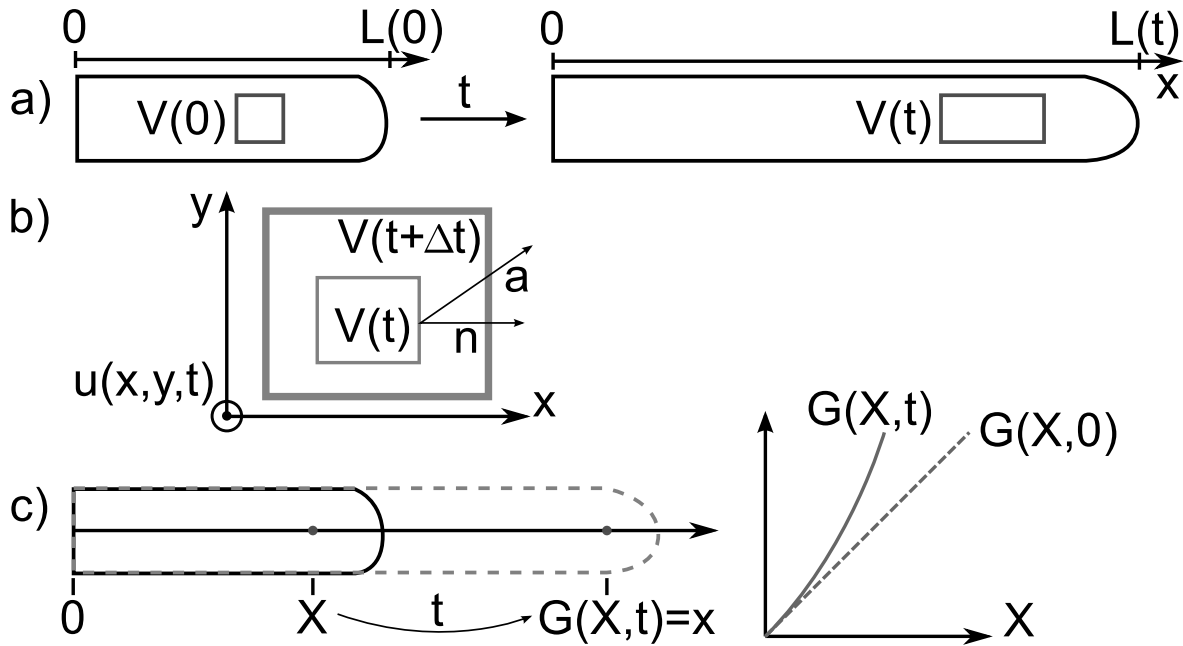
Figure 3.15: **Protein Dynamics in a Growing Pollen Tube.** (a) An arbitrary volume element V is shown at an initial time point and after an amount of time t has elapsed. In this span of time, the pollen tube has grown and the volume element V is distorted from its original configuration $V(0)$ to its final configuration $V(t)$. (b) Growth of an arbitrary volume element, due to growth velocity **a**, from its initial configuration $V(t)$ to its final configuration $V(t+\Delta t)$ in a fixed concentration field $u(x, y, t)$. As this volume element expands, the amount of protein u it encompasses changes. Vector **n** is the outwards normal on the surface of $V(t)$. (c) We are interested in longitudinal growth of the pollen tube allowing us to ignore intricate geometric effects due to bending, ballooning, and other processes that are observed experimentally. This simplification turns our pollen tube system into a one-dimensional spatial domain. We understand growth as a mapping $G(\cdot)$ that maps material points from their initial positions X to positions $G(X, t)$ after time t. The mapping $G(X, t)$ denotes the trajectory of X due to growth and this trajectory is generally unknown as it evolves over time with changing material properties of the pollen tube. We impose however two conditions on $G(\cdot)$: an initial condition $G(X, 0) = X$ (dashed line) and a boundary condition $G(0, t) = 0$. This boundary condition ensures that we model pollen tube elongation and not translation in space. Further note that $G(X, t)$ will never fall below the diagonal dashed line $G(X, 0)$ as we assume that our pollen tube either grows or stalls and never shrinks.

Our Model of Local Growth in the Pollen Tube

Growth of the pollen tube is localised to the tip: In growing pollen tubes, the cytoskeleton is arranged in a way that promotes directed exocytosis of membrane and cell wall material at the tip, Section 3.1.3.

We think of pollen tube growth as the product of two factors: Firstly, is the pollen tube in the condition to grow overall? And secondly, are there local differences in the readiness to grow? We express this mathematically as follows:

$$a(X,t) = a_G(u,v)a_L(X), \tag{3.26}$$

where we represent any local growth differences by a_L and global, systemic properties that determine overall propensity of the pollen tube to grow by a_G. We assume that ROP1 is responsible for downstream effects that place the pollen tube in a growth regime, Section 3.1.3, and we therefore model a_G as a function of the state of our protein system, u and v.

To focus our work on the interplay between the ROP1 system and growth of the pollen tube, we choose the most parsimonious assumption for the spatial contribution to growth by equating a_L=1. This assumption results in uniform growth along the length of our pollen tube model, Figure 3.16. With this assumption our model loses specificity but our results gain generality and we do not need to impose any prior belief on a_L.
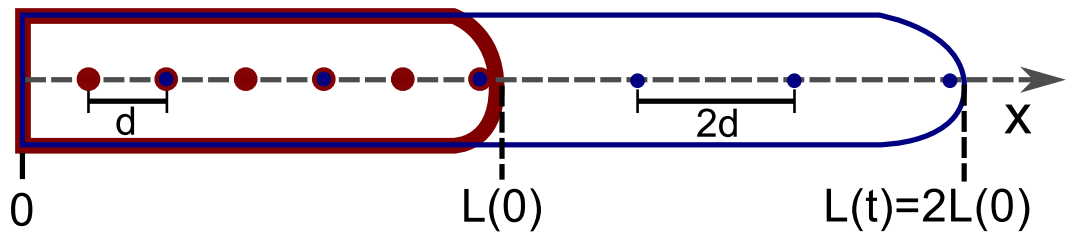
Figure 3.16: **Uniform growth assumption used in our work.** To focus our work on the interplay between the ROP1 system and growth we assume uniform growth along the length of the pollen tube, which we encode by substituting a_L=1 into Equation (3.26). The resultant growth velocity, $a(X,t)=a_G(u,v)$, depends solely on the state of the ROP1 system. This assumption has the effect that the volume elements of the pollen tube move apart from each other uniformly along the length of the pollen tube. Suppose that the pollen tube doubles in length over a period of time t then uniformly spaced volume elements (red disks) will double the distance between themselves over the same period of time (blue disks). While this uniform growth assumption reduces the specificity of our model it does add generality to our model and our predictions.

3.2 ROP1 Localisation as Desired Output

In this section we subject our ROP1 model to growth and observe how its properties of perturbability and polarisability change under these conditions.

Specifically, here we take the point of view that growth is a process that is independent of the ROP1 system. In this simplified context, we refer to the bifurcation

diagram that we constructed for the non-growing case to gain intuition of what we expect to observe, Figure 3.17.

Assuming that the active profile is solely the output of the system, Figure 3.17 shows that an imbalance between growth and production can cause the ROP1 system to shift away from the perturbable regime: (i) when production outperforms growth the system moves towards an unperturbable state with the active form uniformly high, while (ii) when growth outperforms production the system moves towards an unperturbable state with the active form uniformly low.
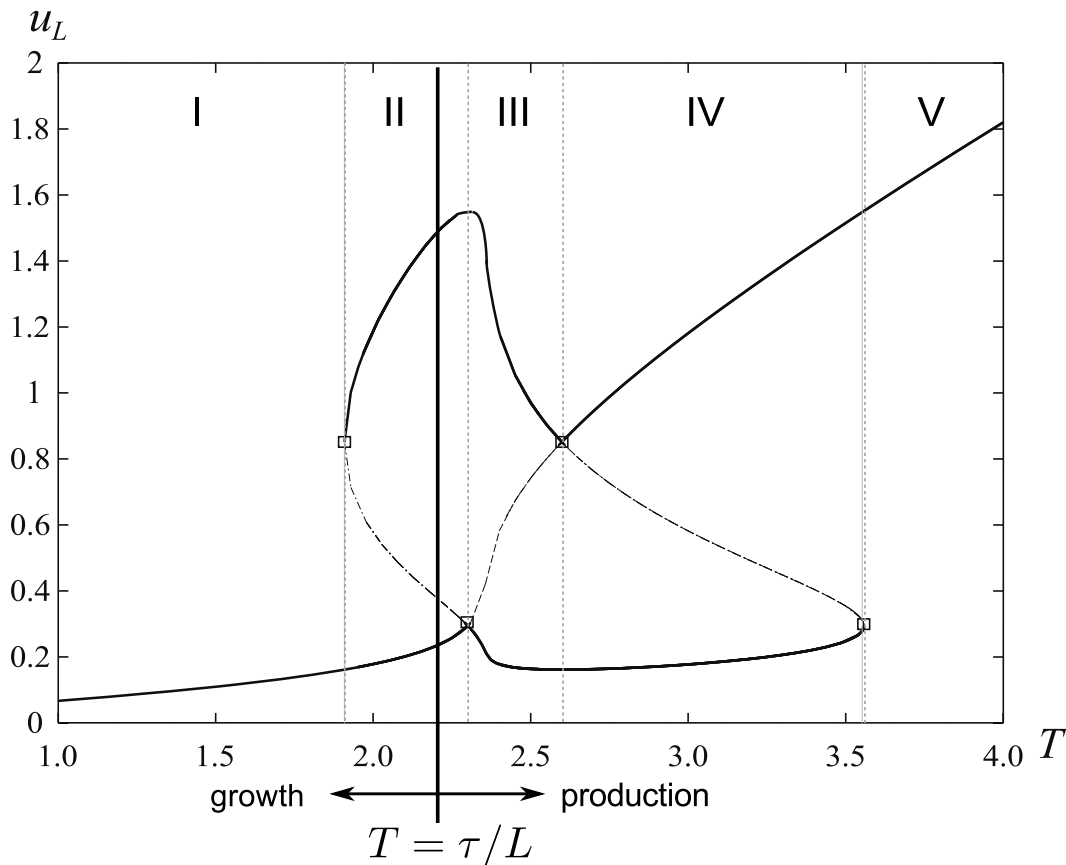
Figure 3.17: **Antagonistic Effect of Pollen Tube Growth and Protein Production on Perturbability.** The total concentration T is a ratio of total protein amounts τ and pollen tube length L, Equation (3.7). As we discussed in Figure 3.13, the total concentration determines if our system, Equation (3.22), is perturbable. As growth increases L and protein production increases τ, these two biological processes antagonise each other and pull T in opposite directions.

3.2.1   The Effect of Growth on ROP1 Localisation

To illustrate the effect of longitudinal growth on ROP1 localisation in the pollen tube, Figure 3.17, we consider Equations (3.22) with constant growth:

$$a(X, t) = a_0, \qquad (3.27)$$

where we substitute $a_G = a_0 > 0$ into Equation (3.26).

We use our standard parameter values, Table 3.1, and apply a stronger stimulus of $k_{s,0} = 0.01 \text{ s}^{-1}$. As our previous results in Figure 3.12 show, the resultant initial total concentration of $T_0 = \tau / L_0$ lies in zone II of Figure 3.17 so that a sufficiently high perturbation in the active form will trigger a travelling interface.

Provided that the constant growth velocity, which we subject our pollen tube to is sufficiently slow, our expectation for this numerical experiment is that our system spends enough time in zone II so that we can trigger a travelling interface in active ROP1. We further expect that as the pollen tube grows, the domain where active ROP1 is localised shrinks in size and vanishes after a sufficient amount of time so that ROP1 displays a low, uniform concentration profile. We draw these expectations from Figure 3.17 that predicts traversal from zone II to zone I for a ROP1 system subjected to growth, and our sharp front analysis in Section 3.1.9 that predicts a reduction in the size of the domain where ROP1 is localised for reduced total concentrations T.

Our expectations are confirmed by our numerical results, Figure 3.18: Our system spends enough time in zone II where our choice of stimulus, Table 3.1, triggers a travelling interface between localised active ROP1 on the right and dilute active ROP1 on the left. This interface retracts and disappears after a sufficient amount of time and gives way to a low, uniform concentration profile of ROP1. We also indicate the position of the interface as predicted by our sharp front approximation described in Section 3.1.9, Figure 3.18 (top, black line). Our predicted interface position is a sensible fit until boundary conditions become dominant towards the end of the simulation.

This result shows that if the growth velocity of our pollen tube is independent from the ROP1 localisation system, then growth will inadvertently prevent the ROP1 system from sustaining apical localisation.

Further, our result in Figure 3.18 highlights an implicit assumption in our reaction-diffusion model of ROP1, Equation (3.4): The inactive form of ROP1 is assumed to

recycle fast between the membrane and the cytoplasm and we therefore do not observe a shift in the domain where active ROP1 is localised but rather gradual dilution of apical ROP1, Figure 3.19.
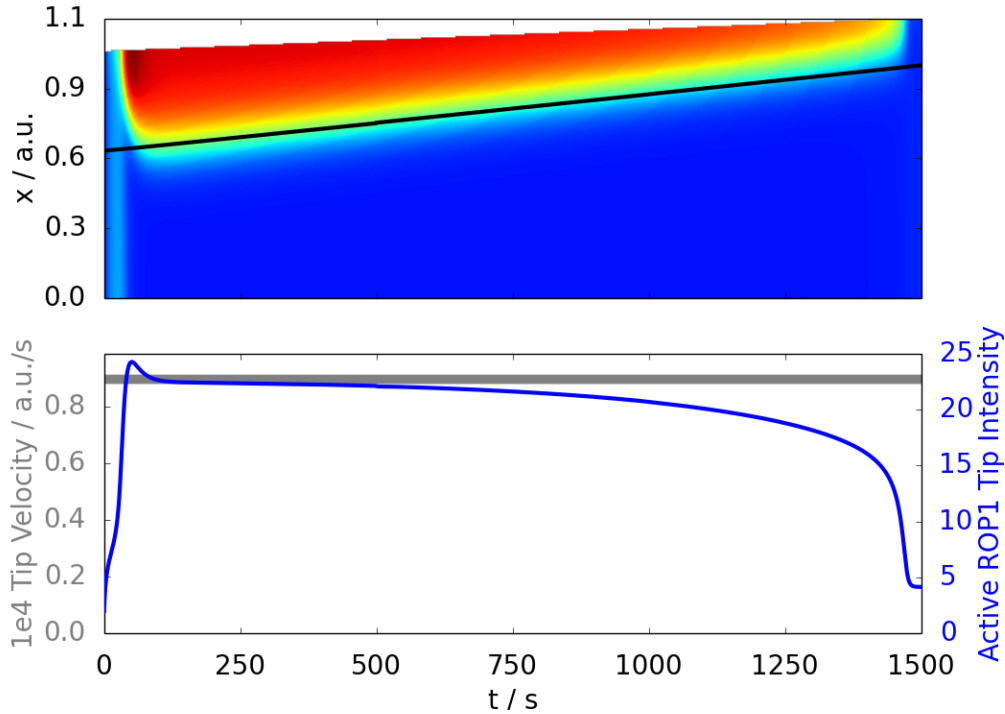
Figure 3.18: **Constant Growth Drives System out of Perturbable Regime.** Simulation of Equations (3.22) with constant growth velocity, Equation (3.27), with a_0=0.00009 a.u./s, initial conditions as in Figure 3.12, $k_{s,0}$=0.01 s^{-1}, and all remaining parameters as in Table 3.1. As predicted by our LPA, Figure 3.17, growth drives our system out of zone II and into zone I where the active form of ROP1 shows inadvertently a low, uniform concentration profile. The magnitude of our choice of a_0 is sufficiently low so that our system spends enough time in zone II where it is stimulated to form a travelling interface between highly localised active ROP1 and dilute active ROP1. The black line indicates the position of the interface between high and low active ROP1 as predicted by our sharp front approximation, Section 3.1.9.

3.2.2 The Effect of Production on ROP1 Localisation

The antagonist of growth in Figure 3.17 is protein production: If our system starts out in zone II and we subject our pollen tube to conditions under which it synthesises protein and does not grow, we expect that our system gradually traverses into zone V where we expect to observe a uniformly high profile of active ROP1.

To illustrate this effect, we amend Equations (3.25) with a zero-order production term that adds inactive ROP1 to our system:

94
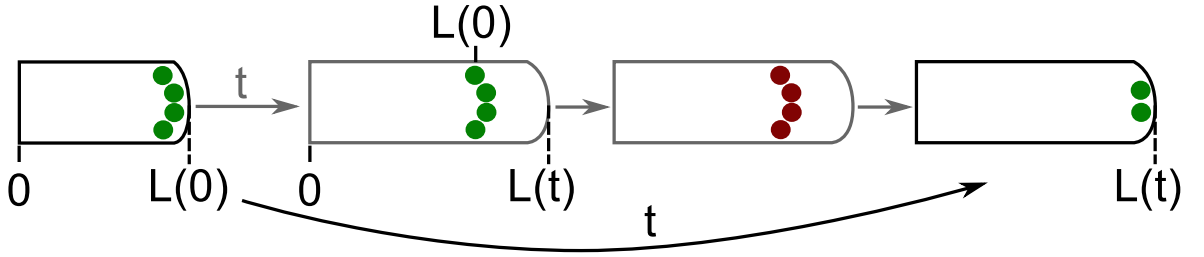
Figure 3.19: **Comparison of Growth Effects Expected for Different ROP1 Recycling Timescales.** Suppose that active ROP1, denoted by green disks, is localised initially to the tip of the pollen tube. (grey arrows) Then we may expect that upon growth, active ROP1 entities remain near the initial apex, $L(0)$, of the pollen tube where they are inactivated and recycled to the cytoplasm (red disks). Since the pollen tube has grown the total concentration of ROP1 is reduced and therefore reactive pressure of cytosolic inactive ROP1 back onto the apical membrane is lower, resulting in a smaller domain size where active ROP1 is localised. (black arrow) Our numerical experiment, Figure 3.18, does not show a shift in localised active ROP1 but rather direct gradual dilution of active ROP1 at the apex. Therefore, our model of the ROP1 system, Equation (3.4), encodes implicitly fast recycling of inactive ROP1.

$$
\frac{\partial u}{\partial t} + u\frac{G_{Xt}}{G_X} + \frac{G_t}{G_X}\frac{\partial u}{\partial X} = D_u\left(\frac{1}{G_X^2}\frac{\partial^2 u}{\partial X^2} - \frac{G_{XX}}{G_X^3}\frac{\partial u}{\partial X}\right) + f(u,v), \tag{3.28}
$$

$$
\frac{\partial v}{\partial t} + v\frac{G_{Xt}}{G_X} + \frac{G_t}{G_X}\frac{\partial v}{\partial X} = D_v\left(\frac{1}{G_X^2}\frac{\partial^2 v}{\partial X^2} - \frac{G_{XX}}{G_X^3}\frac{\partial v}{\partial X}\right) - f(u,v) + \kappa,
$$

$$
\frac{\partial u}{\partial X} = 0, \ (X = 0,1),
$$

$$
\frac{\partial v}{\partial X} = 0, \ (X = 0,1),
$$

$$
\frac{\partial G}{\partial t} = a(X,t),
$$

$$
G(0,t) = 0,
$$

where $\kappa > 0$ is a zero-order rate constant. For our numerical experiment we further choose zero growth velocity

$$
a(X,t) = 0, \tag{3.29}
$$

to model a stationary, non-growing pollen tube.
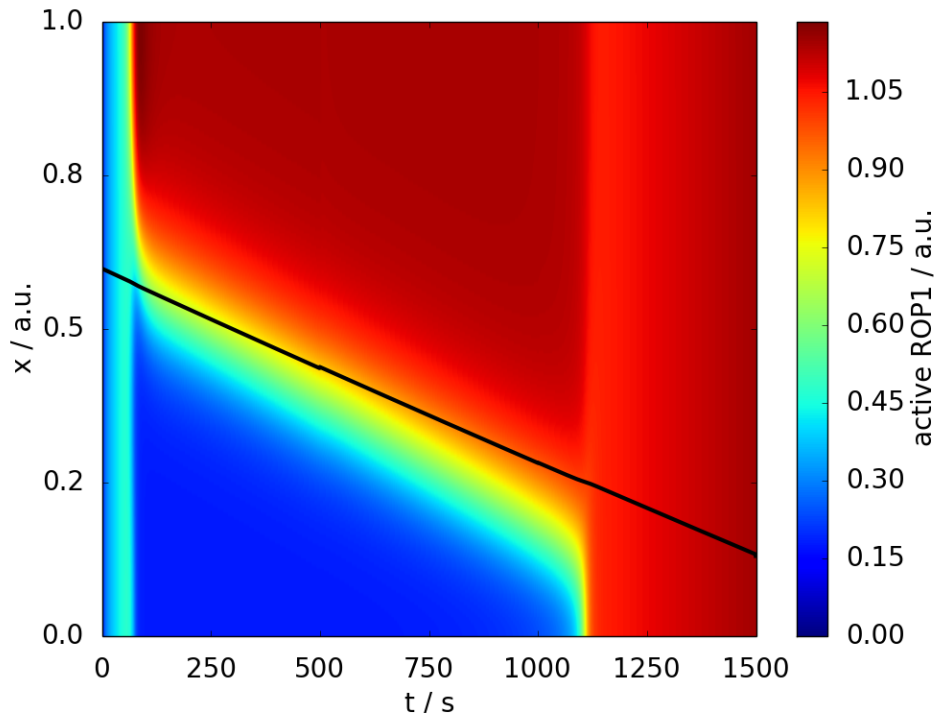
Figure 3.20: **Constant Protein Production by Itself Drives System out of Perturbable Regime.** Simulation of Equations (3.28) with constant, zero-order protein production designated by κ. Our choice of κ=0.0004 a.u./s is sufficiently low so that our system spends enough time in a perturbable regime where our choice of stimulus, Table 3.1, can trigger a moving interface between localised active ROP1 and dilute ROP1. This established interface moves towards the rear of the pollen tube due to production κ before our system steps outside the perturbable regime and displays a uniformly high ROP1 profile.

Our numerical experiment, Figure 3.20, confirms our expectation: For a sufficiently small production rate constant κ our system remains in a perturbable regime for long enough to get stimulated into forming a moving interface between highly localised active ROP1 and dilute active ROP1. This interface moves towards the rear of the pollen tube as production rate κ adds more protein to the system. This effect is also illustrated by our sharp front approximation of the interface position, Figure 3.20 (black line). Eventually our system steps outside the perturbable regime, Figure 3.17, and displays a high-concentration, uniform profile of active ROP1.

This result indicates that synthesis rates of ROP1 that are entirely unchecked will drive our ROP1 system out of the perturbable regime and annihilate tip polarity.

3.2.3 The Combined Effect of Production & Growth on ROP1 Localisation

Our results in Sections 3.2.1 and 3.2.3 confirm that sustained imbalance between growth and protein production will inadvertently move our system out of the perturbable regime – nullifying the possibility for localisation of active ROP1 at the tip.

However, generally we observe apical localisation of active ROP1 in growing pollen tubes, and so it is reasonable to postulate that there is some control mechanism that keeps the total concentration of ROP1 within the perturbable regime, Figure 3.17.

In this numerical experiment, we test the hypothesis that this control mechanism keeps the total concentration constant. The time change of the total concentration T is given by the following expression:

$$\dot{T} \;\; = \;\; \frac{\dot{\tau}}{L} - \frac{\tau}{L}\frac{\dot{L}}{L}, \tag{3.30}$$

where due to our boundary conditions, Equation (3.6), the only source of mass change is κ in Equation (3.28) hence $\dot{\tau}$ is given by the following expression

$$\dot{\tau} = L\kappa. \tag{3.31}$$

Note that if L is a function of time $L \equiv L(t)$ then $\dot{\tau}$ is also a function of time. All other components of Equation (3.30) are either given by the state of our system, τ, L, or by our choice of growth velocity a_G.

Suppose that T lies in the perturbable regime initially, Figure 3.17, and that our control mechanism enforces zero change in T. From Equation (3.30), $\dot{T}=0$ under the following condition:

$$\kappa = \frac{\tau}{L}\frac{\dot{L}}{L}. \tag{3.32}$$

This condition tells us that a hypothetical control mechanism needs to regulate the rate of protein production dynamically and it needs to know the current state of the system, τ and L, and how fast it grows, $\dot{L}$. The results of our numerical experiment

are shown in Figure 3.21.
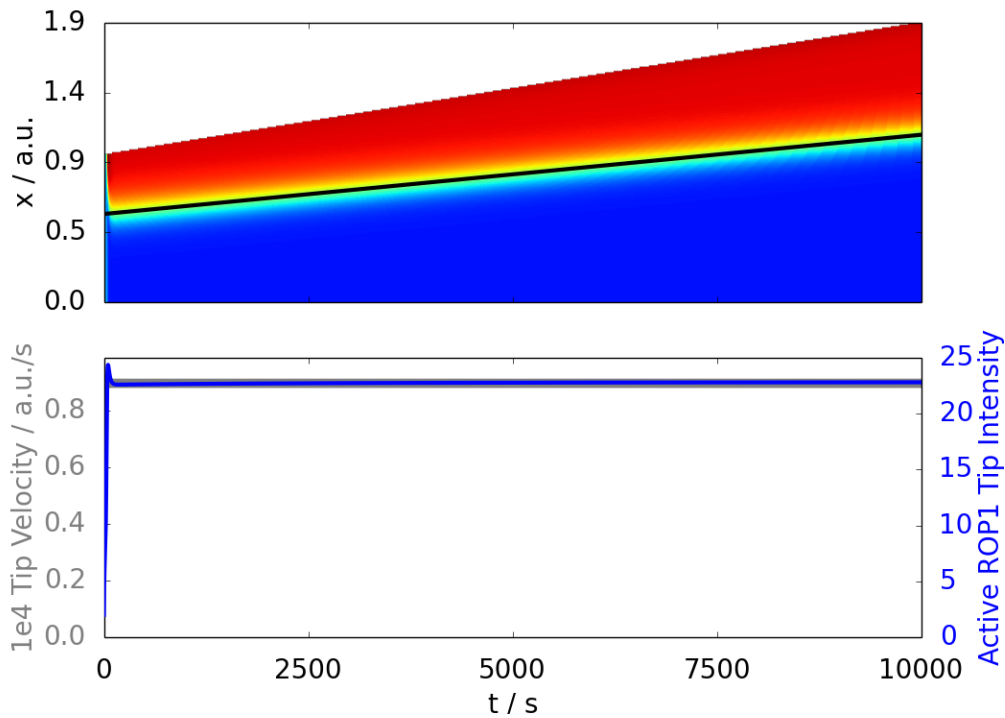
Figure 3.21: **A Presumed Mechanism That Locks the Total Concentration Into the Perturbable Regime Sustains ROP1 Localisation.** Simulation of Equations (3.28) with constant total amounts T and other simulation parameter values as in Table 3.1. The constant growth velocity $a_G=0.00009$ a.u./s and stimulus strength $k_{s,0}=0.01$ s^{-1}. The black line indicates the predicted interface position as in the other results figures.

Our results in Figure 3.21 indicate that keeping the total concentration fixed and in the perturbable regime sustains localisation of ROP1.

We point out that $\dot{T}=0$ if and only if κ fulfills strictly the condition in Equation (3.32). Any deviation, for instance through noise, of κ away from this condition will result in a changing total concentration, $\dot{T}\neq0$, and therefore eventual breakdown of ROP1 localisation.

As a side note regarding our results in Figure 3.21: The width of the high-concentration ROP1 plateau increases over time as is shown both by our full numerical simulation and our sharp front approximation, Figure 3.21 (top, black line). To understand this, we recall our sharp front approximation, Equation (3.12) and use the right-polarised expressions in Equation (3.14). Suppose now that our sharp front approximation is in a stable equilibrium with $u_L=u_L^*$, $u_R=u_R^*$, and $w=w^*$. Further, in our numerical experiment $T=$const. We now rearrange the expression for the inactive form v, Equation (3.14) to obtain

$$w^* = \frac{T - v + Lu_L^*}{u_R^* + u_L^*},$$

which tells us that, everything else kept constant, increasing L through growth increases the width w of the high-concentration plateau.

3.3 ROP1 Localisation as Part of a Feedback System

Our results in Section 3.2 demonstrate that the total concentration of ROP1 is key in apical ROP1 localisation. We have also reviewed literature in Section 3.1.3 that establishes a causal link between apical ROP1 localisation and pollen tube growth. Here we focus on closing the feedback loop between ROP1 localisation, pollen tube growth, and the total concentration of ROP1.

Existing models predict sustained apical ROP1 localisation, Section 3.1.6, and also we describe a control mechanism that retains apical ROP1 localisation, Section 3.2.3. In these existing models, localised active ROP1 is viewed as a desirable outcome and the suggested mechanisms depend on "knowing" a number of intricate properties of the overall system in order to retain ROP1 localisation. As an example, the control mechanism we suggest in Section 3.2.3 needs to be capable of regulating the rate of protein production κ while being fully aware of current protein amounts τ, the length of the pollen tube L, and its growth velocity $\dot{L}$.

These control mechanisms have two aspects in common: Firstly, they view ROP1 localisation as the outcome of successful communication between some control mechanism and growth. Secondly, they are quite complex biologically since a thorough understanding of the current system state needs to be encoded with biological components.

Here we propose a different control mechanism that includes active ROP1 localisation as a cornerstone of control and that reduces the amount of information required to exert this control.

3.3.1 Our Model of Feedback Provided by ROP1 Localisation

To add the concentration profile of active ROP1 to a control mechanism we need to decide what representation of this concentration profile is biologically meaningful.

The simplest and likely most biological model would be to set the local growth velocity proportional to the concentration profile of the active form

$$\mathbf{a}(\mathbf{X}, t) \propto \mathbf{u}(\mathbf{X}, t).$$

This model is sensible biologically as it reproduces the functional role of increasing local growth that we ascribe to localised ROP1, Section 3.1.3. Thus in a two- or three-dimensional model of the pollen tube, this relationship would reproduce tip growth and ballooning naturally through the vectorial growth velocity $\mathbf{a}(\mathbf{X}, t)$. However, for our one-dimensional pollen tube model where the growth velocity is a scalar, representing longitudinal growth, Equation (3.15), this relationship would make unreasonable predictions such as fast longitudinal growth for rear-localised active ROP1.

To gain intuition about the velocity of longitudinal pollen tube growth we look to experimental results reported by Hwang et al. (2010), Figure 3.22. Hwang et al. (2010) bombarded a growing pollen tube with a ROP1 construct for overexpression and observed the pollen tube over a period of 2.5 hours. They observed that greater apical spread of active ROP1 (greater reported cap size) correlated with a slower longitudinal pollen tube growth velocity, Figure 3.22. Further, Hwang et al. (2010) reported that small and medium ROP1 cap sizes can be observed to oscillate together with pollen tube growth under normal conditions, Figure 3.32. The cap sizes reported in Figure 3.32 correspond to those reported for early time points 0:0 and 12:40 in Figure 3.22.

Since the bombarded pollen tube was growing prior to bombardment, Figure 3.22, we assume that the ROP1 cap had a medium size and that the growth velocity was relatively fast prior to bombardment. We assume that the growth velocity was fast since a naturally growing pollen tube would want to reach the ovules as quickly as possible. We therefore assume that the slow growth velocities reported in Figure 3.32 are due to a small ROP1 cap size. Taken together, we assume that

the data reported by Hwang et al. (2010) points to the following relation between the ROP1 cap size and longitudinal growth velocity: A small ROP1 cap size causes slow longitudinal growth, a medium cap size fast growth, and a large cap size slow growth.

Based on this relation we propose to use the centre of mass of active ROP1 as an input into the control mechanism of our one-dimensional pollen tube. The centre of mass of $u(X,t)$ is defined as follows:

$$\widetilde{\varphi}(u(X,t)) = \int_0^{L_0} s u(s,t) \, \mathrm{d}s, \tag{3.33}$$

where the upper bound L_0 is the initial length of the pollen tube as we work with the active form $u(X,t)$ in material coordinates X. Our above expression for the centre of mass, Equation (3.33), introduces the following biases: (i) $\widetilde{\varphi}$ increases with increasing overall levels of $u(X,t)$, and (ii) a uniform concentration profile of $u(X,t)$ results in a centre of mass $\widetilde{\varphi}$=0.5 since our initial domain has length L_0=1 a.u., Table 3.1. To remove these biases from our measure, we define the following modified centre of mass

$$\varphi(u(X,t)) = 2 \left(\frac{\widetilde{\varphi}(u)}{L_0 T_u} - \frac{1}{2} \right), \tag{3.34}$$

where $T_u = \int_0^{L_0} u(s,t) \, \mathrm{d}s$ is the total concentration of the active form $u(X,t)$ on the membrane. We point out that the centre of mass φ in Equation (3.34) is given in material coordinates X and can therefore be thought of as a relative position along the length of the pollen tube.

To gain intuition about the centre of mass we plot four profiles and their $\varphi(\cdot)$ values in Figure 3.23.

We substitute φ into a_G of Equation (3.26) and obtain the following ROP1 dependent expression for our growth velocity:

$$a(X,t) = k_a \varphi(u(X,t)), \tag{3.35}$$

where $k_a \geq 0$ is a proportionality constant.

We plot the behaviour of the longitudinal growth velocity, Equation (3.35), as a

function of simplified concentration profiles, Figure 3.23, where we equate the high ROP1 plateau of these profiles to the ROP1 cap size, Figure 3.24.

The observed behaviour in Figure 3.24 shows that our model of the relation between longitudinal growth velocity and cap size, Equation (3.35), reproduces the relation suggested by experimental data (Hwang et al., 2010): Small and large ROP1 cap sizes result in slow longitudinal growth while a medium cap size induces a maximal longitudinal growth velocity.
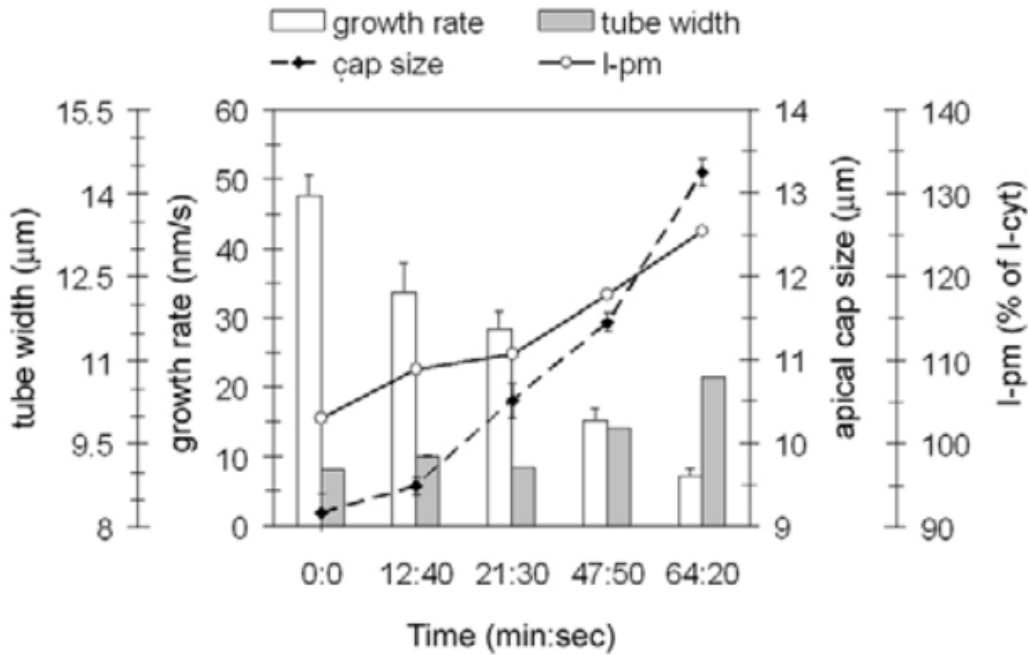
Figure 3.22: **Observed Relation Between Pollen Tube Growth Rate and ROP1 Localisation.** Hwang et al. (2010) bombarded pollen tubes with small amounts of ROP1 construct for overexpression and observed their growth over a period of 2.5 hours. Their results show that a wider spread of active ROP1 (greater cap size) correlates with a reduced longitudinal growth rate. Reproduced from Hwang et al. (2010).

3.3.2   Expectations for Numerical Experiments

With our model for ROP1-driven local growth velocity, Equation (3.35), we can first derive some intuition before discussing the results of our numerical experiments.

Suppose that initially active ROP1 is homogeneous and the pollen tube is not growing with velocity $a(X, 0) = 0$ everywhere. The approximate acceleration of the pollen tube given by our model, Equation (3.35), is

$$\frac{a(X, t + \Delta t) - a(X, t)}{\Delta t} = \frac{k_a}{\Delta t} \left(\varphi(u(X, t + \Delta t)) - \varphi(u(X, t)) \right). \tag{3.36}$$
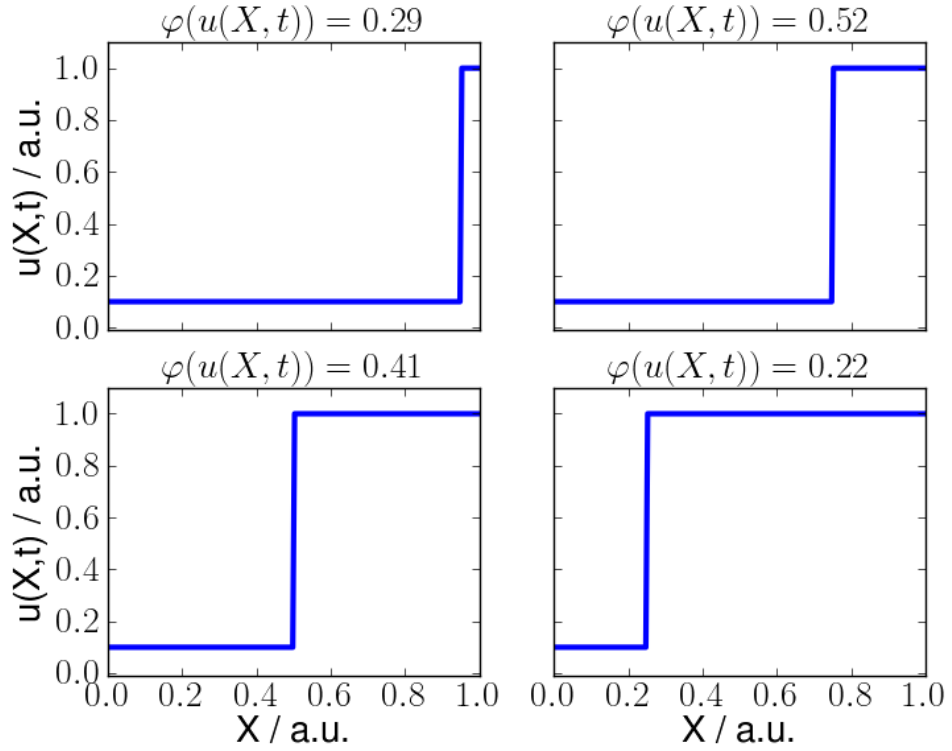
Figure 3.23: **Visualisation of φ for Sample Concentration Profiles.** Our choice of φ, Equation (3.34), produces the following measures of apical ROP1 localisation: Narrow (top left) and wide (bottom right) tip-localised active ROP1 result in low φ values while a medium width of active ROP1 (top right) results in a high value. This behaviour of φ is sensible in light of our one-dimensional approximation of the pollen tube and our focus on longitudinal growth, Figure 3.22.

Starting with uniform active ROP1, $\varphi(\cdot){=}0$, and applying a sufficiently strong stimulus $k_{s,0}$ to localise active ROP1 we expect that $\varphi(u(X,t+\Delta t)){>}\varphi(u(X,t))$, $\Delta t{>}0$. Therefore, since the right-hand side of Equation (3.36) is positive the left-hand side is also positive – hence we expect that inducing tip-localisation of active ROP1 accelerates the pollen tube.

Suppose now that active ROP1 is tip-localised in a growing pollen tube. Then due to growth, Equation (3.35), ROP1 is diluted and using results from our sharp front approximation, Section 3.1.9, we expect that the interface width of localised active ROP1 narrows. If the interface of active ROP1 starts out as in Figure 3.23 (top right) then retraction of the interface towards the pollen tube tip, Figure 3.23 (top left), causes deceleration as given by Equation (3.36). If the interface of active ROP1 starts out as in Figure 3.23 (bottom right) then an intermediate phase of acceleration will be followed by deceleration. Overall, we expect that growth driven by apical localisation of active ROP1 causes retraction of the interface towards the tip and
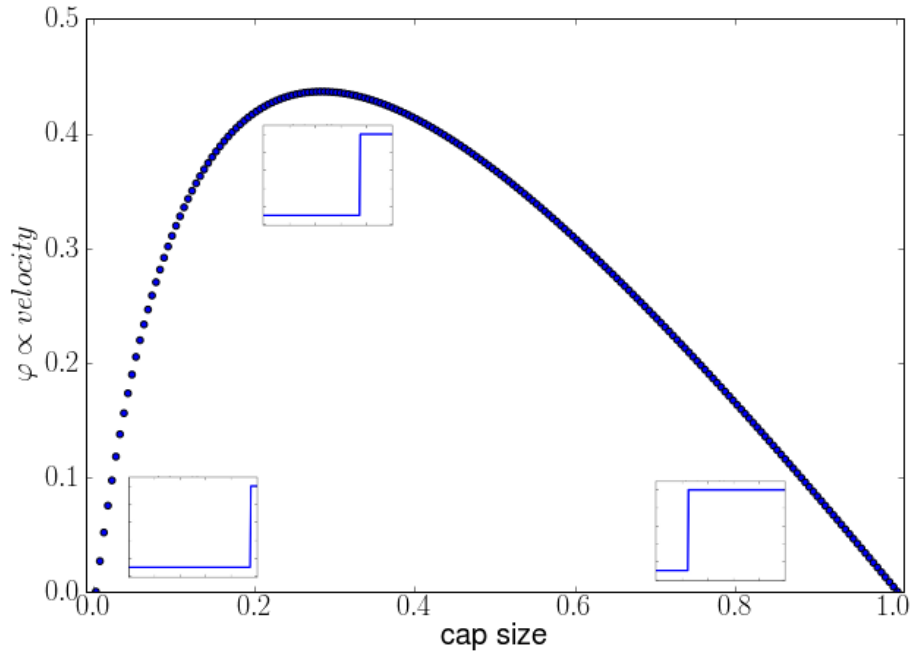
Figure 3.24: **Relation Between Pollen Tube Growth Rate and ROP1 Localisation Encoded by Our Model.** We plot φ (which is proportional to growth velocity) against the width of the high ROP1 plateau of concentration profiles equivalent to those shown in Figure 3.23. We assume that our width of the high concentration profile is equivalent to the cap size reported by Hwang et al. (2010). The behaviour of φ reproduces the relation between ROP1 localisation and growth that we derive from data reported by Hwang et al. (2010): A small ROP1 cap size causes slow longitudinal growth, a medium cap size fast growth, and a large cap size slow growth.

therefore deceleration.

Due to our choice of ROP1-dependent growth velocity, Equation (3.35), growth and ROP1 dilution will only stop once the ROP1 interface has collapsed.

Once the interface in active ROP1 has collapsed, localisation of active ROP1 – and therefore acceleration – can only be triggered anew if the total ROP1 concentration lies in the perturbable regime, Figure 3.17. At this point, protein production would push ROP1 concentrations back into the perturbable regime so that another round of acceleration and deceleration can be triggered. We therefore expect that a combination of ROP1-driven growth and constitutive protein production will generate oscillatory growth.

From these considerations we understand that constant growth of the pollen tube may be achieved by circumventing deceleration once active ROP1 is tip-localised. To avoid deceleration, the pollen tube would need to avoid dilution of ROP1 and therefore the rate of protein production in our model would need to adapt to the

rate of growth via a feedback loop. As a first approximation, such a feedback loop could prevent any change in the total concentration T of our system thus enforcing $\dot{T}=0$. We have studied this example in Section 3.2.3 where we realised that $\dot{T}=0$ can be achieved by adapting dynamically the rate of protein production κ to the current state of our system, Equation (3.32).

In summary, we expect oscillatory growth when ROP1-driven growth velocity and protein production are entirely uncoupled, and constant growth when we link protein production tightly to the state and growth velocity of our system.

3.3.3   ROP1-Driven Growth

In this numerical experiment, we simulate our system, Equations (3.28), with zero protein production, $\kappa=0$ a.u./s, and ROP1-dependent growth rate, Equation (3.35).

Given appropriate parameter values, we expect that our system spends enough time near its initial state (in the perturbable regime) where the applied stimulus is strong enough to trigger apical localisation of active ROP1. We further expect that the pollen tube only starts growing once active ROP1 is localised apically, and continues growing until the total concentration falls below a critical value and the system moves out of the perturbable regime. Hence we expect that the pollen tube in this numerical experiment accelerates and decelerates once without showing sustained oscillatory growth, Section 3.3.2.

Our results for this numerical experiment, Figure 3.25, confirm our expectation: Unopposed ROP1-driven growth prohibits maintenance of apical ROP1 localisation. Further, our estimate of the interface position reproduces the observed interface well until boundary effects become important, Figure 3.25 (black line).

As an aside: The width of our estimated interface does not reach zero but settles in a narrow value towards the end of the simulation, Figure 3.25 (black line). This is due to the fact that the total concentration does not change once ROP1 localisation is lost.
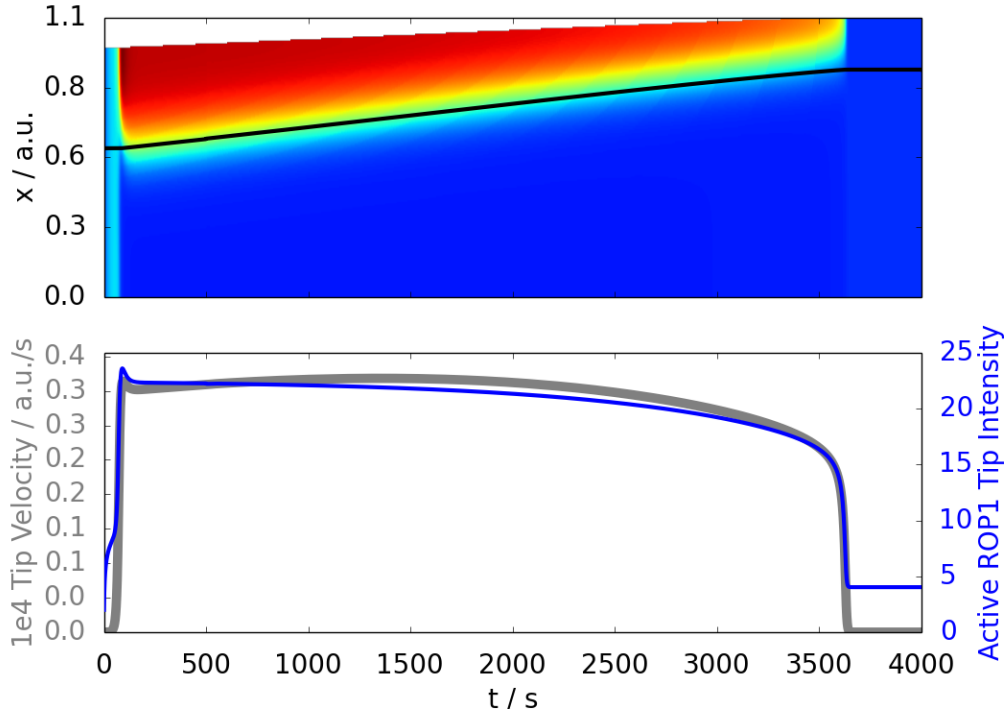
Figure 3.25: **Unopposed ROP1-Driven Growth Destroys ROP1 Localisation.** In this numerical experiment, the growth rate, Equation (3.35), is proportional to the centre of mass of ROP1, $\varphi(u(X,t))$. Therefore, the observed tip velocity (bottom, grey curve) is proportional to $\varphi(u(X,t))$. The extent of the high plateau of active ROP1 (top) retracts continually while the tip velocity is high – this retraction is due to continued reduction in total concentration. The black line traces the approximate position of the interface as predicted by our sharp front approximation, Section 3.1.9. Once the total concentration falls below a critical point, boundary effects draw the remaining high plateau in and the observed interface position deviates from our sharp front approximation. This experiment confirms our previous results, Figure 3.18: Unopposed growth inadvertently lowers the total concentration of ROP1 and prohibits maintenance of apical localisation of ROP1. Parameter values k_a=0.0005 s^{-1}, $k_{s,0}$=0.001 s^{-1} and all remaining parameter values as in Table 3.1.

3.3.4 ROP1-Driven Growth & ROP1 Production

As we note in Section 3.3.2 we expect oscillatory growth when growth velocity and protein production are uncoupled. Here we test this prediction by simulating our system, Equation (3.28), with ROP1-dependent growth velocity and constitutive protein production. We introduce zero-order protein production by setting κ>0 in Equation (3.28).

Our results in Figure 3.26 confirm our expectation laid out in Section 3.3.2: Decoupling ROP1-driven growth from constitutive protein production is sufficient to generate oscillatory growth. The interface position predicted by our sharp front approximation, Figure 3.26, confirms our expectation further: Tip-localised active
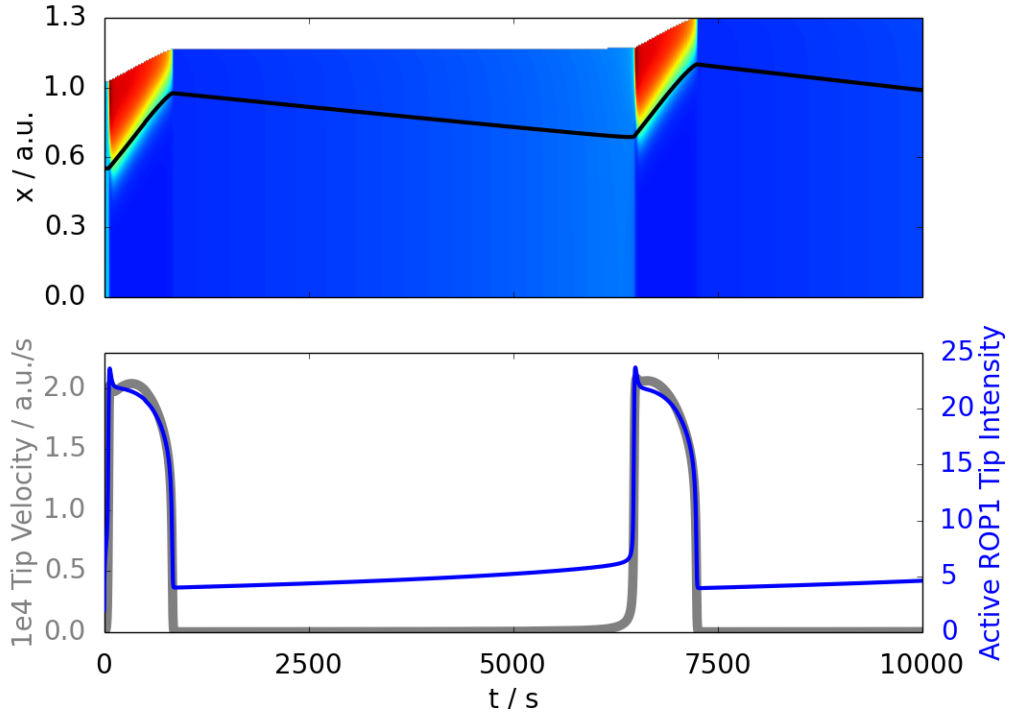
Figure 3.26: **ROP1-Dependent Growth with Zero-Order Protein Production.**
Pollen tube growth with centre-of-mass-dependent growth, Equation (3.35),
and zero-order protein production rate $\kappa=0.00004$ a.u./s. Parameter values
$k_a=0.0005$ s^{-1}, $k_{s,0}=0.001$ s^{-1} and all remaining parameter values as in Table 3.1. Ini-
tially, the pollen tube has zero growth velocity and the rate of protein production is
sufficiently small so that our system remains in the perturbable regime long enough
so that active ROP1 can be stimulated to localise at the tip. Tip-localised active ROP1
induces growth, eventually lowering the total concentration of active ROP1 below
a critical value thus moving the pollen tube from the perturbable regime into zone
I, Figure 3.17. Once active ROP1 is unlocalised, growth ceases and protein produc-
tion moves the system from zone I back into the perturbable regime where a second
cycle of ROP1 localisation and dispersal is triggered. This confirms our expecta-
tion argued in Section 3.3.2: Decoupling ROP1-dependent growth from constitutive
protein production is sufficient for oscillatory growth.

ROP1 triggers growth thus dispersing localised ROP1 by lowering the total concen-

tration – this is confirmed by the narrowing approximate high plateau, Figure 3.26

(black line). Once growth has ceased due to a collapse in the ROP1 interface, consti-

tutive protein production gets a chance to increase the total concentration until the

applied stimulus $k_{s,0}$ is sufficient to trigger ROP1 relocalisation – this is confirmed

by the widening approximate high plateau between the first interface collapse and

second interface rise, Figure 3.26 (black line).

When we run this numerical experiment, Figure 3.26, for longer we observe that

initially oscillatory growth gives way to constant growth, Figure 3.27, and eventual

loss of polarity on long time scales.
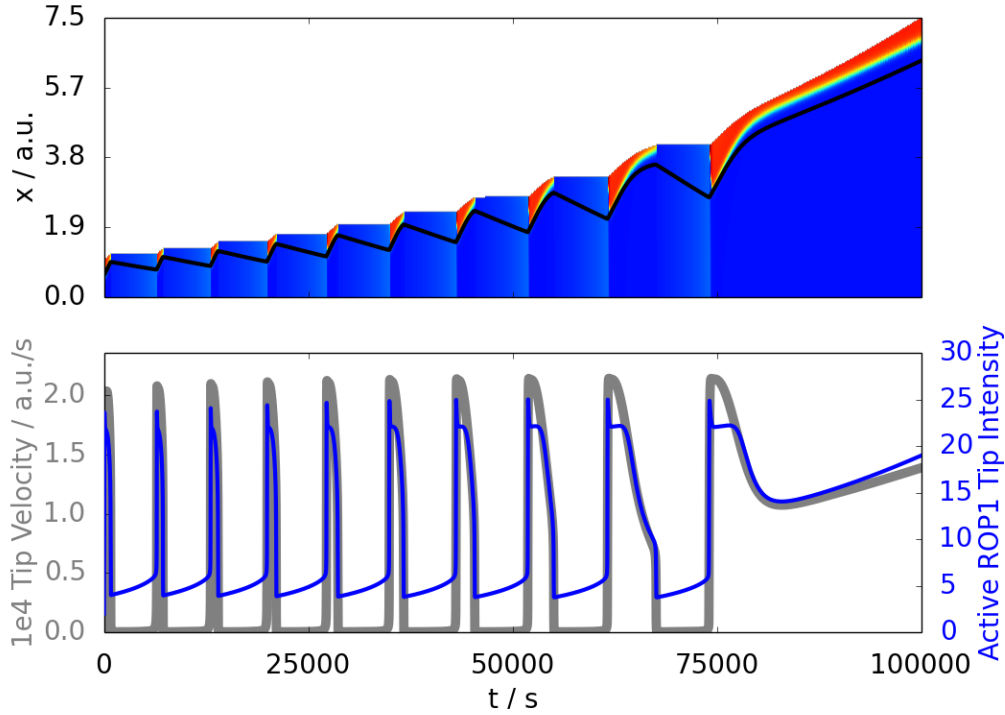
Figure 3.27: **ROP1-Dependent Growth with Zero-Order Protein Production.** Equivalent simulation conditions as in Figure 3.26 but simulated for a longer amount of time. At longer time scales, the oscillatory growth behaviour observed in Figure 3.26 turns into constant growth with stabilised apical ROP1 localisation. This is due to a decrease in peak relative growth velocity as we explain in the main text and Figure 3.29.

To understand this observation we recall our results form Section 3.1.9: The total concentration T of our system is paramount in determining perturbability and maintenance of the interface between high and low active ROP1. We further understand that loss of an established ROP1 interface is due to dilution of ROP1 below a threshold value, Figures 3.25 and 3.26. Hence our result in Figure 3.27 suggests that the total concentration T is not diluted sufficiently to cause interface collapse towards the end of the experiment.

To understand this observation, it is important to realise that while the absolute growth velocity during growth bursts does not change, the relative growth velocity does, Figure 3.28.

The importance of this decline in relative growth velocity, Figure 3.28, becomes apparent when we recall the time change of the total concentration T, Equation (3.30): The term $\frac{\dot{L}}{L}$ in Equation (3.30) is the relative growth velocity, Figure 3.28 (black line), and $\dot{\tau}$ is given by Equation (3.31). We substitute our simulation data shown in Figure 3.27 into the expression for $\dot{T}$ and plot $\dot{T}$ in Figure 3.29.
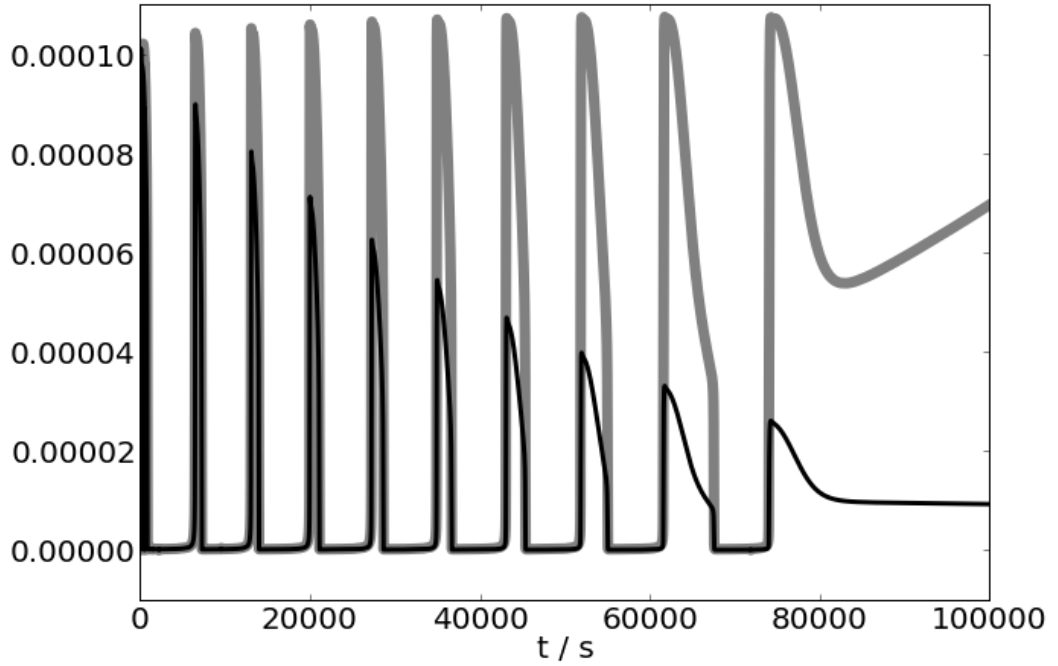
Figure 3.28: **Comparison of Absolute and Relative Growth Velocity.** The relative growth velocity (black), $\dot{L}/L$, decreases due to the increased pollen tube length while the absolute growth velocity (grey), $\dot{L}$, during growth bursts remains constant. Velocities extracted from simulation data shown in Figure 3.27. The units of the relative growth velocity are $[\dot{L}/L]=1/s$, those of the absolute growth velocity are $[\dot{L}]=a.u./s$.

As our plot of $\dot{T}$ in Figure 3.29 demonstrates, suppression of T, denoted by negative $\dot{T}$ values, decreases gradually as the relative growth velocity decreases.

In summary, our model predicts that a pollen tube that shows growth bursts of constant absolute growth velocity will eventually stop oscillating and grow at a constant rate, Figure 3.29. This is due to the fact that a constant absolute peak velocity generates a decreasing relative peak velocity and therefore becomes less capable of diluting the total concentration below a critical value which necessary for oscillations.

3.3.5 ROP1-Driven Growth & Length-Scaled Growth Velocity

We revisit our earlier expression for the growth velocity, Equation (3.35), in light of the following mental model of pollen tube growth: Suppose we divide our one-dimensional pollen tube of length L into ℓ equal intervals of length Δ_x^i, $i=1,\ldots,\ell$. Further, suppose that the growth velocity of each of these intervals is proportional to its length (for instance due to uptake of aqueous solution along its length)
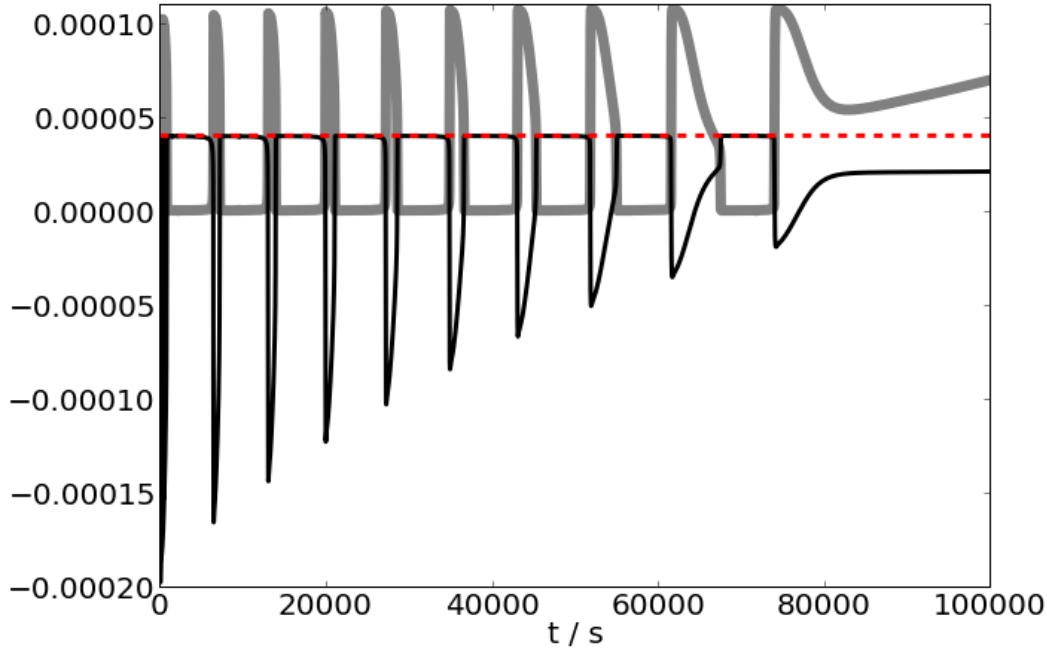
Figure 3.29: **Time Change of Total Concentration.** The black line denotes $\dot{T}$, the grey line is the absolute growth velocity shown in Figure 3.28, the red dashed line denotes κ in Equation (3.28). Growth bursts $\dot{L}>0$, Equation (3.30), suppress T but the rate of suppression decreases gradually together with the magnitude of relative growth, Figure 3.28. After each growth burst, $\dot{T}$ returns to background protein production levels κ to recharge the ROP1 system. Based on simulation data shown in Figure 3.27.

$$\frac{d\Delta_x^i}{dt} \propto \Delta_x^i, \ i=1,\ldots,\ell,$$

then, assuming that all ℓ intervals grow equally fast, and using the relation $L=\sum_{i=1}^{\ell} \Delta_x^i$

we obtain

$$\frac{dL}{dt} \propto L.$$

We modify our previous expression for the longitudinal growth velocity, Equation (3.35), with this proportionality relation and obtain

$$a(X,t) = Lk_a\varphi(u(X,t)), \tag{3.37}$$

where L is the current length in x-coordinates of the pollen tube.

In this numerical experiment we repeat the simulation shown in Figure 3.27 and use our length-scaled velocity term, Equation (3.37). Our results indicate that a length-scaled growth velocity is sufficient to sustain oscillatory growth over an ex-

110

tended period of time, Figure 3.30.
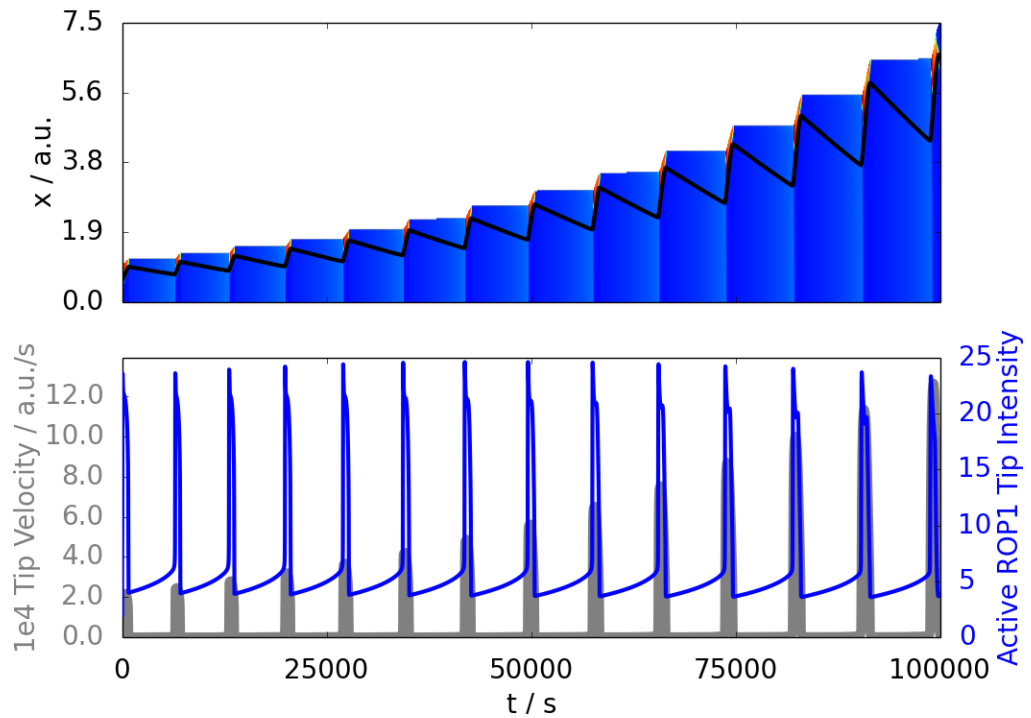
Figure 3.30: **ROP1-Dependent, Length-Scaled Growth.** All simulation parameters are equal to those used in Figure 3.27 except for use of the length-scaled velocity, Equation (3.37). Dilution due to growth remains sufficient to suppress the total concentration below a critical threshold and trigger oscillatory growth repeatedly. The observed absolute growth velocity increases over time indicating that a fast pollen tube has greater propensity to oscillate than a pollen tube with a constant, slow peak velocity, Figure 3.27.

Our model predicts that pollen tubes that scale their peak velocity with their length are capable of sustaining oscillatory growth for longer, Figure 3.30.

Therefore, our model predicts that oscillatory growth can be viewed as the default mode of pollen tube growth: Pollen tubes that are capable of scaling their peak growth velocity would therefore be more likely to show oscillatory growth, Figure 3.30, than those pollen tubes that continue growing at a relatively constant, slow peak velocity, Figure 3.27. This prediction falls in line with reported observations that fast pollen tubes show oscillations whereas slow ones grow at constant velocities (Rojas et al., 2011) and (JA Feijo, personal communication).

3.3.6 ROP1-Driven, Non-Oscillatory Growth Requires Constant Total Concentrations

In Sections 3.3.4 and 3.3.5 we show that our model of ROP1-driven growth predicts oscillatory growth with repeated dilution ($\dot{T}<0$) and recharge ($\dot{T}>0$) of the total concentration. Live pollen tubes are observed to show both oscillatory and non-oscillatory growth, Figure 3.5, and we are therefore interested in conditions under which our model predicts constant growth.

To this end we look at the time change of the growth velocity of our pollen tube, Equation (3.35)

$$\frac{\partial a}{\partial t} = k_a \frac{\partial \varphi}{\partial t} = k_a \frac{\partial \varphi}{\partial u} \frac{\partial u}{\partial t}, \tag{3.38}$$

where the second equality is due to the chain rule of differentiation. Therefore, the time change of the growth velocity equals zero under either of two conditions: (i) $\frac{\partial \varphi}{\partial u}=0$ which implies a ROP1-independent expression for φ, or (ii) $\frac{\partial u}{\partial t}=0$ which holds for steady states of the spatial ROP1 system.

The former condition holds for a choice of φ that does not vary continuously with the concentration profile of the active form of ROP1 and is therefore not interesting. The latter condition holds in steady states of the spatial ROP1 system which can only be reached when all parameters of the system, including the total concentration T, are kept constant. For appropriate parameter values, including T, we can therefore reinterpret the latter condition as $\dot{T}=0$ which is the same condition we established in Section 3.2.3 for sustained ROP1 localisation despite constant pollen tube growth.

In this numerical experiment we implement this condition, $\dot{T}=0$, together with ROP1-driven growth, Equation (3.35), to test our prediction.
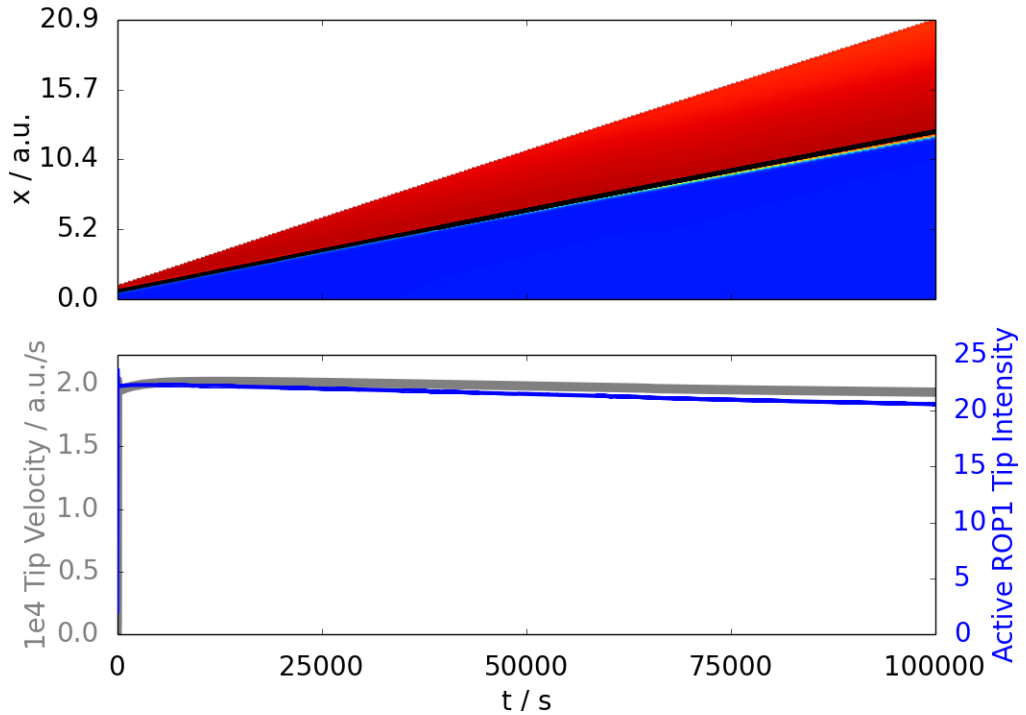
Figure 3.31: **ROP1-Dependent Growth With Tightly Coupled Protein Production.**
We use parameter values equivalent to those used in Figure 3.26 and enforce constant total concentrations by choosing κ as in Equation (3.32). Linking protein production tightly to ROP1-driven growth avoids dilution of ROP1 and maintains constant growth.

Our results in Figure 3.31 confirm our prediction above and show that our model produces non-oscillatory ROP1-driven growth for tightly matched growth and protein production. We point out that the sustained localisation of ROP1, and therefore non-oscillatory growth, in Figure 3.31 is subject to the same restrictive stability properties discussed in Section 3.2.3.

3.3.7 Delayed Feedback Between ROP1 Localisation and Growth Velocity

The oscillations that our model produces are such that apical ROP1 activity and growth rate oscillate in-phase, Figures 3.27 and 3.30. Recently, Hwang et al. (2010) have shown that a peak in apical ROP1 activity is followed by a spike in tip growth velocity, Figure 3.32.

Similar to Yan et al. (2009) we use the observed time correlation between ROP1 activity and growth rate as a first approximation of their causal link and introduce delay in our local growth velocity:

$$a(X,t) = k_a \varphi(u(X, t - \nu)), \tag{3.39}$$

where we have modified Equation (3.35) to include delay $\nu \geq 0$ and for $t<0$ we choose $u(X,t)$ such that $\varphi(u(X,t))=0$ for $t<0$.

The results of our numerical experiment, Figure 3.33, show the expected delay between a peak in apical ROP1 activity and subsequent growth rate spike. It is important to appreciate that the delay introduced by Yan et al. (2009) into their model served to introduce oscillatory growth, Section 3.1.6, whereas our model produces oscillatory growth through a mechanism that is delay-independent.

In future efforts our framework may be extended to include more refined models of molecular interactions which may be a valuable tool in disentangling the time correlations that are observed experimentally, Figure 3.32. As a first step we propose to include the immediate ROP1 effectors RIC3 and RIC4 which, due to their scaffolding activity, have been proposed to cause some delay between active ROP1 and biomechanical changes that lead to growth (Gu et al., 2005; Yan et al., 2009).
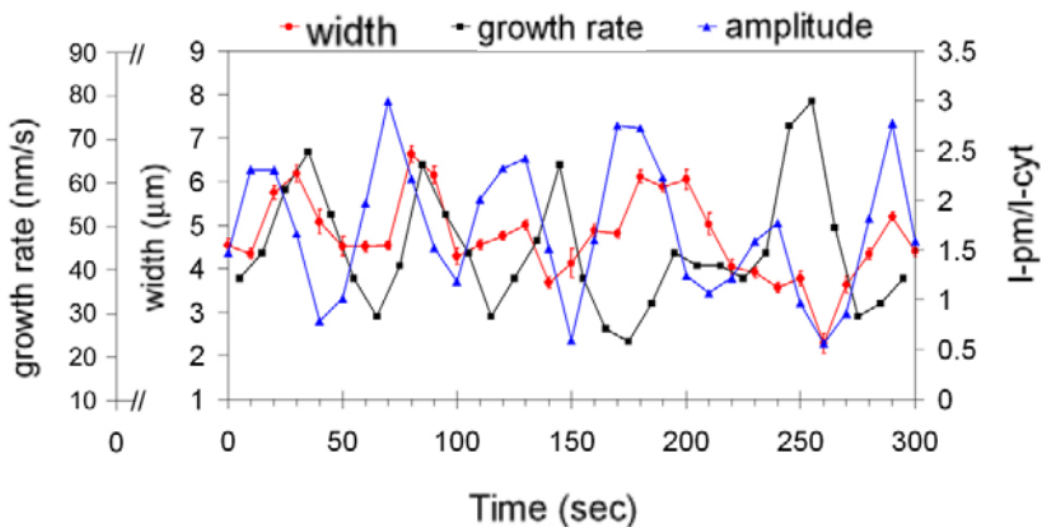
Figure 3.32: **Growth and ROP1 Activity Oscillate Out of Phase.** Experimental evidence suggests that growth velocity and apical activity of ROP1 oscillate out of phase with ROP1 activity likely preceding pollen tube growth (Hwang et al., 2010). Reproduced from Hwang et al. (2010).
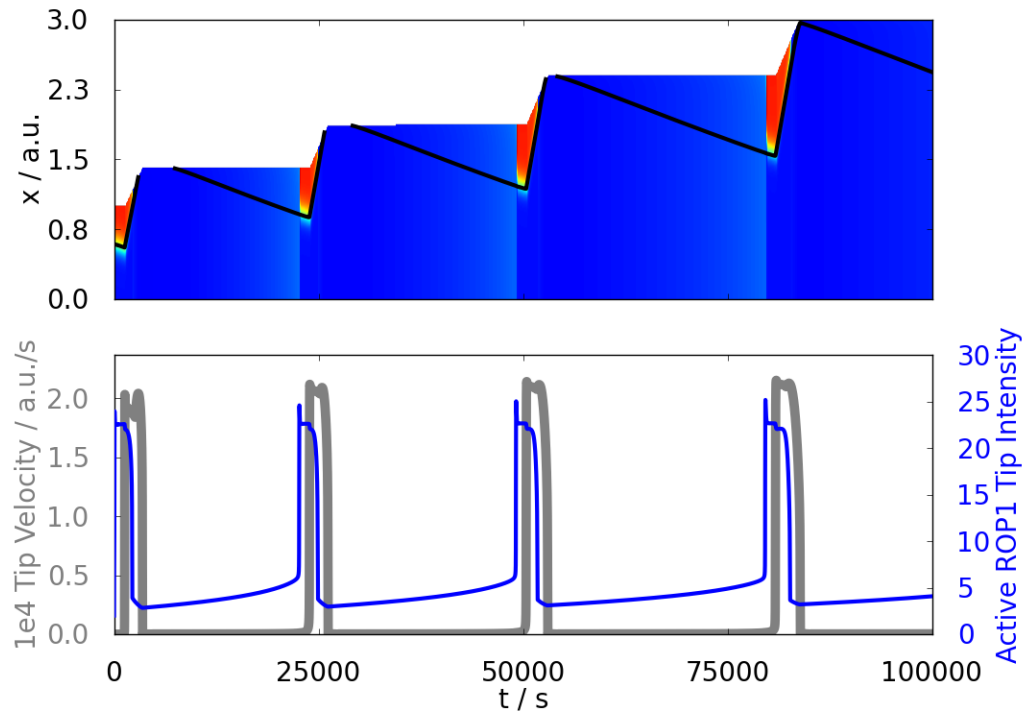
Figure 3.33: **Pollen Tube Dynamics with Delay Between ROP1 Localisation and Growth Rate.** Introduction of time delay between ROP1 localisation and growth rate increase shows peak separation as is observed in experiments, Figure 3.32 We use the same parameter values as in Figure 3.30 and apply a delay of ν=5000 s.

3.4 Discussion

In this manuscript we studied the pollen tube as a model system of tip growth.

This work was motivated by our observation that pollen tube growth is often-times discussed at two different scales that have, to the best of our knowledge, not yet been combined: One set of models focuses on the biophysical properties of pollen tube growth and assumes key properties about the lower, molecular scale involved (Rojas et al., 2011; Kroeger et al., 2011). Another set of models focuses on the dynamics of key molecular components with little effort to scale these models up to the biophysical level (Yan et al., 2009; Kost, 2008).

In this manuscript we described a model of a key molecular component of pollen tube growth and its interplay with higher-level growth processes. The molecular component that we studied is ROP1, a Rho-homologous molecular switch that lo-calises to the tip of growing pollen tubes. To link this molecular entity to higher-level growth processes, we used the phenomenological assumption that active ROP1 in-

duces growth locally (Hwang et al., 2010).

3.4.1 Localisation of ROP1 Depends on the Total Concentration of ROP1

We have shown that the total concentration of ROP1 in the pollen tube is a key determinant of whether ROP1 can localise to induce growth, Figure 3.13.

3.4.2 Our Model Assumes Tip-Localised ROP1 Activation

Our spatial model of the pollen tube is one-dimensional and the ROP1 model can therefore not make use of geometric effects observed in two dimensions: In a two-dimensional, elliptical spatial domain we observe that the mass of localised active ROP1 migrates as a whole towards either of the end points of the major axis. We observe that the mass of localised active ROP1 only stops migrating once the boundary between highly and lowly concentrated active ROP1 lies perpendicular to the major axis of the ellipse.

Since our spatial model of the pollen tube is one-dimensional our simulations are unable to make use of this geometric effect. Therefore, we resort to introducing a tip-localised stimulus term so that we do not observe rear-localised ROP1, Figure 3.11, but rather tip-localised ROP1, Figure 3.12.

Our rationale is that we model the pollen tube after it has emerged from the pollen grain and we assume that the semi-elliptical shape of the apex draws localised active ROP1 into the tip as in the aforementioned geometric effect. The tip-localised stimulus term is intended to mimic this implicit geometric effect.

In future efforts it will be important to extend our spatial model to two dimensions to show the influence of this geometric effect more conclusively.

3.4.3 Our Model Predicts Both Constant and Oscillatory Pollen Tube Growth

Experimentally, pollen tube growth is observed to be both constant and oscillatory, Figure 3.5. Our model predicts that tight matching of the growth velocity and pro-

tein production sustains ROP1 localisation and therefore constant growth, Figure 3.31. This mode of constant growth is, however, unstable to small unmatched deviations in either the growth velocity or protein production rate, Section 3.2.3. At the other extreme, our model predicts oscillatory growth when the growth rate and protein production are "entirely decoupled", Figures 3.27 and 3.30. When decoupled, a growth sprint that reduces the total concentration of ROP1 is followed by a rest phase that allows protein production to replenish ROP1 concentrations to a level where it can re-localise to induce growth.

3.4.4   Our Model May Explain Why Fast Pollen Tubes Oscillate

As we have shown, pollen tubes with constant peak growth velocities will eventually transition from an oscillatory to a constant growth regime, Figure 3.27. On the other hand, pollen tubes that are capable of scaling their peak growth velocity with length are able to sustain oscillatory growth, Figure 3.30. Our results fall in line with observations that fast pollen tubes tend to oscillate while slow ones tend to grow with a constant velocity (JA Feijo, personal communication) and (Rojas et al., 2011).

3.4.5   Our Model Predicts Falsely Complete Disappearance of the ROP1 Peak

We further point out a discrepancy between our numerical experiments, Figures 3.26 and 3.30, and experimental observations, Figure 3.32: Our numerical experiments show intermittent, complete disappearance of ROP1 from the tip while in live pollen tubes some localised ROP1 intensity is always retained. We identify our ROP1-dependent growth velocity, Equation (3.35), as the source of this discrepancy and further propose that this discrepancy can be lessened by amending the growth velocity as follows:

$$a(X,t) = k_a \cdot \max\left(0, \varphi(u(X,t)) - \varphi_0\right), \tag{3.40}$$

where φ_0 denotes an offset, and we use the maximum function $\max(\cdot)$ to only retain

non-negative velocities. The offset φ_0 allows us to specify the ROP1 cap size at which the growth velocity vanishes, Figure 3.34. Thus for an appropriate choice of φ_0 and tip-localised active ROP1, we expect that the pollen tube stops growing, allowing for ROP1 to recharge, while the ROP1 cap has not disappeared completely.
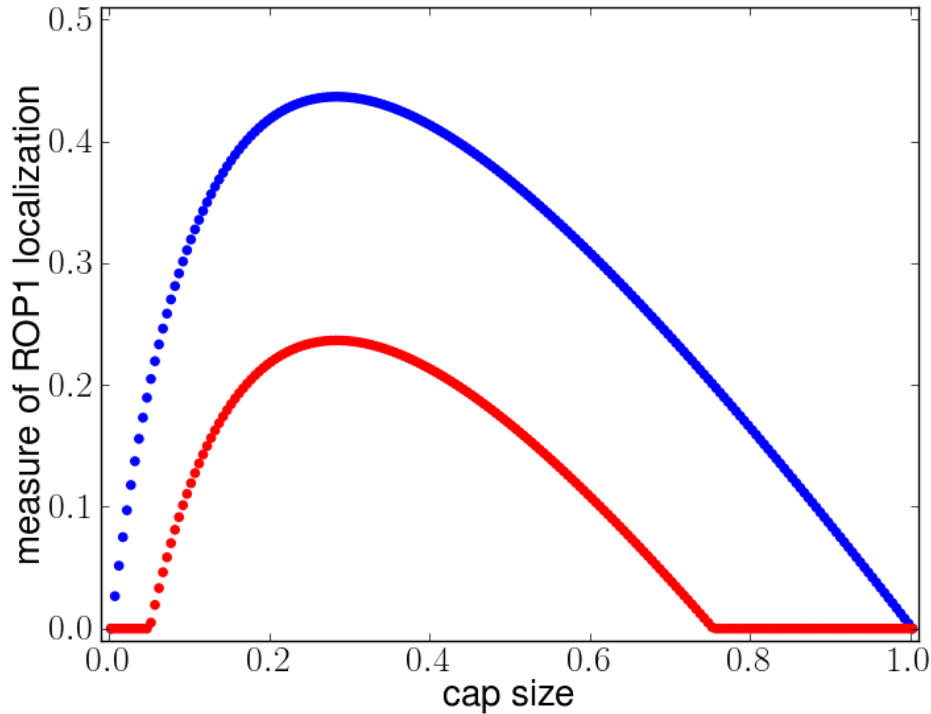
Figure 3.34: **Modified Measure of the Localisation of ROP1 Permitting Non-Vanishing ROP1.** We plot our measure of ROP1 localisation, φ (blue disks), that we use in our numerical experiments as input to the growth velocity, Equation (3.35), together with a modified measure, $\max(0, \varphi - \varphi_0)$ (red disks), proposed in Equation (3.40). The used offset here is $\varphi_0 = 0.2$. The modified measure (red disks) ensures that the ROP1 cap does not disappear completely before the pollen tube stops growing. Since the modified measure (red disks) is lower in absolute value than the unmodified measure (blue disks), the value of k_a in Equation (3.40) may need to be modified to recover peak growth velocities equivalent to our numerical experiments.

3.4.6 Our Model Predicts Falsely Spiky Growth Rate Profiles

Our results predict spiky growth rate profiles for pollen tubes that are in the oscillatory growth regime where the pollen tube grows quickly for a brief period before slowing down for a prolonged period, Figure 3.30. These spikes stand in contrast to experimentally observed growth rate profiles that display a more sinusoidal behaviour, Figure 3.32.

Similar growth rate spikes are produced by the biophysical model described by

Kroeger et al. (2011) although the growth rate in their model does not drop to zero at its lowest point.

In a future endeavour it will be interesting to study the growth rate profile of pollen tubes more thoroughly to better understand whether it displays spiky or sinusoidal behaviour.

3.4.7 Parameter Estimates

Our numerical experiments show that a tight coupling between the growth velocity of the pollen tube and the rate of ROP1 production, κ, permits non-oscillatory growth.

To the best of our knowledge, κ has not been reported in the literature for the biosynthesis of ROP1 in pollen tubes. Here we estimate κ for non-oscillatory pollen tubes, based on the insight we have gained from our numerical experiments.

To offer estimates in sensible units we reformulate the total concentration T as a volume concentration

$$T = \frac{\tau}{V}, \tag{3.41}$$

where τ is the total protein mass and V is the volume of the pollen tube which we represent as a cylinder of length L and radius R, Figure 3.9. Similar to the one-dimensional case, Section 3.2.3, the time change in the total concentration of ROP1 is given by

$$\dot{T} = \frac{\dot{\tau}}{V} - \frac{\tau}{V}\frac{\dot{V}}{V}, \tag{3.42}$$

and assuming that protein is produced everywhere at the same rate κ, $\dot{\tau}$ is given by

$$\dot{\tau} = V\kappa. \tag{3.43}$$

For constant total concentrations ($\dot{T}=0$) and therefore constant growth, κ needs to meet the following condition

$$\kappa = \frac{\tau}{V}\frac{\dot{V}}{V}. \tag{3.44}$$

In this relation we estimate that the total concentration of ROP1 is fixed at $\tau/V{=}5\,\mu M$ (Yan et al., 2009). Further, we estimate for wild-type *Arabidopsis* pollen tubes that the growth rate $a{=}0.05\,\mu m/s$ (Hwang et al., 2010), and that the radius $R{=}3\,\mu m$ (Nishikawa et al., 2005). Since the volume of a cylinder of length L and radius R is $\pi R^2 L$, we find that the volume expansion rate for an *Arabidopsis* pollen tube is $\dot{V}{=}1.4\,\mu m^3/s$. Wild-type *Arabidopsis* pollen tubes are up to 300 μm long six hours after germination (Li, 1999) and we estimate κ so that pollen tubes of corresponding length grow at a constant rate, Figure 3.35.
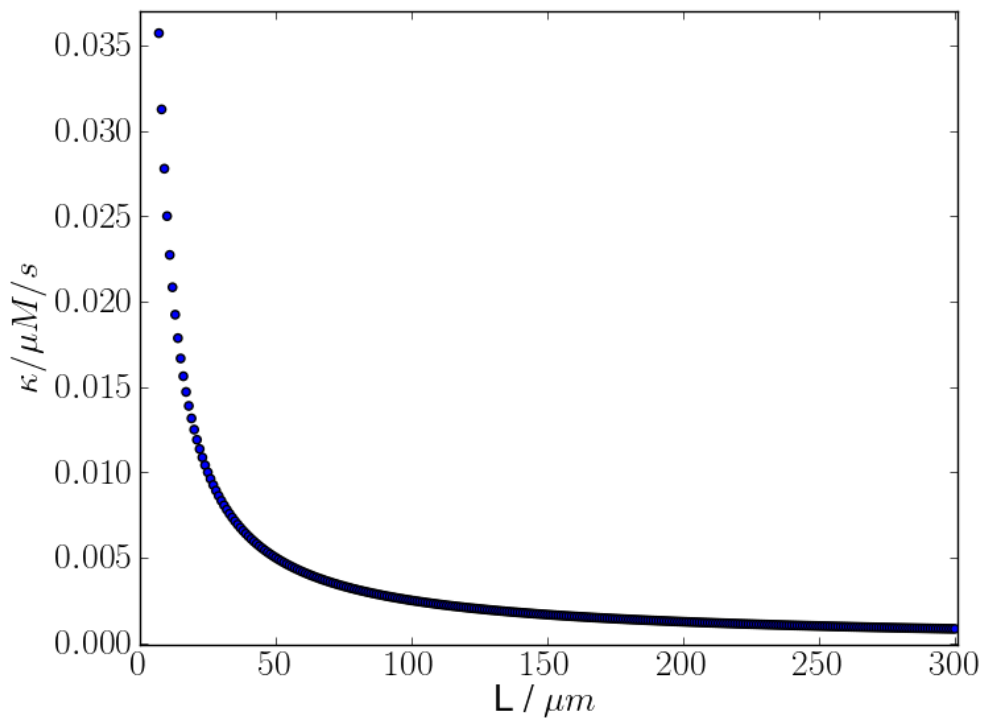
Figure 3.35: **Estimate of κ for Constant Growth of *Arabidopsis* Pollen Tubes up to Six Hours After Germination.** We estimate that these length-dependent values of κ reflect rates of ROP1 production for which *Arabidopsis* pollen tubes of corresponding length grow at a constant velocity. This is due to the fact that these values of κ should keep the total concentration of ROP1 T relatively constant. Our numerical experiments suggest that reducing κ below the indicated values should trigger oscillatory growth through repeated dilution ($\dot{T}{<}0$) and recharge ($\dot{T}{>}0$) of ROP1. The maximal and minimal predicted values of κ are 0.035 $\mu M/s$ (far left) and 0.00084 $\mu M/s$ (far right) respectively. The main text describes details of the derivation.

3.4.8 Outlook

Our current model does not include formation of callose plugs which have been described by some authors as important for keeping the volume of the cytoplasm

manageable (Zhang et al., 2008). Pollen tubes vary wildly in their pattern of callose plug formation and mutants that do not form callose plugs grow almost indistinguishably from wild type (Nishikawa et al., 2005; Qin et al., 2012). However, the inclusion of callose plugs may form another interesting venue for explaining ROP1 dilution, ROP1 replenishment, and regulation of cytosolic volume in general. Our model framework is extensible to include the effect of callose plug formation and future efforts in this direction may prove important in elucidating the purpose and observed variation of callose plug formation. Experimental measurement of cytosolic volume and ROP1 dilution in growing pollen tubes is likely important to reconcile the role of callose plugs: An extended model that includes callose plugs can offer model predictions for wild-type behaviour, whereas the model presented in this manuscript offers predictions for mutants that lack callose plugs.

The model we developed also lends itself well to testing hypotheses about the causal link between different oscillatory components of pollen tube growth. Oftentimes, observations of components that oscillate in- or out-of-phase, such as growth rate and ROP1 intensity, have been drawn together to hypothesise about the causal links between these components (Rounds & Bezanilla, 2013). With our framework we are now able to include hypothesised causal links directly and observe their effect on growth behaviour, Figure 3.33.

In future efforts, our model can be extended to more detailed descriptions of the dynamics and biochemical interactions of ROP1 with biophysical growth processes. Thus our modelling framework can act as a bridge to further link molecular descriptions with macroscopic models of growth.

Chapter 4

GraTeLPy: Graph-Theoretic Linear Stability Analysis

4.1 Abstract

A biochemical mechanism with mass action kinetics can be represented as a directed bipartite graph (bipartite digraph) and modelled by a system of differential equations. If the differential equations of the mechanism can give rise to saddle-node bifurcation, which often leads to multistability, then the bipartite digraph contains a structure referred to as a critical fragment. In GraTeLPy we have implemented a previously published graph-theoretic method that identifies the critical fragments of the bipartite digraph of biochemical mechanisms. We have optimised GraTeLPy so that it can aid in rapid mechanism discrimination which is an important step in unfurling phenomenological to mechanistic models of biological processes.

4.2 Introduction

Throughout this thesis we look at cell polarity conferred by localisation of homologues of the Rho family of proteins. The previous chapters were based in part on a model of Rho localisation that was proposed by Mori et al. (2008), Figure 4.1. Based upon existing molecular understanding of the reactions involved in Rho localisation, Figure 4.1 (left), Mori et al. (2008) proposed and analysed a phenomenological kinetic model that omits all intermediate molecular species and focuses solely on

the active and inactive form of Rho as the key molecular species, Figure 4.1 (right). In this phenomenological model, GAP-driven inactivation and GDI-dependent sequestration, Figure 4.1 (left, red dashed rectangle), are represented by a unimolecular reaction, Figure 4.1 (right and bottom, red solid rectangle). Further, activation involving a GEF-dependent intermediate, Figure 4.1 (left, green dashed rectangle), is represented by a sigmoidal term, Figure 4.1 (right and bottom, green solid rectangle).

$$\frac{\partial a}{\partial t} = D_a \nabla^2 a \boxed{- \delta a} \boxed{+ b \left(\frac{\gamma a^2}{K^2 + a^2} \right)}$$
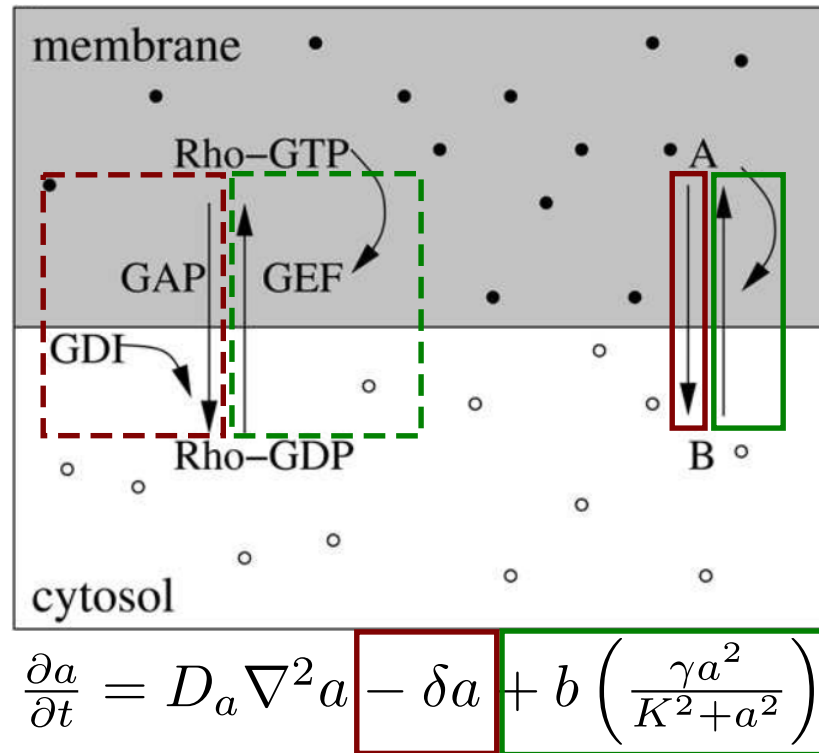
Figure 4.1: **Juxtaposition of a Mechanistic Model and a Phenomenological Model of the Chemical Reactions Underlying Rho Localisation.** (left) A mechanistic model of Rho localisation that includes a number of known activating and inactivating enzymes as well as membrane-binding and -unbinding events. This model includes a number of intermediate chemical species as explicit variables. (right) The simplified, phenomenological model of Rho localisation proposed by Mori et al. (2008). This model includes solely the active and inactive form of Rho, called *A* and *B* respectively, and the presumed mechanistic activities of certain intermediate species are summarised as certain reaction terms: The GAP- and GDI-driven inactivation and sequestration of Rho (left, dashed red rectangle) is summarised as a unimolecular inactivation term (right and bottom, red solid rectangle). Further, the GEF-driven activation of Rho and resulting positive feedback (left, dashed green rectangle) is summarised as a term sigmoidal in the concentration of active Rho (right and bottom, solid green rectangle). Modified from Mori et al. (2008).

This simplified model has several advantages: Firstly, it helps understand the core properties of the Rho system that permit its active form to localise. These core properties include autocatalysis of active Rho, Figure 4.1 (red rectangles), which

is essential for both bistability of the chemical mechanism (Mori et al., 2008; Semplice et al., 2012) and bistability of the chemical mechanism in a limit case of diffusion (Walther et al., 2012; Holmes, 2014) – bistability of either kind is important for Rho localisation and therefore cell polarity. Secondly, this simplified model describes only two concentration variables thus permitting thorough analysis of the conditions under which active Rho localises (Mori et al., 2008; Walther et al., 2012; Holmes, 2014).

Despite its advantages and general value in modelling Rho localisation, this simplified model carries insufficient molecular detail when we are interested in system-specific molecular detail. For instance we may be interested in the distribution and molecular function of various molecular species, such as intermediate complexes. In this context we wish to unfurl a simplified model, Figure 4.1 (right), to a more detailed one, Figure 4.1 (left), that describes the molecular species of interest while retaining key mechanistic properties, such as autocatalysis of active Rho, and overall propensity for Rho to localise.

Unfurling a simplified mechanism is an iterative process where we first propose a number of molecular species and reactions, test whether the proposed mechanism shows desirable behaviour such as Rho localisation, and amend with more detail in case it does not. To test for desirable behaviour, proposed mechanisms are often translated into differential equation models and are studied with methods from bifurcation analysis (Kuznetsov, 1998). Bifurcation analysis methods are easily applied when the differential equations have one or two variables (phase plane analysis) or have a relatively small number of parameters (numerical bifurcation analysis). These methods become difficult or computationally expensive when testing mechanisms with many variables and reactions for desirable behaviour.

When unfurling a phenomenological model, Figure 4.1 (right), to a mechanistic model, Figure 4.1 (left), we generally expect to propose mechanisms with a great number of variables and parameters with uncertain values. We therefore need approaches that can cope with large dynamical systems and that generate results independent of parameter values.

Stoichiometric network analysis is a set of approaches that have been developed for the analysis of the dynamical systems of large mechanisms and have the fol-

lowing capabilities (Clarke, 1988). These approaches help analyse whether a proposed mechanism has the potential to show desirable behaviour, independently of parameter values. Further these approaches help understand how adding and removing reactions changes overall capacity for the desirable behaviour (Clarke, 1988). Both of these capabilities are important since measured rate constants are often off by orders of magnitude and it is common to add and remove mechanistic steps iteratively when unfurling simplified models (Clarke, 1988). Related, recent graph-theoretic methods that represent chemical reaction networks as certain types of graphs can analyse large reaction networks for their potential to show bistable behaviour (Craciun & Feinberg, 2006; Mincheva & Roussel, 2007).

A classical example where stoichiometric network analysis helped study complex mechanisms is the search for the mechanism of the Belousov-Zhabotinsky reaction (Clarke, 1988). Examples where graph-theoretic methods have been used include studies related to cell polarity (Goryachev & Pokhilko, 2008) and photosynthesis (Amin & Roussel, 2014).

Despite their great utility, graph-theoretic methods become challenging to apply by hand to large mechanisms making their computational implementation highly desirable. Existing software packages in this sphere implement graph-theoretic methods that can be used to decide if the overall structure of a given mechanism has the capacity for multistability (Ellison & Feinberg, 2000; Soranzo & Altafini, 2009). In GraTeLPy we have implemented a graph-theoretic method developed by Mincheva & Roussel (2007) that identifies the precise substructures of a mechanism that can convey multistability (Walther et al., 2014). Knowledge of these substructures may help direct experiments and can help estimate parameter values for which the mechanism is multistable (Craciun & Feinberg, 2006; Mincheva & Roussel, 2007).

GraTeLPy has the following advantages over similar software packages: (i) It identifies the sets of species and reactions that can give rise to multistability, (ii) the generated output can guide parameter estimation, and (iii) it is both open source and is based on freely available software. GraTeLPy can also run in parallel on computer clusters which can increase the size of testable mechanisms, a limitation other approaches have tried to circumvent differently (Conradi et al., 2007).

4.3 Mathematical Background

Here we summarise general terminology and graph-theoretic criteria for multistability that GraTeLPy is based on (Mincheva & Roussel, 2007).

4.3.1 Mathematical model

A biochemical mechanism with n species A_i, $i=1,\ldots,n$, and m elementary reactions B_j can be written as

$$B_j : \sum_{i=1}^{n} \alpha_{ji} A_i \xrightarrow{k_j} \sum_{i=1}^{n} \beta_{ji} A_i, \quad j = 1,\ldots,m, \tag{4.1}$$
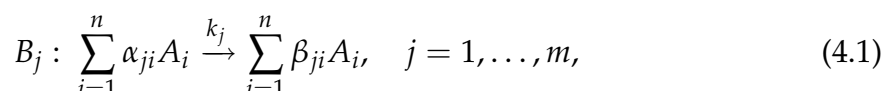
where $k_j > 0$, $j=1,\ldots,m$ are the rate constants. The constants $\alpha_{ji} \geq 0$ and $\beta_{ji} \geq 0$ in Equation (4.1) are small integers called stoichiometric coefficients that account for the number of molecules of species A_i participating in the jth elementary reaction.

Mincheva & Roussel (2007) developed their graph-theoretic criteria for multistability for a generic reaction mechanism, Equation (4.1). For brevity we here summarise their graph-theoretic approach by example and we will refer continually to the following illustrative sample mechanism:

```
# Mechanism file
# for use in GraTeLPy
[A1] ->   ; k1
 -> [A1] ; k2
[A1] + [A2] -> [A3] ; k3
[A3] -> [A2] ; k4
[A1] + [A3] -> [A4] ; k5
[A4] -> [A1] + [A3] ; k6
```

$$
\begin{aligned}
B_1 &: & A_1 &\xrightarrow{k_1} \emptyset, \\
B_2 &: & \emptyset &\xrightarrow{k_2} A_1, \\
B_3 &: A_1 + A_2 &\xrightarrow{k_3} A_3, \\
B_4 &: & A_3 &\xrightarrow{k_4} A_2, \\
B_5 &: A_1 + A_3 &\xrightarrow{k_5} A_4, \\
B_6 &: & A_4 &\xrightarrow{k_6} A_1 + A_3.
\end{aligned}
\tag{4.2}
$$
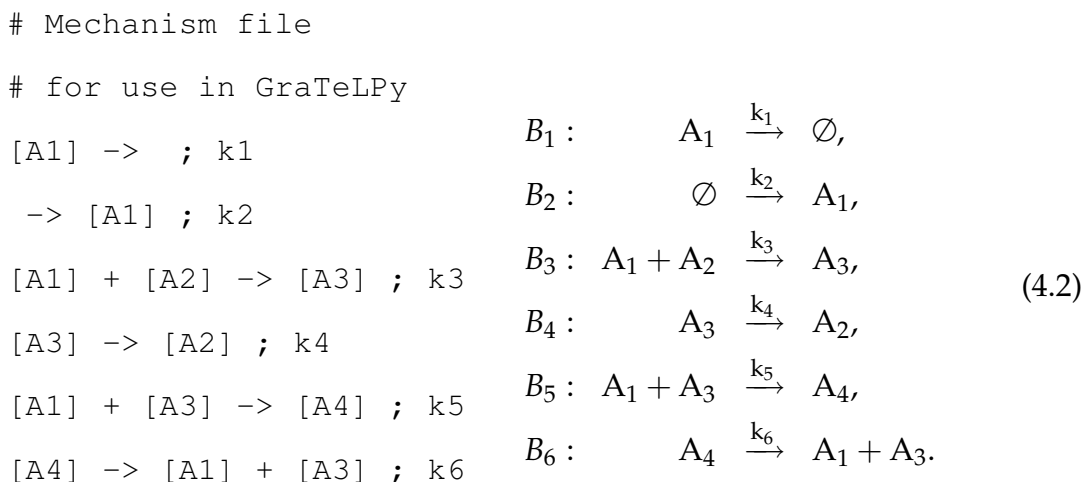
This is the reversible substrate inhibition mechanism which we choose for illustration since it is a well-studied system in the context of the graph-theoretic framework that GraTeLPy implements (Mincheva & Roussel, 2007). Equation (4.2) shows the mechanism in chemical notation on the right-hand side and the same mechanism translated to a format understood by GraTeLPy on the left-hand side.

Generally, we assume that every species A_k in Equation (4.1) is consumed and produced in at least one true reaction, i.e. a reaction which is different from an outflow reaction $A_k \rightarrow \emptyset$ or an inflow reaction $\emptyset \rightarrow A_k$. However, we do not specifically require that all species participate in an inflow and an outflow reaction. This way we ensure that if we removed all outflow reactions, there would be no species in the system that acts as a sink for all other species. In our example, the first two reactions of Equation (4.2) are an inflow and an outflow reaction respectively. Further, in Equation (4.2) species A_1 is consumed in reactions B_3 and B_5, A_2 in B_3, A_3 in B_4 and B_5, and A_4 in B_6.

We further generally assume mass action kinetics for the mechanism, Equation (4.1), so that the rates of reaction become

$$w_j(\mathbf{u}) = k_j u_1^{\alpha_{j1}} \dots u_n^{\alpha_{jn}}, \qquad j = 1, \dots, m, \tag{4.3}$$

where $u_k(t)$ is the concentration at time t of species A_k, $k=1, \dots, n$, and $k_j \geq 0$ is the reaction rate constant. To illustrate, for our example in Equation (4.2) the rate of reaction B_1 is $w_1 = k_1 u_1$ and that of reaction B_5 is $w_5 = k_5 u_1 u_3$.

Commonly we study reaction mechanisms such as Equation (4.2) as a system of ordinary differential equations (ODE) that describe how fast a given species A_i is consumed or produced. The ODE of a generic mass-action mechanism, Equation (4.1), can be written in vector form as

$$\dot{\mathbf{u}}(t) = S\mathbf{w}(\mathbf{u}), \tag{4.4}$$

where $\mathbf{u}(t) = (u_1(t), \dots, u_n(t))$ is the concentration vector, $S_{ji} = \beta_{ji} - \alpha_{ji}$ are the entries of the stoichiometric matrix S, and $\mathbf{w}(\mathbf{u}) = (w_1(\mathbf{u}), \dots, w_m(\mathbf{u}))$ is the vector of reaction rates, Equation (4.3).

For our example, the system of ODEs for Equation (4.2) is

$$\begin{aligned}
\dot{u}_1 &= -k_1 u_1 + k_2 - k_3 u_1 u_2 - k_5 u_1 u_3 + k_6 u_4, \\
\dot{u}_2 &= -k_3 u_1 u_2 + k_4 u_3, \\
\dot{u}_3 &= k_3 u_1 u_2 - k_4 u_3 - k_5 u_1 u_3 + k_6 u_4, \\
\dot{u}_4 &= k_5 u_1 u_3 - k_6 u_4.
\end{aligned} \tag{4.5}$$

The rank of the stoichiometric matrix S of Equation (4.5) equals 3 since there is one conservation relation $u_2 + u_3 + u_4 = c_0$. To see this, we write out the stoichiometric matrix S of Equation (4.5):

$$S = \begin{bmatrix} -1 & 1 & -1 & 0 & -1 & 1 \\ 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix}, \tag{4.6}$$

where the rows of S give us the stoichiometry of the corresponding chemical species in each reaction. As an example, for A_4 we compute the dot product of the fourth row of S with $\mathbf{w}$ and obtain $\dot{u}_4 = w_5 - w_6$. Realising that $\dot{u}_2 + \dot{u}_3 + \dot{u}_4 = 0$ we can rewrite $\dot{u}_4 = -\dot{u}_2 - \dot{u}_3$ hence we can express the fourth row of S as a linear combination of its second and third row.

Conventionally, we write the right-hand side of an ODE such as Equation (4.4) as a function so that, for instance, $\dot{u}_i = f_i(\mathbf{u})$. Given our choice of mass action kinetics these functions are

$$\dot{u}_i(t) = f_i(\mathbf{u}) = \sum_{j=1}^{m} S_{ji} w_j(\mathbf{u}).$$

Suppose that our generic reaction mechanism is in a steady state so that $\mathbf{w}(\mathbf{u}) = (0, \ldots, 0) = \mathbf{0}$. Then we can determine if this steady state is stable or unstable by estimating the time behaviour of small perturbations away from the steady state with a linear approximation (Kuznetsov, 1998). The linear approximation of the time behaviour of Equation (4.4) is given by the Jacobian $J(\mathbf{u}, \mathbf{w})$ whose elements can be written as

$$J_{ik}(\mathbf{u}, \mathbf{w}) = \frac{\partial f_i}{\partial u_k} = \sum_{j=1}^{m} S_{ji} \alpha_{jk} \frac{w_j}{u_k}. \tag{4.7}$$

Note that the concentrations u_k, $k = 1, \ldots, n$ and the reaction rates $w_j(\mathbf{u})$, $j = 1, \ldots, m$ (both considered evaluated at a positive equilibrium) are used as parameters in Equation (4.7), i.e. we evaluate the Jacobian in a specific steady state. Further note that stability of the corresponding steady state is determined by the eigenvalues of $J(\mathbf{u}, \mathbf{w})$ (Kuznetsov, 1998).

The Jacobian matrix of our example, Equation (4.5), parametrised in $(\mathbf{u}, \mathbf{w})$ is

$$J(\mathbf{u}, \mathbf{w}) = \begin{bmatrix} -\frac{w_1+w_3+w_5}{u_1} & -\frac{w_3}{u_2} & -\frac{w_5}{u_3} & \frac{w_6}{u_4} \\ -\frac{w_3}{u_1} & -\frac{w_3}{u_2} & \frac{w_4}{u_3} & 0 \\ \frac{w_3-w_5}{u_1} & \frac{w_3}{u_2} & -\frac{w_3+w_4}{u_3} & -\frac{w_6}{u_4} \\ \frac{w_5}{u_1} & 0 & \frac{w_5}{u_3} & -\frac{w_6}{u_4} \end{bmatrix}. \tag{4.8}$$

The eigenvalues of a generic Jacobian $J(\mathbf{u}, \mathbf{w})$ are the zeros of its characteristic polynomial

$$P(\lambda) = \det(J(\mathbf{u}, \mathbf{w}) - \lambda I) = \sum_{k=0}^{n} a_k(\mathbf{u}, \mathbf{w})(-\lambda)^{n-k}, \tag{4.9}$$

where I is the identity matrix and coefficient a_k ($a_0=1$) is the sum of principal minors of order k (Gantmacher, 2000):

$$a_k = \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} J \begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ i_1 & i_2 & \cdots & i_k \end{pmatrix}. \tag{4.10}$$

As in our example, Equation (4.6), ODE models of general mechanisms, Equation (4.4), often have mass conservation relations lowering the rank r of both the stoichiometric matrix S and the Jacobian $J(\mathbf{u}, \mathbf{w})$. In these cases the last non-zero coefficient in Equation (4.9) is $a_r(\mathbf{u}, \mathbf{w})$. To see this intuitively, in our example where $n=4$ and $r=3$ the coefficient a_4 of Equation (4.9) is

$$a_4 = J \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix} = \det J = 0, \tag{4.11}$$

where the last equality holds due to the fact that row four of J can be expressed as a linear combination of rows two and three. Hence in our example the last non-zero coefficient of the characteristic polynomial of the Jacobian, Equation (4.8), is

$$a_3(\mathbf{u}, \mathbf{w}) = \frac{w_4 w_6(w_1 + w_3)}{u_1 u_3 u_4} + \frac{w_1 w_3 w_6}{u_1 u_2 u_4} + \frac{w_3 w_5(w_1 - w_4)}{u_1 u_2 u_3}. \tag{4.12}$$

4.3.2 The Bipartite Digraph of a Biochemical Mechanism

Further below we will show that the potential of a mechanism for multistability can be judged by the constants of the characteristic equation of its Jacobian matrix (Kuznetsov, 1998). One representation of these constants is the determinant-based approach we described in Section 4.3.1. The framework implemented by GraTeLPy expresses these constants as functions of a graph representation of the dynamical system thus allowing its user to reason structurally about stability behaviour (Mincheva & Roussel, 2007). Further below we will compare these two representations of the coefficients and the statements about multistability that we can derive from them.

Here we summarise terminology necessary to understand the graph-based representation introduced by Volpert & Ivanova (1987); Mincheva & Roussel (2007). To illustrate the definitions in this section we continue referencing the reversible substrate mechanism, Equation (4.2).

Generally, a directed bipartite graph, or bipartite digraph for short, has a node set that consists of two disjoint subsets, V_1 and V_2, and each of its directed arcs has one end in V_1 and the other in V_2 Harary (1969). The **bipartite digraph** G of a biochemical reaction network, Equation (4.1), is defined as follows. The **nodes** are separated into two sets, one for the chemical species $V_1 = \{A_1, A_2, \ldots, A_n\}$ and one for the elementary reactions $V_2 = \{B_1, B_2, \ldots, B_m\}$. We draw an **arc** from A_k to B_j if and only if species A_k is a reactant in reaction j, i.e. if the stoichiometric coefficient $\alpha_{jk} > 0$ in (4.1). Similarly, we draw an arc from B_j to A_i if and only if A_i is a product in reaction j, i.e. if the stoichiometric coefficient $\beta_{ji} > 0$ in (4.1). In our example, Equation (4.2), we draw an arc from A_1 to B_3 and one from B_3 to A_3, Figure 4.2, and so forth.

The bipartite digraph of the reversible substrate inhibition mechanism, Equation (4.2), is shown in Figure 4.2.
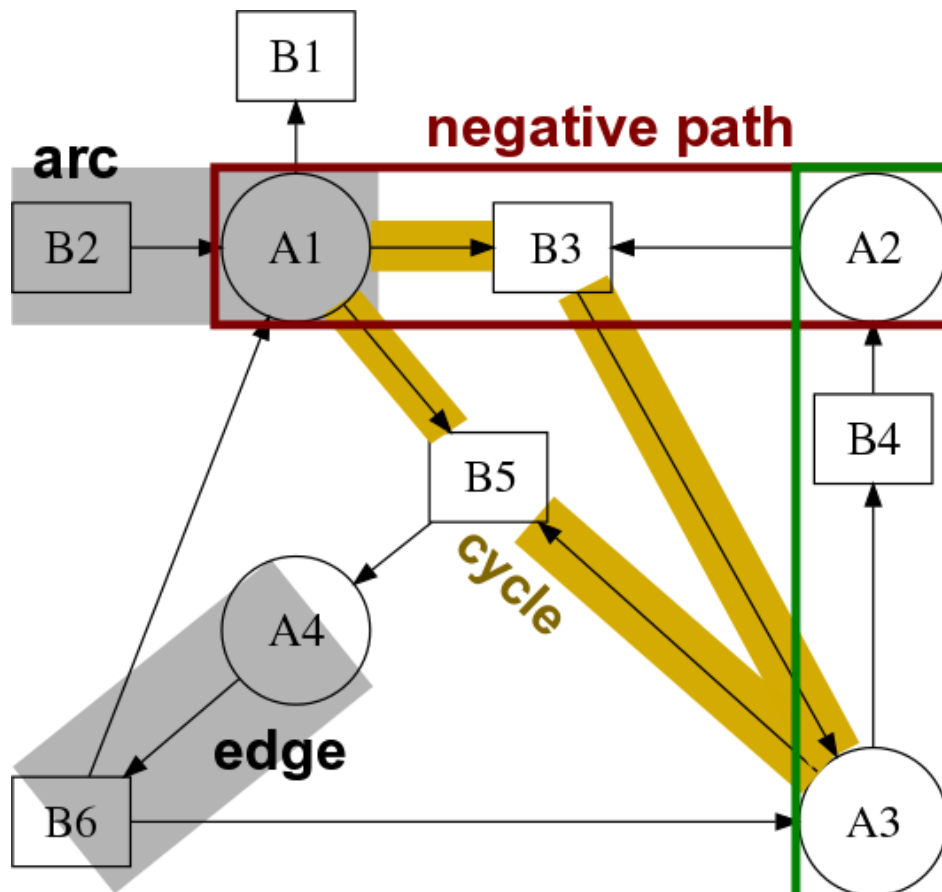
Figure 4.2: **Bipartite Digraph of the Reversible Substrate Inhibition Reaction Mechanism (Mincheva & Roussel, 2007).** Circles denote species nodes and squares denote reaction nodes of the mechanism, Equation (4.2). **Arcs** are the relationships between nodes in the bipartite digraph. **Edges** are those arcs that denote flow of molecules into a reaction such as the consumption of A_4 in reaction B_6 through the highlighted edge. **Negative paths** denote elements in the bipartite digraph where two chemical species flow into the same reaction. **Positive paths** denote elements where a reaction turns one chemical species into another. **Cycles** denote contiguous chains of conversion (positive paths) and molecular interaction (negative paths) between chemical species.

The arc $[A_k, B_j]$ is an **edge** if $\alpha_{jk} > 0$, i.e. if species A_k is a reactant in reaction j. For instance, the arc pointing from A_4 to B_6 is an edge whereas the arc pointing from B_2 to A_1 is not an edge, Figure 4.2 (grey rectangles). The **weight of an edge** $E=[A_k, B_j]$ is defined as

$$K_E = -\alpha_{jk}^2. \tag{4.13}$$

For example, the edge $E=[A_4, B_6]$ in Figure 4.2 has weight $K_E = -1$.

If $\alpha_{jk}\beta_{ji} > 0$, then the arcs (A_k, B_j) and (B_j, A_i) form a **positive path** $[A_k, B_j, A_i]$ that corresponds to the production of A_i from A_k in a reaction j. The **weight of the positive path** $[A_k, B_j, A_i]$ is defined as $\alpha_{jk}\beta_{ji}$. For example, the positive path

$[A_3, B_4, A_2]$ in Figure 4.2 (green frame) has weight 1 as can be deduced from Equation (4.2).

If $\alpha_{jk}\alpha_{ji}>0$, then the arcs (A_k, B_j) and (A_i, B_j) form a **negative path** $[\overline{A_k, B_j, A_i}]$ that corresponds to A_k and A_i interacting as reactants in reaction j. The weight of the negative path $[\overline{A_k, B_j, A_i}]$ is defined as $-\alpha_{jk}\alpha_{ji}$. Note that the negative paths $[\overline{A_k, B_j, A_i}]$ and $[\overline{A_i, B_j, A_k}]$ are considered to be different since they start at different species nodes. For example, both $[\overline{A_1, B_3, A_2}]$ and $[\overline{A_2, B_3, A_1}]$ in Figure 4.2 (red frame) are negative paths with weight -1, where the weight can be deduced from the stoichiometric coefficients in Equation (4.2).

A **cycle** C of G is a sequence of distinct paths with the last species node of each path being the same as the first species node of the next path. For example, the positive path $[A_1, B_3, A_3]$ and negative path $[\overline{A_3, B_5, A_1}]$ form a cycle, Figure 4.2 (yellow highlight). For brevity we denote cycles by their species and reaction nodes so that the cycle in our example is written as $C=\binom{A_1, A_3}{B_3, B_5}$. Generally, cycles are denoted by $C=\binom{A_{i_1}, A_{i_2}, \ldots, A_{i_k}}{B_{j_1}, B_{j_2}, \ldots, B_{j_k}}$, where the number of species nodes k defines their **order**. The species nodes in a cycle are unique, but there may be a repetition among the reaction nodes. This is because negative paths containing the same nodes are considered different depending on the starting species node – hence the corresponding reaction node is repeated.

A cycle is **positive** if it contains an even number of negative paths and **negative** if it contains an odd number of negative paths. In our example the cycle $C=\binom{A_1, A_3}{B_3, B_5}$ is negative, Figure 4.2 (yellow highlight). The sign of a cycle C can also be determined from its **cycle weight** which is the product of the weights of its negative and positive paths:

$$K_C = \prod_{[\overline{A_k, B_j, A_i}]\in C}(-\alpha_{jk}\alpha_{ji}) \prod_{[A_k, B_j, A_i]\in C}\alpha_{jk}\beta_{ji}. \tag{4.14}$$

For example, $C=\binom{A_1, A_3}{B_3, B_5}$ in Figure 4.2 (yellow highlight) has order 2 with weight $K_C=-1$ and therefore is a negative cycle. The cycle $C=\binom{A_2, A_3}{B_3, B_4}$ in Figure 4.2 has order 2 with weight $K_C=1$ and therefore is a positive cycle.

A **subgraph** $g=\{L_1, L_2, \ldots, L_s\}$ of G consists of edges or cycles L_i, $i=1, \ldots, s$, where each participating species is the beginning of exactly one edge or path in

a cycle. In other words, the edges and cycles in a subgraph are species mutually disjoint. The number of species nodes in a subgraph is defined as its **order** and equals s for the above subgraph g. The **subgraph weight** of g is defined by the cycle and edge weights, Equations (4.14) and (4.13) respectively, of the cycles C and edges E in g:

$$K_g = (-1)^c \prod_{C \in g} K_C \prod_{E \in g} (-K_E), \tag{4.15}$$

where c is the number of cycles in g. For our example, the subgraph $g = \{[A_1, B_5], \binom{A_2, A_3}{B_3, B_4}\}$ with weight $K_g = -1$ is shown in Figure 4.3 (bottom right).
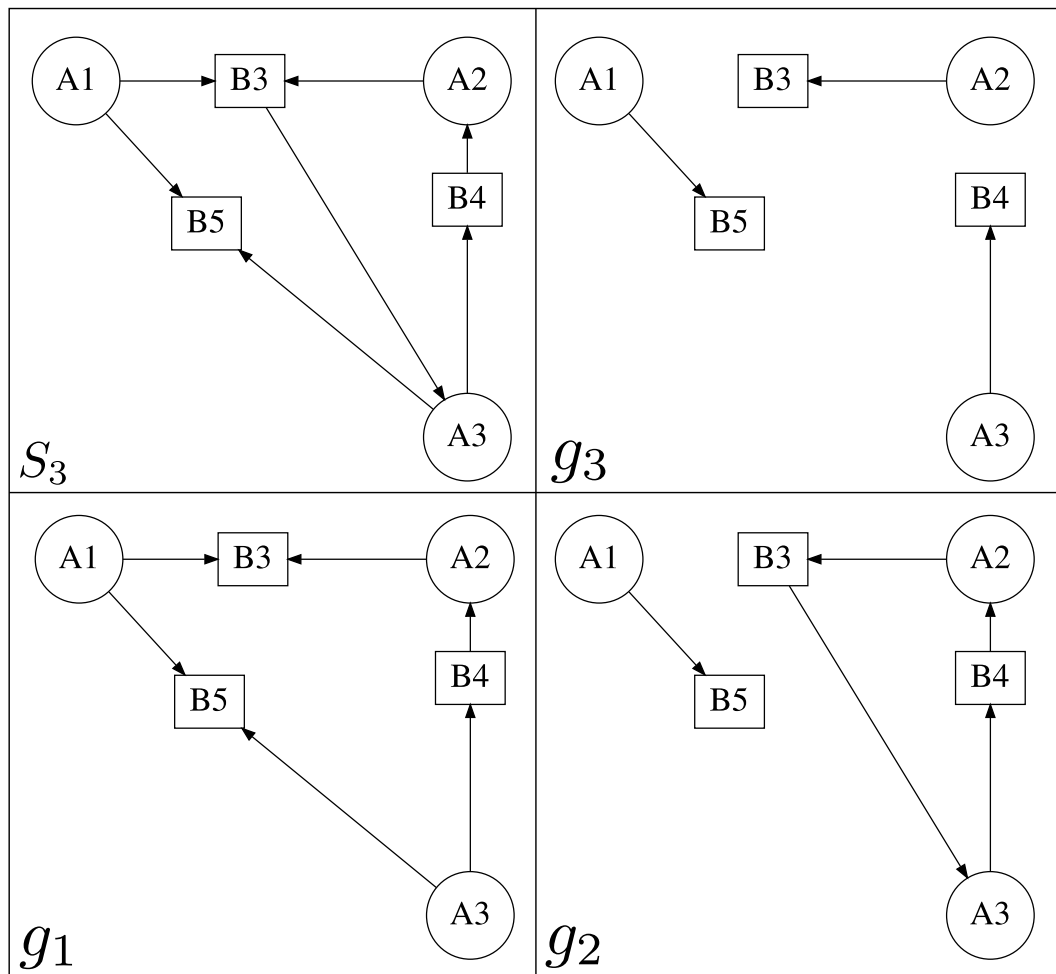
Figure 4.3: **Critical Fragment and Subgraphs of the Reversible Substrate Inhibition Mechanism.** Critical fragment $S_3\binom{1,2,3}{5,3,4}$ and constituent subgraphs of the reversible substrate inhibition mechanism computed by GraTeLPy. (top left) Critical fragment $S_3 = \binom{1,2,3}{5,3,4}$. (top right) Subgraph $g_3 = \{[A_1, B_5], [A_2, B_3], [A_3, B_4]\}$. (bottom left) Subgraph $g_1 = C_2 = \binom{A_1, A_3, A_2}{B_5, B_4, B_3}$. (bottom right) Subgraph $g_2 = \{[A_1, B_5], C_1 = \binom{A_2, A_3}{B_3, B_4}\}$.

Since more than one path can exist between species nodes via different reaction

nodes in a bipartite digraph, the number of subgraphs through the same node sets may be greater than one. The set of all subgraphs g of order k with the same species nodes $\bar{V}_1 = \{A_{i_1}, \ldots, A_{i_k}\}$ and reaction nodes $\bar{V}_2 = \{B_{j_1}, \ldots, B_{j_k}\}$ sets is called a **fragment** of order k and is denoted by $S_k\binom{i_1, \ldots, i_k}{j_1, \ldots, j_k}$. The **weight** of a fragment $S_k\binom{i_1, \ldots, i_k}{j_1, \ldots, j_k}$ is defined as:

$$K_{S_k} = \sum_{g \in S_k} K_g. \tag{4.16}$$

If $K_{S_k} < 0$, then S_k is a **critical fragment** and we clarify the significance of this classification below.

For example, the fragment $S_3\binom{1,2,3}{5,3,4}$ is shown in Figure 4.3 (top left) together with its three subgraphs $g_1 = C_2 = \binom{A_1, A_3, A_2}{B_5, B_4, B_3}$, Figure 4.3 (bottom left), $g_2 = \{[A_1, B_5],$ $C_1 = \binom{A_2, A_3}{B_3, B_4}\}$, Figure 4.3 (bottom right), and $g_3 = \{[A_1, B_5], [A_2, B_3], [A_3, B_4]\}$, Figure 4.3 (top right). Each of the first two subgraphs, g_1 and g_2, contain a positive cycle, and therefore $S_3\binom{1,2,3}{5,3,4}$ is a critical fragment since

$$K_{S_3} = \sum_{g \in S_3} K_g = K_{g_1} + K_{g_2} + K_{g_3} = -1 - 1 + 1 = -1 < 0.$$

Both Volpert & Ivanova (1987) and Mincheva & Roussel (2007) have shown that the coefficients of the characteristic polynomial, Equation (4.9), have the following graph-theoretic representation:

$$a_k(\mathbf{u}, \mathbf{w}) = \sum_{S_k\binom{i_1, \ldots, i_k}{j_1, \ldots, j_k}} K_{S_k} \frac{w_{j_1} \ldots w_{j_k}}{u_{i_1} \ldots u_{i_k}}, \quad k = 1, \ldots, n. \tag{4.17}$$

Note that here we sum over all subgraphs contained in all fragments of order k.

There is a one-to-one correspondence between any one fragment $S_k\binom{i_1, \ldots, i_k}{j_1, \ldots, j_k}$ in Equation (4.17) and a corresponding non-zero term in $a_k(\mathbf{u}, \mathbf{w})$ in Equation (4.9) (Mincheva & Roussel, 2007). For our example, the negative summand $-w_3 w_4 w_5$ $/(u_1 u_2 u_3)$ of $a_3(\mathbf{u}, \mathbf{w})$, Equation (4.12), corresponds to the critical fragment $S_3\binom{1,2,3}{5,3,4}$ shown in Figure 4.3 (top left).

4.3.3 Instability Criteria for the Jacobian and the Bipartite Digraph

So far we have mentioned that the stability properties of dynamical systems can be deduced from the coefficients of the characteristic polynomial of their Jacobian matrix, Section 4.3.1. We have further seen that these coefficients can be expressed as certain components of the bipartite digraph of a given dynamical system, Section 4.3.2.

Multistability in ODE models often arises by destabilising a stable steady state through a saddle-node bifurcation (Tyson, 1975; Kuznetsov, 1998). When a saddle-node bifurcation occurs a real eigenvalue $\lambda(\mathbf{u}, \mathbf{w})$ of $J(\mathbf{u}, \mathbf{w})$ changes sign as the parameters $(\mathbf{u}, \mathbf{w})$ change values – the saddle-node bifurcation occurs where exactly one $\lambda(\mathbf{u}, \mathbf{w}){=}0$ (Kuznetsov, 1998). Hence, a necessary condition for multistability arising from a saddle-node bifurcation is $a_n(\mathbf{u}, \mathbf{w}){=}\det(-J(\mathbf{u}, \mathbf{w})){=}0$ for some parameter values of $(\mathbf{u}, \mathbf{w})$ (Kuznetsov, 1998). If the rank r of the Jacobian matrix is reduced to $r{<}n$ due to the existence of mass conservation relations, Section 4.3.1, then $a_r(\mathbf{u}, \mathbf{w}){=}0$ for some values of $(\mathbf{u}, \mathbf{w})$ is necessary (Mincheva & Roussel, 2007).

To understand this intuitively for our example, Equation (4.2), we expand the general form of the characteristic polynomial, Equation (4.9), up to the $n{=}4$ species in our example:

$$P(\lambda) = (-\lambda)^4 + a_1(-\lambda)^3 + a_2(-\lambda)^2 + (-\lambda)a_3 + a_4, \tag{4.18}$$

where $a_4{=}0$ since the rank r of the Jacobian is $r{=}3$, Equation (4.11). Hence it is straightforward to see that $\lambda{=}0$ solves the characteristic equation $P(\lambda){=}0$ if $a_3{=}0$. Therefore, a_3 needs to contain at least one negative summand so that its overall sum can vanish for some values of $(\mathbf{u}, \mathbf{w})$. The presence of a negative summand does however not guarantee that we will be able to find values of $(\mathbf{u}, \mathbf{w})$ so that a_3 vanishes. Fragments S_k whose weight K_{S_k} is negative are referred to as critical fragments because they provide negative summands to the corresponding coefficients of the characteristic polynomial.

Generally, for a system with rank r a critical fragment $S_r\binom{i_1,\dots,i_r}{j_1,\dots,j_r}$ of order r, corresponding uniquely to a negative summand in Equation (4.17) for $k{=}r$, is required for a saddle-node bifurcation, and thus for multistability (Mincheva & Roussel, 2007;

Volpert & Ivanova, 1987). Thus the potential of a biochemical mechanism, Equation (4.1), for multistability depends on the structure of its bipartite digraph.

4.4   Implementation

The purpose of GraTeLPy is to enumerate all critical fragments of user-defined order k and, based on these fragments, inform the user of the potential bifurcation properties of their mechanism (Walther et al., 2014). Here we focus on the potential for multistability stemming from a saddle-node bifurcation so that the fragment order we are interested in is $k=r$.

Schematically, GraTeLPy achieves this purpose as follows, Figure 4.4: First we read in the mechanism from the user-provided input file and generate the bipartite digraph from it. We then enumerate all fragments S_k of user-defined order k and place them in a computational queue that allows us to further process each fragment, Figure 4.4 (top diamond nodes).

For each fragment S_k in this queue, we carry out a linear sequence of operations, Figure 4.4 (central rectangular nodes), to enumerate all subgraphs g of order k contained in S_k. We generate the weight K_g of each subgraph g and sum up the weights of all subgraphs contained in S_k to compute the weight K_{S_k} of that fragment. It is at this point that we can decide whether or not fragment S_k is a critical fragment of order k: if $K_{S_k}<0$ then S_k is critical, otherwise it is not critical.

Once we have processed all fragments in the queue in this manner we can output all collected results for the user and draw conclusions based upon these results, Figure 4.4 (bottom diamond nodes).

Processing the queue of enumerated fragments, Figure 4.4 (central rectangular nodes), is inherently parallel as each fragment may be handled independently of all other fragments. To make use of this parallelism GraTeLPy provides two scripts that implement a server role and a client role respectively.

The server script takes care of the actions in the top and bottom diamond nodes in Figure 4.4: It creates the bipartite digraph, enumerates all fragments, places them in a queue Figure 4.4 (top diamond nodes), and collects all computed data returned by the client processes before displaying them for the user Figure 4.4 (bottom dia-

mond nodes).

The client script deals with actions in the central rectangular nodes in Figure 4.4: It fetches a fragment from the queue presented by the server, generates all subgraphs for the given fragment, weighs both these subgraphs and the fragment, and reports all computed results back to the server. If the fragment queue is not exhausted the client fetches another fragment and repeats these steps.

This server-client architecture allows the user to run one or multiple instances of the client script to analyse the many fragments of a large mechanism in parallel.

Note that the format of the mechanism file read by the fragment server is close to how we usually describe chemical mechanisms. An example of the expected format is provided in Equation (4.2). At present GraTeLPy expects that the user converts their model to this format. However, we plan to include functionality to read mechanisms provided in SBML (Hucka & *et al.*, 2003) and other formats in a future release of GraTeLPy.

We now elaborate on our implementation of the steps carried out by the server, Section 4.4.1, and the client, Section 4.4.2.

GraTeLPy Fragment Server

Read mechanism and generate bipartite digraph.

Generate all fragments of order k.

Place all generated fragments in the fragment queue.

Fragment 1 fetched by a client.   Fragment j fetched by a client.   Fragment N fetched by a client.

GraTeLPy Client

Generate all subgraphs for fragment j.

...

Score all subgraphs and produce fragment score.

...

Store subgraphs and fragment score received from client.

Generate output from all results collected.

GraTeLPy Fragment Server

Figure 4.4: **Flowchart That Summarises the Steps Taken by GraTeLPy to Find all Critical Fragments of a Given Order.** The division of tasks between a single server and one or more clients is highlighted. (top diamonds) The fragment server reads the user-specified mechanism file and generates the bipartite digraph. The server generates all fragments of user-defined order k and places them in a queue. (centre rectangles) One or more client scripts fetch fragments off the queue and process them independently. For each fragment S_k, a client generates all subgraphs and computes the weight of each subgraph. The subgraph weights are then added to compute the weight of the corresponding fragment. The client passes the computed data back to the server and fetches another fragment off the queue if the queue is not yet exhausted. (bottom diamonds) After preparing the fragment queue, the server waits for the results sent by the clients. Upon receipt of the client-computed results for a fragment, the server stores these results if the fragment is found to have non-zero weight. Once the fragment queue, created initially by the server, is exhausted the server informs the user of the number of critical fragments discovered and generates other informative output.

4.4.1 Fragment Enumeration

To build some intuition about fragment enumeration, let us think of a fragment of order k as a collection of k unique substance nodes and k possibly repeated reaction nodes.

For our example, Figure 4.2, we look for fragments of order three, Equation (4.12). A randomly selected collection $\{A_1, A_2, A_3, B_1, B_2, B_3\}$ cannot be combined to form a subgraph since it does not contain a reaction node that A_3 can form an edge or path with. However, the set $\{A_1, A_2, A_3, B_3, B_4, B_5\}$ is a fragment since its nodes can form a subgraph consisting of edges $[A_1, B_5]$, $[A_2, B_3]$, and $[A_3, B_4]$, Figure 4.3 (top right). As mentioned above, we denote this fragment by the shorthand $S_3\binom{1,2,3}{5,3,4}$.

Generalising this example to a mechanism with n species and m reactions, we generate all fragments of order k combinatorially by pairing the $\binom{n}{k}$ unique combinations of species with m^k combinations of reaction nodes. This approach generates $\binom{n}{k} \cdot m^k$ collections of species and reaction nodes (fragment candidates) where, as in the above example, many of these are not fragments since their species and reaction nodes do not form a subgraph. Hence this approach requires verification of each fragment candidate generated.

The computational cost of this combinatorial approach becomes prohibitive even for relatively small values for n and m as it generates a large number of fragment candidates that then require verification. As a different approach that saves computational time by generating fewer fragment candidates, we use the fact that every fragment $S_k\binom{i_1,i_2,\ldots,i_k}{j_1,j_2,\ldots,j_k}$ has one subgraph that contains only edges $[A_{i_s}, B_{j_s}]$, $s=1,\ldots,k$. Based upon this fact we generate all $\binom{n}{k}$ species combinations and extend each combination to zero, one, or multiple subgraphs of the form $g=\{E_{i_1}, \cdots, E_{i_k}\}$ where $E_{i_s}=[A_{i_s}, B_{j_s}]$, $s=1,\ldots,k$, is an edge. The number of subgraphs, and equivalently fragment candidates, that we extend each species combination to is mechanism-dependent. Since fragments correspond uniquely to the subgraphs containing edges only, using this method we can generate all possible fragments. In our example, Figure 4.2, one of $\binom{4}{3}$ combinations of three substance nodes is the collection $\{A_1, A_2, A_3\}$ which contains the subgraph $g=\{[A_1, B_5], [A_2, B_3], [A_3, B_4]\}$, Figure 4.3 (top right).

Let us denote by $|E_i|$ $(i=1,\ldots,n_i)$ the number of edges that a species A_i is the starting node of. If we assume that each species A_i is, on average, the starting node of $|E|=\text{Avg}(E_i)$ edges, then this approach generates approximately $\binom{n}{k}\cdot|E|^k$ fragments. Empirically, we observe that $|E|$ is usually considerably less than some common values for the number of reactions m. In our example, Figure 4.2, A_1 is the starting node of two edges, A_2 one edge, A_3 two edges, and A_4 one edge: Here $|E|=1.5$ compared with a much greater $m=6$. It is reasonable to assume that this empirical value for $|E|$ is relatively constant across many different mechanisms while m may grow substantially.

GraTeLPy implements the latter approach which we observe to generate orders of magnitude fewer fragments than the former, combinatorial approach. This implemented approach for fragment enumeration is an important development in GraTeLPy that permits analysis of larger mechanisms. To highlight the achieved reduction in computational cost we plot the number of fragment candidates (of varying order k) generated by both methods for the double-layer mitogen-activated protein kinase (MAPK) mechanism with $n=12$ species and $m=18$ reactions (Walther et al., 2014) in Figure 4.5.

4.4.2 Subgraph Enumeration

To generate all subgraphs contained within a given fragment $S_k\binom{i_1,i_2,\ldots,i_k}{j_1,j_2,\ldots,j_k}$ we first need to determine what "building blocks" we have available. These building blocks are the edges $[A_{i_s}, B_{j_s}]$, positive paths $[A_{i_s}, B_{j_s}, A_{i_l}]$, and negative paths $\overline{[A_{i_s}, B_{j_s}, A_{i_l}]}$ $(l,s\in\{1,\ldots,k\})$, that are induced by the species and reactions of fragment S_k. We refer collectively to edges, positive paths, and negative paths of a fragment as its **subgraph components**.

To relate to our example, Figure 4.2, consider fragment $S_3\binom{1,2,3}{5,3,4}$, Figure 4.3 (top left). For this fragment we identify the following subgraph components: edges $[A_1, B_5]$, $[A_2, B_3]$, and $[A_3, B_4]$, positive paths $[A_2, B_3, A_3]$ and $[A_3, B_4, A_2]$, and negative paths $\overline{[A_1, B_5, A_3]}$ and $\overline{[A_2, B_3, A_1]}$. Note that we only include paths whose first half can be constructed from an edge in the corresponding fragment. We do so to preserve node multiplicity: All subgraphs constructed from a set of subgraph com-
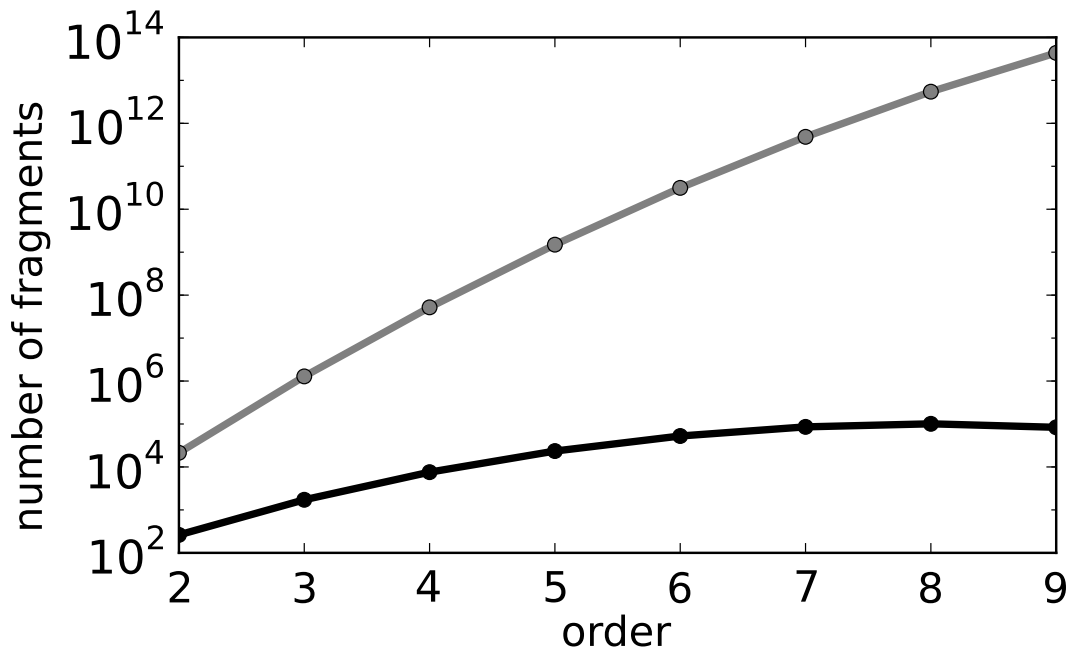
Figure 4.5: **Comparison of the Number of Fragment Candidates Generated by the two Described Algorithms.** Number of fragment candidates of different orders generated for the double-layer MAPK network (Walther et al., 2014) (i) combinatorially (grey) and (ii) generated from the unique correspondence between fragments and edges-only subgraphs (black). The latter algorithm is implemented in GraTeLPy, generates orders of magnitude fewer fragment candidates than the combinatorial algorithm, and therefore achieves a substantial decrease in computational cost. The double-layer MAPK network has $n=12$ species and $m=18$ reactions (Walther et al., 2014).

ponents need to contain the exact same species and reaction nodes with the same multiplicity as the corresponding fragment.

GraTeLPy stores the subgraph components of a fragment S_k in a lookup table which lists for each species A_{i_s} all subgraph components that lie in S_k and that A_{i_s} is the beginning of. We generate the subgraph components of fragment $S_k \binom{i_1, i_2, \ldots, i_k}{j_1, j_2, \ldots, j_k}$ as follows: (i) We extract all edges $[A_{i_s}, B_{j_s}]$ ($s=1, \ldots, k$) and store each in our lookup table keyed with A_{i_1}, (ii) each edge $[A_{i_s}, B_{j_s}]$ in the resultant table we combine with all arcs (B_{j_s}, A_{i_l}) ($l=1, \ldots, k$) that lie in both the bipartite digraph and S_k to generate all positive paths, and (iii) each edge $[A_{i_s}, B_{j_s}]$ in the table, we combine with all arcs (A_{i_l}, B_{j_s}) that lie in both the bipartite digraph and S_k to generate all negative paths. All positive and negative paths that we discover are added in our lookup table to the record belonging to species A_{i_s}.

Once our lookup table contains all subgraph components we need to combine these components to enumerate exhaustively all subgraphs contained in the corresponding fragment. Suppose that in fragment S_k species A_{i_1} induces L_{i_1} subgraph

components, A_{i_2} induces L_{i_2}, and so forth. We can attempt to generate symbolically a subgraph g of S_k by selecting at random one subgraph component per species since each species must be the starting node of exactly one component (Mincheva & Roussel, 2007). Using this approach we would generate

$$\#\{\text{subgraph candidates}\} = L_{i_1} \cdot L_{i_2} \cdots L_{i_k} \tag{4.19}$$

subgraph candidates for fragment $S_k\binom{i_1,i_2,\dots,i_k}{j_1,j_2,\dots,j_k}$.

Even though this approach guarantees that each species is the starting node of exactly one subgraph component, we generate combinations that contain paths that do not form cycles as defined in Section 4.3.2. Paths that are not part of a cycle do not contribute to valid subgraphs (Mincheva & Roussel, 2007). Hence this approach generates subgraph candidates that need to be filtered to remove those candidates that do not satisfy the definition of a subgraph given in Section 4.3.2.

While this combinatorial approach combined with a filtering step allows us to enumerate all subgraphs, this process generates unnecessary computational work and its computational cost becomes prohibitive for larger mechanisms. We develop novel methodology in the following subsections that reduces substantially the computational cost of enumerating subgraphs.

Cycle Detection: The Path Graph

We can avoid generating many invalid subgraph candidates if we do not join paths combinatorially but rather only join those paths that connect to form cycles, Section 4.3.2.

We first introduce **expanded paths**, where each negative path $\overline{[A_i, B_m, A_j]}$ is expanded into two positive paths $[A_i, B_m, A_j]$ and $[A_j, B_m, A_i]$. This is necessary in the following methodology because negative paths can be traversed in two directions. Note that we can only expand a negative path into positive paths that respect the node multiplicity of the fragment under consideration. In our example, Figure 4.3 (top left), fragment $S_3\binom{1,2,3}{5,3,4}$ contains the negative path $\overline{[A_1, B_5, A_3]}$ that can only be expanded into the positive path $[A_1, B_5, A_3]$ – since the edge $[A_3, B_5]$ is not included in $S_3\binom{1,2,3}{5,3,4}$ we cannot expand to $[A_3, B_5, A_1]$.

To enumerate all cycles in a fragment S_k we construct the directed graph (digraph) Φ, which we refer to as **path graph**. The nodes of Φ correspond uniquely to the expanded negative paths and the positive paths of fragment S_k. The links in Φ are drawn so that their origins represent paths whose endpoints are the beginning of the paths pointed to: Suppose that Φ contains nodes $[A_{i_1}, B_{j_s}, A_{i_2}]$ and $[A_{i_2}, B_{j_l}, A_{i_3}]$ then there is a link pointing from the former to the ladder node. Self-loops in Φ are also included so that a node corresponding to the path $[A_{i_1}, B_{j_1}, A_{i_1}]$ points to itself.

For illustration we show the path graph Φ that we construct for fragment $S_3\binom{1,2,3}{5,3,4}$, Figure 4.3 (top left), in Figure 4.6.

$C_1: \quad \{[A_3, B_4, A_2], [A_2, B_3, A_3]\}$
$C_2: \quad \{[A_3, B_4, A_2], \overline{[A_2, B_3, A_1]}, \overline{[A_1, B_5, A_3]}\}$
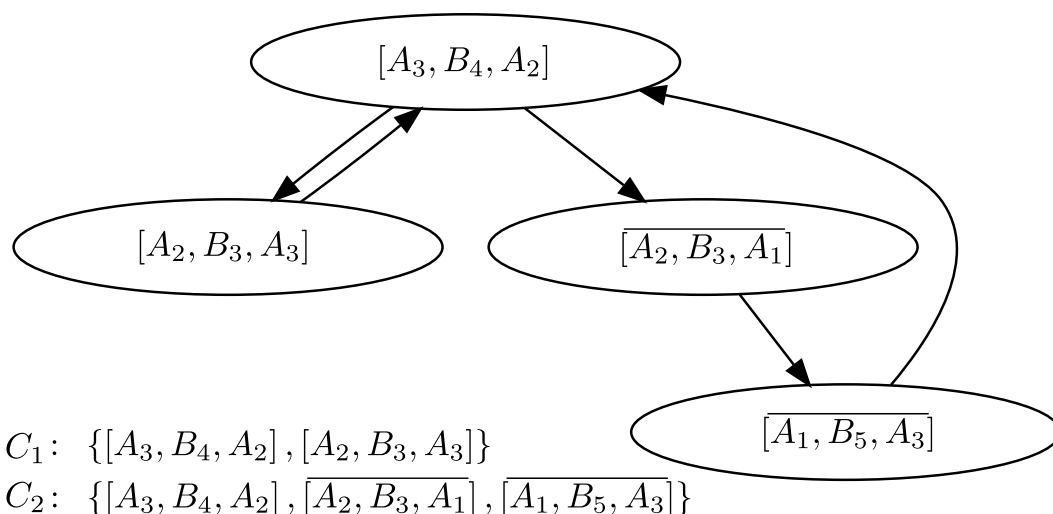
Figure 4.6: **Path Graph for the Critical Fragment of the Reversible Substrate Inhibition Mechanism.** The path graph Φ and the detected cycles that do not have repeated species as starting nodes of paths of the critical fragment $S_3\binom{1,2,3}{5,3,4}$, Figure 4.3 (top left). The two cycles C_1 and C_2 that were reported previously for this fragment by Mincheva & Roussel (2007) are discovered by our method implemented in GraTeLPy.

To detect cycles in the path graph Φ, and ultimately in the subgraphs of a given fragment S_k, we use an implementation of Johnson's algorithm (Johnson, 1975) provided by NetworkX (Hagberg et al., 2008).

Some of the cycles in Φ enumerated by Johnson's algorithm revisit species and are therefore not cycles of the bipartite graph G of the mechanism (Mincheva & Roussel, 2007). This is the case because Johnson's algorithm returns all cycles in Φ without further validation of the properties of these cycles. Computationally it is straightforward to filter out generated cycles that revisit species – however this extra filtering step incurs computational cost. Adaptation of Johnson's algorithm so that only cycles that do not revisit species are generated is likely possible and left

for future work.

Cycle Combinations: The Cycle Graph

We now have enumerated all cycles contained in a given fragment S_k. To proceed with the enumeration of all subgraphs in S_k we remind ourselves that S_k includes exactly one subgraph that contains edges only, Section 4.4.1. Therefore all subgraphs that remain to be discovered must contain at least one cycle. Hence we proceed by enumerating all possible cycle combinations in S_k – each of these combinations will then serve as the basis of one or multiple subgraph candidates.

The following combinatorial approach is easy to devise and implement and offers some intuition of the task. Suppose we discovered $|\Gamma|$ cycles in S_k then we can generate all combinations of c cycles by drawing without replacement and without regard for the order of the c draws. Combining $c=1, 2, \ldots, |\Gamma|$ cycles there are at most $\sum_{c=1}^{|\Gamma|} \binom{|\Gamma|}{c}$ possible ways to combine all $|\Gamma|$ cycles. At most since a valid combination of cycles needs to respect the species and reaction node multiplicity of S_k. For instance, in Figure 4.6 cycles C_1 and C_2 cannot be combined as this would repeat species A_2 and A_3.

To form a subgraph we amend, if necessary, each valid cycle combination with edges whose species are in S_k, but not in any member of the cycle combination. In our example, Figure 4.6, we need to amend C_1 with one edge, $[A_1, B_5]$, to generate a subgraph of order 3, Figure 4.3 (bottom right). However C_2, Figure 4.6, does not need to be amended to form a subgraph of order 3, Figure 4.3 (bottom left).

While this combinatorial approach is easy to implement its computational cost becomes prohibitive as the initial draw without replacement contributes to a large number of subgraph candidates that require subsequent verification of their node multiplicity.

To lower the computational cost, we reduce the problem of generating cycle combinations to a classical graph-theoretic problem that generates fewer subgraph candidates and can be solved with available algorithms from the literature. To this end we generate for S_k an undirected graph Γ whose nodes correspond uniquely to the $|\Gamma|$ cycles in S_k. We refer to the undirected graph Γ as **cycle graph**. We draw a line between two nodes of Γ if a combination of these two cycles respects the node mul-

tiplicity of S_k. Such a set of cycles can, if necessary, be amended with edges to form a subgraph. In classical graph theory, a **clique** is a set of nodes of an undirected graph such that every node is connected to every other node in the set. Given the cycle graph Γ we can enumerate all combinations of c cycles by generating all cliques in Γ of cardinality (node count) c. This is a standard problem in graph theory, known as clique enumeration, and can be solved using existing algorithms from the literature (Zhang et al., 2005).

Intuitively this approach of combining cycles that are linked in Γ instead of combining them by drawing without replacement has the benefit of locality: In Γ we only combine those cycles that are linked, and therefore combinable, and we do not bother combining cycles that are far from each other in Γ.

In our example, Figure 4.6, the cycle graph Γ we construct from C_1 and C_2 consists of these two cycles as unconnected nodes. Since in this example Γ contains no lines between nodes, the two cliques present are of cardinality one and equal $\{C_1\}$ and $\{C_2\}$ respectively.

Available library methods focus on solving the maximal clique problem of enumerating all or some cliques of maximal cardinality in a given undirected graph (Hagberg et al., 2008). However, for our problem we need to enumerate all cliques of all sizes. Therefore, we implemented in GraTeLPy an algorithm by Zhang et al. (2005) that enumerates all cliques of all sizes of an undirected graph.

Performance of Subgraph Enumeration

At the beginning of Section 4.4.2 we introduced a combinatorial approach for subgraph enumeration. While straightforward to implement, the computational cost of this approach is prohibitive as it generates a large number of subgraph candidates, Equation (4.19), that each require verification of their node multiplicity. In Sections 4.4.2 and 4.4.2 we split the problem of subgraph enumeration into cycle enumeration and cycle combination and described efficient, graph-based solutions for both subproblems.

Here we compare the computational cost of the combinatorial algorithm with the graph-based algorithms that we developed and implemented in GraTeLPy. To this end we count the number of subgraph candidates generated by both approaches

for 100 randomly selected fragments (of varying size) of the double-layer MAPK mechanism (Walther et al., 2014). Our results, Figure 4.7, show that the graph-based algorithms implemented in GraTeLPy generate multiple orders of magnitude fewer subgraph candidates than the naive, combinatorial algorithm.
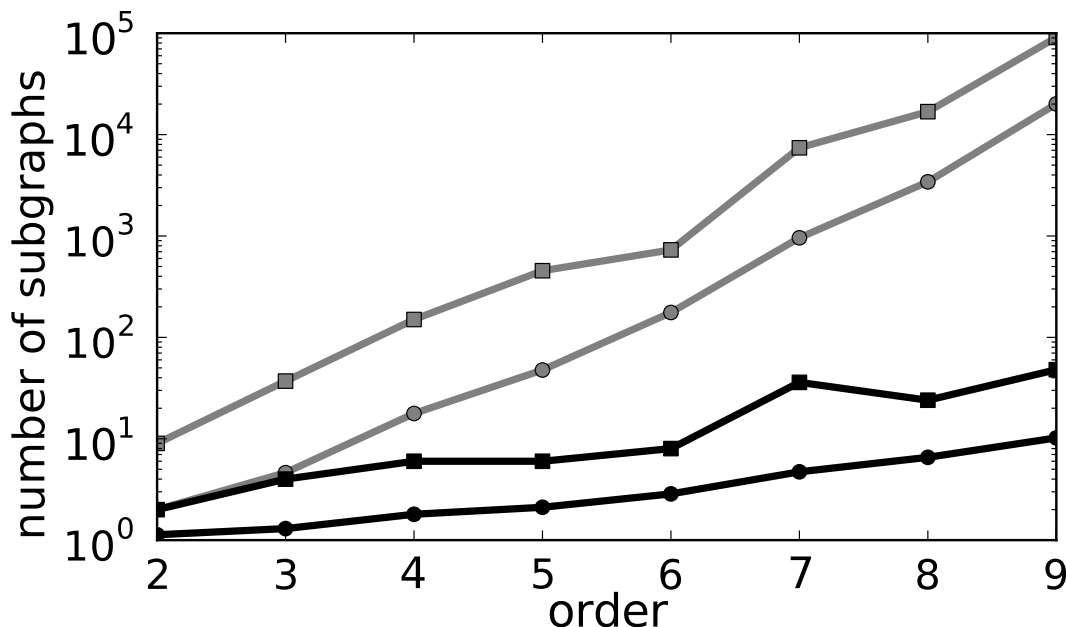
Figure 4.7: **Number of Subgraph Candidates Generated by the Combinatorial Algorithm and the Graph-Based Algorithm Implemented in GraTeLPy.** Number of subgraph candidates generated for 100 randomly selected fragments (of indicated order) of the double-layer MAPK network (Walther et al., 2014). (grey) Combinatorial approach, (black) graph-based approach implemented in GraTeLPy that uses the path graph, Section 4.4.2, and cycle graph, Section 4.4.2. Circles denote averages, squares denote maxima (maximal number of subgraph candidates generated for any one of 100 randomly selected fragments). The algorithm implemented by GraTeLPy (black) generates orders of magnitude fewer subgraph candidates than the combinatorial approach (grey) would thus greatly reducing computational cost.

4.5 Sample Application: Cdc42 in Yeast

Our motivation for developing GraTeLPy was our interest in unfurling a phenomenological mechanism of cell polarity to study the spatial distribution of the involved membrane-bound protein species, Section 4.2. To this end we developed and optimised GraTeLPy so that it can serve as a software tool for rapid mechanism discrimination.

In this section we discuss a sample analysis of a related chemical mechanism with the current implementation of GraTeLPy and point to future extensions of GraTeLPy that will offer a more complete analysis pipeline.

As we mentioned in Chapter 1, *Saccharomyces cerevisiae* (budding yeast) divides by formation of a bud that is separated from the parent cell through mitosis. Bud formation is preceded by the localisation of active Cdc42, a Rho homologue, which directs outgrowth of the bud (Slaughter et al., 2009). A mathematical model of Cdc42 localisation in budding yeast described by Goryachev & Pokhilko (2008) relies on a Turing-type mechanism, Chapter 1. While its underlying mathematical mechanism and, therefore, predicted localisation patterns differ from the wave-pinning mechanism, the molecular species and reactions described in the model of Goryachev & Pokhilko (2008) bear close resemblance to those in our motivational example, Section 4.2.

4.5.1 Cdc42 Network in Yeast

Goryachev & Pokhilko (2008) have proposed the following reaction network as a model of the dynamics of Cdc42 in budding yeast:

$$
\begin{aligned}
B_1 &: \quad E_c \xrightarrow{k_1} E_m, \\
B_2 &: \quad E_m \xrightarrow{k_2} E_c, \\
B_3 &: \quad RT + E_m \xrightarrow{k_3} M, \\
B_4 &: \quad M \xrightarrow{k_4} RT + E_m, \\
B_5 &: \quad E_m + RD \xrightarrow{k_5} E_m + RT, \\
B_6 &: \quad M + RD \xrightarrow{k_6} M + RT, \\
B_7 &: \quad RT \xrightarrow{k_7} RD, \\
B_8 &: \quad E_c + RT \xrightarrow{k_8} M, \\
B_9 &: \quad RDI_m \xrightarrow{k_9} I + RD, \\
B_{10} &: \quad RD + I \xrightarrow{k_{10}} RDI_m, \\
B_{11} &: \quad RDI_c \xrightarrow{k_{11}} RDI_m, \\
B_{12} &: \quad RDI_m \xrightarrow{k_{12}} RDI_c,
\end{aligned}
\tag{4.20}
$$
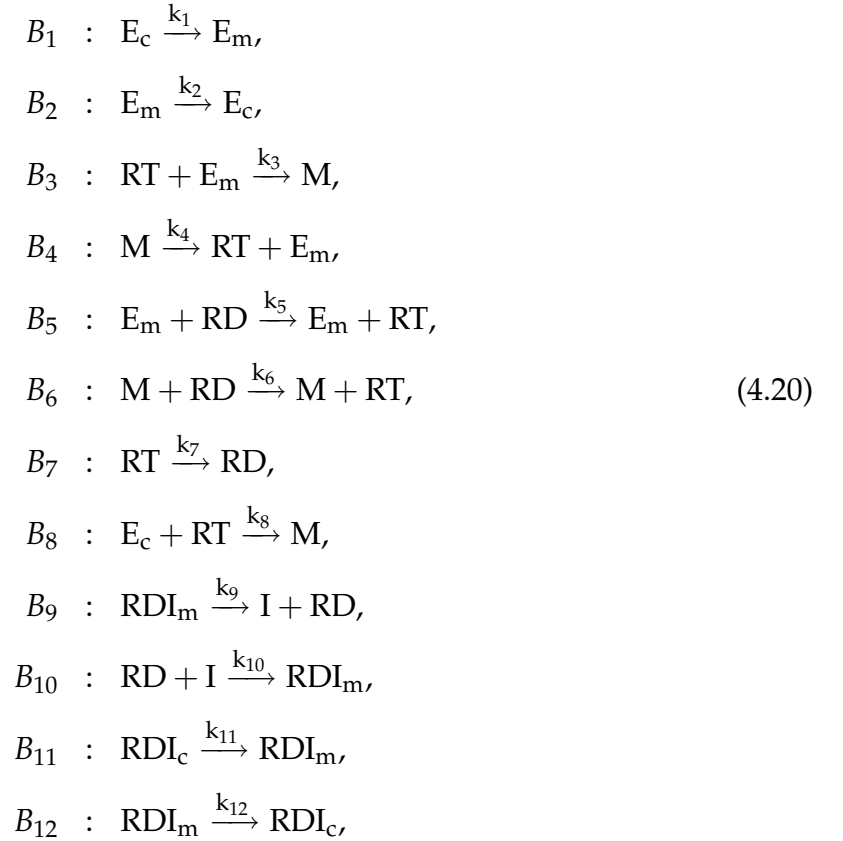
where RD and RT denote the membrane-bound inactive and active form of Cdc42

respectively, I denotes cytoplasmic GDI that forms a membrane-bound complex with RD, RDI_m, that detaches from the membrane and diffuses as RDI_c in the cytoplasm. All of these species have equivalents in the mechanistic model described earlier, Figure 4.1 (left): RD corresponds to Rho-GDP, RT to Rho-GTP, I to GDI, and the two RDI species are equivalent to membrane-bound and cytoplasmic GDI-Rho-GDP complexes.

The enzyme E is a complex that contains Cdc42-activating Cdc24 and exists in both a cytoplasmic and membrane-bound form, E_c and E_m, respectively. Cdc24 corresponds to the more generic GEF in Figure 4.1 (left). However, complex E also contains a yeast-specific adaptor protein Bem1 (Goryachev & Pokhilko, 2008), hence species and reactions that depend on E are unlikely to play a role in more generic mechanistic models, Figure 4.1 (left). When E is on the membrane, it can form a catalytic complex M together with RT, that aids activation of membrane-bound RD.

The bipartite digraph of the Cdc42 mechanism is shown in Figure 4.8 and the mechanism has a stoichiometric matrix of rank $r=5$. This mechanism, Equation (4.20), shows Cdc42 localisation through a Turing-type mechanism (Goryachev & Pokhilko, 2008) and GraTeLPy can inform the user about necessary conditions pertaining to this type of instability (Walther et al., 2014). However, for brevity we limit our discussion to multistability of this mechanism: Both multistability of reaction mechanisms (Semplice et al., 2012) and spatially extended mechanisms in a diffusion limit (Walther et al., 2012; Holmes, 2014) have been shown to give rise to protein localisation. To this end we limit our analysis to fragments S_k of order $k=r=5$.

4.5.2 Fragment Detection with GraTeLPy

GraTeLPy detects 35 critical fragments of order five among which we find the two critical fragments reported by Goryachev & Pokhilko (2008) and shown in Figure 4.9. Overall, GraTeLPy detects 152 fragments that each contribute one non-zero summand to the corresponding constant, a_5, of the characteristic polynomial of the Jacobian, Equation (4.17). The median running time of this analysis with one processor is 9.7 seconds and with two processors 6.1 seconds.
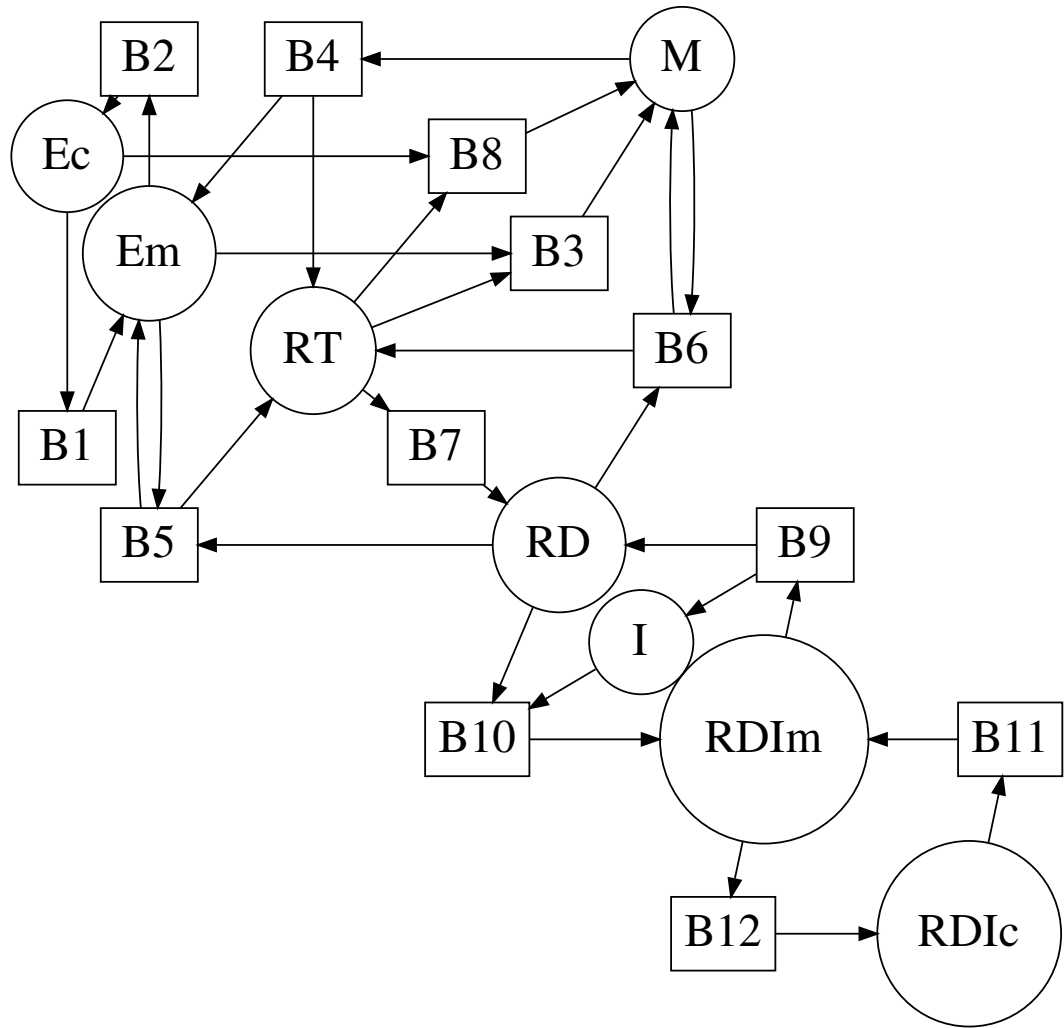
Figure 4.8: **Bipartite Digraph of the Yeast Cdc42 Network Described by Goryachev & Pokhilko (2008).** RD and RT are the membrane-bound inactive and active form of Cdc42 respectively, I is cytoplasmic GDI that forms a membrane-bound complex with RD, RDI_m, which detaches from the membrane and diffuses as RDI_c in the cytoplasm. The enzyme E is a complex that contains Cdc42-activating Cdc24 and exists in both a cytoplasmic and membrane-bound form, E_c and E_m, respectively. When E is on the membrane, it can form a catalytic complex M together with RT, that aids activation of membrane-bound RD.

4.5.3 Putting the Fragments to Use

GraTeLPy enumerates all fragments that contribute non-zero terms to the corresponding coefficient of the characteristic polynomial, Equation (4.17): For the Cdc42 mechanism, Equation (4.20), GraTeLPy enumerates 152 such fragments of order $k=r=5$: Of these, 35 are critical and 117 are non-critical fragments, therefore a_5 is a sum over 35 negative summands and 117 positive summands.

We remind ourselves that we are interested in the behaviour of the steady-states of the Cdc42 mechanism which we can express by adapting Equation (4.4):
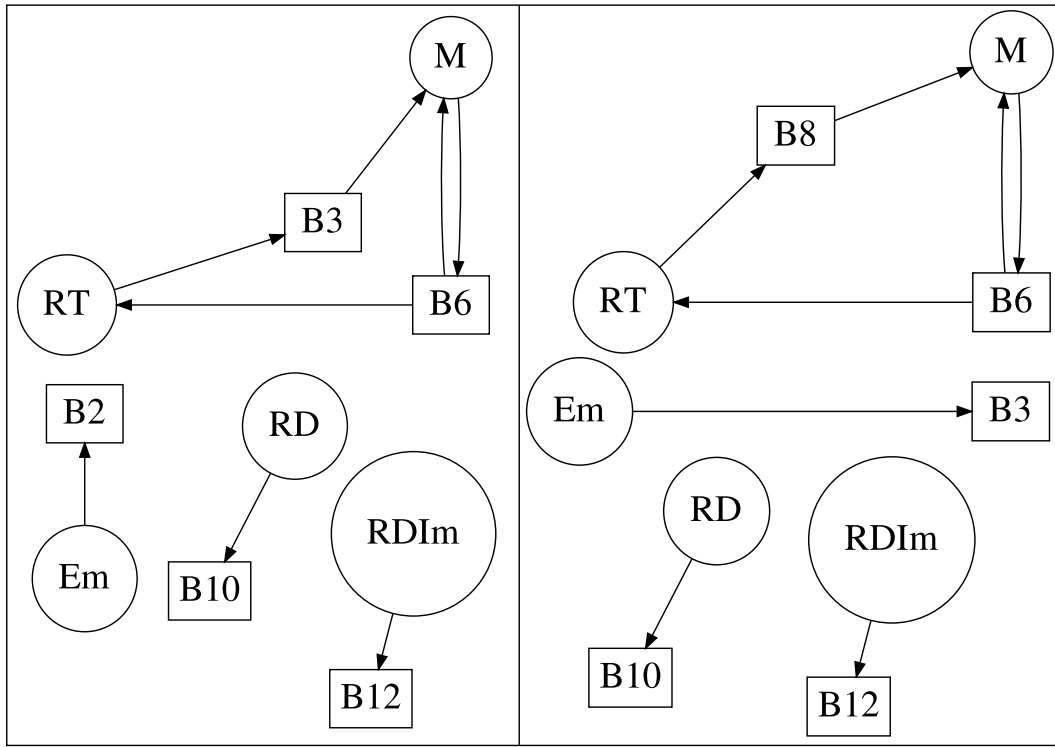
Figure 4.9: **Critical Fragments of the Yeast Cdc42 Mechanism.** Critical fragments of order 5 of the yeast Cdc42 network reported by Goryachev & Pokhilko (2008) and reproduced by GraTeLPy. These critical fragments show that a key component of this mechanism, Equation (4.20), is the catalytic reactivity of M: The membrane-bound complex M upregulates production of RT but is itself not consumed in this reaction. It is this catalytic reactivity that has the potential to generate multistability and therefore contribute to protein localisation (Semplice et al., 2012).

$$\mathbf{0} = S\mathbf{w}^*, \tag{4.21}$$

where S is the stoichiometric matrix of Equation (4.20) and all flux rates are non-negative

$$w_j^* \geq 0, \ j = 1, \ldots, 12. \tag{4.22}$$

The 12 non-negativity conditions in Equation (4.22) and the 8 steady-state flux conditions in Equation (4.21) constrain the permissible steady-state flux vectors $\mathbf{w}^*$ to a convex, polyhedral cone that extends into the positive orthant of flux space (Clarke, 1988; Orth et al., 2010). Flux space is the space that rate vectors $\mathbf{w}$ lie in and in our example, Equation (4.20), this space has dimension 12. This cone is generally referred to as the allowable solution space (Orth et al., 2010) and is polyhedral because its base is a polygon.

The fragments generated by GraTeLPy allow us to create a symbolic represen-

tation of a_5, based on Equation (4.17), that is a sum of terms involving the steady-state flux rates w_j^* and steady-state concentrations u_i^*. If we can generate a set of permissible values for w_j^* and u_i^* so that the corresponding $a_5(\mathbf{u}^*, \mathbf{w}^*) < 0$ then the steady state $(\mathbf{u}^*, \mathbf{w}^*)$ is unstable (Clarke, 1988). Given these 12 steady-state flux rates $w_1^*, \ldots, w_{12}^*$ and 8 steady-state concentration values $u_1^*, \ldots, u_8^*$ we can then compute the 12 reaction rate constants $k_1, \ldots, k_{12}$. Further, since this mechanism, Equation (4.20), has 8 species but is of rank $r=5$ there are 3 conservation constants that we can also compute from the above choice for $\mathbf{u}^*$ and $\mathbf{w}^*$. We are then in a position to vary numerically one of the rate constants k_i to detect the occurrence of the predicted saddle-node bifurcation ($a_5=0$) (Clarke, 1988).

Computing permissible values w_j^* and u_i^* so that $a_5(\mathbf{u}, \mathbf{w}) < 0$ is not trivial and we propose two approaches that may be explored in future endeavours. Here we assume that we either have measurements or literature estimates of the steady-state concentrations, $\mathbf{u}^*$, and that our main goal here is to compute rate constants k_j so that the Cdc42 mechanism, Equation (4.20), shows the predicted saddle-node bifurcation. The rate constants k_j can be computed with Equation (4.3).

Non-Linear Optimisation Inside the Flux Cone

Generally, the problem of computing vectors $\mathbf{w}^*$ so that $a_5(\mathbf{u}^*, \mathbf{w}^*) < 0$ may be regarded as a minimisation problem where we minimise a_5 over flux space.

Before we can employ one of multiple standard strategies to minimise a_5, we require a fast method for generating vectors $\mathbf{w}^*$ that lie in the polyhedral cone defined by Equations (4.21) and (4.22).

A brute force approach where we generate vectors $\mathbf{w}^*$ in the positive orthant of flux space randomly is unlikely to perform well since in 12-dimensional space we are unlikely to hit random points so that both $S\mathbf{w}^*=0$ and $a_5<0$.

A more viable approach is to sample points $\mathbf{w}^*$ directly inside the polyhedral cone spanned by the allowable solutions space (Clarke, 1988). To sample a_5 inside this cone we need the basis vectors that span the cone. Construction of these basis vectors is an important problem known in the field of systems biology as extreme current enumeration and software exists that can aid in this (Orth et al., 2010).

Once we can sample repeatedly points $\mathbf{w}^*$ inside this cone, we can employ a

variety of strategies to minimise a_5. Since *a priori* a_5 does not possess an obvious structure over the cone we suggest to explore the use of genetic algorithms for this optimisation problem.

We leave expansion of GraTeLPy to include extreme current enumeration and optimisation for future work.

Linear Optimisation Inside the Flux Cone

Instead of minimising the full expression of a_5 an alternative approach is to maximise the negative summands of a_5. To help with this task it would be beneficial if we could identify variables w_j that appear predominantly or exclusively in the negative summands of a_5. For simplicity, suppose that we identified flux rates $w_{j_1}^-, \ldots, w_{j_n}^-$ that appear exclusively in the negative summands of a_5. Then to minimise a_5 we can maximise the following linear cost function:

$$a_c(\mathbf{w}) = \sum_{m=1}^{n} w_{j_m}^-. \tag{4.23}$$

Maximising $a_c(\mathbf{w}^*)$ over flux space with the constraints $S\mathbf{w}^* = \mathbf{0}$ and $w_j^* \geq 0$ defines a linear program which can be solved with well-established computer software.

The output generated by GraTeLPy makes it straightforward to test whether any variables w_j appear exclusively in the negative summands of a_5. In general the summands of a coefficient a_k can be thought of as products of ratios of the elements of the edges in the corresponding fragments, Equation (4.17). We can therefore look for variables w_j^- that occur exclusively in negative summands by identifying those edges that appear exclusively in critical fragments. GraTeLPy identifies the following set of edges in the negative summands of a_5

$$
\begin{aligned}
E_n \;=\; & \{ [E_c, B_1], [E_c, B_8], [E_m, B_2], [E_m, B_3], \tag{4.24} \\
& [E_m, B_5], [I, B_{10}], [M, B_4], [M, B_6], \\
& [RD, B_{10}], [RDI_c, B_{11}], [RDI_m, B_{12}], \\
& [RDI_m, B_9], [RT, B_3], [RT, B_8] \},
\end{aligned}
$$

and the following set of edges in the positive summands of a_5

$$E_p = \{[E_c, B_1], [E_c, B_8], [E_m, B_2], [E_m, B_3], \qquad (4.25)$$
$$[E_m, B_5], [I, B_{10}], [M, B_4], [RD, B_{10}],$$
$$[RD, B_5], [RD, B_6], [RDI_c, B_{11}], [RDI_m, B_{12}],$$
$$[RDI_m, B_9], [RT, B_3], [RT, B_7], [RT, B_8]\}.$$

The difference between these two sets reveals one edge that occurs exclusively in the negative summands of a_5:

$$E_n \backslash E_p = [M, B_6]. \qquad (4.26)$$

Given this edge, Equation (4.26), our linear cost function becomes $a_c(\mathbf{w}) = w_6$ as the node B_6 corresponds to reaction rate w_6. This allows us to formulate the following linear program:

$$
\begin{aligned}
\text{maximise} \quad & \mathbf{e}_6 \cdot \mathbf{w}, \\
\text{subject to} \quad & S\mathbf{w} \geq \mathbf{0}, \\
& -S\mathbf{w} \geq \mathbf{0}, \\
& \mathbf{w} \geq \mathbf{0},
\end{aligned}
\qquad (4.27)
$$

where $\mathbf{e}_6$ is the unit vector whose sixth element equals one and all other elements equal zero, and we encode the equality $S\mathbf{w} = \mathbf{0}$ with the two inequalities $S\mathbf{w} \geq \mathbf{0}$ and $-S\mathbf{w} \geq \mathbf{0}$ which are easier to handle for numerical algorithms. Once we have found a maximal value for w_6 under the indicated constraints we can then test whether this value and appropriately chosen values for the remaining flux rates w_j, $j \neq 6$, render a_5 negative.

We note that an advantage of linear programs is that they define naturally constrained optimisation problems: The constraints $\mathbf{w} \geq \mathbf{0}$ and $S\mathbf{w} = \mathbf{0}$ are natural parts of the linear program in Equation (4.27) and we therefore do not need sample specifically flux rates $\mathbf{w}$ inside the polyhedral cone.

We leave implementation and validation of this method for future endeavours.

4.6   Discussion

In GraTeLPy we have implemented a graph-theoretic method that allows for parameter-free mechanism discrimination.

While our discussion here focused on discovering mechanisms with multistability through saddle-node bifurcation, conditions for other types of bifurcations leading to oscillations and Turing instability can also be detected with GraTeLPy (Walther et al., 2014). A promising route here is also a combination of GraTeLPy with the local perturbation analysis (LPA) method discussed in Chapters 2 and 3 as well as in Holmes (2014); Walther et al. (2012). In a diffusivity limit the LPA splits the reaction network of a given mechanism into global and local species, and bistability of the resulting mechanism is usually associated with pattern formation and cell polarity (Holmes, 2014; Walther et al., 2012). In GraTeLPy the bipartite digraph can be expanded similarly and an analysis of the resulting expanded graph is likely equivalent to the results obtained with the LPA.

As opposed to similar software packages, GraTeLPy is both open-source and is based on free software. Further, GraTeLPy enables users to identify the specific elements of mechanisms that can lead to certain types of bifurcations. This level of detail can provide mechanistic understanding of bifurcations and can aid in determining critical parameter values.

We plan to leverage this level of detail in future extensions of GraTeLPy so that critical parameter values for which certain types of bifurcations occur are computed automatically. We have discussed two possible strategies to tackle this problem, one based on non-linear optimisation inside a high-dimensional cone and another based on linear programming.

The mathematical framework implemented in the current version of GraTeLPy assumes that the rates of all chemical reactions follow the law of mass action (Mincheva & Roussel, 2007). An extension of GraTeLPy to a framework based on multigraphs would lift this restriction (Mincheva & Craciun, 2008). This extension would enable rapid model discrimination of reaction networks that are traditionally large and assume reaction rates different from mass action, such as gene regulatory networks (Karlebach & Shamir, 2008).

Chapter 5

Discussion and Outlook

In this thesis I studied the interplay between morphology and cell polarity. I focused on an existing, phenomenological model of a class of proteins that lie at the basis of many biological phenomena of polarity, Chapter 1: The wave-pinning mini-model of Rho dynamics described by Mori et al. (2008) lends itself well to numerical and analytic studies and has proved amenable to the specific cellular conditions I am interested in.

My work was motivated by my survey of the range of cell morphologies and the mathematical models that are used commonly to describe polarity-inducing biochemical processes, Chapter 1. As my survey outlined there is a vast size range in effective reaction volumes that emerge in different cell types and even within individual cells.

To the best of my knowledge, there have not been any attempts at describing the influence of cell size and cell size change on the biochemical reactions of the wave-pinning mini-model (Mori et al., 2008) and resulting influence on predicted cell polarity behaviour.

With my work in this thesis I filled a gap in the literature by investigating the influence of extremely small, extremely large, and growing reaction volumes on the cell polarity behaviour predicted by the wave-pinning mini-model (Mori et al., 2008). I further presented a novel computational tool that helps to unfurl phenomenological terms in mathematical models to precise mechanisms: In future endeavours this tool can enhance our molecular understanding of the interplay between cell morphology and cell polarity.

5.1 Discussion

My work showed that subjecting the wave-pinning mini-model to sufficiently small reaction volumes inhibits formation and maintenance of cell polarity due to stochastic effects, Chapter 2. However, my work also demonstrated that a medium level of stochasticity can also contribute to fine-tuning of cell polarity, Chapter 2.

I have further shown that large and growing reaction volumes prevent cell polarity through dilution of the involved biochemical species, Chapter 3. My work in this area also demonstrated that dilution can be countered by balancing growth and protein production. My results show that tight coupling of growth and production can sustain constant growth, while loose coupling can only sustain oscillatory growth.

To the best of my knowledge, my work on the interplay between the wave-pinning mini-model and cell morphology is the first study that investigates the wave-pinning mechanism under these cellular conditions. Here I contrast my results with the behaviour of other mathematical models of cell polarity under similar cellular conditions as described in the literature.

Meyers et al. (2006) performed a computational study in which they investigated the behaviour of a LEGI model, Chapter 1, in different cell shapes and sizes. Their model consisted of a membrane-bound receptor and a cytoplasmic signal that is activated by the receptor, diffuses away from the point of activation, and is inactivated by an intracellular reaction. This induces an intracellular gradient in the active form of the signal with its maximum at the membrane where the receptor is located. The distance over which the concentration in the active signal decays depends on its diffusivity and rate of inactivation (Meyers et al., 2006)

$$L_{\text{gradient}} = \sqrt{\frac{D}{k}}, \tag{5.1}$$

where D is the diffusion coefficient of the active signal and k is the rate constant of the first-order reaction that inactivates the signal. Assuming that a sufficiently high concentration of the signal is necessary to upregulate target pathways inside the cell, Meyers et al. (2006) found that pathways in thin and small cells are more likely to get activated than in large cells. As an example of the effect of differences in the

effective reaction volume within cells, Meyers et al. (2006) pointed out that the small height of the leading edge of migrating cells contributes locally to the abundance of cytoplasmic signal and therefore upregulation of downstream pathways.

The results obtained by Meyers et al. (2006) for their LEGI model stand in contrast to my observations with the wave-pinning model under similar cellular conditions. While Meyers et al. (2006) predict an increase in intracellular signalling in small reaction volumes, my observations in Chapter 2 suggested loss of intracellular signalling below a threshold volume. Interestingly, the observed change in intracellular signalling was gradual in the study by Meyers et al. (2006) whereas it happened suddenly at a critical point in my work in Chapter 2. At the other extreme, at large cell sizes, there is more agreement between the work by Meyers et al. (2006) and my work in Chapter 3: Under this cellular condition Meyers et al. (2006) predicted low intracellular signalling near the centre of the cell, while I also predicted loss of intracellular signalling. It is interesting to point out that low intracellular signalling in large cells is due to the interplay between diffusion, cell size, and signal inactivation in the LEGI model (Meyers et al., 2006) whereas it is due to dilution of the biochemical species in the wave-pinning model, Chapter 3.

Crampin et al. (1999) studied a Turing-type mechanism of cell polarity in growing cells. Turing-type mechanisms are susceptible to displaying multiple concentration maxima unless the diffusivity of the involved biochemical species and the size of the simulated cell are matched closely (Jilkine & Edelstein-Keshet, 2011). In agreement with this property, Crampin et al. (1999) observed the emergence of multiple maxima out of one maximum in their simulations of cell growth. Importantly, Crampin et al. (1999) did not observe loss of intracellular signalling but rather an emergence of multiple foci of signalling.

The observations made by Crampin et al. (1999) for a Turing-type mechanism in a growing cell stand in contrast with my observations for the wave-pinning mechanism. I observed that growth diluted the relevant biochemical species and therefore prevented cell polarity, Chapter 3, whereas growth induces multiple peaks for a Turing-type mechanism (Crampin et al., 1999). It is important to point out that Turing-type mechanisms do generally not conserve protein mass and therefore growth increases the amount of available protein in the work by Crampin et al.

(1999).

5.2   Outlook

While in this thesis I focused on the interplay between cell size and cell polarity, in future endeavours it will be interesting to include cell shape as an important manifestation of cell morphology. Important examples where intricate cell shapes and cell polarity occur concurrently include pavement cells where multiaxial shapes develop under guidance of localised Rho-homologues (Fu et al., 2005), and pollen tubes where molecular exchange between a dome-shaped apex and a straight base is believed to sustain cell polarity under growth (Kost, 2008).

As the molecular exchange suggested by Kost (2008) indicated, to further understand the interplay between cell shape and cell polarity it is likely important to better understand the molecular species and events involved in maintaining cell polarity. In the pollen tube, a fuller picture of these species and events would likely enable us to better understand the observed separation of involved molecular species (Kost, 2008) and would open new venues for matching theory and molecular experiments. A molecular model of localisation of a Rho homologue in the growing pollen tube tip would permit more detailed predictions for mutant behaviour that could then be used to drive an iterative process of simulation and experiment. To this end, I suggest that the model discrimination software I developed in Chapter 4 would be a useful tool for unfurling existing phenomenological models of cell polarity and could be part of a pipeline that involves model prediction and experimental validation.

Bibliography

Altschuler, S. J., Angenent, S. B., Wang, Y., & Wu, L. F. (2008). On the spontaneous emergence of cell polarity. *Nature*, *454*, 886–9.

Amin, M. R., & Roussel, M. R. (2014). Graph-theoretic analysis of a model for the coupling between photosynthesis and photorespiration. *Can. J. Chemistry*, *92*, 85–93.

Andrews, S. S., & Bray, D. (2004). Stochastic simulation of chemical reactions with spatial resolution and single molecule detail. *Phys. Biol.*, *1*, 137–51.

Britton, N. (1985). Travelling wave front solutions of a differential-difference equation arising in the modelling of myelinated nerve axon. In B. Sleeman & R. Jarvis (Eds.), *Ordinary and Partial Differential Equations*, vol. 1151 of *Lecture Notes in Mathematics*, pp. 77–89. Springer Berlin / Heidelberg.

Charest, P. G., & Firtel, R. A. (2007). Big roles for small GTPases in the control of directed cell movement. *Biochem. J.*, *401*, 377–90.

Chen, Z., Lessey, E., Berginski, M. E., Cao, L., Li, J., Trepat, X., Itano, M., Gomez, S. M., Kapustina, M., Huang, C., Burridge, K., Truskey, G., & Jacobson, K. (2013). Gleevec, an Abl family inhibitor, produces a profound change in cell shape and migration. *PLOS One*, *8*, e52233.

Clarke, B. L. (1988). Stoichiometric network analysis. *Cell Biophys.*, *12*, 237–53.

Conradi, C., Flockerzi, D., Raisch, J., & Stelling, J. (2007). Subnetwork analysis reveals dynamic features of complex (bio)chemical networks. *P. Natl. Acad. Sci. USA*, *104*, 19175–80.

Craciun, G., & Feinberg, M. (2006). Multiple equilibria in complex chemical reaction networks. II. *SIAM J. Appl. Math., 66*, 1321–38.

Crampin, E. J., Gaffney, E. A., & Maini, P. K. (1999). Reaction and diffusion on growing domains: scenarios for robust pattern formation. *Bull. Math. Biol., 61*, 1093–120.

Crampin E. J. (2000). *Reaction-Diffusion Patterns on Growing Domains*. Ph.D. thesis, University of Oxford.

Cross, F. R. (1988). DAF1, a mutant gene affecting size control, pheromone arrest, and cell cycle kinetics of Saccharomyces cerevisiae. *Mol. Cell. Biol., 8*, 4675–84.

Ellis, R. J. (2001). Macromolecular crowding: an important but neglected aspect of the intracellular environment. *Curr. Opin. Struct. Biol., 11*, 114–9.

Ellison, P., & Feinberg, M. (2000). How catalytic mechanisms reveal themselves in multiple steady-state data. I. *J. Mol. Catal. A-Chem., 154*, 155–67.

Erban, R., Chapman, J., & Maini, P. (2007). A practical guide to stochastic simulations of reaction-diffusion processes. *arxiv:0704.1908v2*.

Erban, R., & Chapman, S. J. (2007). Reactive boundary conditions for stochastic simulations of reaction-diffusion processes. *Phys. Biol., 4*, 16–28.

Erban, R., & Chapman, S. J. (2009). Stochastic modelling of reaction-diffusion processes: algorithms for bimolecular reactions. *Phys. Biol., 6*, 046001.

Fáth, G. (1998). Propagation failure of traveling waves in a discrete bistable medium. *Physica D, 116*, 176–90.

Franklin-Tong, V. (1999). Signaling and the modulation of pollen tube growth. *Plant Cell, 11*, 727–38.

Fu, Y., Gu, Y., Zheng, Z., Wasteneys, G., & Yang, Z. (2005). Arabidopsis interdigitating cell growth requires two antagonistic pathways with opposing action on cell morphogenesis. *Cell, 120*, 687–700.

Gamba, A., De Candia, A., Di Talia, S., Coniglio, A., Bussolino, F., & Serini, G. (2005). Diffusion-limited phase separation in eukaryotic chemotaxis. *Proc. Natl. Acad. Sci. U.S.A.*, *102*, 16927–32.

Gantmacher, F. (2000). *Matrix Theory Vol. 1*. Chelsea Publishing Company.

Geitmann, A. (2010). How to shape a cylinder: pollen tube as a model system for the generation of complex cellular geometry. *Sex. Plant Reprod.*, *23*, 63–71.

Gierer, A., & Meinhardt, H. (1972). A theory of biological pattern formation. *Kybernetik*, *12*, 30–9.

Gillespie, D. (1976). A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comput. Phys.*, *22*, 403–34.

Goryachev, A., & Pokhilko, A. (2008). Dynamics of cdc42 network embodies a turing-type mechanism of yeast cell polarity. *FEBS Lett.*, *582*, 1437–43.

Grima, R. (2010). Intrinsic biochemical noise in crowded intracellular conditions. *J. Chem. Phys.*, *132*, 185102.

Gu, Y., Fu, Y., Dowd, P., Li, S., Vernoud, V., Gilroy, S., & Yang, Z. (2005). A Rho family GTPase controls actin dynamics and tip growth via two counteracting downstream pathways in pollen tubes. *J. Cell. Biol.*, *169*, 127–38.

Hagberg, A. A., Schult, D. A., & Swart, P. J. (2008). Exploring network structure, dynamics, and function using NetworkX. In *Proceedings of the 7th Python in Science Conference (SciPy2008)*, pp. 11–15.

Harary, F. (1969). *Graph Theory*. Addison-Wesley, Reading, MA.

Hepler, P. K., Lovy-Wheeler, A., McKenna, S. T., & Kunkel, J. G. (2006). Ions and pollen tube growth. In *The Pollen Tube*, vol. 3 of *Plant Cell Monographs*, pp. 47–69. Springer Berlin Heidelberg.

Holdaway-Clarke, T. L., Feijo, J. A., Hackett, G. R., Kunkel, J. G., & Hepler, P. K. (1997). Pollen Tube Growth and the Intracellular Cytosolic Calcium Gradient Oscillate in Phase while Extracellular Calcium Influx Is Delayed. *Plant Cell*, *9*, 1999–2010.

Holdaway-Clarke, T. L., & Hepler, P. K. (2003). Control of pollen tube growth: role of ion gradients and fluxes. *New Phytol.*, *159*, 539–63.

Holmes, W. R. (2014). An efficient, nonlinear stability analysis for detecting pattern formation in reaction diffusion systems. *Bull. Math. Biol.*, *76*, 157–83.

Holmes, W. R., Carlsson, A. E., & Edelstein-Keshet, L. (2012a). Regimes of wave type patterning driven by refractory actin feedback: transition from static polarization to dynamic wave behaviour. *Phys. Biol.*, *9*, 1–21.

Holmes, W. R., Lin, B., Levchenko, A., & Edelstein-Keshet, L. (2012b). Modelling cell polarization driven by synthetic spatially graded Rac activation. *PLOS Comput. Biol.*, *8*, e1002366.

Houk, A. R., Jilkine, A., Mejean, C. O., Boltyanskiy, R., Dufresne, E. R., Angenent, S. B., Altschuler, S. J., Wu, L. F., & Weiner, O. D. (2012). Membrane Tension Maintains Cell Polarity by Confining Signals to the Leading Edge during Neutrophil Migration. *Cell*, *148*, 175–88.

Howard, M., & Rutenberg, A. (2003). Pattern Formation inside Bacteria: Fluctuations due to the low copy number of proteins. *Phys. Rev. Lett.*, *90*, 128102.

Hucka, M., & *et al.* (2003). The systems biology markup language (sbml): a medium for representation and exchange of biochemical network models. *Bioinformatics*, *19*, 524–31.

Hwang, J. U., Gu, Y., & Lee, Y. J. (2005). Oscillatory ROP GTPase activation leads the oscillatory polarized growth of pollen tubes. *Mol. Biol. Cell*, *16*, 5385–99.

Hwang, J.-U., Wu, G., Yan, A., Lee, Y.-J., Grierson, C. S., & Yang, Z. (2010). Pollen-tube tip growth requires a balance of lateral propagation and global inhibition of Rho-family GTPase activity. *J. Cell Sci.*, *123*, 340–50.

Isaacson, S. A., McQueen, D. M., & Peskin, C. S. (2011). The influence of volume exclusion by chromatin on the time required to find specific DNA binding sites by diffusion. *Proc. Natl. Acad. Sci. U.S.A.*, *108*, 3815–20.

Jilkine, A., & Edelstein-Keshet, L. (2011). A comparison of mathematical models for polarization of single eukaryotic cells in response to guided cues. *PLOS Comput. Biol.*, *7*, e1001121.

Jilkine, A., Marée, A. F. M., & Edelstein-Keshet, L. (2007). Mathematical model for spatial segregation of the Rho-family GTPases based on inhibitory crosstalk. *Bull. Math. Biol.*, *69*, 1943–78.

Johnson, D. (1975). Finding all the elementary circuits of a directed graph. *SIAM J. Comput.*, *4*, 77–84.

Johnson, D. I. (1999). Cdc42: an essential Rho-type GTPase controlling eukaryotic cell polarity. *Microbiol. Mol. Biol. Rev.*, *63*, 54–105.

Karlebach, G., & Shamir, R. (2008). Modelling and analysis of gene regulatory networks. *Nat. Rev. Mol. Cell Bio.*, *9*, 770–80.

Keener, J. P. (1987). Propagation and Its Failure in Coupled Systems of Discrete Excitable Cells. *SIAM J. Appl. Math.*, *47*, 556–72.

Khain, E., Lin, Y. T., & Sander, L. M. (2011). Fluctuations and stability in front propagation. *Europhys. Lett.*, *93*, 28001.

Kost, B. (2008). Spatial control of Rho (Rac-Rop) signaling in tip-growing plant cells. *Trends Cell Biol.*, *18*, 119–27.

Kroeger, J. H., Zerzour, R., & Geitmann, A. (2011). Regulator or driving force? The role of turgor pressure in oscillatory plant cell growth. *PLOS One*, *6*, e18549.

Kuznetsov, Y. (1998). *Elements of Applied Bifurcation Theory*, vol. 112. Springer.

Levchenko, A., & Iglesias, P. A. (2002). Models of eukaryotic gradient sensing: application to chemotaxis of amoebae and neutrophils. *Biophys. J.*, *82*, 50–63.

Li, H. (1999). Control of Pollen Tube Tip Growth by a Rop GTPase Dependent Pathway That Leads to Tip-Localized Calcium Influx. *Plant Cell Online*, *11*, 1731–42.

Li, Z., Hannigan, M., Mo, Z., Liu, B., Lu, W., Wu, Y., Smrcka, A. V., Wu, G., Li, L., & Liu, M. (2003). Directional Sensing Requires G$\beta\gamma$-Mediated PAK1 and PIXα-Dependent Activation of Cdc42. *Cell*, *114*, 215–27.

Lin, Y., & Yang, Z. (1997). Inhibition of Pollen Tube Elongation by Microinjected Anti-Rop1Ps Antibodies Suggests a Crucial Role for Rho-Type GTPases in the Control of Tip Growth. *Plant Cell*, *9*, 1647–59.

Marée, A. F. M., Grieneisen, V. A., & Edelstein-Keshet, L. (2012). How cells integrate complex stimuli: the effect of feedback from phosphoinositides and cell shape on cell polarization and motility. *PLOS Comput. Biol.*, *8*, e1002402.

Marée, A. F. M., Jilkine, A., Dawes, A., Grieneisen, V. A., & Edelstein-Keshet, L. (2006). Polarization and movement of keratocytes: a multiscale modelling approach. *Bull. Math. Biol.*, *68*, 1169–211.

McAdams, H. H., & Arkin, A. (1997). Stochastic mechanisms in gene expression. *Proc. Natl. Acad. Sci. U.S.A.*, *94*, 814–9.

McQuarrie, D. A. (1967). Stochastic approach to chemical kinetics. *J. Appl. Probab.*, *4*, 413–78.

Meinhardt, H. (1999). Orientation of chemotactic cells and growth cones: models and mechanisms. *J. Cell Sci.*, *112*, 2867–74.

Messerli, M. A., Créton, R., Jaffe, L. F., & Robinson, K. R. (2000). Periodic increases in elongation rate precede increases in cytosolic Ca2+ during pollen tube growth. *Dev. Biol.*, *222*, 84–98.

Meyers, J., Craig, J., & Odde, D. J. (2006). Potential for control of signaling pathways via cell size and shape. *Curr. Biol.*, *16*, 1685–93.

Miller, D. D., Callaham, D. A., Gross, D. J., & Hepler, P. K. (1992). Free Ca2+ gradient in growing pollen tubes of Lilium. *J. Cell Sci.*, *101*, 7–12.

Mincheva, M., & Craciun, G. (2008). Multigraph conditions for multistability, oscillations and pattern formation in biochemical reaction networks. *Proc. IEEE*, *96*, 1281–91.

Mincheva, M., & Roussel, M. R. (2007). Graph-theoretic methods for the analysis of chemical and biochemical networks. I. *J. Math. Biol.*, *55*, 61–86.

Mogilner, A., & Keren, K. (2009). The shape of motile cells. *Curr. Biol.*, *19*, 762–71.

Mori, Y., Jilkine, A., & Edelstein-Keshet, L. (2008). Wave-pinning and cell polarity from a bistable reaction-diffusion system. *Biophys. J.*, *94*, 3684–97.

Mori, Y., Jilkine, A., & Edelstein-Keshet, L. (2011). Asymptotic and bifurcation analysis of wave-pinning in a reaction-diffusion model for cell polarization. *SIAM J. Appl. Math.*, *71*, 1401–27.

Narang, A. (2006). Spontaneous polarization in eukaryotic gradient sensing: a mathematical model based on mutual inhibition of frontness and backness pathways. *J. Theor. Biol.*, *240*, 538–53.

Nishikawa, S.-I., Zinkl, G. M., Swanson, R. J., Maruyama, D., & Preuss, D. (2005). Callose (beta-1,3 glucan) is essential for Arabidopsis pollen wall patterning, but not tube growth. *BMC Plant Biol.*, *5*, 22.

Orth, J. D., Thiele, I., & Palsson, B. Ø. (2010). What is flux balance analysis? *Nature Biotechnol.*, *28*, 245–8.

Palin, R., & Geitmann, A. (2012). The role of pectin in plant morphogenesis. *Biosystems*, *109*, 397–402.

Paulsson, J., Berg, O. G., & Ehrenberg, M. (2000). Stochastic focusing: fluctuation-enhanced sensitivity of intracellular regulation. *Proc. Natl. Acad. Sci. U.S.A.*, *97*, 7148–53.

Postma, M., Bosgraaf, L., M., Loovers H., & van Haastert, P. J. M. (2004). Chemotaxis: signalling modules join hands at front and tail. *EMBO Rep.*, *5*, 35–40.

Postma, M., & van Haastert, P. J. M. (2001). A diffusion-translocation model for gradient sensing by chemotactic cells. *Biophys. J.*, *81*, 1314–23.

Qin, P., Ting, D., Shieh, A., & McCormick, S. (2012). Callose plug deposition patterns vary in pollen tubes of Arabidopsis thaliana ecotypes and tomato species. *BMC Plant Biol.*, *12*, 178.

Raftopoulou, M., & Hall, A. (2004). Cell migration: Rho GTPases lead the way. *Dev. Biol.*, *265*, 23–32.

Rao, C. V., Wolf, D. M., & Arkin, A. P. (2002). Control, exploitation and tolerance of intracellular noise. *Nature*, *420*, 231–7.

Richards, F. J. (1959). A flexible growth function for empirical use. *J. Exp. Biol.*, *10*, 290–300.

Ridley, A. J. (2006). Rho GTPases and actin dynamics in membrane protrusions and vesicle trafficking. *Trends Cell. Biol.*, *16*, 522–9.

Ridley, A. J., Schwartz, M. A., Burridge, K., Firtel, R. A., Ginsberg, M. H., Borisy, G., Parsons, J. T., & Horwitz, A. R. (2003). Cell migration: integrating signals from front to back. *Science*, *302*, 1704–9.

Rojas, E. R., Hotton, S., & Dumais, J. (2011). Chemically mediated mechanical expansion of the pollen tube cell wall. *Biophys. J.*, *101*, 1844–53.

Rounds, C. M., & Bezanilla, M. (2013). Growth mechanisms in tip-growing plant cells. *Annu. Rev. Plant Biol.*, *64*, 243–65.

Schnell, S., & Turner, T. E. (2004). Reaction kinetics in intracellular environments with macromolecular crowding: simulations and rate laws. *Prog. Biophys. Mol. Biol.*, *85*, 235–60.

Semplice, M., Veglio, A., Naldi, G., Serini, G., & Gamba, A. (2012). A bistable model of cell polarity. *PLOS One*, *7*, e30977.

Slaughter, B. D., Das, A., Schwartz, J. W., Rubinstein, B., & Li, R. (2009). Dual modes of cdc42 recycling fine-tune polarized morphogenesis. *Dev. Cell*, *17*, 823–35.

Smith, H. (2011). *An Introduction to Delay Differential Equations with Applications to the Life Sciences*, vol. 57 of *Texts in Applied Mathematics*. Springer New York.

Soranzo, N., & Altafini, C. (2009). ERNEST: a toolbox for chemical reaction network theory. *Bioinformatics*, *25*, 2853–4.

Strychalski, W., Adalsteinsson, D., & Elston, T. C. (2010). Simulating Biochemical Signaling Networks in Complex Moving Geometries. *SIAM J. Sci. Comput.*, *32*, 3039–70.

Turing, A. M. (1952). The Chemical Basis of Morphogenesis. *Philos. Trans. R. Soc. B Biol. Sci.*, *237*, 37–72.

Tyson, J. J. (1975). Classification of instabilities in chemical reaction systems. *J. Chem. Phys.*, *62*, 1010–15.

Volpert, A., & Ivanova, A. (1987). Mathematical modeling (in Russian).

Walther, G. R., Hartley, M., & Mincheva, M. (2014). GraTeLPy: graph-theoretic linear stability analysis. *BMC Syst. Biol.*, *8*, 22.

Walther, G. R., Marée, A. F. M., Edelstein-Keshet, L., & Grieneisen, V. A. (2012). Deterministic versus stochastic cell polarisation through wave-pinning. *Bull. Math. Biol.*, *74*, 2570–99.

Yalovsky, S., Bloch, D., Sorek, N., & Kost, B. (2008). Regulation of membrane trafficking, cytoskeleton dynamics, and cell polarity by ROP/RAC GTPases. *Plant Physiol.*, *147*, 1527–43.

Yan, A., Xu, G., & Yang, Z. (2009). Calcium participates in feedback regulation of the oscillating ROP1 Rho GTPase in pollen tubes. *Proc. Natl. Acad. Sci. U. S. A.*, *106*, 22002–7.

Yang, Z. (2008). Cell polarity signaling in Arabidopsis. *Annu. Rev. Cell Dev. Biol.*, *24*, 551–75.

Zhang, D., Wengier, D., Shuai, B., Gui, C.-P., Muschietti, J., McCormick, S., & Tang, W.-H. (2008). The pollen receptor kinase LePRK2 mediates growth-promoting signals and positively regulates pollen germination and tube growth. *Plant Physiol.*, *148*, 1368–79.

Zhang, Y., Abu-Khzam, F. N., Baldwin, N. E., Chesler, E. J., Langston, M. A., & Samatova, N. F. (2005). Genome-scale computational approaches to memory-intensive applications in systems biology. In *Supercomputing, 2005. Proceedings of the ACM/IEEE SC 2005 Conference*, p. 12.

van Zon, J. S., & ten Wolde, P. R. (2005). Green's-function reaction dynamics: a particle-based approach for simulating biochemical networks in time and space. *J. Chem. Phys.*, *123*, 234910.